(19) Евразийское патентное ведомство

(21) 202192547 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.02.01

(22) Дата подачи заявки
2020.03.19

(51) Int. Cl. ***C12N 15/11*** (2006.01)
C12N 9/12 (2006.01)
C12N 9/22 (2006.01)
C12N 15/10 (2006.01)

(54) СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ РЕДАКТИРОВАНИЯ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

(31) **62/820,813; 62/858,958; 62/889,996; 62/922,654; 62/913,553; 62/973,558; 62/931,195; 62/974,537; 62/944,231; 62/991,069; 63/100,548**
(32) **2019.03.19; 2019.06.07; 2019.08.21; 2019.08.21; 2019.10.10; 2019.10.10; 2019.11.05; 2019.12.05; 2019.12.05; 2020.03.17; 2020.03.17**
(33) **US**
(86) **PCT/US2020/023713**
(87) **WO 2020/191234 2020.09.24**

(71) Заявитель:
ЗЕ БРОУД ИНСТИТЬЮТ, ИНК.; ПРЕЗИДЕНТ ЭНД ФЕЛЛОУЗ ОФ ГАРВАРД КОЛЛЕДЖ (US)

(72) Изобретатель:
Лиу Дэвид Р., Анзалоун Эндрю Вито, Рэндольф Пейтон, Нельсон Джеймс (US)

(74) Представитель:
Гизатуллина Е.М., Угрюмов В.М., Христофоров А.А., Строкова О.В., Гизатуллин Ш.Ф., Костюшенкова М.Ю., Парамонова К.В., Прищепный С.В., Джермакян Р.В. (RU)

(57) Описаны композиции и способы для осуществления праймерного редактирования молекулы-мишени ДНК (например генома), которые обеспечивают встраивание нуклеотидного изменения и/или осуществление нацеленного мутагенеза. Нуклеотидное изменение может включать изменение одного нуклеотида (например любую транзицию или трансверсию), вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов. Более конкретно, описаны слитые белки, включающие программируемые нуклеиновой кислотой ДНК-связывающие белки ("napDNAbp") и полимеразу (например обратную транскриптазу), которые направляются к определенной последовательности ДНК с помощью модифицированной направляющей РНК, обозначенной как PEgRNA. PEgRNA модифицирована (относительно стандартной направляющей РНК) для включения удлиненной части, которая включает последовательность матрицы для синтеза ДНК, которая кодирует одноцепочечный ДНК-флэп, который является гомологичным цепи являющейся мишенью подвергаемой редактированию эндогенной последовательности ДНК, но который содержит требуемое одно или несколько нуклеотидных изменений и который за счет синтеза на основе полимеразы (например обратной транскриптазы) становится встроенным в молекулу-мишень ДНК. Также описаны различные способы, включающие использование праймерного редактирования, включая лечение заболеваний, связанных с экспансией тринуклеотидных повторов, внесение нацеленных пептидных меток, лечение прионного заболевания с помощью внесения защитных мутаций, манипулирование РНК-кодирующими генами для внесения РНК-меток для управления функцией и экспрессией РНК, использование праймерного редактирования для конструирования сложных библиотек генов, использование праймерного редактирования для вставки иммуноэпитопов в белки, использование праймерного редактирования для внесения индуцируемых доменов димеризации в белки-мишени, и способы доставки, помимо прочего.

202192547 A1

202192547 A1

202192547 A1

202192547 A1

СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ РЕДАКТИРОВАНИЯ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

ГОСУДАРСТВЕННАЯ ПОДДЕРЖКА

[0001] Данное изобретение было осуществлено при государственной поддержке в соответствии с грантами U01AI142756, RM1HG009490, R01EB022376, и R35GM118062, выданными Национальными институтами здравоохранения США. Государство обладает определенными правами на данное изобретение.

РОДСТВЕННЫЕ ЗАЯВКИ И УКАЗАНИЕ В КАЧЕСТВЕ ССЫЛКИ

[0002] Данная предварительная заявка на патент США ссылается и приводит в качестве ссылки следующие заявки, а именно: предварительную заявку на патент США 62/820,813, поданную 19 марта 2019 г. (внутренний номер дела поверенного B1195.70074US00), предварительную заявку на патент США 62/858,958 (внутренний номер дела поверенного B1195.70074US01), поданную 7 июня 2019 г., предварительную заявку на патент США 62/889,996 (внутренний номер дела поверенного B1195.70074US02), поданную 21 августа 2019 г., предварительную заявку на патент США 62/922,654, поданную 21 августа 2019 г. (внутренний номер дела поверенного B1195.70083US00), предварительную заявку на патент США 62/913,553 (внутренний номер дела поверенного B1195.70074US03), поданную 10 октября 2019 г., предварительную заявку на патент США 62/973,558 (внутренний номер дела поверенного B1195.70083US01), поданную 10 октября 2019 г., предварительную заявку на патент США 62/931,195 (внутренний номер дела поверенного B1195.70074US04), поданную 5 ноября 2019 г., предварительную заявку на патент США 62/944,231 (внутренний номер дела поверенного B1195.70074US05), поданную 5 декабря 2019 г., предварительную заявку на патент США 62/974,537 (внутренний номер дела поверенного B1195.70083US02), поданную 5 декабря 2019 г., предварительную заявку на патент США 62/991,069 (внутренний номер дела поверенного B1195.70074US06), поданную 17 марта 2020 г., и предварительную заявку на патент США (номер не доступен на момент подачи данной заявки) (внутренний номер дела поверенного B1195.70083US03), поданную 17 марта 2020 г.

УРОВЕНЬ ТЕХНИКИ ИЗОБРЕТЕНИЯ

[0003] В соответствии с некоторыми оценками, патогенные одиночные нуклеотидные мутации связаны приблизительно с 50% заболеваний человека, при которых

присутствует генетический компонент[7]. К сожалению, варианты лечения для пациентов с данными генетическими заболеваниями остаются крайне ограниченными, несмотря на десятилетия исследований в области генной терапии[8]. Вероятно наиболее разумным подходом для решения данной терапевтической проблемы является прямое корректирование одиночных нуклеотидных мутаций в геномах пациентов, которое будет направлено на первопричину заболевания и вероятно обеспечит долговременную пользу. Несмотря на то, что подобная стратегия ранее являлась немыслимой, недавние достижения в плане возможностей редактирования генома, принесенные разработкой системы CRISPR/Cas[9], сделали данный терапевтический подход возможным. Путем прямого конструирования последовательности направляющей РНК (нРНК), которая содержит ~20 нуклеотидов, являющихся комплементарными последовательности-мишени ДНК, связанные с CRISPR (Cas) нуклеазы могут получать доступ практически к любому необходимому участку генома[1,2]. На сегодняшний день были идентифицированы и адаптированы для применения в рамках геномного редактирования несколько систем на основе мономерной бактериальной нуклеазы Cas[10]. Данное естественное разнообразие нуклеаз Cas, наряду с растущим набором модифицированных вариантов[11-14], обеспечивает благоприятную почву для разработки новых технологий геномного редактирования.

[0004] В то время как манипулирование генами с помощью CRISPR теперь является зрелым способом, точное редактирование одиночных пар оснований в геноме человека остается серьезной проблемой[3]. Направленная гомологичная репарация (HDR) долго применялась в клетках человека и других организмов для вставки, коррекции, или замены ДНК-последовательностей в участках двуцепочечных разрывов (DSB) с помощью донорских матриц для репарации ДНК, которые кодируют необходимые изменения[15]. Однако традиционная HDR имеет очень низкую эффективность в большинстве типов клеток человека, в частности в неделящихся клетках, при этом конкурентное негомологичное соединение концов (NHEJ) в основном приводит к побочным продуктам в виде вставок-делеций (инделям)[16]. Другие проблемы связаны с осуществлением DSB, которые могут вызывать крупные хромосомные перестроения и делеции в локусах-мишенях[17], или активировать ось p53, что ведет к остановке роста и апоптозу[18,19].

[0005] Были исследованы некоторые подходы для устранения данных недостатков HDR. Например, было показано, что репарация одноцепочечных разрывов ДНК

олигонуклеотидными донорами снижает образование инделей, однако выходы желаемых продуктов репарации остаются низкими[20]. В рамках других стратегий осуществляются попытки направить репарацию к HDR вместо NHEJ с помощью низкомолекулярных и биологических реагентов[21-23]. Однако эффективность данных способов вероятно зависит от типа клеток, при этом нарушение нормального состояния клеток может вести к нежелательным и непредсказуемым эффектам.

[0006] Недавно исследователи, возглавляемые профессором David Liu et al., разработали редактирование оснований в качестве технологии, которая обеспечивает редактирование нуклеотидов-мишеней без создания DSB или использования HDR[4–6,24–27]. Прямое модифицирование оснований ДНК с помощью слитых с Cas ферментов в виде дезаминазы обеспечивает C•G на T•A или A•T на G•C конверсии пар оснований в рамках короткого целевого окна (~5-7 оснований) с очень высокой эффективностью. В результате этого редакторы оснований стали быстро использоваться научным сообществом. Однако следующие факторы ограничивают их общую применимость в плане точного геномного редактирования: (1) наблюдается "побочное редактирование" не являющихся мишенями C или A оснований в рамах целевого окна; (2) наблюдаются смеси продуктов в виде нуклеотидов-мишеней; (3) основания-мишени должны быть расположены на 15±2 нуклеотида выше от последовательности PAM; и (5) репарация малых мутаций в виде вставки и делеции является невозможной.

[0007] Таким образом, разработка программируемых редакторов, которые гибким образом способны вносить любое необходимое однонуклеотидное изменение и/или которые способны вносить вставки или делеции пар оснований (например, по меньшей мере, 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100 или более вставок или делеций пар оснований) и/или которые могут изменять или модифицировать нуклеотидную последовательность в участке-мишени с высокой специфичностью и эффективностью, значительно расширит диапазон и терапевтический потенциал технологий геномного редактирования на основе CRISPR.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[0008] В рамках настоящего изобретения описана полностью новая платформа геномного редактирования, называемая "праймерным редактированием". Праймерное редактирование представляет собой универсальный и точный способ геномного редактирования, в рамках которого осуществляется прямая запись новой генетической информации в определенный участок ДНК с помощью программируемого

нуклеиновой кислотой ДНК-связывающего белка ("napDNAbp"), работающего совместно с полимеразой (то есть в форме слитого белка или иным образом поступающего *in trans* с napDNAbp), при этом система праймерного редактирования программируется направляющей РНК ("PEgRNA") для праймерного редактирования (PE), которая как указывает участок-мишень, так и служит матрицей для синтеза требуемого изменения в форме замещающей цепи ДНК в виде удлинения (ДНК или РНК), встроенного в направляющую РНК (например на 5′ или 3′-конце, или во внутренней части направляющей РНК). Замещающая цепь, содержащая требуемое изменение (например замещение одного азотистого основания), имеет ту же последовательность, что и эндогенная цепь подвергаемого редактированию участка-мишени (за исключением того, что она включает требуемое изменение). С помощью механизма репарации и/или репликации ДНК эндогенная цепь участка-мишени замещается новой синтезированной замещающей цепью, содержащей требуемое изменение. В некоторых случаях праймерное редактирование может быть описано как технология геномного редактирования типа "поиск-и-замена", поскольку праймерные редакторы, как описано здесь, не только обеспечивают поиск и нахождение требуемого участка-мишени для редактирования, но в то же время кодируют замещающую цепь, содержащую требуемое изменение, которая встраивается вместо соответствующей эндогенной цепи ДНК участка-мишени.

[0009] Праймерные редакторы по настоящему изобретению частично связаны с открытием того, что механизм мишень-праймированной обратной транскрипции (TPRT) или "праймерное редактирование" может быть использован или адаптирован для осуществления точного геномного редактирования на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью (например как показано в рамках различных вариантов осуществления на Фиг.1A-1F). TPRT естественным образом применяется в рамках подвижных элементов ДНК, например ретротранспозонов млекопитающих без длинных концевых повторов (LTR) и бактериальных интронов Группы II[28,29]. Здесь авторы использовали продукты слияния белок Cas-обратная транскриптаза или родственные системы для нацеливания на определенную последовательность ДНК с помощью направляющей РНК, осуществления одноцепочечного разрыва в участке-мишени, и применения никированной ДНК в качестве праймера для обратной транскрипции сконструированной матрицы в виде обратной транскриптазы, которая встраивается в направляющую РНК. Однако в то

время как данная концепция начинается с праймерных редакторов, в рамках которых используются обратные транскриптазы в качестве компонента ДНК-полимеразы, описанные здесь праймерные редакторы не ограничиваются обратными транскриптазами, но могут включать использование практически любой ДНК-полимеразы. Действительно, в то время как по тексту заявки могут указываться праймерные редакторы с "обратными транскриптазами", здесь устанавливается, что обратные транскриптазы являются лишь одним из типов ДНК-полимераз, которые могут работать в рамках праймерного редактирования. Таким образом, при указании в данном описании "обратных транскриптаз" специалист в данной области должен понимать, что вместо обратной транскриптазы могут использоваться любые подходящие ДНК-полимеразы. Таким образом, в одном аспекте праймерные редакторы могут включать Cas9 (или эквивалентный napDNAbp), который программируется для нацеливания на ДНК-последовательность путем его связывания со специализированной направляющей РНК (то есть PEgRNA), содержащей спейсерную последовательность, которая отжигается на комплементарном протоспейсере в ДНК-мишени. Специализированная направляющая РНК также содержит новую генетическую информацию в форме удлинения, которое кодирует замещающую цепь ДНК, содержащую требуемое генетическое изменение, которая используется для замены соответствующей эндогенной цепи ДНК в участке-мишени. Для переноса информации из PEgRNA в ДНК-мишень механизм праймерного редактирования включает никирование участка-мишени в одной цепи ДНК для высвобождения 3′-гидроксильной группы. Затем свободная 3′-гидроксильная группа может использоваться для праймирования полимеризации ДНК кодирующего изменение удлинения на PEgRNA напрямую в участок-мишень. В различных вариантах осуществления удлинение, которое обеспечивает матрицу для полимеризации замещающей содержащей изменение цепи, может быть образовано из РНК или ДНК. В случае РНК удлинения полимераза праймерного редактора может представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). В случае ДНК удлинения полимераза праймерного редактора может представлять собой ДНК-зависимую ДНК-полимеразу.

[0010] Новая синтезированная цепь (то есть замещающая цепь ДНК, содержащая требуемое изменение), которая образуется описанными здесь праймерными редакторами, будет являться гомологичной относительно последовательности

геномной мишени (то есть иметь такую же последовательность) за исключением включения требуемого нуклеотидного изменения (например изменения отдельного нуклеотида, делеции, или вставки, или их комбинации). Новая синтезированная (или замещающая) цепь ДНК также может быть обозначена как одноцепочечный ДНК-флэп, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. В определенных вариантах осуществления система может комбинироваться с использованием фермента в виде подверженной ошибкам обратной транскриптазы (например представленной в виде слитого с Cas9 доменом белка или представленной с Cas9 доменом *in trans*). Фермент в виде подверженной ошибкам обратной транскриптазы может вносить изменения при синтезе одноцепочечного ДНК-флэпа. Таким образом, в определенных вариантах осуществления подверженная ошибкам обратная транскриптаза может использоваться для внесения нуклеотидных изменений в ДНК-мишень. В зависимости от подверженной ошибкам обратной транскриптазы, которая используется в системе, изменения могут являться случайными или неслучайными.

[0011] Коррекция гибридизированного промежуточного продукта (включающего одноцепочечный ДНК-флэп, синтезированный обратной транскриптазой, гибридизированной с эндогенной цепью ДНК) может включать удаление образованного вытесненного флэпа эндогенной ДНК (например с помощью эндонуклеазы 5′-концевого ДНК-флэпа, FEN1), лигирование синтезированного одноцепочечного ДНК-флэпа с ДНК-мишенью, и ассимилирование требуемого нуклеотидного изменения в результате процессов репарации и/или репликации клеточной ДНК. Поскольку синтез ДНК на основе матриц обеспечивает точность на уровне отдельных нуклеотидов при модификации любого нуклеотида, включая вставки и делеции, объем данного подхода является очень широким и, как представляется, он может использоваться в рамках множества вариантов применения в фундаментальной науке и терапии.

[0012] В одном аспекте изобретение предусматривает слитый белок, включающий программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу. В различных вариантах осуществления слитый белок способен осуществлять геномное редактирование с помощью мишень-инициированной обратной транскрипции в присутствии удлиненной направляющей РНК.

[0013] В определенных вариантах осуществления napDNAbp обладает никазной активностью. napDNAbp также может представлять собой белок Cas9 или его функциональный эквивалент, например имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[0014] В определенных вариантах осуществления napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[0015] В других вариантах осуществления слитый белок при образовании комплекса с удлиненной направляющей РНК способен связываться с последовательностью-мишенью ДНК.

[0016] В других вариантах осуществления последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[0017] В других вариантах осуществления связывание слитого белка в виде комплекса с удлиненной направляющей РНК образует R-петлю. R-петля может включать (i) РНК-ДНК гибрид, включающий удлиненную направляющую РНК и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

[0018] В других вариантах осуществления комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности обратной транскриптазы, имеющей свободный 3′-конец.

[0019] В различных вариантах осуществления удлиненная направляющая РНК включает (a) направляющую РНК и (b) удлинение РНК на 5′ или 3′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК. Удлинение РНК может включать (i) последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера. В различных вариантах осуществления последовательность матрицы обратной транскрипции может кодировать одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, примыкающей к участку разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[0020] В различных вариантах осуществления удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по

меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[0021] В других вариантах осуществления одноцепочечный ДНК-флэп может гибридизироваться с последовательностью эндогенной ДНК, примыкающей к участку разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения. В других вариантах осуществления одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, примыкающую к участку разрыва и имеющую свободный 5′-конец. В определенных вариантах осуществления вытесненная эндогенная ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

[0022] В различных вариантах осуществления клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[0023] В различных других вариантах осуществления требуемое нуклеотидное изменение вносится в окне редактирования, которое соответствует от -4 до +10 относительно последовательности PAM.

[0024] В других вариантах осуществления требуемое нуклеотидное изменение вносится в окне редактирования, которое соответствует приблизительно от -5 до +5 относительно участка разрыва, приблизительно от -10 до +10 относительно участка разрыва, приблизительно от -20 до +20 относительно участка разрыва, приблизительно от -30 до +30 относительно участка разрыва, приблизительно от -40 до +40 относительно участка разрыва, приблизительно от -50 до +50 относительно участка разрыва, приблизительно от -60 до +60 относительно участка разрыва, приблизительно от -70 до +70 относительно участка разрыва, приблизительно от -80 до +80 относительно участка разрыва, приблизительно от -90 до +90 относительно участка разрыва, приблизительно от -100 до +100 относительно участка разрыва, приблизительно от -200 до +200 относительно участка разрыва.

[0025] В различных вариантах осуществления napDNAbp включает аминокислотную последовательность SEQ ID NO: 18. В различных других вариантах осуществления

napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 26-39, 42- 61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18-25 и 87-88 (SpCas9); и SEQ ID NO: 62-72(Cas12).

[0026] В других вариантах осуществления обратная транскриптаза описанных слитых белков и/или композиций может включать любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, и 766. В других вариантах осуществления обратная транскриптаза может включать аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентичная аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, и 766. Данные последовательности могут являться встречающимися в природе последовательностями обратной транскриптазы, например из ретровируса или ретротранспозона, или данные последовательности могут являться рекомбинантными.

[0027] В различных других вариантах осуществления описанные здесь слитые белки могут включать различные структурные конфигурации. Например, слитые белки могут включать структуру NH_2-[napDNAbp]-[обратная транскриптаза]-COOH; или NH_2-[обратная транскриптаза]-[napDNAbp]-COOH, при этом каждый экземпляр "]-[" обозначает наличие необязательной последовательности линкера.

[0028] В различных вариантах осуществления последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769, или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, или 90%, или 95%, или 99% идентична любой одной аминокислотной последовательности линкера SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[0029] В различных вариантах осуществления требуемое нуклеотидное изменение, которое вносится в ДНК-мишень, может представлять собой изменение одного нуклеотида (например транзицию или трансверсию), вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

[0030] В определенных случаях вставка имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей

мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 30, по меньшей мере, 40, по меньшей мере, 50, по меньшей мере, 60, по меньшей мере, 70, по меньшей мере, 80, по меньшей мере, 90, по меньшей мере, 100, по меньшей мере, 200, по меньшей мере, 300, по меньшей мере, 400, или, по меньшей мере, 500 нуклеотидов.

[0031] В определенных случаях делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 30, по меньшей мере, 40, по меньшей мере, 50, по меньшей мере, 60, по меньшей мере, 70, по меньшей мере, 80, по меньшей мере, 90, по меньшей мере, 100, по меньшей мере, 200, по меньшей мере, 300, по меньшей мере, 400, или, по меньшей мере, 500 нуклеотидов.

[0032] В другом аспекте настоящее изобретение предусматривает удлиненную направляющую РНК, включающую направляющую РНК и, по меньшей мере, одно удлинение РНК. Удлинение РНК может быть расположено на 3′-конце направляющей РНК. В других вариантах осуществления удлинение РНК может быть расположено на 5′-конце направляющей РНК. В других вариантах осуществления удлинение РНК может быть расположено во внутримолекулярном положении в направляющей РНК, однако, предпочтительно, внутримолекулярное положение удлиненной части не нарушает функционирование протоспейсера.

[0033] В различных вариантах осуществления удлиненная направляющая РНК способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК. Последовательность-мишень ДНК может включать цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[0034] В различных вариантах осуществления удлиненной направляющей РНК указанное, по меньшей мере, одно удлинение РНК может включать последовательность матрицы обратной транскрипции. В различных других вариантах осуществления удлинение РНК может дополнительно включать участок связывания праймера обратной транскрипции. В других вариантах осуществления удлинение РНК

может включать линкерную или спейсерную последовательность, которая присоединяет удлинение РНК к направляющей РНК.

[0035] В различных вариантах осуществления удлинение РНК может иметь длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 150 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0036] В других вариантах осуществления последовательность матрицы обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0037] В других вариантах осуществления последовательность участка связывания праймера обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0038] В других вариантах осуществления необязательная линкерная или спейсерная последовательность имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0039] В различных вариантах осуществления удлиненных направляющих РНК последовательность матрицы обратной транскрипции может кодировать одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, примыкающей к участку разрыва, при этом одноцепочечный ДНК-флэп

включает требуемое нуклеотидное изменение. Одноцепочечный ДНК-флэп может вытеснять эндогенную одноцепочечную ДНК в участке разрыва. Вытесненная эндогенная одноцепочечная ДНК в участке разрыва может иметь 5′-конец и образовывать эндогенный флэп, который может быть подвергнут эксцизии клеткой. В различных вариантах осуществления эксцизия 5′-концевого эндогенного флэпа может обеспечивать стимулирование образования продукта, поскольку удаление 5′-концевого эндогенного флэпа стимулирует гибридизирование одноцепочечного 3′ ДНК-флэпа с соответствующей комплементарной цепью ДНК и встраивание или ассимилирование требуемого нуклеотидного изменения, которое несет одноцепочечный 3′ ДНК-флэп, в ДНК-мишени.

[0040] В различных вариантах осуществления удлиненных направляющих РНК клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[0041] В определенных вариантах осуществления PEgRNA включает нуклеотидную последовательность SEQ ID NO: 131 , 222 , 394, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 641, 642, 643, 644, 645, 646, 647, 648, 649, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 738, 2997, 2998, 2999, 3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3305, 3306, 3307, 3308, 3309, 3310, 3311, 3312, 3313, 3314, 3315, 3316, 3317, 3318, 3319, 3320, 3321, 3322, 3323, 3324, 3325, 3326, 3327, 3328, 3329, 3330, 3331, 3332, 3333, 3334, 3335, 3336, 3337, 3338, 3339, 3340, 3341, 3342, 3343, 3344, 3345, 3346, 3347, 3348, 3349, 3350, 3351, 3352, 3353, 3354, 3355, 3356, 3357, 3358, 3359, 3360, 3361, 3362, 3363, 3364, 3365, 3366, 3367, 3368, 3369, 3370, 3371, 3372, 3373, 3374, 3375, 3376, 3377, 3378, 3379, 3380, 3381, 3382, 3383, 3384, 3385, 3386, 3387, 3388, 3389, 3390, 3391, 3392, 3393, 3394, 3395, 3396, 3397, 3398, 3399, 3400, 3401, 3402, 3403, 3404, 3405, 3406, 3407, 3408, 3409, 3410, 3411, 3412, 3413, 3414, 3415, 3416, 3417, 3418, 3419, 3420, 3421, 3422, 3423, 3424, 3425, 3426, 3427, 3428, 3429, 3430, 3431, 3432,

3433, 3434, 3435, 3436, 3437, 3438, 3439, 3440, 3441, 3442, 3443, 3444, 3445, 3446, 3447, 3448, 3449, 3450, 3451, 3452, 3453, 3454, 3455, 3479, 3480, 3481, 3482, 3483, 3484, 3485, 3486, 3487, 3488, 3489, 3490, 3491, 3492, 3493, 3522, 3523, 3524, 3525, 3526, 3527, 3528, 3529, 3530, 3531, 3532, 3533, 3534, 3535, 3536, 3537, 3538, 3539, 3540, 3549, 3550, 3551, 3552, 3553, 3554, 3555, 3556, 3628, 3629, 3630, 3631, 3632, 3633, 3634, 3635, 3636, 3637, 3638, 3639, 3640, 3641, 3642, 3643, 3644, 3645, 3646, 3647, 3648, 3649, 3650, 3651, 3652, 3653, 3654, 3655, 3656, 3657, 3658, 3659, 3660, 3661, 3662, 3663, 3664, 3665, 3666, 3667, 3668, 3669, 3670, 3671, 3672, 3673, 3674, 3675, 3676, 3677, 3678, 3679, 3680, 3681, 3682, 3683, 3684, 3685, 3686, 3687, 3688, 3689, 3690, 3691, 3692, 3693, 3694, 3695, 3696, 3697, 3698, 3755, 3756, 3757, 3758, 3759, 3760, 3761, 3762, 3763, 3764, 3765, 3766, 3767, 3768, 3769, 3770, 3771, 3772, 3773, 3774, 3775, 3776, 3777, 3778, 3779, 3780, 3781, 3782, 3783, 3784, 3785, 3786, 3787, 3788, 3789, 3790, 3791, 3792, 3793, 3794, 3795, 3796, 3797, 3798, 3799, 3800, 3801, 3802, 3803, 3804, 3805, 3806, 3807, 3808, 3809, и 3810 , или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 131 , 222 , 394, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 641, 642, 643, 644, 645, 646, 647, 648, 649, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 738, 2997, 2998, 2999, 3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3305, 3306, 3307, 3308, 3309, 3310, 3311, 3312, 3313, 3314, 3315, 3316, 3317, 3318, 3319, 3320, 3321, 3322, 3323, 3324, 3325, 3326, 3327, 3328, 3329, 3330, 3331, 3332, 3333, 3334, 3335, 3336, 3337, 3338, 3339, 3340, 3341, 3342, 3343, 3344, 3345, 3346, 3347, 3348, 3349, 3350, 3351, 3352, 3353, 3354, 3355, 3356, 3357, 3358, 3359, 3360, 3361, 3362, 3363, 3364, 3365, 3366, 3367, 3368, 3369, 3370, 3371, 3372, 3373, 3374, 3375, 3376, 3377, 3378, 3379, 3380, 3381, 3382, 3383, 3384, 3385, 3386, 3387, 3388, 3389, 3390, 3391, 3392, 3393, 3394, 3395, 3396, 3397, 3398, 3399, 3400, 3401, 3402, 3403, 3404, 3405,

3406, 3407, 3408, 3409, 3410, 3411, 3412, 3413, 3414, 3415, 3416, 3417, 3418, 3419, 3420, 3421, 3422, 3423, 3424, 3425, 3426, 3427, 3428, 3429, 3430, 3431, 3432, 3433, 3434, 3435, 3436, 3437, 3438, 3439, 3440, 3441, 3442, 3443, 3444, 3445, 3446, 3447, 3448, 3449, 3450, 3451, 3452, 3453, 3454, 3455, 3479, 3480, 3481, 3482, 3483, 3484, 3485, 3486, 3487, 3488, 3489, 3490, 3491, 3492, 3493, 3522, 3523, 3524, 3525, 3526, 3527, 3528, 3529, 3530, 3531, 3532, 3533, 3534, 3535, 3536, 3537, 3538, 3539, 3540, 3549, 3550, 3551, 3552, 3553, 3554, 3555, 3556, 3628, 3629, 3630, 3631, 3632, 3633, 3634, 3635, 3636, 3637, 3638, 3639, 3640, 3641, 3642, 3643, 3644, 3645, 3646, 3647, 3648, 3649, 3650, 3651, 3652, 3653, 3654, 3655, 3656, 3657, 3658, 3659, 3660, 3661, 3662, 3663, 3664, 3665, 3666, 3667, 3668, 3669, 3670, 3671, 3672, 3673, 3674, 3675, 3676, 3677, 3678, 3679, 3680, 3681, 3682, 3683, 3684, 3685, 3686, 3687, 3688, 3689, 3690, 3691, 3692, 3693, 3694, 3695, 3696, 3697, 3698, 3755, 3756, 3757, 3758, 3759, 3760, 3761, 3762, 3763, 3764, 3765, 3766, 3767, 3768, 3769, 3770, 3771, 3772, 3773, 3774, 3775, 3776, 3777, 3778, 3779, 3780, 3781, 3782, 3783, 3784, 3785, 3786, 3787, 3788, 3789, 3790, 3791, 3792, 3793, 3794, 3795, 3796, 3797, 3798, 3799, 3800, 3801, 3802, 3803, 3804, 3805, 3806, 3807, 3808, 3809, и 3810.

[0042] В другом аспекте изобретения предусмотрены комплексы, включающие описанный здесь слитый белок и любую описанную выше удлиненную направляющую РНК.

[0043] В других аспектах изобретения предусмотрен комплекс, включающий napDNAbp и удлиненную направляющую РНК. napDNAbp может представлять собой никазу Cas9, или может представлять собой аминокислотную последовательность SEQ ID NO: 42-57 (никаза Cas9) и 65 (никаза AsCas12a), или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 42-57 (никаза Cas9) и 65 (никаза AsCas12a).

[0044] В различных вариантах осуществления, включающих комплекс, удлиненная направляющая РНК способна направлять napDNAbp к последовательности-мишени ДНК. В различных вариантах осуществления обратная транскриптаза может поступать *in trans*, то есть поступать из другого источника, отличного от самого комплекса. Например, обратная транскриптаза может поступать в ту же клетку, имеющую комплекс, путем введения отдельного вектора, который отдельным образом кодирует обратную транскриптазу.

[0045] В другом аспекте предусмотрены полинуклеотиды. В определенных вариантах осуществления полинуклеотиды могут кодировать любой из описанных здесь слитых белков. В определенных других вариантах осуществления полинуклеотиды могут кодировать любой из описанных здесь napDNAbp. В других вариантах осуществления полинуклеотиды могут кодировать любую из описанных здесь обратных транскриптаз. В других вариантах осуществления полинуклеотиды могут кодировать любую из описанных здесь удлиненных направляющих РНК, любую из последовательностей матрицы обратной транскрипции, или любой из праймерных участков обратной транскрипции, или любую из необязательных последовательностей линкера.

[0046] В других аспектах в рамках изобретения предусмотрены векторы, включающие описанные здесь полинуклеотиды. Таким образом, в определенных вариантах осуществления векторы включают полинуклеотиды для кодирования слитых белков, включающих napDNAbp и обратную транскриптазу. В других вариантах осуществления векторы включают полинуклеотиды, которые по отдельности кодируют napDNAbp и обратную транскриптазу. В других вариантах осуществления векторы могут включать полинуклеотиды, которые кодируют удлиненные направляющие РНК. В различных вариантах осуществления векторы могут включать один или несколько полинуклеотидов, которые кодируют napDNAbp, обратную транскриптазу, и удлиненные направляющие РНК на одних и тех же или различных векторах.

[0047] В других аспектах в рамках изобретения предусмотрены клетки, включающие слитый белок, как описано здесь, и удлиненную направляющую РНК. Клетки могут быть трансформированы векторами, включающими слитые белки, napDNAbp, обратную транскриптазу, и удлиненные направляющие РНК. Данные генетические элементы могут располагаться на одних и тех же векторах или на различных векторах.

[0048] В другом аспекте в рамках изобретения предусмотрены фармацевтические композиции. В определенных вариантах осуществления фармацевтические композиции включают один или несколько napDNAbp, слитый белок, обратную транскриптазу, и удлиненную направляющую РНК. В определенных вариантах осуществления описанный здесь слитый белок и фармацевтически приемлемый наполнитель. В других вариантах осуществления фармацевтические композиции включают любую описанную здесь удлиненную РНК и фармацевтически приемлемый наполнитель. В других вариантах осуществления фармацевтические композиции включают любую описанную здесь удлиненную РНК в комбинации с любым

описанным здесь слитым белком и фармацевтически приемлемым наполнителем. В других вариантах осуществления фармацевтические композиции включают любую полинуклеотидную последовательность, кодирующую один или несколько napDNAbp, слитый белок, обратную транскриптазу, и удлиненную направляющую РНК. В других вариантах осуществления различные описанные здесь компоненты могут быть разделены на одну или несколько фармацевтических композиций. Например, первая фармацевтическая композиция может включать слитый белок или napDNAbp, вторая фармацевтическая композиция может включать обратную транскриптазу, а третья фармацевтическая композиция может включать удлиненную направляющую РНК.

[0049] В другом аспекте настоящее изобретение предусматривает наборы. В одном варианте осуществления набор включает один или несколько полинуклеотидов, кодирующих один или несколько компонентов, включая слитый белок, napDNAbp, обратную транскриптазу, и удлиненную направляющую РНК. Наборы могут также включать векторы, клетки, и выделенные препараты полипептидов, включая любой описанный здесь слитый белок, napDNAbp, или обратную транскриптазу.

[0050] В другом аспекте настоящее изобретение предусматривает способы применения описанных композиций.

[0051] В одном варианте осуществления способы относятся к способу внесения требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК. Способ сначала включает контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и удлиненную направляющую РНК, при этом слитый белок включает napDNAbp и обратную транскриптазу, и при этом удлиненная направляющая РНК включает последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение. Далее способ включает никирование последовательности двуцепочечной ДНК в не являющейся мишенью цепи, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3′-конец. Затем способ включает гибридизацию 3`конца свободной одноцепочечной ДНК с последовательностью матрицы обратной транскрипции, тем самым обеспечивая праймирование домена обратной транскриптазы. Затем способ включает полимеризацию цепи ДНК с 3′-конца, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение. Затем способ включает замещение цепи эндогенной ДНК, примыкающей к участку

разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

[0052] В других вариантах осуществления в изобретении предусмотрен способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и направляющей РНК, которая нацеливает napDNAbp на локус-мишень, при этом направляющая РНК включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение. Далее способ включает образование свободного 3′-конца в цепи ДНК в локусе-мишени и затем гибридизацию свободного 3′-конца с последовательность RT матрицы для праймирования обратной транскрипции. Далее одноцепочечный ДНК-флэп, включающий, по меньшей мере, одно требуемое нуклеотидной изменение на основе последовательности RT матрицы, синтезируется или полимеризируется обратной транскриптазой. Наконец, указанное, по меньшей мере, одно требуемое нуклеотидное изменение встраивается в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

[0053] В других вариантах осуществления изобретение предусматривает способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение; и (c) встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[0054] В определенных вариантах осуществления этап замещения цепи эндогенной ДНК включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию

несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

[0055] В различных вариантах осуществления требуемое нуклеотидное изменение может представлять собой замещение отдельного нуклеотида (например изменение в виде транзиции или трансверсии), делецию, или вставку. Например, требуемое нуклеотидное изменение может представлять собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[0056] В других вариантах осуществления требуемое нуклеотидное изменение может обеспечивать конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[0057] В других вариантах осуществления способ обеспечивает внесение требуемого нуклеотидного изменения, которое представляет собой вставку. В определенных случаях вставка имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 30, по меньшей мере, 40, по меньшей мере, 50, по меньшей мере, 60, по меньшей мере, 70, по меньшей мере, 80, по меньшей мере, 90, по меньшей мере, 100, по меньшей мере, 200, по меньшей мере, 300, по меньшей мере, 400, или, по меньшей мере, 500 нуклеотидов.

[0058] В других вариантах осуществления способ обеспечивает внесение требуемого нуклеотидного изменения, которое представляет собой делецию. В определенных других случаях делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по

меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 30, по меньшей мере, 40, по меньшей мере, 50, по меньшей мере, 60, по меньшей мере, 70, по меньшей мере, 80, по меньшей мере, 90, по меньшей мере, 100, по меньшей мере, 200, по меньшей мере, 300, по меньшей мере, 400, или, по меньшей мере, 500 нуклеотидов.

[0059] В различных вариантах осуществления требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена. Связанный с заболеванием ген может быть связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса. В других вариантах осуществления связанный с заболеванием ген может быть связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[0060] Описанные здесь способы могут включать слитые белки, имеющие napDNAbp, который представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9. В других вариантах осуществления napDNAbp и обратная транскриптаза не кодируются в виде единого слитого белка, но могут поступать в рамках отдельных конструктов. Таким образом, в некоторых вариантах осуществления обратная транскриптаза может поступать *in trans* относительно napDNAbp (а не в виде слитого белка).

[0061] В различных вариантах осуществления, включающих способы, napDNAbp может включать аминокислотную последовательность SEQ ID NO: 26- 61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18 -25 и 87-88 (SpCas9); и SEQ ID NO: 62-72(Cas12). napDNAbp может также включать аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательность любой одной из SEQ ID NO: 26- 61, 75-76, 126, 130, 137, 141, 147,

153, 157, 445, 460, 467, и 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18 -25 и 87-88 (SpCas9); и SEQ ID NO: 62-72(Cas12).

[0062] В различных вариантах осуществления, включающих способы, обратная транскриптаза может включать любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, и 766. Обратная транскриптаза может также включать аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, и 766.

[0063] Способы могут включать применение PEgRNA, включающей нуклеотидную последовательность SEQ ID NO: 131, 222, 394, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 641, 642, 643, 644, 645, 646, 647, 648, 649, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 738, 2997, 2998, 2999, 3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3305, 3306, 3307, 3308, 3309, 3310, 3311, 3312, 3313, 3314, 3315, 3316, 3317, 3318, 3319, 3320, 3321, 3322, 3323, 3324, 3325, 3326, 3327, 3328, 3329, 3330, 3331, 3332, 3333, 3334, 3335, 3336, 3337, 3338, 3339, 3340, 3341, 3342, 3343, 3344, 3345, 3346, 3347, 3348, 3349, 3350, 3351, 3352, 3353, 3354, 3355, 3356, 3357, 3358, 3359, 3360, 3361, 3362, 3363, 3364, 3365, 3366, 3367, 3368, 3369, 3370, 3371, 3372, 3373, 3374, 3375, 3376, 3377, 3378, 3379, 3380, 3381, 3382, 3383, 3384, 3385, 3386, 3387, 3388, 3389, 3390, 3391, 3392, 3393, 3394, 3395, 3396, 3397, 3398, 3399, 3400, 3401, 3402, 3403, 3404, 3405, 3406, 3407, 3408, 3409, 3410, 3411, 3412, 3413, 3414, 3415, 3416, 3417, 3418, 3419, 3420, 3421, 3422, 3423, 3424, 3425, 3426, 3427, 3428, 3429, 3430, 3431, 3432, 3433, 3434, 3435, 3436, 3437, 3438, 3439, 3440, 3441, 3442, 3443, 3444, 3445, 3446, 3447, 3448, 3449, 3450, 3451, 3452, 3453, 3454, 3455, 3479, 3480, 3481, 3482, 3483, 3484, 3485, 3486, 3487, 3488, 3489, 3490, 3491, 3492, 3493, 3522, 3523, 3524, 3525, 3526, 3527, 3528,

3529, 3530, 3531, 3532, 3533, 3534, 3535, 3536, 3537, 3538, 3539, 3540, 3549, 3550, 3551, 3552, 3553, 3554, 3555, 3556, 3628, 3629, 3630, 3631, 3632, 3633, 3634, 3635, 3636, 3637, 3638, 3639, 3640, 3641, 3642, 3643, 3644, 3645, 3646, 3647, 3648, 3649, 3650, 3651, 3652, 3653, 3654, 3655, 3656, 3657, 3658, 3659, 3660, 3661, 3662, 3663, 3664, 3665, 3666, 3667, 3668, 3669, 3670, 3671, 3672, 3673, 3674, 3675, 3676, 3677, 3678, 3679, 3680, 3681, 3682, 3683, 3684, 3685, 3686, 3687, 3688, 3689, 3690, 3691, 3692, 3693, 3694, 3695, 3696, 3697, 3698, 3755, 3756, 3757, 3758, 3759, 3760, 3761, 3762, 3763, 3764, 3765, 3766, 3767, 3768, 3769, 3770, 3771, 3772, 3773, 3774, 3775, 3776, 3777, 3778, 3779, 3780, 3781, 3782, 3783, 3784, 3785, 3786, 3787, 3788, 3789, 3790, 3791, 3792, 3793, 3794, 3795, 3796, 3797, 3798, 3799, 3800, 3801, 3802, 3803, 3804, 3805, 3806, 3807, 3808, 3809, и 3810, или нуклеотидную последовательность, имеющую, по меньшей мере, 80%, или, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно них. Способы могут включать применение удлиненных направляющих РНК, которые включают удлинение РНК на 3′-конца, при этом удлинение РНК включает последовательность матрицы обратной транскрипции.

[0064] Способы могут включать применение удлиненных направляющих РНК, которые включают удлинение РНК на 5′-конце, при этом удлинение РНК включает последовательность матрицы обратной транскрипции.

[0065] Способы могут включать применение удлиненных направляющих РНК, которые включают удлинение РНК в внутримолекулярном положении в направляющей РНК, при этом удлинение РНК включает последовательность матрицы обратной транскрипции.

[0066] Способы могут включать применение удлиненных направляющих РНК, имеющих одно или несколько удлинений РНК длиной, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 30, по меньшей мере, 40, по меньшей мере, 50, по меньшей мере, 60, по меньшей мере, 70, по меньшей мере, 80, по меньшей мере, 90, по меньшей мере, 100, по меньшей мере, 200, по меньшей мере, 300, по меньшей мере, 400, или, по меньшей мере, 500 нуклеотидов.

[0067] Следует понимать, что вышеуказанные концепции и дополнительные концепции, которые обсуждаются ниже, могут быть сгруппированы в рамках любой подходящей комбинации, поскольку настоящее изобретение не является ограниченным в данном отношении. Дополнительно, другие преимущества и новые признаки настоящего изобретения станут очевидны на основе нижеследующего подробного описания различных неограничивающих вариантов осуществления при их рассмотрении совместно с прилагаемыми фигурами.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[0068] Нижеследующие фигуры образуют часть настоящего описания и включены для дополнительной демонстрации определенных аспектов настоящего изобретения, которое может быть лучше понятно со ссылкой на одну или несколько из данных фигур в комбинации с подробным описанием конкретных приведенных здесь вариантов осуществления.

[0069] На **Фиг.1A** приведена схема примера процесса внесения изменения одного нуклеотида, и/или вставки, и/или делеции в молекулу ДНК (например геном) с помощью слитого белка, включающего обратную транскриптазу, слитую с белком Cas9, в комплексе с молекулой удлиненной направляющей РНК. В данном варианте осуществления направляющая РНК удлиняется на 3´-конце для включения последовательности матрицы в виде обратной транскриптазы. Схема показывает, как обратная транскриптаза (RT), слитая с никазой Cas9, в комплексе с направляющей РНК (нРНК) связывается с участком-мишенью ДНК и осуществляет разрыв PAM-содержащей цепи ДНК, примыкающей к нуклеотиду-мишени. Фермент RT использует никированную ДНК в качестве праймера для синтеза ДНК из нРНК, которая используется в качестве матрицы для синтеза новой цепи ДНК, которая кодирует требуемое изменение. Показанный процесс редактирования может быть обозначен как редактирование на основе мишень-инициированной обратной транскрипции (TRT редактирование) или, равным образом, как "праймерное редактирование".

[0070] На **Фиг.1B** показано то же, что и на Фиг.1A, за исключением того, что комплекс праймерного редактора обозначен более общим образом как [napDNAbp]-[P]:PEgRNA или [P]-[napDNAbp]:PEgRNA, где "P" обозначает любую полимеразу (например обратную транскриптазу), "napDNAbp" обозначает программируемый нуклеиновой кислотой ДНК-связывающий белок (например SpCas9), а "PEgRNA" обозначает направляющую РНК для праймерного редактирования, а "]-[" обозначает

необязательный линкер. Как описано в других разделах, например на Фиг.3A-3G, PEgRNA включает 5′ плечо удлинения, включающее участок связывания праймера и матрицу для синтеза ДНК. Несмотря на то, что это не показано, подразумевается, что плечо удлинения PEgRNA (то есть которое включает участок связывания праймера и матрицу для синтеза ДНК) может представлять собой ДНК или РНК. Конкретная полимераза, предполагаемая в данной конфигурации, будет зависеть от природы матрицы для синтеза ДНК. Например, если матрица для синтеза ДНК представляет собой РНК, то полимераза будет представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). Если матрица для синтеза ДНК представляет собой ДНК, то полимераза может представлять собой ДНК-зависимую ДНК-полимеразу.

[0071] На **Фиг.1C** приведена схема примера процесса внесения изменения одного нуклеотида, и/или вставки, и/или делеции в молекулу ДНК (например геном) с помощью слитого белка, включающего обратную транскриптазу, слитую с белком Cas9, в комплексе с молекулой удлиненной направляющей РНК. В данном варианте осуществления направляющая РНК удлиняется на 5′-конце для включения последовательности матрицы в виде обратной транскриптазы. Схема показывает, как обратная транскриптаза (RT), слитая с никазой Cas9, в комплексе с направляющей РНК (нРНК) связывается с участком-мишенью ДНК и осуществляет разрыв PAM-содержащей цепи ДНК, примыкающей к нуклеотиду-мишени. Фермент RT использует никированную ДНК в качестве праймера для синтеза ДНК из нРНК, которая используется в качестве матрицы для синтеза новой цепи ДНК, которая кодирует требуемое изменение. Показанный процесс редактирования может быть обозначен как редактирование на основе мишень-инициированной обратной транскрипции (TRT редактирование) или, равным образом, как "праймерное редактирование".

[0072] На **Фиг.1D** показано то же, что и на Фиг.1C, за исключением того, что комплекс праймерного редактора обозначен более общим образом как [napDNAbp]-[P]:PEgRNA или [P]-[napDNAbp]:PEgRNA, где "P" обозначает любую полимеразу (например обратную транскриптазу), "napDNAbp" обозначает программируемый нуклеиновой кислотой ДНК-связывающий белок (например SpCas9), а "PEgRNA" обозначает направляющую РНК для праймерного редактирования, а "]-[" обозначает необязательный линкер. Как описано в других разделах, например на Фиг.3A-3G, PEgRNA включает 3′ плечо удлинения, включающее участок связывания праймера и

матрицу для синтеза ДНК. Несмотря на то, что это не показано, подразумевается, что плечо удлинения PEgRNA (то есть которое включает участок связывания праймера и матрицу для синтеза ДНК) может представлять собой ДНК или РНК. Конкретная полимераза, предполагаемая в данной конфигурации, будет зависеть от природы матрицы для синтеза ДНК. Например, если матрица для синтеза ДНК представляет собой РНК, то полимераза будет представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). Если матрица для синтеза ДНК представляет собой ДНК, то полимераза может представлять собой ДНК-зависимую ДНК-полимеразу. В различных вариантах осуществления PEgRNA может быть модифицирована или синтезирована для включения матрицы для синтеза ДНК на основе ДНК.

[0073] **Фиг.1E** представляет собой схему, на которой показан пример процесса того как синтезированная одиночная цепь ДНК (которая включает требуемое нуклеотидное изменение) подвергается коррекции, так что требуемое нуклеотидное изменение встраивается в ДНК. Как показано, после синтеза подвергнутой редактированию цепи (или "мутагенной цепи") выравнивание с эндогенной цепью, расщепление флэпа эндогенной цепи и лигирование ведет к встраиванию изменения ДНК после коррекция несовпадающего ДНК дуплекса за счет процессов репарации и/или репликации эндогенной ДНК.

[0074] **Фиг.1F** представляет собой схему, показывающую, что "никирование противоположной цепи" может быть встроено в способ коррекции в соответствии с **Фиг.1E** для стимулирования образования требуемого продукта вместо продукта, связанного с возвратом. При никировании противоположной цепи второй комплекс Cas9/нРНК используется для внесения второго разрыва в противоположной цепи относительно исходной подвергнутой никированию цепи. Это индуцирует процессы репарации и/или репликации эндогенной клеточной ДНК для предпочтительного замещения не подвергавшейся редактированию цепи (то есть цепи, содержащей второй подвергнутый разрыву участок).

[0075] На **Фиг.1G** приведена другая схема примера процесса внесения изменения одного нуклеотида, и/или вставки, и/или делеции в локус-мишень молекулы ДНК (например геном) с помощью программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) в виде комплекса с удлиненной направляющей РНК. Данный процесс может быть обозначен как вариант осуществления праймерного

редактирования. Удлиненная направляющая РНК включает удлинение на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК. На этапе (a) комплекс napDNAbp/нРНК контактирует с молекулой ДНК, при этом нРНК направляет napDNAbp для связывания с локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК (цепи R-петли, или PAM-содержащей цепи, или не являющейся мишенью цепи ДНК, или протоспейсерной цепи) локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК. На этапе (c) 3′-концевая цепь ДНК взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции. В определенных вариантах осуществления 3′-концевая цепь ДНК гибридизируется с определенной праймерной последовательностью RT на удлиненной части направляющей РНК. На этапе (d) вносится обратная транскриптаза, которая синтезирует одиночную цепь ДНК от 3′-конца праймированного участка к 3′-концу направляющей РНК. Это образует одноцепочечный ДНК-флэп, включающий требуемое нуклеотидное изменение (например изменение одного основания, вставку, или делецию, или их комбинацию). На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа, так что требуемое нуклеотидное изменение становится встроенным в локус-мишень. Данный процесс может быть смещен в сторону образования требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа, который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с комплементарной последовательностью на другой цепи. Процесс также может быть смещен в сторону образования продукта с никированием второй цепи, как показано в качестве примера на Фиг.1F. Данный процесс может обеспечивать внесение, по меньшей мере, одного или нескольких из следующих генетических изменений: трансверсий, транзиций, делеций, и вставок.

[0076] Фиг.1H представляет собой схему, на которой показаны типы генетических изменений, которые являются возможными в рамках описанных здесь процессов праймерного редактирования. Типы нуклеотидных изменений, которые могут быть достигнуты с помощью праймерного редактирования, выключают делеции (включая

коротки и длинные делеции), однонуклеотидные изменения (включая транзиции и трансверсии), инверсии, и вставки (включая короткие и длинные вставки).

[0077] Фиг.11 представляет собой схему, на которой показано временно́е никирование второй цепи на примере PE3b (PE3b = слитый белок праймерного редактора PE2 + PEgRNA + никирующая вторую цепь направляющая РНК). Временно́е никирование второй цепи представляет собой вариант никирования второй цепи для реализации образования требуемого подвергнутого редактированию продукта. Термин "временно́е" означает то, что второй разрыв не подвергнутой редактированию цепи осуществляется только после встраивания требуемого изменения в подвергнутую редактированию цепь. Это обеспечивает возможность избежать одновременные разрывы на обеих цепях, которые приводят к двуцепочечным разрывам ДНК.

[0078] На **Фиг.1J** показан вариант предполагаемого здесь праймерного редактирования, в рамках которого осуществляется замещение napDNAbp (например никазы SpCas9) любым доменом программируемой нуклеазы, например нуклеазами цинкового пальца (ZFN) или подобными активатору транскрипции эффекторными нуклеазами (TALEN). Таким образом, предполагается, что подходящие нуклеазы не обязательно должны "программироваться" нацеленной на нуклеиновую кислоту молекулой (например направляющей РНК), но вместо этого могут программироваться путем задания специфичности ДНК-связывающего домена, например (и в частности) нуклеазы. Как и при праймерном редактировании с использованием фрагментов napDNAbp, является предпочтительным, чтобы такие альтернативные программируемые нуклеазы были модифицированы таким образом, что только одна цепь ДНК-мишени подвергается разрыву. Другими словами, программируемые нуклеазы предпочтительно должны функционировать в качестве никаз. После выбора программируемой нуклеазы (например ZFN или TALEN) в систему могут быть добавлены дополнительные виды функциональности для обеспечения ее работы в соответствии с подобным праймерному редактированию механизмом. Например, программируемые нуклеазы могут быть модифицированы путем присоединения к ним (например через химический линкер) РНК или ДНК плеча удлинения, при этом плечо удлинения включает участок связывания праймера (PBS) и матрицу для синтеза ДНК. Программируемая нуклеаза также может быть присоединена (например через химический или аминокислотный линкер) к полимеразе, природа которой будет зависеть от того, является ли плечо удлинения ДНК или РНК. В случае плеча

удлинения в виде РНК полимераза может представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). В случае плеча удлинения в виде ДНК полимераза может представлять собой ДНК-зависимую ДНК-полимеразу (например прокариотическую полимеразу, включая Pol I, Pol II, или Pol III, или эукариотическую полимеразу, включая Pol a, Pol b, Pol g, Pol d, Pol e, или Pol z). Система может также включать другие виды функциональности, добавленные в виде продуктов слияния с программируемыми нуклеазами или добавленные *in trans* для реализации реакции в целом (например (a) хеликазу для разматывания ДНК в участке разрыва, чтобы сделать цепь с разрывом с 3'-концом доступной в качестве праймера, (b) флэп-эндонуклеазу (например FEN1) для обеспечения удаления эндогенной цепи на цепи с разрывом для смещения реакции в сторону замещения эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:нРНК для создания второго участка разрыва на противоположной цепи, что может обеспечивать интеграцию синтезированной коррекции за счет предпочтительной клеточной репарации не подвергнутой редактированию цепи). Аналогично праймерному редактированию с использованием napDNAbp, подобный комплекс с иным образом программируемой нуклеазой может использоваться для синтеза и последующего встраивания новой синтезированной замещающей цепи ДНК, несущей представляющее интерес изменение, перманентно в участок-мишень ДНК.

[0079] На **Фиг.1K** в рамках одного варианта осуществления показаны структурные свойства ДНК-мишени, которая может подвергаться редактированию с помощью праймерного редактирования. ДНК-мишень включает "не являющуюся мишенью цепь" и "цепь-мишень". Цепь-мишень представляет собой цепь, которая отжигается на спейсере PEgRNA комплекса праймерного редактора, который распознает участок PAM (в этом случае — NGG, который распознается праймерными редакторами на основе стандартного SpCas9). Цепь-мишень может быть также обозначена как "не включающая PAM цепь" или "не подвергаемая редактированию цепь". С другой стороны, не являющаяся мишенью цепь (то есть цепь, содержащая протоспейсер и PAM-последовательность NGG) может быть обозначена как "включающая PAM цепь" или "подвергаемая редактированию цепь". В различных вариантах осуществления участок разрыва PE комплекса будет располагаться в протоспейсере включающей PAM цепи (например в случае PE на основе SpCas9). Расположение разрыва будет соответствовать конкретному Cas9, который образует PE. Например, в случае PE на

основе SpCas9 участок разрыва в фосфодиэфирной связи располагается между основаниями три (положение "-3" относительно положения 1 последовательности PAM) и четыре (положение "-4" относительно положения 1 последовательности PAM). Участок разрыва в протоспейсере образует свободную 3’ гидроксильную группу, которая, как видно на последующих фигурах, образует комплекс с участком связывания праймера на плече удлинения PEgRNA и обеспечивает субстрат для начала полимеризации одиночной цепи кода ДНК с помощью матрицы для синтеза ДНК плеча удлинения PEgRNA. Данная реакция полимеризации катализируется полимеразой (например обратной транскриптазой) слитого белка PE в направлении от 5’ до 3’. Полимеризации терминируется до достижения ядра нРНК (например за счет включения сигнала терминации полимеризации, или вторичной структуры, которая функционирует для терминирования полимеризационной активности PE), обеспечивая продуцирование одноцепочечного ДНК-флэпа, который простирается от исходной 3’ гидроксильной группы никированной включающей PAM цепи. Матрица для синтеза ДНК кодирует одноцепочечную ДНК, которая гомологична эндогенной 5’-концевой одиночной цепи ДНК, которая немедленно следует за участком разрыва на включающей PAM цепи и обеспечивает встраивание требуемого нуклеотидного изменения (например замещения одного основания, вставки, делеции, инверсии). Положение требуемого изменения может соответствовать любому положению ниже участка разрыва на включающей PAM цепи, которое может включать положение +1, +2, +3, +4 (начало участка PAM), +5 (положение 2 участка PAM), +6 (положение 3 участка PAM), +7, +8, +9, +10, +11, +12, +13, +14, +15, +16, +17, +18, +19, +20, +21, +22, +23, +24, +25, +26, +27, +28, +29, +30, +31, +32, +33, +34, +35, +36, +37, +38, +39, +40, +41, +42, +43, +44, +45, +46, +47, +48, +49, +50, +51, +52, +53, +54, +55, +56, +57, +58, +59, +60, +61, +62, +63, +64, +65, +66, +67, +68, +69, +70, +71, +72, +73, +74, +75, +76, +77, +78, +79, +80, +81, +82, +83, +84, +85, +86, +87, +88, +89, +90, +91, +92, +93, +94, +95, +96, +97, +98, +99, +100, +101, +102, +103, +104, +105, +106, +107, +108, +109, +110, +111, +112, +113, +114, +115, +116, +117, +118, +119, +120, + 121, +122, +123, +124, +125, +126, +127, +128, +129, +130, +131, +132, +133, +134, +135, +136, +137, +138, +139, +140, +141, +142, +143, +144, +145, +146, +147, +148, +149, или +150, или более (относительно положения ниже участка разрыва). При замещении 3'-концевой одноцепочечной ДНК (содержащей представляющее интерес изменение) эндогенной 5’-концевой одноцепочечной ДНК процессы репарации и репликации ДНК

будут приводить к перманентному внесению участка изменения во включающую PAM цепь и затем к коррекции несовпадения в не включающей PAM цепи, которое будет присутствовать в участке редактирования. За счет этого изменение будет присутствовать на обоих цепях ДНК в участке-мишени ДНК. Очевидно, что определения "подвергнутая редактированию цепь" и "не подвергнутая редактированию" цепь служат лишь для обозначения цепей ДНК, затрагиваемых механизмом PE. "Подвергнутая редактированию цепь" представляет собой цепь, которая сначала подвергается редактированию путем замещения 5'-концевой одноцепочеченой ДНК непосредственно ниже участка разрыва синтезированной 3'-концевой одноцепочечной ДНК, содержащей требуемое изменение. "Не подвергаемая редактированию" цепь представляют собой парную цепь подвергаемой редактированию цепи, но которая сама по себе также становится редактированной за счет репарации и/или репликации для сохранения комплементарности относительно подвергнутой редактированию цепи, в частности представляющего интерес изменения.

[0080] На **Фиг.1L** изображен механизм праймерного редактирования, демонстрирующий структурные свойства ДНК-мишени, комплекса праймерного редактирования, и взаимодействие между PEgRNA и ДНК-мишенью. Сначала праймерный редактор, включающий слитый белок, имеющий полимеразу (например обратную траснкриптазу) и napDNAbp (например никазу SpCas9, например SpCas9, имеющую деактивирующую мутацию в домене нуклеазы HNH (например H840A) или деактивирующую мутацию в домене нуклеазы RuvC (D10A)) образует комплекс с PEgRNA и ДНК, имеющей подвергаемую редактированию ДНК-мишень. PEgRNA включает спейсер, ядро нРНК (так известное как каркас нРНК или основа нРНК) (которое связывается с napDNAbp), и плечо удлинения. Плечо удлинения может быть расположено на 3-конце, 5'-конце, или где-то в промежуточном месте в молекуле PEgRNA. Как показано, плечо удлинения расположено на 3'-конце PEgRNA. Плечо удлинения включает в направлении от 3' к 5' участок связывания праймера и матрицу для синтеза ДНК (включающую как представляющее интерес изменение, так и участки гомологии (то есть плечи гомологии), которые являются гомологичными 5'-концевой одноцепочечной ДНК, следующей сразу же за участком разрыва включающей PAM цепи. Как показано, при внесении разрыва с созданием, таким образом, свободной 3' гидроксильной группы непосредственно выше участка разрыва, участок непосредственно выше участка разрыва на включающей PAM цепи отжигается на

комплементарной последовательности на 3'-конце плеча удлинения, которая обозначается как "участок связывания праймера", создавая короткий двуцепочечный участок со свободным 3’ гидроксильным концом, который образует субстрат для полимеразы комплекса праймерного редактора. Полимераза (например обратная транскриптаза) затем полимеризируется в виде цепи ДНК от 3’ карбоксильного конца к концу плеча удлинения. Последовательность одноцепочечной ДНК кодируется матрицей для синтеза ДНК, которая представляет собой часть плеча удлинения (то есть за исключением участка связывания праймера), которая "считывается" полимеразой для синтеза новой ДНК. Данная полимеризация фактически обеспечивает удлинение последовательности исходного 3’ карбоксильного конца первоначального участка разрыва. Матрица для синтеза ДНК кодирует одиночную цепь ДНК, которая включает не только требуемое изменение, но также участки, которые являются гомологичными относительно эндогенной одноцепочечной ДНК непосредственно ниже участка разрыва на включающей PAM цепи. Далее копируемая 3’-концевая одноцепочечная ДНК (то есть 3’ одноцепочечный ДНК-флэп) вытесняет соответствующую гомологичную эндогенную 5’-концевую одиночную цепь ДНК непосредственно ниже участка разрыва на включающей PAM цепи, образуя промежуточную ДНК, имеющую 5’-концевой одноцепочечный ДНК-флэп, который удаляется клеткой (например с помощью флэп-эндонуклеазы). 3’-концевой одноцепочечный ДНК-флэп, который отжигается на комплементе эндогенного 5’-концевого одноцепочечного ДНК-флэпа, лигируется с эндогенной цепью после удаления 5’ ДНК-флэпа. Требуемое изменение в 3’-концевом одноцепочечном ДНК-флэпе, который был подвергнут отжигу и лигированию, образует несовпадение с комплементарной цепью, которое подвергается репарации ДНК и/или раунду репликации, тем самым обеспечивая перманентное внесение требуемого изменение в обе цепи.

[0081] На **Фиг.2** показаны три комплекса Cas (SpCas9, SaCas9, и LbCas12a), которые могут применяться в описанных здесь праймерных редакторах, и их свойства, связанные с PAM, нРНК, и расщеплением ДНК. На фигуре показаны конструкции комплексов, включающих SpCas9, SaCas9, и LbCas12a.

[0082] На **Фиг.3A-3F** показаны конструкции сконструированной нРНК праймерного редактора (**Фиг.3A**), 3´ нРНК праймерного редактора (**Фиг.3B**), и внутримолекулярного удлинения (**Фиг.3C**). Удлиненная направляющая РНК (или удлиненная нРНК) также может быть обозначена здесь как PEgRNA или

"направляющая РНК для праймерного редактирования". На **Фиг.3D** и **Фиг.3E** представлены дополнительные варианты осуществления 3′ и 5′ нРНК праймерного редактора (PEgRNA), соответственно. На **Фиг.3F** показано взаимодействие между 3′-концевой направляющей нРНК праймерного редактора и последовательностью-мишенью ДНК. Варианты осуществления на **Фиг.3A-3C** показывают примеры расположения последовательности матрицы обратной транскрипции (то есть более широко обозначаемой как матрица для синтеза ДНК, как указано, поскольку RT представляет собой лишь один из типов полимеразы, которая может применяться в контексте праймерных редакторов), участок связывания праймера, и необязательную последовательность линкера в удлиненных частях 3′, 5′, и внутримолекулярные версии, а также общее расположение спейсера и участков ядра. Описанный процесс праймерного редактирования не ограничивается данными конфигурациями удлиненных направляющих РНК. В варианте осуществления на **Фиг.3D** показана структура примера предполагаемой здесь PEgRNA. PEgRNA включает три основных компонента в направлении от 5′ к 3′, а именно: спейсер, ядро нРНК, и плечо удлинения на 3′-конце. Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5′ к 3′, а именно: участок связывания праймера (A), матрицу для редактирования (B), и плечо гомологии (C). Кроме того, PEgRNA может включать необязательный 3′-концевой модификаторный участок (e1) и необязательный 5′-концевой модификаторный участок (e2). Кроме того, PEgRNA может включать сигнал терминации транскрипции на 3′-конце PEgRNA (не показан). Данные структурные элементы дополнительно определены здесь. Изображение структуры PEgRNA не является ограничивающим и охватывает вариации в плане расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми из показанных участков без ограничения в виде расположения на 3′- или 5′-концах. PEgRNA может включать в определенных вариантах осуществления вторичную структуру РНК, например, но без ограничения, шпильки, шпильки/петли, опорные петли, домены рекрутирования РНК-связывающего белка (например аптамер MS2, который рекрутирует и связывает белок MS2cp). Например, подобные вторичные структуры могут быть расположены в спейсере, ядре нРНК, или плече удлинения, и, в частности, в модификаторных участках e1 и/или e2. Помимо вторичных структур РНК, PEgRNA могут включать (например в модификаторных участках e1 и/или e2)

химический линкер или поли(N) линкер или хвост, где "N" может представлять собой любое азотистое основание. В некоторых вариантах осуществления (например как показано на Фиг.72(c)), химический линкер может функционировать для предотвращения обратной транскрипции каркаса или ядра sgРНК. Кроме того, в определенных вариантах осуществления (например см. Фиг.72(c)), плечо удлинения (3) может состоять из РНК или ДНК, и/или может включать один или несколько аналогов азотистых оснований (например которые могут добавлять функциональность, например температурную устойчивость). Кроме того, ориентация плеча удлинения (3) может соответствовать естественному направлению от 5′ к 3′, или оно может быть синтезировано с противоположной ориентаций в направлении от 3′ к 5′ (относительно общей ориентации молекулы PEgRNA). Следует также отметить, что средний специалист в данной области будет способен выбрать подходящую ДНК-полимеразу в зависимости от природы материалов нуклеиновой кислоты плеча удлинения (то есть ДНК или РНК) для применения в рамках праймерного редактирования, которая может быть реализована либо в виде продукта слияния с napDNAbp или поступать *in trans* в виде отдельного фрагмента для синтеза требуемого кодируемого матрицей 3′ одноцепочечного ДНК-флэпа, который включает требуемое изменение. Например, если плечо удлинения представляет собой РНК, то ДНК-полимераза может представлять собой обратную транскриптазу или другую подходящую РНК-зависимую ДНК-полимеразу. Однако если плечо удлинения представляет собой ДНК, то ДНК-полимераза может представлять собой ДНК-зависимую ДНК-полимеразу. В различных вариантах осуществления поступление ДНК-полимеразы может осуществляться *in trans*, например путем применения домена рекрутирования РНК-белка (например MS2 шпильки, расположенной на PEgRNA (например в участке e1 или e2, или где-либо еще), и белка MS2cp, слитого с ДНК-полимеразой, тем самым обеспечивая солокализацию ДНК-полимеразы с PEgRNA). Следует также отметить, что участок связывания праймера в целом не образует часть матрицы, которая используется ДНК-полимеразой (например обратной транскриптазой) для кодирования полученного 3′ одноцепочечного ДНК-флэпа, который включает требуемое изменение. Таким образом, обозначение "матрица для синтеза ДНК" обозначает участок или часть плеча удлинения (3), которая используется в качестве матрицы ДНК-полимеразой для кодирования требуемого 3′ одноцепочечного ДНК-флэпа, содержащего изменение и участки гомологии относительно 5’ эндогенного одноцепочечного ДНК-флэпа,

который замещается продуктом синтеза ДНК в рамках праймерного редактирования в виде 3′ одноцепочечной цепи ДНК. В некоторых вариантах осуществления матрица для синтеза ДНК включает "матрицу для редактирования" и "плечо гомологии", или одно или несколько плечей гомологии, например до и после матрицы для редактирования. Размеры матрицы для редактирования могут соответствовать замещению одного нуклеотида, или вставке, или инверсии ДНК. Кроме того, матрица для редактирования также может включать делецию, которая может быть внесена путем кодирования плеча гомологии, которое содержит требуемую делецию. В других вариантах осуществления матрица для синтеза ДНК также может включать участок e2 или его часть. Например, если участок e2 включает вторичную структуру, которая вызывает терминацию активности ДНК-полимеразы, то является возможным, что функция ДНК-полимеразы будет терминироваться до фактического кодирования в ДНК какой-либо части участка e2. Также возможно, что некоторая часть или даже весь участок e2 будет кодироваться в ДНК. То, какое количество e2 будет фактически использовано в качестве матрицы, будет зависеть от его структуры и того, нарушает ли эта структура функцию ДНК-полимеразы.

[0083] В варианте осуществления на **Фиг.3E** показана структура примера другой предполагаемой здесь PEgRNA. PEgRNA включает три основных компонента в направлении от 5′ к 3′, а именно: спейсер, ядро нРНК, и плечо удлинения на 3′-конце Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5′ к 3′, а именно: участок связывания праймера (A), матрицу для редактирования (B), и плечо гомологии (C). Кроме того, PEgRNA может включать необязательный 3′-концевой модификаторный участок (e1) и необязательный 5′-концевой модификаторный участок (e2). Кроме того, PEgRNA может включать сигнал терминации транскрипции на 3′-конце PEgRNA (не показан). Данные структурные элементы дополнительно определены здесь. Изображение структуры PEgRNA не является ограничивающим и охватывает вариации в плане расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми из показанных участков, без ограничения в виде расположения на 3′- или 5′-концах. PEgRNA может включать в определенных вариантах осуществления вторичную структуру РНК, например, но без ограничения, шпильки, шпильки/петли, опорные петли, домены рекрутирования РНК-связывающего белка (например аптамер MS2, который рекрутирует и связывает белок

MS2cp). Данные вторичные структуры могут быть расположены в любом месте в рамках молекулы PEgRNA. Например, подобные вторичные структуры могут быть расположены в спейсере, ядре нРНК, или плече удлинения, и, в частности, в модификаторных участках e1 и/или e2. Помимо вторичных структур РНК, PEgRNA могут включать (например в модификаторных участках e1 и/или e2) химический линкер или поли(N) линкер или хвост, где "N" может представлять собой любое азотистое основание. В некоторых вариантах осуществления (например как показано на Фиг.72(C)), химический линкер может функционировать для предотвращения обратной транскрипции каркаса или ядра sgРНК. Кроме того, в определенных вариантах осуществления (например см. Фиг.72(c)), плечо удлинения (3) может состоять из РНК или ДНК, и/или может включать один или несколько аналогов азотистых оснований (например которые могут добавлять функциональность, например температурную устойчивость). Кроме того, ориентация плеча удлинения (3) может соответствовать естественному направлению от 5′ к 3′, или оно может быть синтезировано с противоположной ориентаций в направлении от 3′ к 5′ (относительно общей ориентации молекулы PEgRNA). Следует также отметить, что средний специалист в данной области будет способен выбрать подходящую ДНК-полимеразу в зависимости от природы материалов нуклеиновой кислоты плеча удлинения (то есть ДНК или РНК) для применения в рамках праймерного редактирования, которая может быть реализована либо в виде продукта слияния с napDNAbp или поступать *in trans* в виде отдельного фрагмента для синтеза требуемого кодируемого матрицей 3′ одноцепочечного ДНК-флэпа, который включает требуемое изменение. Например, если плечо удлинения представляет собой РНК, то ДНК-полимераза может представлять собой обратную транскриптазу или другую подходящую РНК-зависимую ДНК-полимеразу. Однако если плечо удлинения представляет собой ДНК, то ДНК-полимераза может представлять собой ДНК-зависимую ДНК-полимеразу. В различных вариантах осуществления поступление ДНК-полимеразы может осуществляться in trans, например путем применения домена рекрутирования РНК-белка (например MS2 шпильки, расположенной на PEgRNA (например в участке e1 или e2, или где-либо еще), и белка MS2cp, слитого с ДНК-полимеразой, тем самым обеспечивая солокализацию ДНК-полимеразы с PEgRNA). Следует также отметить, что участок связывания праймера в целом не образует часть матрицы, которая используется ДНК-полимеразой (например обратной транскриптазой) для кодирования

полученного 3′ одноцепочечного ДНК-флэпа, который включает требуемое изменение. Таким образом, обозначение "матрица для синтеза ДНК" обозначает участок или часть плеча удлинения (3), которая используется в качестве матрицы ДНК-полимеразой для кодирования требуемого 3′ одноцепочечного ДНК-флэпа, содержащего изменение и участки гомологии относительно 5’ эндогенного одноцепочечного ДНК-флэпа, который замещается продуктом синтеза ДНК в рамках праймерного редактирования в виде 3′ одноцепочечной цепи ДНК. В некоторых вариантах осуществления матрица для синтеза ДНК включает "матрицу для редактирования" и "плечо гомологии", или одно или несколько плечей гомологии, например до и после матрицы для редактирования. Размеры матрицы для редактирования могут соответствовать замещению одного нуклеотида, или вставке, или инверсии ДНК. Кроме того, матрица для редактирования также может включать делецию, которая может быть внесена путем кодирования плеча гомологии, которое содержит требуемую делецию. В других вариантах осуществления матрица для синтеза ДНК также может включать участок e2 или его часть. Например, если участок e2 включает вторичную структуру, которая вызывает терминацию активности ДНК-полимеразы, то является возможным, что функция ДНК-полимеразы будет терминироваться до фактического кодирования в ДНК какой-либо части участка e2. Также возможно, что некоторая часть или даже весь участок e2 будет кодироваться в ДНК. То, какое количество e2 будет фактически использовано в качестве матрицы, будет зависеть от его структуры и того, нарушает ли эта структура функцию ДНК-полимеразы

[0084] Схема на **Фиг.3F** показывает взаимодействие стандартной PEgRNA с участком-мишенью двуцепочечной РНК и сопутствующее получение 3′ одноцепочечного ДНК-флэпа, содержащего представляющее интерес генетическое изменение. Двуцепочечная ДНК показана с верхней цепью (то есть цепью-мишенью) в ориентации от 3′ к 5′ и нижней цепью (то есть включающей PAM цепью или не являющейся мишенью цепью) в направлении от 5′ к 3′. Верхняя цепь включает комплемент "протоспейсера" и комплемент последовательности PAM и обозначается как "цепь-мишень", поскольку она представляет собой цепь, на которую нацеливается и отжигается спейсер PEgRNA. Комплементарная нижняя цепь обозначается как "не являющаяся мишенью цепь" или "включающая PAM цепь" или "протоспейсерная цепь", поскольку она содержит последовательность PAM (например NGG) и протоспейсер. Хотя это и не показано, изображенная PEgRNA будет образовывать

комплекс с Cas9 или эквивалентным доменом слитого белка праймерного редактора. Как показано на схеме, спейсер PEgRNA отжигается на комплементарном участке протоспейсера на цепи-мишени. Это взаимодействие образует ДНК/РНК гибрид между спейсерной РНК и комплементом протоспейсерной ДНК и индуцирует образование R-петли в протоспейсере. Как описано в других разделах, затем белок Cas9 (не показан) индуцирует разрыв в не являющейся мишенью цепи, как показано. Это затем ведет к образованию участка в виде 3′ оцДНК-флэпа (одноцепочечной ДНК) непосредственно выше участка разрыва, который, в соответствии с *z*, взаимодействует с 3′-концом PEgRNA в участке связывания праймера. 3′-конец оцДНК-флэпа (то есть последовательность праймера в виде обратной транскриптазы) отжигается на участке связывания праймера (A) на PEgRNA, тем самым обеспечивая праймирование обратной транскриптазы. Далее обратная транскриптаза (например поступающая *in trans* или поступающая *cis* в виде слитого белка, присоединенного к конструкту Cas9) затем обеспечивает полимеризацию одиночной цепи ДНК, которая кодируется матрицей для синтеза ДНК (включая матрицу для редактирования (B) и плечо гомологии (C)). Полимеризация продолжается по направлению к 5′-концу плеча удлинения. Полимеризованная цепь оцДНК образует 3′-концевой оцДНК-флэп, который, как описано в других местах (например на **Фиг.1G**) вносится в эндогенную ДНК, вытесняя соответствующую эндогенную цепь (которая удаляется в виде 5′-концевого ДНК-флэпа эндогенной ДНК) и обеспечивая внесение требуемого нуклеотидного изменения (изменения одной пары азотистых оснований нуклеотида, делеций, вставок (включая целые гены) с помощью раундов естественной репарации/репликации ДНК).

[0085] На **Фиг.3G** показан другой вариант осуществления рассматриваемого здесь праймерного редактирования. В частности, на верхней схеме изображен один вариант осуществления праймерного редактора (PE), который включает слитый белок napDNAbp (например SpCas9) и полимеразу (например обратную транскриптазу), которые соединены линкером. PE образует комплекс с PEgRNA путем связывания с ядром нРНК PEgRNA. В показанном варианте осуществления PEgRNA имеет 3′ плечо удлинения, которое включает, начиная от 3′-конца, участок связывания праймера (PBS) и далее матрицу для синтеза ДНК. На нижней схеме изображен вариант праймерного редактора, обозначенный как "*trans* праймерный редактор (tPE)". В данном варианте осуществления матрица для синтеза ДНК и PBS отсоединяются от PEgRNA и

презентируются на отдельной молекуле, обозначаемой как *trans* РНК матрица праймерного редактора ("tPERT"), которая включает домен рекрутирования РНК-белка (например шпильки MS2). Сам PE дополнительно модифицируется для включения продукта слияния с rPERT-рекрутирующим белком ("RP"), который представляет собой белок, который специфично распознает и связывает домен рекрутирования РНК-белка. В примере, где домен рекрутирования РНК-белка представляет собой шпильку MS2, соответствующий rPERT-рекрутирующий белок может представлять собой MS2cp системы мечения MS2. Система мечения MS2 основана на естественном взаимодействии белка оболочки бактериофага MS2 ("MCP" или "MS2cp") со структурой в виде шпильки, присутствующей в геноме фага, то есть "шпилькой MS2" или "аптамером MS2". В случае *trans* праймерного редактирования комплекс RP-PE:нРНК "рекрутирует" tPERT, имеющую подходящий домен рекрутирования РНК-белка для солокализации с комплексом PE:нРНК, тем самым предоставляя PBS и матрицу для синтеза ДНК *in trans* для использования в рамках праймерного редактирования, как показано в примере, изображенном на Фиг.3H.

[0086] На **Фиг.3H** показан процесс *trans* праймерного редактирования. В данном варианте осуществления *trans* праймерный редактор включает праймерный редактор "PE2" (то есть продукт слияния Cas9(H840A) и варианта RT MMLV (вируса лейкоза мышей Молони)), слитый с белком MS2cp (то есть типом рекрутирующего белка, который распознает и связывает аптамер MS2), и который образует комплекс с sgРНК (то есть стандартной направляющей РНК вместо PEgRNA). *Trans* праймерный редактор связывается с ДНК-мишенью и никирует не являющуюся мишенью цепь. Белок MS2cp рекрутирует tPERT *in trans* за счет специфичного взаимодействия с доменом рекрутирования РНК-белка на молекуле tPERT. tPERT становится солокализованной с *trans* праймерным редактором, тем самым предоставляя функции PBS и матрицы для синтеза ДНК *in trans* для использования полимеразой в виде обратной транскриптазы для синтеза одноцепочечного ДНК-флэпа, имеющего 3′-конец и содержащего требуемую генетическую информацию, кодируемую матрицей для синтеза ДНК.

[0087] На **Фиг.4A-4E** демонстрируются *in vitro* анализы TPRT (то есть анализы праймерного связывания). **Фиг.4A** представляет собой схему удлинения на основе нРНК матрицы флуоресцентно меченных субстратов ДНК с помощью фермента RT, PAGE. На **Фиг.4B** показана TPRT (то есть праймерное редактирование) с

использованием подвергнутых предварительному никированию субстратов, dCas9, и 5′-удлиненных нРНК с матрицей для синтеза различной длины. На **Фиг.4C** показана RT реакция с использованием подвергнутых предварительному никированию субстратов ДНК при отсутствии Cas9. На **Фиг.4D** показана TPRT (то есть праймерное редактирование) полных субстратов дцДНК с использованием Cas9(H840A) и 5′-удлиненных нРНК. На **Фиг.4E** показана 3′-удлиненная нРНК матрица с подвергнутыми предварительному никированию и полными субстратами дцДНК. Все реакции осуществлялись с использованием RT M-MLV.

[0088] На **Фиг.5** показаны *in vitro* анализы с использованием 5′-удлиненных нРНК с матрицами для синтеза различной длины. В данном наборе экспериментов флуоресцентно меченные (Cy5) ДНК-мишени использовали в качестве субстратов и подвергали предварительному никированию. Cas9, использованный в этих экспериментах, представляет собой каталитически мертвый Cas9 (dCas9), а использованная RT представляет собой Superscript III, коммерческую RT, полученную из вируса лейкоза мышей Молони (M-MLV). Комплексы dCas9:нРНК были образованы из очищенных компонентов. Затем добавляли флуоресцентно меченный субстрат ДНК совместно с dNTP и ферментом RT. Спустя 1 час инкубирования при 37 °C продукты реакции анализировали с помощью денатурирующего электрофореза на основе полиакриламидного геля с применением мочевины (PAGE). Изображение на геле показывает удлинение исходной цепи ДНК до показателей длины, которые соответствуют длине матрицы обратной транскрипции.

[0089] На **Фиг.6** показаны *in vitro* анализы с использованием 5′-удлиненных нРНК с матрицами для синтеза различной длины, аналогичные показанным на **Фиг.5**. Однако субстраты ДНК не подвергались предварительному никированию в данном наборе экспериментов. Cas9, используемый в этих экспериментах, представляет собой никазу Cas9 (мутант SpyCas9 H840A), использованная RT представляет собой Superscript III, коммерческую RT, полученную из вируса лейкоза мышей Молони (M-MLV). Продукты реакции анализировали с помощью денатурирующего электрофореза на основе полиакриламидного геля с применением мочевины (PAGE). Как показано на геле, никаза эффективно расщепляет цепь ДНК при использовании стандартной нРНК (нРНК_0, дорожка 3).

[0090] На **Фиг.7** показано, что 3′ удлинения поддерживают синтез ДНК и не оказывают значительного влияния на активность никазы Cas9. Подвергнутые

предварительному никированию субстраты (черная стрелка) практически полностью конвертируются в продукты RT при использовании dCas9 или Cas9 никазы (дорожки 4 и 5). Более чем 50% конверсия в продукт RT (красная стрелка) наблюдается с полными субстратами (дорожка 3). Использовали никазу Cas9 (мутант SpyCas9 H840A), каталитически мертвый Cas9 (dCas9), и Superscript III, коммерческую RT, полученную из вируса лейкоза мышей Молони (M-MLV).

[0091] На **Фиг.8** показаны эксперименты на основе двойного цвета, которые использовались для определения того, осуществляется ли предпочтительным образом RT реакция с нРНК in cis (связывание с тем же комплексом). Два отдельных эксперимента проводили для 5′-удлиненных и 3′-удлиненных нРНК. Продукты анализировали с помощью PAGE. Соотношение продуктов вычисляли как (Cy3cis/Cy3trans) / (Cy5trans/Cy5cis).

[0092] На **Фиг.9A-9D** показана модель в виде флэп-субстрата**.** На **Фиг.9A** показан репортер с двойным флуоресцентным белком для флэп-направленного мутагенеза. На **Фиг.9B** показана репарация стоп-кодона в клетках HEK. На **Фиг.9C** показаны секвенированные клоны дрожжей после репарации флэпа. На **Фиг.9D** показано тестирование различных свойств флэпа в клетках человека.

[0093] На **Фиг.10** демонстрируется праймерное редактирование на плазмидных субстратах. Репортерную плазмиду с двойной флуоресценцией конструировали для экспрессии в дрожжах (S. cerevisiae). Экспрессия данного конструкта в дрожжах продуцирует только GFP. *In vitro* реакция праймерного редактирования обеспечивает внесение точечной мутации и трансформирует дрожжи исходной плазмидой или *in vitro* Cas9(H840A) никированной плазмидой. Колонии визуализированы с помощью флуоресцентной визуализации. Показаны дрожжевые плазмидные трансформанты с двойным флуоресцентным белком. Трансформирование исходной плазмидой или *in vitro* Cas9(H840A) никированной плазмидой приводит к получению колоний, экспрессирующих только зеленый GFP. Реакция праймерного редактирования с 5′-удлиненными или 3′-удлиненными нРНК продуцирует смесь зеленых и желтых колоний. Последние экспрессируют как GFP, так и mCherry. Еще больше желтых колоний наблюдается при использовании 3′-удлиненной нРНК. Также показан положительный контроль, который не содержит стоп-кодон.

[0094] На **Фиг.11** показано праймерное редактирование на плазмидных субстратах аналогично эксперименту на **Фиг.10**, но вместо внесения точечной мутации в стоп-

кодон праймерное редактирование обеспечивает внесение вставки (слева) или делеции (справа) в виде одного нуклеотида, которая корректирует мутацию сдвига рамки считывания и обеспечивает возможность синтеза mCherry ниже. В обоих экспериментах использовались 3′-удлиненные нРНК.

[0095] На **Фиг.12** показаны продукты редактирования в рамках праймерного редактирования на плазмидных субстратах, проанализированные с помощью секвенирования по Сэнгеру. Колонии из TRT трансформантов отдельным образом отбирали и анализировали с помощью секвенирования по Сэнгеру. Точные изменения наблюдали на основе секвенирования выбранных колоний. Зеленые колонии содержали плазмиды с исходной последовательность ДНК, в то время как желтые колонии содержали точную мутацию, осуществленную с помощью нРНК для праймерного редактирования. Другие точечные мутации или индели не наблюдались.

[0096] На **Фиг.13** показан потенциальный объем новой технологии праймерного редактирования и сравнен с технологиями редактирования оснований на основе дезаминазы.

[0097] На **Фиг.14** показана схема редактирования в клетках человека.

[0098] На **Фиг.15** демонстрируется удлинение участка связывания праймера в нРНК.

[0099] На **Фиг.16** показаны усеченные нРНК для смежного нацеливания.

[0100] **Фиг.17A-17C** представляют собой графики, показывающие % конверсии T в A в нуклеотиде-мишени после трансфекции компонентов в эмбриональные клетки почки человека (HEK). На **Фиг.17A** показаны данные, которые демонстрируют результаты с использованием N-концевого слияния обратной транскриптазы MLV дикого типа и никазы Cas9(H840A) (32-аминокислотный линкер). **Фиг.17B** аналогична **Фиг.17A**, но для C-концевого слияния фермента RT. **Фиг.17C** аналогична **Фиг.17A**, но линкер между RT MLV и Cas9 имеет длину 60 аминокислот вместо 32 аминокислот.

[0101] На **Фиг.18** показана высокая чистота редактирования T в A в участке HEK3 на основе высокопроизводительного секвенирования ампликонов. Вывод анализа на основе секвенирования демонстрирует наиболее часто встречающиеся генотипы подвергнутых редактированию клеток.

[0102] На **Фиг.19** показа эффективность редактирования в нуклеотиде-мишени (синие столбцы) наряду с показателями инделей (оранжевые столбцы). WT обозначает фермент RT MLV дикого типа. Мутантные ферменты (от M1 до M4) содержат перечисленные справа мутации. Показатели редактирования количественно

анализировали с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[0103] На **Фиг.20** показана эффективность редактирования нуклеотида-мишени при внесении одноцепочечного разрыва в цепь комплементарной ДНК поблизости от нуклеотида-мишени. Было протестировано никирование на различном расстоянии от нуклеотида-мишени (треугольники). Показана эффективность редактирования в паре азотистых оснований-мишеней (синие столбцы) наряду с показателями образования инделей (оранжевые столбцы). Пример "нет" не содержит никирующей комплементарную цепь направляющей РНК. Показатели редактирования количественно анализировали с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[0104] На **Фиг.21** демонстрируются обработанные данные высокопроизводительного секвенирования, показывающие требуемую мутацию в виде конверсии T в A и общее отсутствие других значительных побочных продуктов редактирования генома.

[0105] На **Фиг.22** показана схема примера процесса осуществления направленного мутагенеза с использованием подверженной ошибкам обратной транскриптазы в локусе-мишени с помощью программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) в виде комплекса с удлиненной направляющей РНК, то есть праймерное редактирование с использованием подверженной ошибкам RT. Данный процесс может быть обозначен как вариант осуществления праймерного редактирования для направленного мутагенеза. Удлиненная направляющая РНК включает удлинение на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК. На этапе (a) комплекс napDNAbp/нРНК контактирует с молекулой ДНК, при этом нРНК направляет napDNAbp для связывания с подвергаемым мутагенезу локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК. На этапе (c) 3′-концевая цепь ДНК взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции. В определенных вариантах осуществления 3′-концевая цепь ДНК гибридизируется с определенной

праймерной последовательностью RT на удлиненной части направляющей РНК. На этапе (d) вносится подверженная ошибкам обратная транскриптаза, которая синтезирует подвергнутую мутагенезу одиночную цепь ДНК от 3′-конца праймированного участка к 3′-концу направляющей РНК. Примеры мутаций указаны с помощью звездочки "*". Это обеспечивает образование одноцепочечного ДНК-флэпа, включающего требуемый подвергнутый мутагенезу участку. На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа (включающего подвергнутый мутагенезу участок), так что требуемый подвергнутый мутагенезу участок становится встроенным в локус-мишень. Данный процесс может быть смещен в сторону образования требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа, который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с комплементарной последовательностью на другой цепи. Процесс также может быть смещен в сторону образования продукта с никированием второй цепи, как показано в качестве примера на **Фиг.1F**. После процессов репарации и/или репликации эндогенной ДНК мутагенизированный участок становится встроенным в обе цепи ДНК локуса ДНК.

[0106] Фиг.23 представляет собой схему конструкции нРНК для сокращения последовательностей тринуклеотидных повторов и сокращения тринуклеотидных повторов с помощью редактирования генома на основе TPRT (то есть праймерного редактирования). Экспансия тринуклеотидных повторов связана с рядом заболеваний человека, включая болезнь Хантингтона, синдром ломкой Х-хромосомы и атаксию Фридрейха. Наиболее часто встречающийся тринуклеотидный повтор содержит триплеты CAG, хотя также встречаются триплеты GAA (атаксия Фридрейха), и триплеты CGG (синдром ломкой Х-хромосомы). Унаследование предрасположенности к экспансии или приобретение уже экспансированного родительского аллеля повышает вероятность возникновения заболевания. Патогенные экспансии тринуклеотидных повторов гипотетически могут быть скорректированы с помощью праймерного редактирования. Участок выше участка повтора может быть никирован с помощью РНК-направляемой нуклеазы и затем использован для праймирования синтеза новой цепи ДНК, которая содержит здоровое количество повторов (которое зависит от конкретного гена и заболевания). После последовательности повтора добавляется короткий фрагмент гомологии, который соответствует последовательности,

примыкающей к другому концу повтора (красная цепь). Внесение новой синтезированной цепи и последующее замещение эндогенной ДНК новым синтезированным флэпом ведет к сокращению повторяющегося аллеля.

[0107] Фиг.24 представляет собой схему, показывающую точную 10-нуклеотидную делецию с помощью праймерного редактирования. Направляющая РНК, нацеленная на локус HEK3, была сконструирована с помощью матрицы обратной транскрипции, которая кодирует 10-нуклеотидную делецию после участка разрыва. Эффективность редактирования в трансфицированных HEK клетках анализировали с помощью секвенирования ампликонов.

[0108] Фиг.25 представляет собой схему, показывающую конструкцию нРНК для генов пептидного мечения в эндогенных геномных локусах и пептидное мечение с помощью геномного редактирования на основе TPRT (то есть праймерного редактирования). Системы мечения FlAsH и ReAsH включают две части: (1) флуорофор-бимышьяковый зонд и (2) генетически кодируемый пептид, содержащий мотив тетрацистеина, например последовательность FLNCCPGCCMEP (SEQ ID NO: 1). При экспрессии в клетках белки, содержащие мотив тетрацистеина, могут подвергаться флуоресцентному окрашиванию флуорофор-мышьяковыми зондами (см. J. Am. Chem. Soc., 2002, 124 (21), стр. 6063–6076. DOI: 10.1021/ja017687n). В рамках системы "sortagging" (на основе опосредованного сортазой лигирования) используются бактериальные ферменты сортазы, которые ковалентно конъюгируют меченые пептидные зонды с белками, содержащими подходящие пептидные субстраты (см. Nat. Chem. Biol. 2007 Nov;3(11):707-8. DOI: 10.1038/nchembio.2007.31). FLAG-метка (DYKDDDDK (SEQ ID NO: 2)), V5-метка (GKPIPNPLLGLDST (SEQ ID NO: 3)), GCN4-метка (EELLSKNYHLENEVARLKK (SEQ ID NO: 4)), HA-метка (YPYDVPDYA (SEQ ID NO: 5)), и Myc-метка (EQKLISEEDL (SEQ ID NO: 6)) обычно используются в качестве эпитопных меток для иммуноанализов. π-клэмп кодирует пептидную последовательность (FCPF), которая может быть помечена пентафтор-ароматическими субстратами (Nat. Chem. 2016 Feb;8(2):120-8. doi: 10.1038/nchem.2413).

[0109] На **Фиг.26A** показано точное внесение His_6-метки и FLAG-метки в геномную ДНК. С помощью матрицы обратной транскрипции была сконструирована направляющая РНК, нацеленная на локус HEK3, которая кодирует либо вставку 18-nt His-метки, или вставку 24-nt FLAG-метки. Эффективность редактирования в трансфицированных HEK клетках анализировали с помощью секвенирования

ампликонов. Следует обратить внимание, что полная 24-nt последовательность FLAG-метки располагается вне наблюдаемого изображения (секвенирование подтвердило осуществление полной и точной вставки). На **Фиг.26B** показана схема, описывающая различные варианты использования, включающие белковые/пептидные метки, включая (a) превращение белков в растворимые или нерастворимые, (b) изменение или отслеживание клеточной локализации белка, (c) удлинение времени полужизни белка, (d) обеспечение очистки белка, и (e) обеспечение детектирования белков.

[0110] На **Фиг.27** показан обзор праймерного редактирования в виде внесения защитной мутации в *PRNP*, которая предотвращает или останавливает прогрессирование прионного заболевания. Последовательности PEgRNA соответствуют SEQ ID NO: 351 слева (то есть 5′ каркаса sgРНК) и SEQ ID NO: 3864 справа (то есть 3′ каркаса sgРНК).

[0111] **Фиг.28A** представляет собой схему вставки последовательностей, кодирующих мотивы РНК, на основе PE. **Фиг.28B** представляет собой перечень (не исчерпывающий) примеров некоторых мотивов, которые потенциально могут быть вставлены, и их функций.

[0112] **Фиг.29A** представляет собой изображение праймерного редактора. На **Фиг.29B** показаны возможные модификации геномной, плазмидной, или вирусной ДНК на основе PE. На **Фиг.29C** показан пример схемы для вставки библиотеки пептидных петель в определенный белок (в данном случае — GFP) с помощью библиотеки PEgRNA. На **Фиг.29D** показан пример возможных программируемых делеций кодонов или N-, или C-концевых усечений белка с помощью различных PEgRNA. Прогнозируется, что делеции будут возникать с минимальным образованием мутаций сдвига рамки считывания.

[0113] На **Фиг.30** показана возможная схема итеративной вставки кодонов в системе непрерывной эволюции, такой как PACE.

[0114] **Фиг.31** представляет собой иллюстрацию сконструированной нРНК, демонстрирующую ядро нРНК, ~20-nt спейсер, соответствующий гену-мишени, матрицу обратной транскрипции с нуклеотидной последовательностью иммуногенного эпитопа, и участок связывания праймера, соответствующий последовательности гена-мишени.

[0115] **Фиг.32** представляет собой схему, демонстрирующую использование праймерного редактирования для вставки известных иммуногенных эпитопов в

эндогенную или чужеродную геномную ДНК, что приводит к модификации соответствующих белков.

[0116] Фиг.33 представляет собой схему, демонстрирующую конструкцию PEgRNA для вставок последовательности связывания праймера и вставку для связывания праймера в геномную ДНК с помощью праймерного редактирования для определения нецелевого редактирования. В данном варианте осуществления праймерное редактирование осуществляется внутри живой клетки, ткани, или модели с животными. На первом этапе конструируют подходящую PEgRNA. На верхней схеме показан пример PEgRNA, которая может использоваться в данном аспекте. Спейсер в PEgRNA (обозначенный как "протоспейсер") комплементарен одной из цепей геномной мишени. Комплекс PE:PEgRNA (то есть комплекс PE) обеспечивает внесение одноцепочечного 3′-концевого флэпа в участке разрыва, который содержит кодируемую последовательность связывания праймера и участок гомологии (кодируемый плечом гомологии PEgRNA), который комплементарен участку непосредственно ниже участка разрыва (отмечен красным). За счет внесения флэпа и процессов репарации/репликации ДНК синтезированная цепь становится встроенной в ДНК, тем самым обеспечивая внесение участка связывания праймера. Данный процесс может осуществляться в требуемой геномной мишени, но также в других геномных участках, которые могут взаимодействовать с PEgRNA нецелевыми образом (то есть PEgRNA направляет комплекс PE к другим нецелевым участкам вследствие комплементарности спейсерного участка в отношении других геномных участков, которые не являются требуемым геномным участком). Таким образом, последовательность связывания праймера может быть внесена не только в требуемую геномную мишень, но также нецелевые геномные участки в других местах генома. Для детектирования вставки данных участков связывания праймера как в требуемых геномных участках-мишенях, так и нецелевых геномных участках, геномная ДНК (после PE) может быть выделена, фрагментирована, и лигирована с нуклеотидами адаптеров (отмечены красным). Далее может быть выполнена ПЦР с использованием ПЦР олигонуклеотидов, которые отжигаются на адаптерах и вставленной последовательности связывания праймера, для амплификации целевых и нецелевых участков геномной ДНК, в которые с помощью PE был вставлен участок связывания праймера. Затем может быть выполнено высокопроизводительное секвенирование и выравнивание последовательностей для идентификации точек вставки вставленных с

помощью PE последовательностей связывания праймера на целевом участке или нецелевых участках.

[0117] Фиг.34 представляет собой схему, демонстрирующую точную вставку гена с помощью PE.

[0118] Фиг.35A представляет собой схему, демонстрирующую естественный путь передачи сигнала инсулина. **Фиг.35B** представляет собой схему, демонстрирующую активацию меченного FKBP12 рецептора инсулина, контролируемую с помощью FK1012.

[0119] На **Фиг.36** показаны низкомолекулярные мономеры. Ссылки: имеющий "выступы" миметик FK506 (**2**)[107]

[0120] На **Фиг.37** показаны низкомолекулярные димеры. Ссылки: FK1012 4[95,96]; FK1012 5[108]; FK1012 6[107]; AP1903 7[107]; димер циклоспорина A 8[98]; FK506–димер циклоспорина A (FkCsA) 9[100].

[0121] На **Фиг.38A-38F** показан обзор праймерного редактирования и исследования осуществимости *in vitro* и с использованием клеток дрожжей. На **Фиг.38A** показаны 75122 известных патогенных генетических варианта человека в ClinVar (проанализировано в июле 2019) с классификацией по типу. На **Фиг.38B** показано, что комплекс для праймерного редактирования состоит из белка праймерного редактора (PE), содержащего РНК-направляемый ДНК-никирующий домен, например никазу Cas9, слитый со сконструированным доменом обратной транскриптазы и образующий комплекс с направляющей РНК для праймерного редактирования (PEgRNA). Комплекс PE:PEgRNA связывает участок-мишень ДНК и обеспечивает возможность осуществления большого количества точных изменений ДНК в широком диапазоне положений ДНК до или после смежного с протоспейсером мотива (PAM) участка-мишени. На **Фиг.38C** показано, что при связывании ДНК-мишени комплекс PE:PEgRNA осуществляет разрыв содержащей PAM цепи ДНК. Полученный свободный 3′-конец гибридизируется со связывающим праймер участком PEgRNA. Домен обратной транскриптазы катализирует праймерное удлинение с помощью RT матрицы PEgRNA, что обеспечивает образование новой синтезированной цепи ДНК, содержащей требуемое изменение (3′ флэп). Выравнивание между подвергнутым редактированию 3′ флэпом и не подвергнутым редактированию 5′ флэпом, содержащим исходную ДНК, с последующим клеточным расщеплением и лигированием 5′ флэпа, а также репарация или репликация ДНК для коррекции

гетеродуплексной ДНК приводит к получению стабильно отредактированной ДНК. На **Фиг.38D** показаны *in vitro* анализы праймерного удлинения 5′-удлиненной PEgRNA с подвергнутыми предварительному никированию субстратами дцДНК, содержащими меченные 5′-Cy5 цепи PAM, dCas9, и коммерческий вариант RT M-MLV (RT, Superscript III). dCas9 образовывал комплексы с PEgRNA, содержащими RT матрицу различной длины, затем добавлялся в субстраты ДНК наряду с указанными компонентами. Реакции инкубировали при 37 °C в течение 1 часа, затем анализировали с помощью денатурирующего электрофореза PAGE с применением мочевины и визуализировали на предмет флуоресценции Cy5. На **Фиг.38E** показаны анализы праймерного удлинения, выполненные как на Фиг.38D, с применением 3′-удлиненных PEgRNA, предварительно образовавших комплексы с никазой dCas9 или Cas9 H840A, и предварительно подвергнутых разрыву или не подвергнутых разрыву меченных 5′-Cy5 субстратов дцДНК. На **Фиг.38F** показаны колонии дрожжей, трансформированные с помощью репортерных плазмид в виде слитых GFP-mCherry, подвергнутых редактированию *in vitro* с помощью PEgRNA, никазы Cas9, и RT. Плазмиды, содержащие несмысловые или связанные со сдвигом рамки считывания мутации между GFP и mCherry, подвергались редактированию с помощью 5′-удлиненных или 3′-удлиненных PEgRNA, которые восстанавливают трансляцию mCherry за счет трансверсивной мутации, 1 bp вставки, или 1 bp делеции. Двойные положительные в плане GFP и mCherry клетки (желтые) соответствуют успешному редактированию.

[0122] На **Фиг.39A-39D** показано праймерное редактирование геномной ДНК в клетках человека с помощью PE1 и PE2. На **Фиг.39A** показаны PEgRNA, содержащие спейсерную последовательность, каркас sgРНК, и 3′ удлинение, содержащее участок связывания праймера (зеленым) и матрицу обратной транскрипции (RT) (фиолетовый), которая содержит подвергнутое редактированию основание (основания) (красный). Связывающий праймер участок гибридизируется с PAM-содержащей последовательностью ДНК непосредственно выше участка никирования. RT матрица является гомологичной последовательности ДНК ниже разрыва за исключением кодируемого изменения. На **Фиг.39B** показано внесение изменения в виде трансверсии T•A на A•T в участке HEK3 в клетках HEK293T с помощью никазы Cas9 H840A, слитой с обратной транскриптазой M-MLV дикого типа (PE1), и PEgRNA с различной длиной участка связывания праймера. На **Фиг.39C** показано, что использование

сконструированной пентамутантной обратной транскриптазы M-MLV (D200N, L603W, T306K, W313F, T330P) в PE2 значительно повышает эффективность трансверсии в рамках праймерного редактирования в пяти геномных участках в клетках HEK293T, а также изменений в виде малых вставок и малых делеций в HEK3. **Фиг.39D** представляет собой сравнение эффективности редактирования PE2 при различной длине RT матрицы в пяти геномных участках в клетках HEK293T. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0123] На **Фиг.40A-40C** показано никирование PE3 и PE3b системами не подвергаемой редактированию цепи для повышения эффективности праймерного редактирования. **Фиг.40A** представляет собой обзор праймерного редактирования с помощью PE3. После первоначального синтеза подвергаемой редактированию цепи репарация ДНК обеспечит удаление либо новой синтезированной цепи, содержащей изменение (эксцизию 3′ флэпа), или исходной цепи геномной ДНК (эксцизию 5′ флэпа). Эксцизия 5′ флэпа оставляет после себя гетеродуплекс ДНК, содержащий одну подвергнутую редактированию цепь и одну не подвергнутую редактированию цепь. Механизм репарации несовпадений или репликации ДНК может обеспечить коррекцию гетеродуплекса с получением либо подвергнутых редактированию, либо не подвергнутых редактирований продуктов. Никирование не подвергнутой редактированию цепи способствует репарации этой цепи, что приводит к предпочтительному образованию стабильной дуплексной ДНК, содержащей требуемое изменение. На **Фиг.40B** показано влияние никирования комплементарной цепи на эффективность опосредуемого PE3 праймерного редактирования и образование инделей. "Нет" обозначает контроли PE2, которые не приводят к образования разрыва комплементарной цепи. На **Фиг.40C** приведено сравнение эффективности редактирования с использованием PE2 (без разрыва комплементарной цепи), PE3 (произвольный разрыв комплементарной цепи), и PE3b (разрыв комплементарной цепи с учетом вносимого изменения). Все показатели редактирования отражают процентную долю всех считываний в рамках секвенирования, которые содержат требуемое изменение и не содержат инделей, среди всех обработанных клеток без сортировки. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0124] На **Фиг.41A-41K** показаны нацеленные вставки, делеции, и все 12 типов точечных мутаций с применением PE3 в семи эндогенных геномных локусах человека в клетках HEK293T. **Фиг.41A** представляет собой график, показывающий все 12 типов однонуклеотидных правок в виде транзиции и трансверсии от положения +1 до +8 (считая расположение индуцированного PEgRNA разрыва как находящееся между положением +1 и -1) участка HEK3 с помощью 10-nt RT матрицы. **Фиг.41B** представляет собой график, показывающий правки в виде трансверсии с помощью PE3 на большой дистанции в участке HEK3 с помощью 34-nt RT матрицы. **Фиг.41C-41H** представляют собой графики, показывающие все 12 типов изменений в виде транзиции и трансверсии в различных положениях в окне праймерного редактирования для (**Фиг. 41C**) RNF2, (**Фиг. 41D**) FANCF, (**Фиг. 41E**) EMX1, (**Фиг. 41F**) RUNX1, (**Фиг. 41G**) VEGFA, и (**Фиг. 41H**) DNMT1. **Фиг.41I** представляет собой график, показывающий направленные 1- и 3-bp вставки, и 1- и 3-bp делеции с помощью PE3 в семи эндогенных геномных локусах. **Фиг.41J** представляет собой график, показывающий нацеленные точные делеции длиной от 5 до 80 bp в участке-мишени HEK3. **Фиг.41K** представляет собой график, показывающий комбинацию правок в виде вставок и делеций, вставок и точечных мутаций, делеций и точечных мутаций, и двойных точечных мутаций в трех эндогенных геномных локусах. Все показатели редактирования отражают процентную долю всех считываний в рамках секвенирования, которые содержат требуемое изменение и не содержат инделей, среди всех обработанных клеток без сортировки. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0125] На **Фиг.42A-42H** показано сравнение праймерного редактирования и редактирования оснований, и нецелевого редактирования с помощью Cas9 и PE3 в известных нецелевых участках Cas9. На **Фиг.42A** показана общая эффективность редактирования в виде C•G на T•A в тех же нуклеотидах-мишенях для PE2, PE3, BE2max, и BE4max в эндогенных участках HEK3, FANCF, и EMX1 в клетках HEK293T. На **Фиг.42B** показана частота инделей после обработки в соответствии с **Фиг.42A**. На **Фиг.42C** показана эффективность редактирования точных изменений в виде C•G на T•A (без побочных изменений или инделей) для PE2, PE3, BE2max, и BE4max в HEK3, FANCF, и EMX1. Для EMX1 также показаны точные изменения с помощью PE в виде всех возможных комбинаций конвертации C•G на T•A в трех

нуклеотидах-мишенях. На **Фиг.42D** показана общая эффективность редактирования в виде A•T на G•C для PE2, PE3, ABEdmax, и ABEmax в HEK3 и FANCF. На **Фиг.42E** показана точная эффективность редактирования в виде A•T на G•C без побочных изменений или инделей для HEK3 и FANCF. На **Фиг.42F** показана частота инделей после обработки в соответствии с **Фиг.42D**. На **Фиг.42G** показана средняя эффективность редактирования трипликатов (процентная доля считываний в рамках секвенирования с инделями) в клетках HEK293T для нуклеазы Cas9 в четырех целевых и 16 известных нецелевых участках. Проанализированные 16 нецелевых участков представляли собой четыре основных ранее описанных нецелевых участка[118,159] для каждого из четырех целевых участков. Для каждого целевого участка Cas9 соединяли с sgРНК или с каждой из четырех PEgRNA, которые распознают тот же протоспейсер. На **Фиг.42H** показана средняя эффективность целевого и нецелевого редактирования трипликатов и эффективность в плане инделей (ниже в скобках) в HEK293T клетках для PE2 или PE3 относительно каждой PEgRNA на **Фиг.42G**. Показатели целевого редактирования отражают процентную долю всех считываний в рамках секвенирования, которые содержат требуемое изменение и не содержат инделей, среди всех обработанных клеток без сортировки. Показатели нецелевого редактирования отражают модификацию нецелевого локуса, соответствующую праймерному редактированию. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0126] На **Фиг.43A-43I** показано праймерное редактирования в различных линиях клеток человека и первичных нейронах коры головного мозга мыши, внесение и коррекция патогенных мутаций в виде трансверсии, вставки, или делеции, и сравнение праймерного редактирования и HDR**. Фиг.43A** представляет собой график, на котором показано внесение (путем трансверсии в виде T•A на A•T) и коррекция (путем трансверсии A•T на T•A) патогенной мутации E6V в HBB в клетках HEK293T. Показана коррекция либо к HBB дикого типа, либо к HBB, содержащему молчащую мутацию, которая разрушает PAM PEgRNA. **Фиг.43B** представляет собой график, на котором показано внесение (путем 4-bp вставки) и коррекция (путем 4-bp делеции) патогенного HEXA 1278+TATC аллеля в клетках HEK293T. Показана коррекция либо к HEXA дикого типа, либо к HEXA, содержащему молчащую мутацию, которая разрушает PAM PEgRNA. **Фиг.43C** представляет собой график, на котором показано внесение защитного варианта G127V в PRNP в клетках HEK293T с помощью

трансверсии в виде G•C на T•A. **Фиг.43D** представляет собой график, на котором показано праймерное редактирование в других линиях клеток человека, включая K562 (лейкозные клетки костного мозга), U2OS (клетки остеосаркомы), и HeLa (клетки рака шейки матки). **Фиг.43E** представляет собой график, показывающий внесение мутации в виде трансверсии G•C на T•A в DNMT1 первичных нейронов коры головного мозга мыши с помощью лентивирусной системы PE3 на основе двойного разделенного интеина, в рамках которой N-концевая половина представляет собой Cas9 (1-573), слитый с N-интеином, и через саморасщепляемый пептид P2A — с GFP–KASH, а C-концевая половина представляет собой C-интеин, слитый с остатком PE2. Половины PE2 экспрессируются из промотора синапсина человека, который является высокоспецифичным в отношении зрелых нейронов. Отсортированные значения отражают редактирование или индели из GFP-положительных клеток, в то время как неотсортированные значения относятся ко всем ядрам. **Фиг.43F** представляет собой сравнение эффективности PE3 и Cas9-опосредованного HDR редактирования в эндогенных геномных локусах в клетках HEK293T. **Фиг.43G** представляет собой сравнение эффективности PE3 и Cas9-опосредованного HDR редактирования в эндогенных геномных локусах в клетках K562, U2OS, и HeLa. **Фиг. 43H** представляет собой сравнение PE3 и Cas9-опосредованного образования побочных продуктов в виде инделей в клетках HEK293T, K562, U2OS, и HeLa. На **Фиг.43I** показана направленная вставка His6 метки (18 bp), метки в виде FLAG эпитопа (24 bp), или удлиненного участка LoxP (44 bp) в клетках HEK293T с помощью PE3. Все показатели редактирования отражают процентную долю всех считываний в рамках секвенирования, которые содержат требуемое изменение и не содержат инделей, среди всех обработанных клеток. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0127] На **Фиг.44A-44G** показаны *in vitro* исследования для анализа праймерного редактирования с флуоресцентно меченными субстратами ДНК. На **Фиг.44A** показаны анализы сдвига электрофоретической подвижности с dCas9**,** 5′-удлиненными PEgRNA и мечеными 5′ Cy5 субстратами ДНК. PEgRNA 1-5 содержат 15-nt последовательность линкера (линкер A для PEgRNA 1, линкер B для PEgRNA 2-5) между спейсером и PBS, 5-nt последовательность PBS, и RT матрицы длиной 7 nt (PEgRNA 1 и 2), 8 nt (PEgRNA 3), 15 nt (PEgRNA 4), и 22 nt (PEgRNA 5). PEgRNA соответствуют тем, которые указаны на **Фиг.44E** и **44F**; полные последовательности

перечислены в Таблицах 2A-2C. На **Фиг.44B** показаны *in vitro* анализы никирования Cas9 H840A с помощью 5′-удлиненных и 3′-удлиненных PEgRNA. На **Фиг.44C** показано опосредованное Cas9 образование инделей в клетках HEK293T в HEK3 с помощью 5′-удлиненных и 3′-удлиненных PEgRNA. На **Фиг.44D** показан обзор *in vitro* биохимических анализов праймерного редактирования. Тестировали меченные 5′ Cy5 субстраты дцДНК с предварительным никированием и без предварительного никирования. sgРНК, 5′-удлиненные PEgRNA или 3′-удлиненные PEgRNA подвергали предварительному образованию комплексов с никазой dCas9 или Cas9 H840A, затем комбинировали с субстратом дцДНК, RT M-MLV, и dNTP. Реакции оставляли для развития при 37 °C в течение 1 часа с последующим разделением с помощью денатурирующего электрофореза PAGE с применением мочевины и визуализацией на основе флуоресценции Cy5. На **Фиг.44E** показано, что реакции праймерного удлинения с использованием 5′-удлиненных PEgRNA, субстратов ДНК с предварительным никированием, и dCas9 приводят к значительной конверсии в продукты RT. На **Фиг.44F** показаны реакции праймерного удлинения с использованием 5′-удлиненных PEgRNA, как на **Фиг.44B**, с субстратом ДНК без предварительного никирования и никазой Cas9 H840A. Показатели выхода продукта значительно снижены по сравнению с субстратом с предварительным никированием. На **Фиг.44G** показано, что *in vitro* реакция праймерного удлинения с использованием 3′ PEgRNA приводит к одному значительному продукту на основе денатурирующего электрофореза PAGE с применением мочевины. Полосу RT продукта выделяли, элюировали из геля, затем подвергали добавлению гомополимерного хвоста с использованием концевой трансферазы (TdT) с помощью dGTP или dATP. Продукты с хвостом удлиняли с использованием поли-T или поли-C праймеров и секвенировали полученную ДНК. Результаты секвенирования по Сэнгеру показывают, что три нуклеотида, полученные из каркаса нРНК, подвергались обратной транскрипции (добавлены в виде конечных 3′ нуклеотидов в ДНК продукт). Следует обратить внимание, что в рамках экспериментов на основе праймерного редактирования в клетках млекопитающих вставка каркаса PEgRNA является намного более редкой, чем *in vitro* (Фиг.**56A-56D**), потенциально вследствие неспособности связанной обратной транскриптазы осуществлять доступ к Cas9-связанному каркасу направляющей РНК и/или клеточной эксцизии несовпадающих 3′-концов 3′ флэпов, содержащих последовательности каркаса PEgRNA.

[0128] На **Фиг.45A-45G** показана клеточная репарация 3′ ДНК-флэпов в дрожжах из *in vitro* реакций праймерного редактирования. На **Фиг.45A** показано, что репортерные плазмиды с двойным флуоресцентным белком содержат открытые рамки считывания GFP и mCherry, разделенные участком-мишенью, кодирующим стоп-кодон рамки считывания, +1 сдвиг рамки считывания, или -1 сдвиг рамки считывания. Реакции праймерного редактирования осуществляли *in vitro* с использованием никазы Cas9 H840A, PEgRNA, dNTP, и обратной транскриптазы M-MLV, и затем трансформировали в дрожжи. Колонии, которые содержат неотредактированные плазмиды, продуцируют GFP, но не mCherry. Колонии дрожжей, содержащие отредактированные плазмиды, продуцируют как GFP, так и mCherry в виде слитого белка. На **Фиг.45B** показано совмещение флуоресценции GFP и mCherry для колоний дрожжей, трансформированных с помощью репортерных плазмид, содержащих стоп-кодон между GFP и mCherry (неотредактированный отрицательный контроль, сверху), или не содержащих стоп-кодон или сдвиг рамки считывания между GFP и mCherry (предварительно отредактированный положительный контроль, снизу). На **Фиг.45C-45F** показана визуализация флуоресценции mCherry и GFP из колоний дрожжей, трансформированных с использованием продуктов *in vitro* реакции праймерного редактирования. На **Фиг.45C** показана коррекция стоп-кодона в виде трансверсии T•A на A•T с использованием 3′-удлиненной PEgRNA или 5′-удлиненной PEgRNA, как показано на **Фиг.45D**. На **Фиг.45E** показана коррекция +1 сдвига рамки считывания в виде 1-bp делеции с использованием 3′-удлиненной PEgRNA. На **Фиг.45F** показана коррекция -1 сдвига рамки считывания в виде 1-bp вставки с использованием 3′-удлиненной PEgRNA. На **Фиг.45G** показаны результаты секвенирования ДНК по Сэнгеру плазмид, выделенных из колоний только с GFP на **Фиг.45B**, и двойных положительных в плане GFP и mCherry колоний на **Фиг.45C**.

[0129] На **Фиг.46A-46F** показано корректное редактирование по сравнению с образованием инделей при использовании PE1. На **Фиг.46A** показана эффективность редактирования в виде трансверсии T•A на A•T и образование инделей при использовании PE1 в положении +1 HEK3 при использовании PEgRNA, содержащих 10-nt RT матрицы и последовательности PBS длиной 8-17 nt. На **Фиг.46B** показана эффективность редактирования в виде трансверсии G•C на T•A и образование инделей при использовании PE1 в положении +5 EMX1 при использовании PEgRNA, содержащих 13-nt RT матрицы и последовательности PBS длиной 9-17 nt. На **Фиг.46C**

показана эффективность редактирования в виде трансверсии G•C на T•A и образование инделей при использовании PE1 в положении +5 FANCF при использовании PEgRNA, содержащих 17-nt RT матрицы и последовательности PBS длиной 8-17 nt. На **Фиг.46D** показана эффективность редактирования в виде трансверсии C•G на A•T и образование инделей при использовании PE1 в положении +1 RNF2 при использовании PEgRNA, содержащих 11-nt RT матрицы и последовательности PBS длиной 9-17 nt. На **Фиг.46E** показана эффективность редактирования в виде трансверсии G•C на T•A и образование инделей при использовании PE1 в положении +2 HEK4 при использовании PEgRNA, содержащих 13-nt RT матрицы и последовательности PBS длиной 7-15 nt. На **Фиг.46F** показана опосредованная PE1 +1 делеция T, +1 вставка A, и +1 вставка CTT в участке HEK3 с использованием 13-nt PBS и 10-nt RT матрицы. Последовательности PEgRNA соответствуют тем, которые использованы на **Фиг.39C** (см. Таблицы 3A-3R). Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0130] На **Фиг.47A-47S** показан анализ вариантов RT M-MLV для праймерного редактирования. На **Фиг.47A** показаны аббревиатуры для вариантов праймерного редактора, используемого на данной фигуре. На **Фиг.47B** показаны изменения в виде направленной вставки и делеции с использованием PE1 в локусе HEK3. На **Фиг.47C-47H** показано сравнение 18 конструктов праймерного редактора, содержащих варианты RT M-MLV для их способности вносить изменение в виде +2 трансверсии G•C на C•G в HEK3, как показано на **Фиг.47C**, вставки 24-bp FLAG в HEK3, как показано на **Фиг.47D**, изменение в виде +1 трансверсии C•G на A•T в RNF2, как показано на **Фиг.47E**, изменение в виде +1 трансверсии G•C на C•G в EMX1, как показано на **Фиг.47F**, изменение в виде +2 трансверсии T•A на A•T в *HBB*, как показано на **Фиг.47G**, и изменение в виде +1 трансверсии G•C на C•G в FANCF, как показано на **Фиг.47H**. На **Фиг.47I-47N** показано сравнение четырех конструктов праймерного редактора, содержащих варианты M-MLV для их способности вносить изменения, показанные на **Фиг.47C-47H**, в рамках второго раунда независимых экспериментов. На **Фиг.47O-47S** показана эффективность редактирования PE2 в пяти геномных локусах с различной длиной PBS. На **Фиг.47O** показана +1 вариация T•A на A•T в HEK3. На **Фиг.47P** показана +5 вариация G•C на T•A в EMX1. На **Фиг.47Q** показана +5 вариация G•C на T•A в FANCF. На **Фиг.47R** показана +1 вариация C•G на A•T в RNF2. На **Фиг.47S** показана +2 вариация G•C на T•A в HEK4. Значения и планки

ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0131] На **Фиг.48A-48C** показаны свойства конструкции PBS PEgRNA и последовательности RT матрицы. На **Фиг.48A** показана эффективность опосредованного PE2 редактирования в виде +5 трансверсии G•C на T•A (синяя линия) в VEGFA в клетках HEK293T в виде функции длины RT матрицы. Для сравнения указаны индели (серая линия). Последовательность ниже графика показывает последний нуклеотид на основе матрицы для синтеза PEgRNA. Выделены G нуклеотиды (которым соответствуют матричные C в PEgRNA); для максимизации эффективности праймерного редактирования следует избегать RT матрицы, оканчивающиеся на C. На **Фиг.48B** показано изменение в виде +5 трансверсии G•C на T•A и индели для DNMT1, как на **Фиг.48A**. На **Фиг.48C** показано изменение в виде +5 трансверсии G•C на T•A и индели для RUNX1, как на **Фиг.48A**. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0132] На **Фиг.49A-49B** показаны эффекты PE2, PE2 R110S K103L, никазы Cas9 H840A, и dCas9 на жизнеспособность клеток. HEK293T клетки трансфицировали плазмидами, кодирующими PE2, PE2 R110S K103L, никазу Cas9 H840A, или dCas9, совместно с нацеленной на HEK3 PEgRNA плазмидой. Жизнеспособность клеток измеряли каждые 24 часа после трансфекции в течение 3 дней с помощью анализа CellTiter-Glo 2.0 (Promega). На **Фиг.49A** показана жизнеспособность, определенная на основе люминесценции на 1, 2, или 3 день после трансфекции. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов, каждый из которых был реализован в трех технических повторах. На **Фиг.49B** показано процентное значение редактирования и инделей для PE2, PE2 R110S K103L, никазы Cas9 H840A, или dCas9, совместно с нацеленной на HEK3 PEgRNA плазмидой, которая кодирует изменение в виде +5 G на A. Значения эффективности редактирования измеряли на 3 день после трансфекции на основе обработанных клеток наряду с клетками, которые использовали для анализа жизнеспособности на **Фиг.49A**. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0133] На **Фиг.50A-50B** показана опосредованная PE3 коррекция *HBB* E6V и коррекция *HEXA* 1278+TATC с использованием различных PEgRNA. На **Фиг.50A**

показан скрининг 14 PEgRNA на предмет коррекции аллеля *HBB* E6V в клетках HEK293T с использованием PE3. Все проанализированные PEgRNA обеспечивают конвертацию аллеля *HBB* E6V обратно в *HBB* дикого типа без внесения какой-либо молчащей мутации PAM. На **Фиг.50B** показан скрининг 41 PEgRNA на предмет коррекции аллеля *HEXA* 1278+TATC в клетках HEK293T с использованием PE3 или PE3b. Меченные PEgRNA *HEXA* обеспечивают коррекцию патогенного аллеля за счет 4-bp делеции со сдвигом, которая разрушает PAM и оставляет молчащую мутацию. Данные меченные PEgRNA *HEXA* обеспечивают коррекцию патогенного аллеля обратно в дикий тип. В записях, оканчивающихся на "b", используется специфичная в план изменения никирующая sgРНК в комбинации с PEgRNA (система PE3b). Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0134] На **Фиг.51A-51F** показана активность PE3 в линиях клеток человека и сравнение PE3 и Cas9-инициированной HDR. Эффективность осуществления корректного изменения (без инделей) и частота инделей для PE3 и Cas9-инициированной HDR в клетках HEK293T, как показано на **Фиг.51A**, клетках K562, как показано на **Фиг.51B**, клетках U2OS, как показано на **Фиг.51C**, и клетках HeLa, как показано на **Фиг.51D**. Каждый обведенный скобкой сравниваемый элемент обеспечивает внесение идентичных изменений при использовании PE3 и Cas9-инициированной HDR. Контроли без нацеливания представляют собой PE3 и PEgRNA, которые нацелены на не являющийся мишенью локус. На **Фиг.51E** показаны контрольные эксперименты с использованием PEgRNA+PE3 без нацеливания и с использованием dCas9+sgРНК при сравнении с экспериментами с использованием HDR на основе Cas9 дикого типа, которые подтверждают, что донорская HDR матрица оцДНК, являющаяся обычным контаминантом, который искусственно повышает видимые показатели эффективности HDR, не влияет на измерение HDR на **Фиг.51A-51D**. На **Фиг.51F** показаны примеры таблиц аллелей участка HEK3 из образцов геномной ДНК, выделенных из клеток K562 после редактирования с использованием PE3 или Cas9-инициированной HDR. Аллели секвенировали на Illumina MiSeq и анализировали с помощью CRISPResso2[178]. Контрольная последовательность HEK3 из данного участка показана сверху. Показаны таблицы аллелей для отрицательного контроля в виде PEgRNA без нацеливания, +1 вставки CTT в HEK3 с использованием PE3, и +1 вставки CTT в HEK3 с использованием Cas9-инициированной HDR. Частоты

аллелей и соответствующие показатели количества считываний секвенирования Illumina показаны для каждого аллеля. Показаны все аллели, наблюдаемые с частотой ≥0,20%. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0135] На **Фиг.52A-52D** показано распределение по длине патогенных вставок, дубликатов, делеций, и инделей в базе данных ClinVar. Обзор вариантов ClinVar скачивали из NCBI (Национального центра биотехнологической информации) 15 июля 2019 г. Показатели длины указанных вставок, делеций, и дубликатов вычисляли с помощью исходных и измененных аллелей, положений начала и конца вариантов, или подходящей идентификаторной информации в названии варианта. Варианты без указания перечисленной выше информации исключали из анализа. Показатели длины указанных инделей (отдельные варианты, которые включают как вставки, так и делеции, по сравнению с исходным геномом) вычисляли путем определения количества несовпадений или промежутков при наилучшем попарном выравнивании между исходными и измененными аллелями.

[0136] На **Фиг.53A-53B** показаны образцы FACS гейтирования для сортинга GFP-положительных клеток. Ниже приведены примеры исходных файлов, относящихся к пакетному анализу, описывающих стратегию сортинга, использованную для получения линий клеток *HEXA* 1278+TATC и *HBB* E6V HEK293T. Данные для визуализации были получены на цитометре Sony LE-MA900 с использованием программного обеспечения Cell Sorter v. 3.0.5. На Графике 1 показано гейтирование для клеток, которые не экспрессируют GFP. На График 2 показан пример сортинга экспрессирующих P2A–GFP клеток, используемый для выделения линий клеток *HBB* E6V HEK293T. Клетки HEK293T сначала гейтировали на популяции с использованием FSC-A/BSC-A (гейт A), затем сортировали на предмет синглетов с использованием FSC-A/FSC-H (гейт B). Живые клетки сортировали путем гейтирования DAPI-отрицательных клеток (гейт C). Клетки с уровнями флуоресценции GFP, которые были выше уровней соответствующих отрицательному контролю клеток, сортировали с помощью EGFP в качестве флуорохрома (гейт D). На **Фиг.53A** показаны клетки HEK293T (GFP-отрицательные). На **Фиг.53B** показан соответствующий график FACS гейтирования для клеток, экспрессирующих PE2–P2A–GFP. На **Фиг.53C** показаны генотипы для HEXA 1278+TATC гомозиготных клеток HEK293T. На **Фиг.53D** показаны таблицы аллелей для линий гомозиготных клеток HBB E6V HEK293T.

[0137] **Фиг.54** представляет собой схему, показывающую процедуру клонирования PEgRNA.

[0138] **Фиг.55A-55G** представляют собой схемы конструкций PEgRNA. На **Фиг.55A** показана простая диаграмма PEgRNA с доменами, отмеченными (слева) и связанными с nCas9 в участке генома (справа). На **Фиг.55B** показаны различные типы модификаций PEgRNA, которые, как ожидается, будут повышать активность. На **Фиг.55C** показаны модификации PEgRNA для повышения транскрипции более длинных РНК за счет выбора промотора и 5′, 3′ процессинга и терминации. На **Фиг.55D** показано удлинение системы P1, которое представляет собой пример модификации каркаса. На **Фиг.55E** показано, что встраивание синтетических модификаций в участок матрицы, или где-либо еще в рамках PEgRNA, может повышать активность. На **Фиг.55F** показано, что сконструированное встраивание минимальной вторичной структуры в матрицу может предотвращать образование более длинной более ингибиторной вторичной структуры. На **Фиг.55G** показано разделение PEgRNA с использованием второй последовательности матрицы, соединенной с РНК элементом на 3′-конце PEgRNA (слева). Встраивание элементов на 5′ или 3′-концах PEgRNA может стимулировать связывание RT.

[0139] На **Фиг.56A-56D** показано встраивание последовательности каркаса PEнHPK в локусы-мишени. HTS данные анализировали на предмет вставки последовательности каркаса PEgRNA, как описано на Фиг.60A-60B. На **Фиг.56A** показан анализ локуса EMX1. Показан % общих считываний в рамках секвенирования, содержащих один или несколько нуклеотидов последовательности каркаса PEgRNA в рамках вставки, примыкающей к RT матрице (слева); процентная величина общих считываний в рамках секвенирования, содержащих вставку последовательности каркаса PEgRNA указанной длины (в середине); и кумулятивная общая процентная величина вставки PEgRNA вплоть до и включая длину, указанную на оси X. На **Фиг.56B** показано то же, что и на **Фиг.56A**, но для FANCF. На **Фиг.56C** показано то же, что и на **Фиг.56A**, но для HEK3. На **Фиг.56D** показано то же, что и на **Фиг.56A**, но для RNF2. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических повторов.

[0140] На **Фиг.57A-57I** показаны эффекты PE2, PE2-dRT, и никазы Cas9 H840A на количество РНК в рамках транскриптома. Анализ клеточной РНК с уменьшенным содержанием рибосомной РНК, выделенной из клеток HEK293T, экспрессирующих

PE2, PE2-dRT, или никазу Cas9 H840A и нацеленную на PRNP или нацеленную на HEXA PEgRNA. РНК, соответствующие 14410 генам и 14368 генам детектировали в образцах PRNP и HEXA, соответственно. На **Фиг.57A-57F** показана диаграмма Volcano, демонстрирующая -log10 FDR-скорректированное p-значение против log2-кратного изменения содержания транскрипта для каждой РНК при сравнении (**Фиг.57A**) PE2 против PE2-dRT с нацеленной на PRNP PEgRNA, (**Фиг.57B**) PE2 против Cas9 H840A с нацеленной на PRNP PEgRNA, (**Фиг.57C**) PE2-dRT против Cas9 H840A с нацеленной на PRNP PEgRNA, (**Фиг.57D**) PE2 против PE2-dRT с нацеленной на HEXA PEgRNA, (**Фиг.57E**) PE2 против Cas9 H840A с нацеленной на HEXA PEgRNA, (**Фиг.57F**) PE2-dRT против Cas9 H840A с нацеленной на HEXA PEgRNA. Красные точки обозначают гены, которые демонстрируют ≥2-кратное изменение относительного содержания, которое является статистически значимым (FDR-скорректированное p-значение < 0,05). **Фиг.57G-57I** представляют собой диаграммы Венна транскриптов с повышенной и пониженной регуляцией (≥2-кратное изменение) при сравнении образцов PRNP и HEXA на (**Фиг.57G**) PE2 против PE2-dRT, (**Фиг.57H**) PE2 против Cas9 H840A, и (**Фиг.57I**) PE2-dRT против Cas9 H840A.

[0141] На **Фиг.58** показано репрезентативное FACS гейтирование для сортинга ядер нейронов. Ядра последовательно гейтировали на основе сигнала DyeCycle Ruby, соотношения FSC (переднее рассеивание)/SSC (боковое рассеивание), соотношения ширина SSC/высота SSC, и соотношения GFP/DyeCycle.

[0142] На **Фиг.59A-59F** показан протокол клонирования 3′-удлиненных PEgRNA в векторы экспрессии U6 млекопитающих с помощью сборки Golden Gate. На **Фиг.59A** показан обзор клонирования. На **Фиг.59B** показан "Этап 1: Осуществить расщепление pU6-PEgRNA-GG-векторной плазмиды (компонент 1)". На **Фиг.59C** показаны "Этапы 2 и 3: осуществить упорядочивание и отжиг нуклеотидных частей (компоненты 2, 3, и 4)". На **Фиг.59D** показан "Этап 2.b.ii.: фосфорилирование каркаса sgРНК (не обязательно, если нуклеотиды были приобретены в фосфорилированном виде)". На **Фиг.59E** показан "Этап 4: сборка PEgRNA". На **Фиг.59F** показаны "Этапы 5 и 6: Трансформация собранных плазмид". На **Фиг.59G** показана диаграмма с обзором протокола клонирования PEgRNA.

[0143] На **Фиг.60A-60B** показан скрипт на языке Python для количественного анализа интеграции каркаса PEgRNA. Был написан специальный скрипт на языке Python для описания и количественного анализа вставок PEgRNA в геномных локусах-мишенях.

Скрипт в интерактивном режиме осуществляет сравнение текстовых срок увеличивающейся длины, взятых из исходной последовательности (последовательности каркаса направляющей РНК), со считываниями секвенирования в файлах fastq и обеспечивает подсчет количества считываний секвенирования, которые соответствуют поисковому запросу. Каждая последующая текстовая строка соответствует дополнительному нуклеотиду последовательности каркаса направляющей РНК. Таким образом вычисляли интеграции точной длины и кумулятивные интеграции до определенной длины. В начало исходной последовательности включаются от 5 до 6 оснований 3′-конца новой цепи ДНК, синтезированной с помощью обратной транскриптазы, для обеспечения выравнивания и точного подсчета коротких фрагментов sgРНК.

[0144] **Фиг.61** представляет собой график, показывающий процентное значение всех считываний секвенирования с указанным изменением для SaCas9(N580A)- RT MMLV HEK3 +6 C>A. Показаны значения для корректных изменений, а также инделей.

[0145] На **Фиг.62A-62B** продемонстрирована важность протоспейсера для эффективного внесения требуемого изменения в точном месторасположении с помощью праймерного редактирования. **Фиг.62A** представляет собой график, показывающий процентное значение всех считываний секвенирования с представляющими собой мишени парами оснований T•A, конвертированными в A•T, для различных локусов HEK3. **Фиг.62B** представляет собой анализ последовательности, показывающий то же самое.

[0146] **Фиг.63** представляет собой график, показывающий варианты PAM SpCas9 при редактировании PAM (N=3). Процентное значение всех считываний секвенирования с нацеленным изменением PAM показано для SpCas9(H840A)-VRQR-RT MMLV, где NGA > NTA, и для SpCas9(H840A)-VRER-RT MMLV, где NGCG > NTCG. Приведена длина участка связывания праймера (PBS) PEgRNA , длина RT матрицы (RT), и использованная система PE.

[0147] **Фиг.64** представляет собой схему, показывающую внесение различных специфичных в плане участка мишеней рекомбиназы (SSR) в геном с помощью PE. (a) показана общая схема вставки последовательности-мишени рекомбиназы с помощью праймерного редактора. (b) показано, как отдельная SSR-мишень, вставленная с помощью PE, может применяться в качестве участка для геномного встраивания донорской матрицы ДНК. (c) показывает, как тандемная вставка участков-мишеней

SSR может использоваться для делеции части генома. (d) показано, как тандемная вставка участков-мишеней SSR может использоваться для инверсии части генома. (e) показано, как вставка двух участков-мишеней SSR в двух удаленных хромосомных участках может приводить к хромосомной транслокации. (f) показано, как вставка двух различных участков-мишеней SSR в геноме может использоваться для обмена кассетами из донорской матрицы ДНК. См. Пример 17 для дополнительных подробностей.

[0148] На **Фиг.65** показан 1) PE-опосредованный синтез участка-мишени SSR в геноме клетки человека и 2) использование данного участка-мишени SSR для интеграции донорской матрицы ДНК, включающей маркер экспрессии GFP. После успешной интеграции GFP заставляет клетку флуоресцировать. См. Пример 17 для дополнительных подробностей.

[0149] На **Фиг.66** изображен один вариант осуществления праймерного редактора, представленного в виде двух половин белка PE, которые соединяются в виде цельного праймерного редактора за счет аутосплайсингового действия половин разделенного интеина, расположенных в конце или начале каждой из половин белка праймерного редактора.

[0150] На **Фиг.67** изображен механизм удаления интеина из полипептидной последовательности и восстановление пептидной связи между N-концевыми и C-концевыми последовательностями экстеина. (a) изображен общий механизм двух половин белка, каждая из которых содержит половину последовательности интеина, которые при контакте в клетке приводят к образованию полностью функционального интеина, который затем подвергается аутосплайсингу и эксцизии. Процесс эксцизии приводит к образованию пептидной связи между N-концевой половиной белка (или "N экстеином") и C-концевой половиной белка (или "C экстеином") для образования цельного единого полипептида, включающего части в виде N экстеина и C экстеина. В различных вариантах осуществления N экстеин может соответствовать N-концевой половине разделенного слитого белка праймерного редактора, а C экстеин может соответствовать C-концевой половине разделенного праймерного редактора. (b) показан химический механизм эксцизии интеина и восстановление пептидной связи, которая соединяет половину в виде N экстеина (обозначенная красным половина) и половину в виде C экстеина (обозначенная синим половина). Эксцизия разделенных интеинов (то есть N интеина и C интеина в рамках конфигурации с разделенным

интеином) также может быть обозначена как "транс-сплайсинг", поскольку она включает действие в виде сплайсинга двух отдельных компонентов, поступающих *in trans*.

[0151] На **Фиг.68A** демонстрируется, что доставка обеих половин разделенного интеина SpPE (SEQ ID NO: 762) в линкере обеспечивает сохранение активности в трех тестовых локусах при совместной трансфекции в клетки HEK293T.

[0152] На **Фиг.68B** демонстрируется, что доставка обеих половин разделенного интеина SaPE2 (например SEQ ID NO: 443 и SEQ ID NO: 450) обеспечивает восстановление активности полноразмерного SaPE2 (SEQ ID NO: 134) при совместной трансфекции в клетки HEK293T. Остатки, указанные в кавычках, представляют собой последовательность аминокислот 741-743 в SaCas

[0153] 9 (первые остатки C-концевого экстеина), которые являются важными для реакции транс-сплайсинга интеина. "SMP" представляют собой нативные остатки, которые также подвергаются мутации в "CFN" консенсусную сплайсинговую последовательность. Показано, что консенсусная последовательность обеспечивает наиболее высокое восстановление при измерении на основе процентной величины праймерного редактирования.

[0154] На **Фиг.68C** приведены данные, которые показывают, что различные описанные рибонуклеопротеиновые PE комплексы (PE2 с высокой концентрацией, PE3 с высокой концентрацией и PE3 с низкой концентрацией) могут быть доставлены подобным образом.

[0155] На **Фиг.69** показан анализ на основе бактериофаговых бляшек для определения эффективности PE в рамках PANCE. Бляшки (темные круги) обозначают фаг, способный успешно инфицировать *E. coli*. Повышение концентрации L-рамнозы приводит к повышенной экспрессии PE и повышению образования бляшек. Секвенирование бляшек выявило присутствие внесенного с помощью PE геномного изменения.

[0156] На **Фиг.70A-70I** представлен пример подвергнутой редактированию последовательности-мишени в качестве иллюстрации пошаговой инструкции для конструирования PEgRNA и никирующих sgРНК для праймерного редактирования.

Фиг.70A: Этап 1. Определение последовательности-мишени и изменения. Получить последовательность участка-мишени ДНК (~200bp) с центром вокруг месторасположения требуемого изменения (точечной мутации, вставки, делеции, или

их комбинации). **Фиг.70B: Этап 2. Поиск PAM-мишеней**. Идентифицировать PAM поблизости от месторасположения изменения. Осуществлять поиск PAM на обеих цепях. В то время как предпочтительны PAM поблизости от положения изменения, является возможным внести изменения с помощью протоспейсеров и PAM, которые обеспечивают расположение разрыва на расстоянии ≥ 30 nt от положения изменения. **Фиг.70C: Этап 3. Поиск участков разрыва**. Для каждого рассматриваемого PAM идентифицировать соответствующий участок разрыва. Для никазы Sp Cas9 H840A расщепление осуществляется в PAM-содержащей цепи между 3-м и 4-м основаниями 5′ относительно PAM NGG. Все подвергнутые редактированию нуклеотиды должны быть расположены 3′ относительно участка разрыва, так что подходящие PAM должны обеспечивать расположение разрыва 5′ относительно изменения-мишени на PAM-содержащей цепи. В показанном ниже примере присутствуют два возможных PAM. Для простоты остальные этапы продемонстрируют конструирование PEgRNA с использованием только PAM 1. **Фиг.70D: Этап 4. Конструирование спейсерной последовательности**. Протоспейсер Sp Cas9 соответствует 20 нуклеотидам 5′ относительно PAM NGG на PAM-содержащей цепи. Эффективное инициирование транскрипции Pol III требует, чтобы G был первым транскрибированным нуклеотидом. Если первый нуклеотид протоспейсера представляет собой G, то спейсерная последовательность для PEgRNA будет просто представлять собой протоспейсерную последовательность. Если первый нуклеотид протоспейсера не представляет собой G, то спейсерная последовательность PEgRNA представляет собой G и далее протоспейсерную последовательность. **Фиг.70E: Этап 5. Конструирование участка связывания праймера (PBS)**. Идентифицировать праймер ДНК на PAM-содержащей цепи с помощью последовательности начального аллеля. 3′-конец праймера ДНК представляет собой нуклеотид, расположенный непосредственно выше участка разрыва (то есть 4-го основания 5′ относительно PAM NGG для Sp Cas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера (PBS) PEgRNA, содержащий от 12 до 13 нуклеотидов комплементарности относительно праймера ДНК, может использоваться для последовательностей, которые содержат ~40-60% содержимого в виде GC. Для последовательностей с низким содержанием GC должны быть протестированы более длинные PBS (от 14 до 15 nt). Для последовательностей с высоким содержанием GC должны быть протестированы более короткие PBS (от 8 до 11 nt). Оптимальные последовательности PBS должны

быть определены эмпирически независимо от содержания GC. Для конструирования PBS последовательности длиной *p* следует взять обратный комплемент первых *p* нуклеотидов 5′ участка разрыва в PAM-содержащей цепи с использованием последовательности *начального* аллеля. **Фиг.70F: Этап 6. Конструирование RT матрицы**. RT матрица кодирует сконструированное изменение и гомологию относительно последовательности, смежной с изменением. Оптимальные величины длины RT матрицы варьируются в зависимости от участка-мишени. Для коротких изменений (положения от +1 до +6) рекомендуется тестировать короткую (от 9 до 12 nt), среднюю (от 13 до 16 nt), и длинную (от 17 до 20 nt) RT матрицу. Для длинных изменений (положения +7 и более) рекомендуется использовать RT матрицы, которые имеют длину, по меньшей мере, 5 nt (предпочтительно 10 или более nt) после положения изменения для обеспечения достаточной гомологии 3′ ДНК-флэпа. Для длинных изменений несколько RT матриц должны быть подвергнуты скринингу для идентификации функциональных конструкций. Для более крупных вставок и делеций (≥5 nt) рекомендуется встраивание большей 3′ гомологии (~20 nt или более) в RT матрицу. Эффективность редактирования обычно ухудшается, когда RT матрица кодирует синтез G в качестве последнего нуклеотида в обратно транскрибированном ДНК продукте (соответствует C в RT матрице PEgRNA). Поскольку многие RT матрицы поддерживают эффективное праймерное редактирование, при конструировании RT матриц рекомендуется не использовать G в качестве последнего синтезируемого нуклеотида Для конструирования последовательности RT матрицы длиной *r* следует использовать последовательность *требуемого* аллеля и взять обратный комплемент первых *r* нуклеотидов 3′ участка разрыва в цепи, которая первоначально содержала PAM. Следует обратить внимание, что по сравнению с изменениями SNP, изменения в виде вставки или делеции с использованием RT матриц той же длины не будут содержать идентичную гомологию. **Фиг.70G: Этап 7. Сборка полной последовательности PEgRNA**. Следует конкатенировать компоненты PEgRNA в следующем порядке (от 5′ к 3′): спейсер, каркас, RT матрица и PBS. **Фиг.70H: Этап 8. Конструирование никирующих sgРНК для PE3**. Идентифицировать PAM на не подвергаемой редактированию цепи выше и ниже изменения. Оптимальные положения никирования имеют высокую зависимость от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии от 40 до 90 нуклеотидов 5′ относительно положения PEgRNA-

индуцированного разрыва, ведут к более высоким выходам редактирования и меньшему количеству инделей. Никирующая sgРНК имеет спейсерную последовательность, которая совпадает с протоспейсером длиной 20 nt в *начальном* аллеле, с добавлением 5′ G, если протоспейсер не начинается на G. **Фиг.70I: Этап 9. Конструирование никирующих sgРНК PE3b**. Если в комплементарной цепи присутствует PAM и его соответствующий протоспейсер накладывается на последовательность, являющуюся мишенью редактирования, то такое изменение может являться кандидатом для системы PE3b. В системе PE3b спейсерная последовательность никирующей sgРНК совпадает с последовательностью требуемого подвергаемого редактированию аллеля, но не начального аллеля. Система PE3b работает эффективно, когда подвергаемый редактированию нуклеотид (нуклеотиды) попадают в участок затравки (~10 nt от PAM) протоспейсера никирующей sgРНК.Это предотвращает никирование комплементарной цепи до готовности подвергнутой редактированию цепи, предотвращая конкурирование между PEgRNA и sgРНК за связывание с ДНК-мишенью. PE3b также обеспечивает избежание образования одновременных разрывов на обеих цепях, тем самым значительно снижая образование инделей, в то же время сохраняя высокую эффективность редактирования. sgРНК PE3b должны иметь спейсерную последовательность, которая совпадает с протоспейсером длиной 20 nt в *требуемом* аллеле, с добавлением 5′ G, если это необходимо.

[0157] На **Фиг.71A** показана нуклеотидная последовательность молекулы PEgRNA SpCas9 (сверху), которая терминируется на 3′-конце в "UUU" и не содержит элемент в виде опорной петли. Нижняя часть фигуры изображает ту же молекулу PEgRNA SpCas9, но которая дополнительно модифицирована для включения элемента в виде опорной петли, имеющего последовательность 5′-“GAAANNNNN”-3′, вставленную непосредственно перед "UUU" 3′-концом. "N" может представлять собой любое азотистое основание.

[0158] На **Фиг.71B** показаны результаты Примера 18, который демонстрирует, что эффективность праймерного редактирования в клетках HEK или клетках EMX повышается с применением PEgRNA, содержащей элементы в виде опорной петли, в то время как процентная величина образования инделей в значительной степени не меняется.

[0159] На **Фиг.72** изображены альтернативные конфигурации PEgRNA, которые могут применяться в рамках праймерного редактирования. (a) Изображен вариант

осуществления PE2:PEgRNA для праймерного редактирования. Данный вариант осуществления включает PE2 (слитый белок, включающий Cas9 и обратную транскриптазу) в виде комплекса с PEgRNA (как также описано на Фиг.1A-1I и/или Фиг.3A-3E). В данном варианте осуществления матрица для обратной транскрипции встраивается в 3' плечо удлинения на sgРНК для получения PEgRNA, а фермент в виде ДНК-полимеразы представляет собой обратную транскриптазу (RT), слитую напрямую с Cas9. (b) Изображен вариант осуществления MS2cp-PE2:sgРНК + tPERT. Данный вариант осуществления включает продукт слияния PE2 (Cas9 + обратная транскриптаза), который дополнительно слит с белком оболочки бактериофага MS2 (MS2cp) для образования слитого белка MS2cp-PE2. Для реализации праймерного редактирования слитый белок MS2cp-PE2 образует комплекс с sgРНК, которая нацеливает комплекс на определенный участок-мишень в ДНК. Затем данный вариант осуществления предусматривает внесение РНК матрицы для *trans* праймерного редактирования ("tPERT"), которая действует вместо PEgRNA за счет предоставления участка связывания праймера (PBS) и матрицы для синтеза ДНК на отдельной молекуле, то есть tPERT, которая также имеет аптамер MS2 (структуру типа "стебель-петля"). Белок MS2cp рекрутирует tPERT за счет связывания с аптамером MS2 молекулы. (c) Изображены альтернативные варианты конструирования PEgRNA, которые могут быть реализованы с помощью известных способов химического синтеза молекул нуклеиновой кислоты. Например, химический синтез может быть использован для синтеза гибридной РНК/ДНК молекулы PEgRNA для применения в рамках праймерного редактирования, при этом плечо удлинения гибридной PEgRNA представляет собой ДНК вместо РНК. В подобном варианте осуществления ДНК-зависимая ДНК-полимераза может использоваться вместо обратной транскриптазы для синтеза 3′ ДНК-флэпа, включающего требуемое генетическое изменение, который образован с помощью праймерного редактирования. В другом варианте осуществления плечо удлинения может быть синтезировано для включения химического линкера, который предотвращает использование ДНК-полимеразой (например обратной транскриптазой) каркаса или основы sgРНК в качестве шаблона. В другом варианте осуществления плечо удлинения может включать матрицу для синтеза ДНК, которая имеет обратную ориентацию относительно общей ориентации молекулы PEgRNA. В качестве примера и как показано для PEgRNA в ориентации от 5′ к 3′ и с удлинением, присоединенным к 3′-концу каркаса sgРНК, матрица для синтеза

ДНК ориентируется в противоположном направлении, то есть в направлении от 3′ к 5′. Данный вариант осуществления может представлять собой преимущество для вариантов осуществления PEgRNA с плечами удлинения, расположенными на 3'-конце нРНК. За счет инверсии ориентации плеча удлинения синтез ДНК-полимеразой (например обратной транскриптазой) будет терминирован при достижении ей плеча удлинения с новой 5’ ориентацией и будет, таким образом, обеспечивать устранение риска использования ядра нРНК в качестве матрицы.

[0160] На **Фиг.73** показано праймерное редактирование с использованием tPERT и системы рекрутирования MS2 (также известной как способ мечения MS2). sgРНК, нацеливающая белок праймерного редактора (PE2) на локус-мишень, экспрессируется в комбинации с tPERT, содержащей участок связывания праймера (PBS длиной 13 nt или 17 nt), RT матрицей, кодирующей вставку в виде His6 метки и плечо гомологии, и аптамером MS2 (расположенным на 5′ или 3′-конце молекулы tPERT). Используется белок праймерного редактора (PE2) или MS2cp, слитый с N-концом PE2. Редактирование осуществляли с/без использования никирующей комплементарную цепь sgРНК, как в ранее разработанной системе PE3 (обозначенной на оси X метками "PE2+разрыв" или "PE2", соответственно). Это также обозначается и определяется здесь как "никирование второй цепи".

[0161] На **Фиг.74** показана экспрессия аптамера MS2 обратной транскриптазы *in trans* и его рекрутирование системой аптамера MS2. PEgRNA содержит РНК аптамер MS2, вставленный в любую одну из двух шпилек каркаса sgРНК. Обратная транскриптаза M-MLV дикого типа экспрессируется в виде N-концевого или C-концевого продукта слияния с белком оболочки MS2 (MCP). Редактирование осуществляется в участке HEK3 в клетках HEK293T.

[0162] На **Фиг.75** приведена гистограмма для сравнения эффективности (то есть "% общих считываний секвенирования с указанным изменением или инделями") PE2, PE2-trunc, PE3, и PE3-trunc в различных участках-мишенях в различных линиях клеток. Данные показывают, что праймерные редакторы, включающие усеченные варианты RT, были примерно настолько же эффективны, что и праймерные редакторы, включающие неусеченные RT белки.

[0163] На **Фиг.76** демонстрируется эффективность редактирования праймерных редакторов на основе разделенного интеина в Примере 20. Клетки HEK239T трансфицировали плазмидами, кодирующими полноразмерный PE2 или PE2 на основе

разделенного интеина, PEgRNA и никирующую направляющую РНК. Указана консенсусная последовательность (бóльшая часть амино-концевых остатков С-концевого экстеина). Показана процентная величина редактирования в двух участках: вставка CTT HEK3 +1 и G на T PRNP +6. Повторы независимых трансфекций n=3. См. Пример 20.

[0164] На **Фиг.**77 демонстрируется эффективность редактирования праймерных редакторов на основе разделенного интеина в Примере 20. Редактирование анализировали с помощью направленного глубокого секвенирования в основной части коры головного мозга и GFP+ субпопуляции при доставке 5E10vg в расчете на половину SpPE3 и малого количества 1E10 локализованного в ядрах GFP:KASH P0 мышам с помощью внутрицеребровентрикулярной (ICV) инъекции. Редакторы и GFP упаковывали в AAV9 с использованием EFS промотора. Мышей собирали спустя три недели после инъекции и выделяли GFP+ ядра с помощью проточной цитометрии. Показаны отдельные точки данных с анализом 1-2 мышей на одно состояние. См. Пример 20.

[0165] На **Фиг.78** демонстрируется эффективность редактирования праймерных редакторов на основе разделенного интеина в Примере 20. В частности, фигуры изображают конструкты в виде разделенного AAV-SpPE3, использованные в Примере 20. Совместная трансдукция с использованием AAV частиц, которые по отдельности экспрессируют SpPE3-N и SpPE3-C, восстанавливает активность PE3. Следует отметить, что N-концевой геном содержит U6-sgРНК кассету, экспрессирующую никирующую sgРНК, а С-концевой геном содержит U6-PEgRNA кассету, экспрессирующую PEgRNA. См. Пример 20.

[0166] На **Фиг.79** показана эффективность редактирования определенных оптимизированных линкеров, как описано в Примере 21. В частности, данные показывают эффективность редактирования конструкта PE2 с текущим линкером (обозначенным как PE2 – белый прямоугольник) по сравнению с различными версиями с линкером, замещенным последовательностью, как указано в локусах HEK3, EMX1, FANCF, RNF2 для соответствующих PEgRNA для изменений в виде транзиции, трансверсии, вставки, и делеции. Замещающие линкеры обозначены как "1x SGGS", "2x SGGS", "3x SGGS", "1x XTEN", "линкер отсутствует", "1x Gly", "1x Pro", "1x EAAAK", "2x EAAAK", и "3x EAAAK". Эффективность редактирования измеряется в формате гистограммы относительно "контрольной" эффективности редактирования

PE2. Линкер PE2 представляет собой SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127). Все редактирование осуществлялось в контексте системы PE3, то есть которая соответствует конструкту для редактирования PE2 плюс добавлению оптимальной вторичной направляющей никирования sgРНК. См. Пример 21.

[0167] На **Фиг.80** показан график средней кратной эффективности относительно выхода PE2, указывающий на то, что использование последовательности линкера 1x XTEN повышает эффективность редактирования в среднем в 1,14 раза (n=15). См. Пример 21.

[0168] На **Фиг.81** изображен уровень транскрипции PEgRNA из различных промоторов, как описано в Примере 22.

[0169] **Фиг.82** Как показано в Примере 22, влияние различных типов модификаций структуры PEgRNA на эффективность редактирования по сравнению с немодифицированной PEgRNA.

[0170] **Фиг.83** Показан эксперимент в виде PE, целью которого являлось редактирование гена HEK3, в частности вставка длиной 10 nt в положении +1 относительно участка разрыва, и использование PE3. См. Пример 22.

[0171] На **Фиг.84A** изображен пример PEgRNA, имеющей спейсер, ядро нРНК, и плечо удлинения (RT матрица + участок связывания праймера), которая модифицирована на 3'-конце с использованием молекулы тРНК, присоединенной с помощью линкера в виде UCU. тРНК включает различные посттрансляционные модификации. Однако указанная модификация не является обязательной.

[0172] На **Фиг.84B** изображена структура тРНК, которая может использоваться для модификации структур PEgRNA. См. Пример 22. P1 может иметь различную длину. P1 может быть удлинен для обеспечения предотвращения процессинга РНКазой P продукта слияния в виде PEgRNA-тРНК.

[0173] На **Фиг.85** показан эксперимент в виде PE, целью которого являлось редактирование гена FANCF, в частности конверсия G на T в положении +5 относительно участка разрыва, и использование конструкта PE3. См. Пример 22.

[0174] На **Фиг.86** показан эксперимент в виде PE, целью которого являлось редактирование гена HEK3, в частности вставка FLAG метки длиной 71 nt в положении +1 относительно участка разрыва, и использование конструкта PE3. См. Пример 22.

[0175] **Фиг.87** Результаты скрининга в клетках N2A, где PEgRNA обеспечивает внесение 1412Adel, с подробной информацией о длине участка связывания праймера (PBS) и длине матрицы в виде обратной транскриптазы (RT). (Показаны с/без инделей). См. Пример 23.

[0176] **Фиг.88** Результаты скрининга в клетках N2A, где PEgRNA обеспечивает внесение 1412Adel, с подробной информацией о длине участка связывания праймера (PBS) и длине матрицы в виде обратной транскриптазы (RT). (Показаны с/без инделей). См. Пример 23.

[0177] На **Фиг.89** показаны результаты редактирования в промежуточном локусе в гене β-глобина и в HEK3 в здоровых HSC при различной концентрации редактора относительно PEgRNA и никирующей нРНК. См. Пример 23.

ОПРЕДЕЛЕНИЯ

[0178] Если не указано иное, все используемые здесь технические и научные термины имеют значение, которое обычно понимается специалистом в области, к которой относится данное изобретение. В нижеследующих ссылках для специалиста приведено общее определение многих терминов, используемых в данном изобретении: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); и Hale & Marham, The Harper Collins Dictionary of Biology (1991). В соответствии с использованным здесь значением, нижеследующие термины имеют приписываемое им значение, если не указано иное.

Антисмысловая цепь

[0179] В генетике "антисмысловая" цепь сегмента в рамках двуцепочечной ДНК представляет собой матричную цепь, которая, как считается, проходит в ориентации от 3' к 5'. С другой стороны, "смысловая" цепь представляет собой сегмент в рамках двуцепочечной ДНК, которая проходит от 5' к 3' и которая является комплементарной антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' к 5'. В случае сегмента ДНК, который кодирует белок, смысловая цепь представляет собой цепь ДНК, которая имеет ту же последовательность, что и иРНК, которая использует антисмысловую цепь в качестве ее матрицы при транскрипции и в конечном итоге (как правило, но не всегда) подвергается трансляции в белок. Антисмысловая цепь, таким образом, отвечает за РНК, которая позже транслируется в белок, в то время как

смысловая цепь имеет практически такой же состав, что и иРНК. Следует обратить внимание, что для каждого сегмента дцДНК вероятно будет присутствовать два набора смысловых и антисмысловых цепей в зависимости от направления чтения (поскольку смысл и антисмысл зависят от перспективы). В конечном итоге то, какая цепь одного сегмента дцДНК считается смысловой или антисмысловой, определяется генным продуктом или иРНК.

Биспецифический лиганд

[0180] В соответствии с используемым здесь значением термин "биспецифический лиганд" или "биспецифический фрагмент" обозначает лиганд, который связываются с двумя различными лиганд-связывающими доменами. В определенных вариантах осуществления лиганд представляет собой низкомолекулярное соединение, или пептид, или полипептид. В других вариантах осуществления лиганд-связывающий фрагмент представляет собой "домен димеризации", который может быть внесен в виде пептидной метки в белок. В различных вариантах осуществления два белка, каждый из которых включает один и тот же или различные домены димеризации, могут быть индуцированы для димеризации за счет связывания каждого домена димеризации с биспецифическим лигандом. В соответствии с используемым здесь значением "биспецифические лиганды" могут быть равным образом обозначены как "химические индукторы димеризации" или "CID".

Cas9

[0181] Термин "Cas9" или "нуклеаза Cas9" обозначает РНК-направляемую нуклеазу, включающую домен Cas9 или его фрагмент (например белок, включающий активный или неактивный домен расщепления ДНК Cas9, и/или домен связывания нРНК Cas9). Термин "домен Cas9" в соответствии с используемым здесь значением обозначает фрагмент белка, включающий активный или неактивный домен расщепления Cas9 и/или домен связывания нНРК Cas9. "Белок Cas9" представляет собой полноразмерный белок Cas9. Нуклеазу Cas9 также иногда обозначают как нуклеаза casn1 или CRISPR (Короткие Регулярно Расположенные Группами Палиндромные Повторы, Clustered Regularly Interspaced Short Palindromic Repeat)-ассоциированная нуклеаза. CRISPR представляет собой адаптивную иммунную систему, которая обеспечивает защиту против подвижных генетических элементов (вирусов, элементов с непостоянной локализацией, и конъюгативных плазмид). Кластеры CRISPR содержат спейсеры, которые представляют собой последовательности, комплементарные предшествующим

подвижным элементам, атакующие мишени нуклеиновые кислоты. Кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В системах CRISPR типа II корректный процессинг пре-crРНК требует транс-кодируемую малую РНК (tracrРНК), эндогенную рибонуклеазу 3 (rnc) и домен Cas9. tracrРНК служит в качестве направляющей для осуществляемого с помощью рибонуклеазы 3 процессинга пре-crРНК. Далее Cas9/crРНК /tracrРНК обеспечивает расщепление эндонуклеазой линейной или циркулярной дцДНК-мишени, являющейся комплементарной спейсеру. Цепь-мишень, которая не является комплементарной crРНК, сначала разрезается эндонуклеазой, затем подрезается 3′-5′ экзонуклеазой. В природе связывание и расщепление ДНК обычно требует белок и обе РНК. Однако одиночные направляющие РНК ("sgРНК" или просто "нРНК") могут быть сконструированы для включения аспектов как crРНК, так и tracrРНК в виде одиночной РНК. См., например, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полное содержание которого приведено здесь в качестве ссылки. Cas9 распознает короткий мотив в последовательностях повторов CRISPR (PAM или смежный с протоспейсером мотив), чтобы помогать отличать себя от чужого. Последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной области (см., например, , “Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.” Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); “CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.” Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); и “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полное содержание каждого из которых приведено здесь в качестве ссылки). Ортологи Cas9 были описаны у различных видов, включая, но без ограничения, *S. pyogenes* и *S. thermophilus*. Дополнительные подходящие нуклеазы Cas9 и последовательности будут очевидны специалистам в данной области на основе данного описания, при этом такие нуклеазы Cas9 и последовательности включают последовательности Cas9 из организмов и локусов, описанных в Chylinski, Rhun, and Charpentier, “The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems”

(2013) *RNA Biology* 10:5, 726-737; полное содержание которого приведено здесь в качестве ссылки. В некоторых вариантах осуществления нуклеаза Cas9 включает одну или несколько мутаций, которые частично нарушают или деактивируют домен расщепления ДНК.

[0182] Деактивированный нуклеазой домен Cas9 может взаимозаменяемым образом быть обозначен как белок "dCas9" (что означает "мертвый" вследствие нуклеазы Cas9). Способы получения домена Cas9 (или его фрагмента), имеющего деактивированный домен расщепления ДНК, являются известными (см., например, Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression" (2013) *Cell*. 28;152(5):1173-83, полное содержание каждого из которых приведено здесь в качестве ссылки). Например, известно, что домен расщепления ДНК Cas9 включает два субдомена: субдомен нуклеазы HNH и субдомен RuvC1. Субдомен HNH расщепляет цепь, комплементарную нРНК, в то время как субдомен RuvC1 расщепляет некомплементарную цепь. Мутации в данных субдоменах могут заглушать нуклеазную активность Cas9. Например, мутации D10A и H840A полностью деактивируют нуклеазную активность Cas9 *S. pyogenes* (Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, *Cell*. 28;152(5):1173-83 (2013)). В некоторых вариантах осуществления представлены белки, включающие фрагменты Cas9. Например, в некоторых вариантах осуществления белок включает один из двух доменов Cas9: (1) нРНК-связывающий домен Cas9; или (2) домен расщепления ДНК Cas9. В некоторых вариантах осуществления белки, включающие Cas9 или его фрагменты, обозначаются как "варианты Cas9". Вариант Cas9 имеют такую же гомологию, что и Cas9 или его фрагмент. Например, вариант Cas9 является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, по меньшей мере, приблизительно на 99,8% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным Cas9 дикого типа (например SpCas9 согласно SEQ ID NO: 18). В некоторых вариантах осуществления вариант Cas9 может иметь 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 или больше аминокислотных изменений по сравнению с Cas9 дикого типа (например SpCas9 согласно SEQ ID NO: 18). В некоторых вариантах осуществления вариант Cas9 включает фрагмент Cas9 SEQ ID NO: 18 (например нРНК-связывающий домен или домен расщепления ДНК), так что фрагмент является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным соответствующему фрагменту Cas9 дикого типа (например SpCas9 согласно SEQ ID NO: 18). В некоторых вариантах осуществления фрагмент имеет длину, соответствующую, по меньшей мере, на 30%, по меньшей мере, на 35%, по меньшей мере, на 40%, по меньшей мере, на 45%, по меньшей мере, на 50%, по меньшей мере, на 55%, по меньшей мере, на 60%, по меньшей мере, на 65%, по меньшей мере, на 70%, по меньшей мере, на 75%, по меньшей мере, на 80%, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% длине аминокислоты соответствующего Cas9 дикого типа (например SpCas9 согласно SEQ ID NO: 18).

кДНК

[0183] Термин "кДНК" обозначает цепь ДНК, скопированную из РНК матрицы. кДНК является комплементарной РНК матрице.

Циркулярный пермутант

[0184] В соответствии с используемым здесь значением термин "циркулярный пермутант" обозначает белок или полипептид (например Cas9), включающий циркулярную пермутацию, которая представляет собой изменение в структурной конфигурации белка, включающую изменение в порядке аминокислот, следующих в рамках аминокислотной последовательности белка. Другими словами, циркулярные пермутанты представляют собой белки, которые имеют измененные N- и C-концы по сравнению с аналогами дикого типа, например C-концевая половина белка дикого типа

становится новой N-концевой половиной. Циркулярная пермутация (или CP) по существу представляет собой топологическую перестройку первичной последовательности белка, соединяющей его N- и C-конец, зачастую с пептидным линкером, в то же время одновременно обеспечивая разделение его последовательности в другом положении для создания новых смежных N- и C-концов. Результатом является структура белка с иной связностью, но которая зачастую может иметь ту же общую аналогичную трехмерную (3D) форму и возможно включать улучшенные или измененные характеристики, включая сниженную протеолитическую чувствительность, повышенную каталитическую активность, измененное связывание субстрата или лиганда, и/или улучшенную термостабильность. Белки в виде циркулярных пермутантов могут возникать в природе (например конканавалин A и лектин). Кроме того, циркулярная пермутация может возникать в результате посттрансляционных модификций или может являться результатом модификации с использованием рекомбинантных способов.

Циркулярно пермутированный Cas9

[0185] Термин "циркулярно пермутированный Cas9" обозначает любой белок Cas9 или его вариант в виде циркулярного пермутанта, если его N- и C-концы были местно перестроены. Подобные циркулярно пермутированные белки Cas9 ("CP-Cas9") или их варианты сохраняют способность связывать ДНК при образовании комплексов с направляющей РНК (нРНК). См. Oakes et al., “Protein Engineering of Cas9 for enhanced function,” *Methods Enzymol*, 2014, 546: 491–511 and Oakes et al., “CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification,” *Cell*, January 10, 2019, 176: 254-267, каждый из которых приведен здесь в качестве ссылки. Настоящее изобретение охватывает любые ранее известные CP-Cas9 или предусматривает использование нового CP-Cas9 при условии, если полученный циркулярно пермутированный белок сохраняет способность связывать ДНК при образовании комплексов с направляющей РНК (нРНК). Примеры белков CP-Cas9 приведены в SEQ ID NO: 77-86.

CRISPR

[0186] CRISPR представляет собой семейство последовательностей ДНК (то есть кластеров CRISPR) в бактериях и археях, которые соответствуют фрагментам предшествующих инфекций вирусом, который вторгся в прокариоту. Фрагменты ДНК используются прокариотической клеткой для детектирования и устранения ДНК при

последующих атаках аналогичными вирусами и по существу составляет, наряду с массивом CRISPR-ассоциированных белков (включая Cas9 и его гомологи) и CRISPR-ассоциированной РНК, прокариотическую иммунную защитную систему. В природе кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В системах CRISPR определенного типа (например системах CRISPR типа II) корректный процессинг пре-crРНК требует транс-кодируемую малую РНК (tracrРНК), эндогенную рибонуклеазу 3 (rnc) и домен Cas9. tracrРНК служит в качестве направляющей для осуществляемого с помощью рибонуклеазы 3 процессинга пре-crРНК. Далее Cas9/crРНК/tracrРНК обеспечивает расщепление эндонуклеазой линейной или циркулярной дцДНК-мишени, являющейся комплементарной РНК. В частности, цепь-мишень, которая не является комплементарной crРНК, сначала разрезается эндонуклеазой, затем подрезается 3′-5′ экзонуклеазой. В природе связывание и расщепление ДНК обычно требует белок и обе РНК. Однако одиночные направляющие РНК ("sgРНК" или просто "нРНК") могут быть сконструированы для включения аспектов как crРНК, так и tracrРНК в виде одиночной РНК – направляющей РНК. См, например Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полное содержание которого приведено здесь в качестве ссылки. Cas9 распознает короткий мотив в последовательностях повторов CRISPR (PAM или смежный с протоспейсером мотив), чтобы помогать отличать себя от чужого. Биология CRISPR, а также последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной области (см., например, , “Complete genome sequence of an M1 strain of Streptococcus pyogenes.” Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); “CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.” Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); и “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полное содержание каждого из которых приведено здесь в качестве ссылки). Ортологи Cas9 были описаны у различных видов, включая, но без ограничения, *S. pyogenes* и *S. thermophilus*. Дополнительные подходящие нуклеазы Cas9 и последовательности будут очевидны

специалистам в данной области на основе данного описания, при этом такие нуклеазы Cas9 и последовательности включают последовательности Cas9 из организмов и локусов, описанных в Chylinski, Rhun, and Charpentier, “The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems” (2013) RNA Biology 10:5, 726-737; олное содержание которого приведено здесь в качестве ссылки.

[0187] В системах CRISPR определенного типа (например системах CRISPR типа II) корректный процессинг пре-crРНК требует транс-кодируемую малую РНК (tracrРНК), эндогенную рибонуклеазу 3 (rnc) и домен Cas9. tracrРНК служит в качестве направляющей для осуществляемого с помощью рибонуклеазы 3 процессинга пре-crРНК. Далее Cas9/crРНК/tracrРНК обеспечивает расщепление эндонуклеазой линейной или циркулярной нуклеиновой кислоты-мишени, являющейся комплементарной РНК. В частности, цепь-мишень, которая не является комплементарной crРНК, сначала разрезается эндонуклеазой, затем подрезается 3′-5′ экзонуклеазой. В природе связывание и расщепление ДНК обычно требует белок и обе РНК. Однако одиночные направляющие РНК ("sgРНК" или просто "нРНК") могут быть сконструированы для включения вариантов осуществления как crРНК, так и tracrРНК в виде одиночной РНК – направляющей РНК.

[0188] В целом, термин "система CRISPR" обозначает совместно транскрипты и другие элементы, участвующие в экспрессии или направлении активности CRISPR-ассоциированных генов ("Cas"), включая последовательности, кодирующие ген Cas, последовательность tracr (транс-активирующий CRISPR) (например tracrРНК или активную частичную tracrРНК), последовательность tracr mate (включающую "прямой повтор" и tracrРНК-процессированный частичный прямой повтор в контексте эндогенной системы CRISPR), направляющую последовательность (также обозначаемую как "спейсер" в контексте эндогенной системы CRISPR), или другие последовательности и транскрипты из локуса CRISPR. tracrРНК системы комплементарна (полностью или частично) последовательности tracr mate, присутствующей на направляющей РНК.

Матрица для синтеза ДНК

[0189] В соответствии с используемым здесь значением термин "матрица для синтеза ДНК" обозначает участок или часть плеча удлинения PEgRNA, которая используется в качестве матричной цепи полимеразой праймерного редактора для кодирования 3′ одноцепочечного ДНК-флэпа, который содержит требуемое изменение и который

затем за счет механизма праймерного редактирования замещает соответствующую эндогенную цепь ДНК в участке-мишени. В различных вариантах осуществления матрица для синтеза ДНК показана на Фиг.3A (в контексте PEgRNA, включающей 5′ плечо удлинения), Фиг.3B (в контексте PEgRNA, включающей 3′ плечо удлинения), Фиг.3C (в контексте внутреннего плеча удлинения), Фиг.3D (в контексте 3′ плеча удлинения), и Фиг.3E (в контексте 5′ плеча удлинения). Плечо удлинения, включая матрицу для синтеза ДНК, может состоять из ДНК или РНК. В случае плеча удлинения в виде РНК полимераза праймерного редактора может представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). В случае ДНК-полимераза праймерного редактора может представлять собой ДНК-зависимую ДНК-полимеразу. В различных вариантах осуществления (например как показано на Фиг.3D-3E), матрица для синтеза ДНК (4) может включать "матрицу для редактирования" и "плечо гомологии", а также весь или часть необязательного 5'-концевого модификаторного участка e2. То есть в зависимости от природы участка e2 (например того, включает ли он шпильку, опорную петлю, или вторичную структуру в виде шпильки/петли), полимераза может также кодировать ни одну часть, некоторую часть, или весь участок e2. Другими словами, в случае 3′ плеча удлинения матрица для синтеза ДНК (3) может включать часть плеча удлинения (3), которая проходит от 5′-конца участка связывания праймера (PBS) до 3′-конца ядра нРНК, которая может функционировать в качестве матрицы для синтеза одноцепочечной ДНК-полимеразой (например обратной транскриптазой). В случае 5′ плеча удлинения матрица для синтеза ДНК (3) может включать часть плеча удлинения (3), которая проходит от 5′-конца молекулы PEgRNA до 3′-конца матрицы для редактирования. Матрица для синтеза ДНК предпочтительно исключает участок связывания праймера (PBS) PEgRNA, имеющих 3′ плечо удлинения или 5′ плечо удлинения. Определенные описанные здесь варианты осуществления (например Фиг.71A) ссылаются на "RT матрицу", которая включает матрицу для редактирования и плечо гомологии, то есть последовательность плеча удлинения PEgRNA, которая фактически используется в качестве матрицы при синтезе ДНК. Термин "RT матрица" эквивалентен термину "матрица для синтеза ДНК".

[0190] В случае *trans* праймерного редактирования (например Фиг.3G и Фиг.3H) участок связывания праймера (PBS) и матрица для синтеза ДНК могут быть встроены в отдельную молекулу, обозначаемую как РНК-матрица *trans* праймерного редактора.

Домен димеризации

[0191] Термин "домен димеризации" обозначает лиганд-связывающий домен, который связывается со связывающимся фрагментом биспецифического лиганда. "Первый" домен димеризации связывается с первым связывающимся фрагментом биспецифического лиганда, а "второй" домен димеризации связывается со вторым связывающимся фрагментом того же биспецифического лиганда. Когда первый домен димеризации подвергается слиянию с первым белком (например за счет PE, как описано здесь), а второй домен димеризации (например за счет PE, как описано здесь) подвергается слиянию со вторым белком, первый и второй белок димеризуются в присутствии биспецифического лиганда, при этом биспецифический лиганд имеет, по меньшей мере, один фрагмент, который связывается с первым доменом димеризации, и, по меньшей мере, другой фрагмент, который связывается со вторым доменом димеризации.

Ниже

[0192] В соответствии с используемым здесь значением термины "выше" и "ниже" представляют собой термины для указания относительного положения, которые определяют линейное положение, по меньшей мере, двух элементов, расположенных в молекуле нуклеиновой кислоты (одноцепочечной или двуцепочечной), которая ориентирована в направлении от 5′ к 3′. В частности, первый элемент расположен выше второго элемента в молекуле нуклеиновой кислоты, если первый элемент расположен где-либо 5’ относительно второго элемента. Например, SNP расположен выше Cas9-индуцированного участка разрыва, если SNP расположен на 5’ стороне участка разрыва. С другой стороны, первый элемент расположен ниже второго элемента в молекуле нуклеиновой кислоты, если первый элемент расположен где-либо 3’ относительно второго элемента. Например, SNP расположен ниже Cas9-индуцированного участка разрыва, если SNP расположен на 3’ стороне участка разрыва. Молекула нуклеиновой кислоты может представлять собой ДНК (двуцепочечную или одноцепочечную), РНК (двуцепочечную или одноцепочечную), или гибрид ДНК и РНК. Анализ является таким же для одноцепочечной молекулы нуклеиновой кислоты, что и для двуцепочечной молекулы, поскольку термины "выше" и "ниже" обозначают только одну цепь молекулы нуклеиновой кислоты, за исключением случаев, когда необходимо выбрать указываемую цепь двуцепочечной молекулы. Зачастую цепь двуцепочечной ДНК, которая может применяться для

определения относительности положения, по меньшей мере, двух элементов, представляет собой "смысловую" или "кодирующую" цепь. В генетике "смысловая" цепь представляет собой сегмент в рамках двуцепочечной ДНК, которая проходит от 5' к 3' и которая является комплементарной антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' к 5'. Таким образом, в качестве примера, азотистое основание SNP находится ниже последовательности промотора в геномной ДНК (которая является двуцепочечной), если азотистое основание SNP располагается на 3' стороне промотора на смысловой или кодирующей цепи.

Матрица для редактирования

[0193] Термин "матрица для редактирования" обозначает часть плеча удлинения, которая кодирует требуемое изменение в одноцепочечном 3′ ДНК-флэпе, который синтезируется полимеразой, например ДНК-зависимой ДНК-полимеразой, РНК-зависимой ДНК-полимеразой (например обратной транскриптазой). Определенные описанные здесь варианты осуществления (например Фиг.71A) ссылаются на "RT матрицу", которая включает совместно как матрицу для редактирования, так и плечо гомологии, то есть последовательность плеча удлинения PEgRNA, которая фактически используется в качестве матрицы при синтезе ДНК. Термин "RT матрица для редактирования" также эквивалентен термину "матрица для синтеза ДНК", однако "RT матрица для редактирования" обозначает использование праймерного редактора, имеющего полимеразу, которая представляет собой обратную транскриптазу, а "матрица для синтеза ДНК" более широким образом обозначает использование праймерного редактора, имеющего любую полимеразу.

Эффективное количество

[0194] В соответствии с используемым здесь значением термин "эффективное количество" обозначает количество биологически активного агента, которое является достаточным для реализации требуемого биологического ответа. Например, в некоторых вариантах осуществления эффективное количество праймерного редактора (PE) может обозначать количество редактора, которое является достаточным для редактирования участка-мишени нуклеотидной последовательности, например генома. В некоторых вариантах осуществления указанное здесь эффективное количество праймерного редактора (PE), например слитого белка, включающего домен никазы Cas9 и обратную транскриптазу, может обозначать количество слитого белка, которое является достаточным для индуцирования редактирования участка-мишени,

специфично связанного и подвернутого редактирования слитым белком. Как будет очевидно специалисту в данной области, эффективное количество агента, например слитого белка, нуклеазы, гибридного белка, димера белка, комплекса белка (или димера белка) и полинуклеотида, или полинуклеотида может варьироваться в зависимости от различных факторов, например требуемого биологического ответа, например подвигаемого редактированию определенного аллеля, генома, или участка-мишени, являющейся мишенью клетки или ткани, и используемого агента.

Подверженная ошибкам обратная транскриптаза

[0195] В соответствии с используемым здесь значением термин "подверженная ошибкам" обратная транскриптаза (или, более широко, любая полимераза) обозначает обратную транскриптазу (или, более широко, любую полимеразу), которая встречается в природе или которая была получена из другой обратной транскриптазы (например обратной транскриптазы M-MLV дикого типа), которая имеет частоту ошибок, которая меньше, чем частота ошибок обратной транскриптазы M-MLV дикого типа. Сообщается, что частота ошибок обратной транскриптазы M-MLV дикого типа соответствует диапазону в виде одной ошибки на от 15000 (верхняя граница) до 27000 (нижняя граница). Частота ошибок 1 на 15000 соответствует частоте ошибок 6,7 x 10^{-5}. Частота ошибок 1 на 27000 соответствует частоте ошибок 3,7 x 10^{-5}. См. Boutabout et al. (2001) "DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1," *Nucleic Acids Res* 29(11):2217–2222, который полностью приведен здесь в качестве ссылки. Таким образом, для целей данного изобретения термин "подверженная ошибкам" означает RT, которые имеют частоту ошибок в виде более чем одной ошибки на 15000 встраиваний азотистых оснований (6,7 x 10^{-5} или более), например 1 ошибку в 14000 азотистых оснований (7,14 x 10^{-5} или более), 1 ошибку в 13000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 12000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 11000 азотистых оснований или менее (9,1 x 10^{-5} или более), 1 ошибку в 10000 азотистых оснований или менее (1 x 10^{-4} или 0,0001 или более), 1 ошибку в 9000 азотистых оснований или менее (0,00011 или более), 1 ошибку в 8000 азотистых оснований или менее (0,00013 или более), 1 ошибку в 7000 азотистых оснований или менее (0,00014 или более), 1 ошибку в 6000 азотистых оснований или менее (0,00016 или более), 1 ошибку в 5000 азотистых оснований или менее (0,0002 или более), 1 ошибку в 4000 азотистых оснований или менее (0,00025 или более), 1 ошибку в 3000 азотистых оснований или менее (0,00033

или более), 1 ошибку в 2000 азотистых оснований или менее (0,00050 или более), или 1 ошибку в 1000 азотистых оснований или менее (0,001 или более), или 1 ошибку в 500 азотистых основаниях или менее (0,002 или более), или 1 ошибку в 250 азотистых основаниях или менее (0,004 или более).

Экстеин

[0196] В соответствии с используемым здесь значением термин "экстеин" обозначает полипептидную последовательность, которая фланкируется интеином и лигируется с другим экстеином в процессе сплайсинга белка для образования зрелого подвергнутого сплайсингу белка. Как правило, интеин фланкируется двумя последовательностями экстеина, которые лигируется вместе, когда интеин катализирует свою собственную эксцизию. Экстеины, соответственно, представляют собой белковый аналог экзонов в иРНК. Например, полипептид, включающий интеин, может иметь структуру экстеин(N) – интеин – экстеин(C). После эксцизии интеина и сплайсинга двух экстеинов, полученные структуры представляют собой экстеин(N) – экстеин(C) и свободный интеин. В различных конфигурациях экстеины могут представлять собой отдельные белки (например половину Cas9 или слитого белка PE), каждый из которых слит с разделенным интеином, при этом эксцизия разделенных интеинов вызывает совместный сплайсинг последовательностей экстеина.

Плечо удлинения

[0197] Термин "плечо удлинения" обозначает компонент нуклеотидной последовательности PEgRNA, который обеспечивает несколько функций, включая участок связывания праймера и матрицу для редактирования для обратной транскриптазы. В некоторых вариантах осуществления, например на Фиг.3D, плечо удлинения расположено на 3′-конце направляющей РНК. В других вариантах осуществления, например на Фиг.3D, плечо удлинения расположено на 5′-конце направляющей РНК. В некоторых вариантах осуществления плечо удлинения также включает плечо гомологии. В различных вариантах осуществления плечо удлинения включает следующие компоненты в направлении от 5′ к 3′: плечо удлинения, матрицу для редактирования, и участок связывания праймера. Поскольку полимеризационная активность обратной транскриптазы осуществляется в направлении от 5′ к 3′, предпочтительная группировка плеча удлинения, матрицы для редактирования и участка связывания праймера соответствует направлению от 5′ к 3′, так что обратная транскриптаза после праймирования отожженной праймерной последовательностью

обеспечивает полимеризацию одиночной цепи ДНК с использованием матрицы для редактирования в качестве цепи комплементарной матрицы. Дополнительные подробности, такие как длина плеча удлинения, описаны в других разделах описания.

[0198] Плечо удлинения также может быть описано как включающее в целом два участка: участок связывания праймера (PBS) и матрицу для синтеза ДНК, как показано, например, на Фиг.3G (сверху). Участок связывания праймера связывается с праймерной последовательностью, которая образуется из цепи эндогенной ДНК участка-мишени, когда он подвергается никированию комплексом праймерного редактора, тем самым высвобождая 3′-конец на эндогенной никированной цепи. Как описано здесь, связывание праймерной последовательности с участком связывания праймера на плече удлинения PEgRNA создает дуплексный участок со свободным 3′-концом (то есть 3′ праймерной последовательности), который затем обеспечивает субстрат для полимеразы для начала полимеризации одиночной цепи ДНК из свободного 3′-конца в рамках длины матрицы для синтеза ДНК. Последовательность продукта в виде одиночной цепи ДНК представляет собой комплемент матрицы для синтеза ДНК. Полимеризация продолжается по направлению к 5′ матрицы для синтеза ДНК (или плеча удлинения) до терминации полимеризации. Таким образом, матрица для синтеза ДНК соответствует части плеча удлинения, которая закодирована в продукте в виде одноцепочечной ДНК (то есть 3′ одноцепочечного ДНК-флэпа, содержащего информацию в виде требуемого генетического изменения) полимеразой комплекса праймерного редактора и которая в конечном итоге замещает соответствующую эндогенную цепь ДНК участка-мишени, который располагается непосредственно ниже PE-индуцированного участка разрыва. Без связи с теорией, полимеризация матрицы для синтеза ДНК продолжается по направлению к 5′-концу плеча удлинения до осуществления терминации. Полимеризация может терминироваться рядом способов, включая, но без ограничения, (a) достижение 5′-конца РЕнНРК (например в случае 5′ плеча удлинения, где у ДНК-полимеразы просто заканчивается матрица), (b) достижение непроходимой вторичной структуры РНК (например шпильки или шпильки/петли), или (c) достижение сигнала терминации репликации, например определенной нуклеотидной последовательности, которая блокирует или ингибирует полимеразу, или топологического сигнала нуклеиновой кислоты, например суперспиральной ДНК или РНК.

<u>Флэп-эндонуклеаза (например FEN1)</u>

[0199] В соответствии с используемым здесь значением термин "флэп-эндонуклеаза" обозначает фермент, который катализирует удаление 5′ одноцепочечных ДНК-флэпов. Они представляют собой встречающиеся в природе ферменты, которые процессируют удаление 5′ флэпов, образованных в рамках клеточных процессов, включая репликацию ДНК. В рамках описанных здесь способов праймерного редактирования могут применяться эндогенно или *in trans* поступающие флэп-эндонуклеазы для удаления 5′ флэпа эндогенной ДНК, образованного в участке-мишени при праймерном редактировании. Флэп-эндонуклеазы известны в данной области и описаны в Patel et al., "Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends," *Nucleic Acids Research*, 2012, 40(10): 4507-4519, Tsutakawa et al., "Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily," *Cell*, 2011, 145(2): 198-211, и Balakrishnan et al., "Flap Endonuclease 1," *Annu Rev Biochem*, 2013, Vol 82: 119-138 (каждый из которых приведен здесь в качестве ссылки). Пример флэп-эндонуклеазы представляет собой FEN1, которая может быть описана в виде следующей аминокислотной последовательности:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ	SEQ ID NO:
FEN1 ДИКОГО ТИПА	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSI YQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMME NGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQ AQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGI PYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGS PVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVDL CILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPN KYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPNE EELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLD DFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	SEQ ID NO: 7

<u>Функциональный эквивалент</u>

[0200] Термин "функциональный эквивалент" обозначает вторую биомолекулу, которая является эквивалентной в плане функции, но не обязательно эквивалентной в плане структуры относительно первой биомолекулы. Например, "эквивалент Cas9" обозначает белок, который имеет такие же или по существу такие же функции, что и Cas9, но не обязательно такую же аминокислотную последовательность. В контексте изобретения в тексте описания указывается "белок X или его функциональный

эквивалент". В данном контексте "функциональный эквивалент" белка X охватывает любую гомологичную, паралогичную, представляющую собой фрагмент, встречающуюся в природе, сконструированную, мутированную или синтетическую версию белка X, которая несет эквивалентную функцию.

Слитый белок

[0201] В соответствии с используемым здесь значением термин "слитый белок" обозначает гибридный полипептид, который включает домены белка из, по меньшей мере, двух различных белков. Один белок может быть расположен на амино-концевой (N-концевой) части слитого белка или карбокси-концевой (C-концевой) части слитого белка, тем самым образуя "амино-концевой слитый белок" или "карбокси-концевой слитый белок", соответственно. Белок может включать различные домены, например домен связывания нуклеиновой кислоты (например домен связывания нРНК Cas9, который направляет связывание белка с участком-мишенью) и домен расщепления нуклеиновой кислоты или каталитический домен белка для редактирования нуклеиновой кислоты. Другой пример включает Cas9 или его эквивалент обратной транскриптазы. Любой из указанных здесь белков может быть получен с помощью любого способа, известного в данной области. Например, указанные здесь белки могут быть получены с помощью экспрессии рекомбинантного белка и очистки, которые являются особенно подходящими для слитых белков, включающих пептидный линкер. Способы экспрессии рекомбинантных белков и очистки хорошо известны и включают способы, описанные Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), полное содержание которого приведено здесь в качестве ссылки.

Представляющий интерес ген (GOI)

[0202] Термин "представляющий интерес ген" или "GOI" обозначает ген, который кодирует представляющую интерес биомолекулу (например белок или молекулу РНК). Представляющий интерес белок может включать любой внутриклеточный белок, белок мембраны, или внеклеточный белок, например ядерный белок, фактор транскрипции, транспортер ядерной мембраны, ассоциированный с внутриклеточной органеллой белок, мембранный рецептор, каталитический белок, и фермент, терапевтический белок, мембранный белок, мембранный транспортерный белок, белок передачи сигнала, или иммунологический белок (например IgG или другой белок в виде антитела), и так далее. Представляющий интерес ген может также кодировать

молекулу РНК, включая, но без ограничения, информационную РНК (иРНК), транспортную РНК (тРНК), рибисомную РНК (рРНК), малую ядерную РНК (мяРНК), антисмысловую РНК, направляющую РНК, микроРНК (мРНК), малую интерферирующую РНК (миРНК) и бесклеточную РНК (бкРНК).

Направляющая РНК ("нРНК")

[0203] В соответствии с используемым здесь значением термин "направляющая РНК" обозначает определенный тип направляющей нуклеиновой кислоты, которая наиболее часто ассоциирована с белком Cas CRISPR-Cas9 и которая ассоциируется с Cas9, направляя белок Cas9 к определенной последовательности в молекуле ДНК, которая включает комплементарность относительно последовательности протоспейсера направляющей ДНК. Однако данный термин также охватывает эквивалентные молекулы направляющей нуклеиновой кислоты, которые ассоциируется с эквивалентами Cas9, гомологами, ортологами, или паралогами, которые встречаются или не встречаются в природе (например являются сконструированными или рекомбинантными) и которые иным образом программируют эквивалент Cas9 для локализации в определенной нуклеотидной последовательности-мишени. Эквиваленты Cas9 могут включать другой napDNAbp из системы CRISPR любого типа (например типа II, V, VI), включая Cpf1 (системы CRISPR-Cas типа V), C2c1 (системы CRISPR-Cas типа V), C2c2 (системы CRISPR-Cas типа VI) и C2c3 (системы CRISPR-Cas типа V). Дополнительные эквиваленты Cas описаны в Makarova et al., “C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector,” Science 2016; 353(6299), содержание которого приведено здесь в качестве ссылки. Здесь указаны примеры последовательностей и структуры направляющих РНК. Кроме того, здесь указаны способы конструирования подходящих последовательностей направляющих РНК. В соответствии с используемым здесь значением термин "направляющая РНК" может также обозначать "традиционную направляющую РНК" для ее противопоставления модифицированным формам направляющей РНК, обозначаемым как "направляющие РНК для праймерного редактирования" (или "PEgRNA"), которые были изобретены для использования в рамках описанных здесь способов праймерного редактирования и композиций.

[0204] Направляющие РНК или PEgRNA могут включать различные структурные элементы, которые включают, но без ограничения:

[0205] Спейсерную последовательность – последовательность в направляющей РНК или PEgRNA (имеющую длину около 20 nt), которая связывается с протоспейсером в ДНК-мишени.

[0206] ядро нРНК (или последовательность каркаса или основы нРНК) – обозначает последовательность в нРНК, которая отвечает за связывание Cas9, она не включает 20 bp спейсер/нацеливающую последовательность, которая используется для направления Cas9 к ДНК-мишени.

[0207] Плечо удлинения – одноцепочечное удлинение на 3′-конце или 5′-конце PEgRNA, которое включает последовательность участка связывания праймера и матрицы для синтеза ДНК, которая кодирует за счет полимеразы (например обратной транскриптазы) одноцепочечный ДНК-флэп, содержащий представляющее интерес изменение, которое затем встраивается в эндогенную ДНК за счет замещения соответствующей эндогенной цепи, тем самым обеспечивая внесение требуемого генетического изменения.

[0208] Терминатор транскрипции – направляющая РНК или PEgRNA может включать последовательность терминации транскрипции на 3′-конце молекулы.

Плечо гомологии

[0209] Термин "плечо гомологии" обозначает часть плеча удлинения, которая кодирует часть получаемого кодируемого обратной транскриптазой одноцепочечного ДНК-флэпа, который подвергается встраиванию в участок-мишень ДНК за счет замещения эндогенной цепи. Часть одноцепочечного ДНК-флэпа, кодируемая плечом гомологии, является комплементарной не подвергаемой редактированию цепи последовательности-мишени ДНК, которая обеспечивает вытеснение эндогенной цепи и отжиг одноцепочечного ДНК-флэпа в его месте, тем самым обеспечивая внесение изменения. Данный компонент дополнительно описан здесь. Плечо гомологии представляет собой часть матрицы для синтеза ДНК, поскольку оно по определению кодируется полимеразой описанных здесь праймерных редакторов.

Клетка-хозяин

[0210] В соответствии с используемым здесь значением термин "клетка-хозяин" обозначает клетку, которая может являться хозяином, обеспечивать репликацию и экспрессию описанного здесь вектора, например вектора, включающего молекулу нуклеиновой кислоты, кодирующей слитый белок, включающий Cas9 или эквивалент Cas9 и обратную транскриптазу.

Интеины

[0211] В соответствии с используемым здесь значением термин "интеин" обозначает автопроцессирующие полипептидные домены, располагающиеся в организмах всех живых видов. Интеин (*int*ervening prot*ein*, промежуточный белок) обеспечивает реализацию уникального автопроцессирующего события, известного как сплайсинг белка, при котором он эксцизирует себя из более крупного прекурсорного полипептида за счет расщепления двух пептидных связей и, в рамках данного процесса, обеспечивает лигирование фланкирующих последовательностей экстеина (внешнего белка) за счет образования новой пептидной связи. Данное перестроение осуществляется после трансляции (или возможно одновременно с трансляцией). поскольку гены интеина являются встроенными в каркас в рамках других кодирующих белок генов. Более того, опосредованный интеином сплайсинг белков является спонтанным; он не требует внешнего фактора или источника энергии, а лишь фолдинг домена интеина. Данный процесс также известен как цис-сплайсинг белка, в отличие от природного процесса транс-сплайсинга белка с использованием "разделенных интеинов". Интеины представляют собой эквивалент белка аутосплайсирующих РНК интронов (см. Perler et al., Nucleic Acids Res. 22:1125-1127 (1994)), которые катализируют свою собственную эксцизию из прекурсорного белка с одновременным слиянием фланкирующих последовательностей белка, известных как экстеины (описаны в Perler et al., Curr. Opin. Chem. Biol. 1:292-299 (1997); Perler, F. B. Cell 92(1):1-4 (1998); Xu et al., EMBO J. 15(19):5146-5153 (1996)).

[0212] В соответствии с используемым здесь значением термин "сплайсинг белка" обозначает процесс, при котором внутренний участок прекурсорного белка (интеин) эксцизируется и фланкирующие участки белка (экстеины) лигируются для образования зрелого белка. Данный естественный процесс был описан для множества белков как от прокариотов, так и эукариотов (Perler, F. B., Xu, M. Q., Paulus, H. Current Opinion in Chemical Biology 1997, 1, 292-299; Perler, F. B. Nucleic Acids Research 1999, 27, 346-347). Единица интеина содержит необходимые компоненты, требуемые для катализирования сплайсинга белка и зачастую содержит домен эндонуклеазы, который принимает участие в обеспечении мобильности интеина (Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thomer, J., Belfort, M. Nucleic Acids Research 1994, 22, 1127-1127). Однако получаемые белки связываются, а не экспрессируются в виде отдельных белков. Сплайсинг белка может быть также

осуществлен *in trans* с использованием разделенных интеинов, экспрессируемых на отдельных полипептидах, которые спонтанным образом комбинируются для образования единого интеина, который затем подвергается процессу сплайсинга белка для соединения с отдельными белками.

[0213] Описание механизма сплайсинга белка привело к ряду вариантов его использования на основе интеина (Comb, et al., патент США 5496714; Comb, et al., патент США 5834247; Camarero and Muir, J. Amer. Chem. Soc., 121:5597-5598 (1999); Chong, et al., Gene, 192:271-281 (1997), Chong, et al., Nucleic Acids Res., 26:5109-5115 (1998); Chong, et al., J. Biol. Chem., 273:10567-10577 (1998); Cotton, et al. J. Am. Chem. Soc., 121:1100-1101 (1999); Evans, et al., J. Biol. Chem., 274:18359-18363 (1999); Evans, et al., J. Biol. Chem., 274:3923-3926 (1999); Evans, et al., Protein Sci., 7:2256-2264 (1998); Evans, et al., J. Biol. Chem., 275:9091-9094 (2000); Iwai and Pluckthun, FEBS Lett. 459:166-172 (1999); Mathys, et al., Gene, 231:1-13 (1999); Mills, et al., Proc. Natl. Acad. Sci. USA 95:3543-3548 (1998); Muir, et al., Proc. Natl. Acad. Sci. USA 95:6705-6710 (1998); Otomo, et al., Biochemistry 38:16040-16044 (1999); Otomo, et al., J. Biolmol. NMR 14:105-114 (1999); Scott, et al., Proc. Natl. Acad. Sci. USA 96:13638-13643 (1999); Severinov and Muir, J. Biol. Chem., 273:16205-16209 (1998); Shingledecker, et al., Gene, 207:187-195 (1998); Southworth, et al., EMBO J. 17:918-926 (1998); Southworth, et al., Biotechniques, 27:110-120 (1999); Wood, et al., Nat. Biotechnol., 17:889-892 (1999); Wu, et al., Proc. Natl. Acad. Sci. USA 95:9226-9231 (1998a); Wu, et al., Biochim Biophys Acta 1387:422-432 (1998b); Xu, et al., Proc. Natl. Acad. Sci. USA 96:388-393 (1999); Yamazaki, et al., J. Am. Chem. Soc., 120:5591-5592 (1998)). Каждая ссылка приведена здесь в качестве ссылки.

Лиганд-зависимый интеин

[0214] В соответствии с используемым здесь значением термин "лиганд-зависимый интеин" обозначает интеин, который включает лиганд-связывающий домен. Как правило, лиганд-связывающий домен вставляется в аминокислотную последовательность интеина, что приводит к структуре интеин (N) – лиганд-связывающий домен – интеин (C). Как правило, лиганд-зависимые интеины демонстрируют отсутствие или лишь минимальную активность в виде сплайсинга белка при отсутствии подходящего лиганда, а также заметное повышение активности в виде сплайсинга белка в присутствии лиганда. В некоторых вариантах осуществления лиганд-зависимый интеин не демонстрирует наблюдаемой активности в виде

сплайсинга при отсутствии лиганда, но демонстрирует активность в виде сплайсинга в присутствии лиганда. В некоторых вариантах осуществления лиганд-зависимый интеин демонстрирует заметную активность в виде сплайсинга белка при отсутствии лиганда и активность в виде сплайсинга белка в присутствии подходящего лиганда, которая по меньшей мере, в 5 раз, по меньшей мере, в 10 раз, по меньшей мере, в 50 раз, по меньшей мере, в 100 раз, по меньшей мере, в 150 раз, по меньшей мере, в 200 раз, по меньшей мере, в 250 раз, по меньшей мере, в 500 раз, по меньшей мере, в 1000 раз, по меньшей мере, в 1500 раз, по меньшей мере, в 2000 раз, по меньшей мере, в 2500 раз, по меньшей мере, в 5000 раз, по меньшей мере, в 10000 раз, по меньшей мере, в 20000 раз, по меньшей мере, в 25000 раз, по меньшей мере, в 50000 раз, по меньшей мере, в 100000 раз, по меньшей мере, в 500000 раз, или, по меньшей мере, в 1000000 раз больше, чем активность, наблюдаемая при отсутствии лиганда. В некоторых вариантах осуществления повышение активности зависит от дозы и повышается, по меньшей мере, на 1 порядок, по меньшей мере, на 2 порядка, по меньшей мере, на 3 порядка, по меньшей мере на 4 порядка, или, по меньшей мере, на 5 порядков, обеспечивая возможность тонкой настройки активности интеина за счет корректирования концентрации лиганда. Подходящие лиганд-зависимые интеины известны в данной области и включают указанные ниже и описанные в опубликованной заявке на патент США 2014/0065711 A1; Mootz *et al.*, “Protein splicing triggered by a small molecule.” *J. Am. Chem. Soc.* 2002; **124**, 9044–9045; Mootz *et al.*, “Conditional protein splicing: a new tool to control protein structure and function *in vitro* and *in vivo*.” *J. Am. Chem. Soc.* 2003; **125**, 10561–10569; Buskirk *et al.*, *Proc. Natl. Acad. Sci. USA.* 2004; **101**, 10505-10510); Skretas & Wood, “Regulation of protein activity with small-molecule-controlled inteins.” *Protein Sci.* 2005; **14**, 523-532; Schwartz, *et al.*, “Post-translational enzyme activation in an animal via optimized conditional protein splicing.” *Nat. Chem. Biol.* 2007; **3**, 50-54; Peck *et al.*, *Chem. Biol.* 2011; **18** (5), 619-630; полное содержание которых приведено здесь в качестве ссылки. Примеры последовательностей представляют собой:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
2-4 ИНТЕИН:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA

	LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 8)
3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDR VAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASM MGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEIL MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATS SRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRAL DKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHL YSMKYTNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDDK FLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHN C (SEQ ID NO: 9)
30R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 10)
30R3-2 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 11)
30R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 12)
37R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL

	ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC ((SEQ ID NO: 13)
37R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 14)
37R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 15)

Линкер

[0215] В соответствии с используемым здесь значением термин "линкер" обозначает молекулу, связывающую две другие молекулы или фрагмента. Линкер может представлять собой любую аминокислотную последовательность в случае, если линкер соединяет два слитых белка. Например, Cas9 может быть слит с обратной транскриптазой с помощью аминокислотной последовательности линкера. Линкер также может представлять собой нуклеотидную последовательность в случае соединения вместе двух нуклеотидных последовательностей. Например, в данном случае традиционная направляющая РНК соединяется через спейсерную или нуклеотидную последовательность линкера с РНК удлинением направляющей РНК для праймерного редактирования, которая может включать последовательность RT матрицы и участок связывания праймера RT. В других вариантах осуществления линкер представляет собой органическую молекулу, группу, полимер, или химический фрагмент. В некоторых вариантах осуществления линкер имеет длину 5-100 аминокислот, например длину 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-

100, 100-150, или 150-200 аминокислот. Также охватываются более длинные или короткие линкеры.

Выделенный

[0216] Термин "выделенный" обозначает измененный или удаленный из естественной среды. Например, нуклеиновая кислота или пептид, естественным образом присутствующий в живом животном не является "выделенным", но та же нуклеиновая кислота или пептид, который частично или полностью отделен от сопутствующих материалов его естественной среды, является "выделенным". Выделенная нуклеиновая кислота или белок может существовать в по существу очищенной форме или может существовать в ненативной среде, такой как, например, клетка-хозяин.

[0217] В некоторых вариантах осуществления представляющий интерес ген кодируется выделенной нуклеиновой кислотой. В соответствии с используемым здесь значением термин "выделенный" обозначает характеристику материала, как указано здесь, являющегося удаленным из его исходной или нативной среды (например природной среды, если он встречается в природе). Таким образом, встречающийся в природе полинуклеотид или белок или полипептид, присутствующий в живом животном, не является выделенным, но тот же полинуклеотид или полипептид, отделенный за счет вмешательства человека от некоторых или всех сопутствующих материалов в рамках природной системы, является выделенным. Искусственный или сконструированный материал, например не встречающийся в природе конструкт нуклеиновой кислоты, например описанные здесь конструкты экспрессии и векторы, соответственно также обозначается в качестве выделенного. Материал не обязательно должен быть очищенным, чтобы быть выделенным. Соответственно, материал может представлять собой часть вектора и/или часть композиции и все равно являться выделенным, поскольку данный вектор или композиция не является частью среды, в которой материал находится в природе.

Способ мечения MS2

[0218] В различных вариантах осуществления (например как изображено в рамках вариантов осуществления на Фиг.72-73 и в Примере 19) термин "способ мечения MS2" обозначает комбинацию "домена взаимодействия РНК-белок" (также известного как домен или белок рекрутирования РНК-белка"), соединенного с РНК-связывающим белком, которые специфически распознает и связывается с доменом взаимодействия РНК-белок, например определенной структурой в виде шпильки. Данные типы систем

могут использоваться для рекрутирования ряда видов функциональности к комплексу праймерного редактора, который связывается с участком-мишенью. Система мечения MS2 основана на естественном взаимодействии белка оболочки бактериофага MS2 ("MCP" или "MS2cp") со структурой в виде шпильки, присутствующей в геноме фага, то есть "шпилькой MS2". В случае праймерного редактирования способ мечения MS2 включает внесение шпильки MS2 в требуемую молекулу РНК, участвующую в праймерном редактировании (например PEgRNA или tPERT), которая затем составляет специфическую способную подвергаться взаимодействию мишень связывания для РНК-связывающего белка, который распознает и связывается с данной структурой. В случае шпильки MS2 она распознается и связывается белком оболочки бактериофага MS2 (MCP). При этом если MCP соединяется с другим белком (например обратной транскриптазой или другой ДНК-полимеразой), то шпилька MS2 может быть использована для "рекрутирования" данного другого белка *in trans* к участку-мишени, занятому комплексом для праймерного редактирования.

[0219] Описанные здесь праймерные редакторы могут включать в виде аспекта любой известный домен взаимодействия РНК-белок для рекрутирования или "солокализации" определенных представляющих интерес функций в комплексе праймерного редактора. Другие модулярные домены взаимодействия РНК-белок описаны в данной области, например в Johansson et al. , “RNA recognition by the MS2 phage coat protein,” *Sem Virol.*, 1997, Vol. 8(3): 176-185; Delebecque et al., “Organization of intracellular reactions with rationally designed RNA assemblies,” *Science*, 2011, Vol. 333: 470-474; Mali et al., “Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering,” *Nat. Biotechnol.*, 2013, Vol.31: 833-838; и Zalatan et al., “Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds,” *Cell*, 2015, Vol.160: 339-350, каждый из которых полностью приведен здесь в качестве ссылки. Другие системы включают шпильку PP7, которая специфически рекрутирует белок PCP, и шпильку "com", которая специфически рекрутирует белок Com. См. Zalatan et al.

[0220] Нуклеотидная последовательность шпильки MS2 (или также обозначаемой как "аптамер MS2") представляет собой:
GCCAACATGAGGATCACCCATGTCTGCAGGGCC (SEQ ID NO: 763).

[0221] Аминокислотная последовательность MCP или MS2cp представляет собой:
GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ

NRKYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIVKAMQGL LKDGNPIPSAIAANSGIY (SEQ ID NO: 764).

[0222] Шпилька MS2 (или "аптамер MS2") также может быть обозначена как тип " эффекторного домена рекрутирования РНК" (или равным образом как "домен рекрутирования РНК-связывающего белка" или просто как "рекрутирующий домен"), поскольку она представляет собой физическую структуру (например шпильку), которая встраивается в PEgRNA или tPERT, которая эффективно рекрутирует другие эффекторные функции (например РНК-связывающие белки, имеющие различные функции, например ДНК-полимеразы или другие ДНК-модифицирующие ферменты) для PEgRNA или tPERT, которые модифицированы соответствующим образом, и, таким образом, обеспечивает солокализацию эффекторных функций *in trans* для механизма праймерного редактирования. Данный вариант использования не является ограниченным каким-либо образом в плане каких-либо конкретных эффекторных доменов рекрутирования РНК и может включать любой доступный подобный домен, включая шпильку MS2. В Примере 19 и на Фиг.72(b) показано применение аптамера MS2, соединенного с доменом синтеза ДНК (то есть молекулой tPERT) и праймерным редактором, который включает белок MS2cp, слитый с PE2 для обеспечения солокализации комплекса праймерного редактора (комплекса MS2cp-PE2:sgРНК), связанного с участком-мишенью ДНК и доменом синтеза ДНК молекулы rPERT для обеспечения

napDNAbp

[0223] В соответствии с используемым здесь значением термин "программируемый нуклеиновой кислотой ДНК-связывающий белок" или "napDNAbp", примером которого является Cas9, обозначает белки, включающие использование РНК:ДНК гибридизации для нацеливания и связывания определенных последовательностей в молекуле ДНК. Каждый napDNAbp связан, по меньшей мере, с одной направляющей нуклеиновой кислотой (например направляющей РНК), которая локализует napDNAbp в последовательности ДНК, которая включает цепь ДНК (то есть цепь-мишень), которая комплементарна направляющей нуклеиновой кислоте или ее части (например протоспейсеру направляющей РНК). Другими словами, направляющая нуклеиновая кислота "программирует" napDNAbp (например Cas9 или его эквивалент) для локализации и связывания с комплементарной последовательностью.

[0224] Без связи с теорией, механизм связывания комплекса napDNAbp-направляющая РНК в целом включает этап образования R-петли, при котором napDNAbp индуцирует разматывание двуцепочечной ДНК-мишени, тем самым разделяя цепи в участке, связываемом napDNAbp. Затем протоспейсер направляющей РНК гибридизируется с "цепью-мишенью". Это обеспечивает вытеснение "не являющейся мишенью цепи", которая комплементарна цепи-мишени, которая образует одноцепочечный участок R-петли. В некоторых вариантах осуществления napDNAbp включает один или несколько видов нуклеазной активности, которые затем обеспечивают разрезание ДНК, оставляя различные типы разрезов. Например napDNAbp может включать нуклеазную активность, обеспечивающую разрезание не являющейся мишенью цепи в первом месторасположении и/или разрезание цепи-мишени во втором месторасположении. В зависимости от нуклеазной активности, ДНК-мишень может быть разрезана с образованием "двуцепочечного разрыва", при котором обе цепи являются разрезанными. В других вариантах осуществления ДНК-мишень может быть разрезана только в одном участке, то есть ДНК "подвергается никированию" на одной цепи. Пример napDNAbp с различными видами нуклеазной активности включает "никазу Cas9" ("nCas9") и деактивированный Cas9 без нуклеазной активности ("мертвый Cas9" или "dCas9"). Здесь приведены примеры последовательностей для данных и других napDNAbp.

Никазы

[0225] Термин "никаза" обозначает Cas9 с деактивированным одним из двух нуклеазных доменов. Данный фермент способен расщеплять только одну цепь ДНК-мишени.

Последовательность ядерной локализации (NLS)

[0226] Термин "последовательность ядерной локализации" или "NLS" обозначает аминокислотную последовательность, которая промотирует импорт белка в ядро клетки, например за счет ядерного транспорта. Последовательности ядерной локализации известны в данной области и будут очевидны специалисту в данной области. Например, последовательности NLS описаны Plank *et al.*, в международной PCT заявке PCT/EP2000/011690, поданной 23 ноября 23 2000 г., опубликованной как WO/2001/038547 31 мая 2001 г., содержание которой приведено здесь в качестве ссылки в части описания примеров последовательностей ядерной локализации. В некоторых вариантах осуществления NLS включает аминокислотную

последовательность PKKKRKV (SEQ ID NO: 16) или MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 17).

Молекула нуклеиновой кислоты

[0227] В соответствии с используемым здесь значением термин "нуклеиновая кислота" обозначает полимер нуклеотидов. Полимер может включать природные нуклеозиды (то есть аденозин, тимидин, гуанозин, цитидин, уридин, дезоксиаденозин, дезокситимидин, дезоксигуанозин и дезоксицитидин), аналоги нуклеозидов (например 2-аминоаденозин, 2-тиотимидин, инозин, пирроло-пиримидин, 3-метиладенозин, 5-метилцитидин, C5-бромуридин, C5-фторуридин, C5-йодоуридин, C5-пропинилуридин, C5-пропинилцитидин, C5-метилцитидин, 7-деазааденозин, 7-деазагуанозин, 8-оксоаденозин, 8-оксогуанозин, O(6)метилгуанин, 4-ацетилцитидин, 5-(карбоксигидроксиметил)уридин, дигидроуридин, метилпсевдоуридин, 1-метиладенозин, 1-метилгуанозин, N6-метиладенозин и 2-тиоцитидин), химически модифицированные основания, биологически модифицированные основания (например, метилированные основания), интеркалированные основания, модифицированные сахара (например, 2'-фторорибозу, рибозу, 2'-дезоксирибозу, 2′-O-метилцитидин, арабинозу и гексозу) или модифицированные фосфатные группы (например, фосфоротиоаты и 5′ N фосфорамидитные связи).

PEgRNA

[0228] В соответствии с используемым здесь значением термин "направляющая РНК для праймерного редактирования" или "PEgRNA" или "удлиненная направляющая РНК" обозначает специализированную форму направляющей РНК, которая была модифицирована для включения одной или нескольких дополнительных последовательностей для реализации описанных здесь способов праймерного редактирования и композиций. Как описано здесь, направляющая РНК для праймерного редактирования включает один или несколько "удлиненных участков" последовательности нуклеиновой кислоты. Удлиненные участки могут включать, но без ограничения, одноцепочечную РНК или ДНК. Дополнительно, удлиненные участки могут присутствовать на 3′-конце традиционной направляющей РНК. В рамках других вариантов удлиненные участки могут присутствовать на 5′-конце традиционной направляющей РНК. В других вариантах удлиненный участок может располагаться во внутримолекулярном участке традиционной направляющей РНК, например в участке в виде ядра нРНК, который ассоциируется и/или связывается с

napDNAbp. Удлиненный участок включает "матрицу для синтеза ДНК", которая кодирует (за счет полимеразы праймерного редактора) одноцепочечную ДНК, которая, в свою очередь, была сконструирована для того, чтобы являться (a) гомологичной подвергаемой редактированию эндогенной ДНК-мишени и (b) которая включает, по меньшей мере, одно требуемое нуклеотидное изменение (например транзицию, трансверсию, делецию, или вставку) для внесения или интеграции в эндогенную ДНК-мишень. Удлиненный участок также может включать другие функциональные элементы последовательности, например, но без ограничения, "участок связывания праймера" и "спейсерную или линкерную" последовательность или другие структурные элементы, например, но без ограничения, аптамеры, структуры типа "стебель-петля", шпильки, опорные петли (например 3' петлю) или домен рекрутирования РНК-белка (например шпильку MS2). В соответствии с используемым здесь значением "участок связывания праймера" включает последовательность, которая гибридизируется с одноцепочечной последовательностью ДНК, имеющей 3′-конец, образованный из никированной ДНК R-петли.

[0229] В определенных вариантах осуществления PEgRNA соответствуют Фиг.3A, на которой показана PEgRNA, имеющая 5′ плечо удлинения, спейсер, и ядро нРНК. 5′ удлинение дополнительно включает в направлении от 5′ к 3′ матрицу в виде обратной транскриптазы, участок связывания праймера, и линкер. Как показано, матрица в виде обратной транскриптазы также может быть обозначена более широко как "матрица для синтеза ДНК", где полимераза описанного здесь праймерного редактора является не RT, а другим типом полимеразы.

[0230] В определенных других вариантах осуществления PEgRNA соответствуют Фиг.3B, на которой показана PEgRNA, имеющая 5′ плечо удлинения, спейсер, и ядро нРНК. 5′ удлинение дополнительно включает в направлении от 5′ к 3′ матрицу в виде обратной транскриптазы, участок связывания праймера, и линкер. Как показано, матрица в виде обратной транскриптазы также может быть обозначена более широко как "матрица для синтеза ДНК", где полимераза описанного здесь праймерного редактора является не RT, а другим типом полимеразы.

[0231] В других вариантах осуществления PEgRNA соответствуют Фиг.3D, на которой показана PEgRNA, имеющая в направлении от 5′ к 3′ спейсер (1), ядро нРНК (2), и плечо удлинения (3). Плечо удлинения (3) расположено на 3'-конце PEgRNA. Плечо удлинения (3) дополнительно включает в направлении от 5′ к 3′ "участок связывания

праймера" (A), "матрицу для редактирования" (B), и "плечо гомологии" (C). Плечо удлинения (3) также может включать необязательный модификаторный участок на 3′ и 5′-концах, который может представлять собой одинаковые или различающиеся последовательности. Кроме того, 3′-конец PEgRNA может включать последовательность терминатора транскрипции. Данные элементы последовательности PEgRNA дополнительно описаны и определены здесь.

[0232] В других вариантах осуществления PEgRNA соответствуют Фиг.3E, на которой показана PEgRNA, имеющая в направлении от 5′ к 3′ плечо удлинения (3), спейсер (1), и ядро нРНК (2). Плечо удлинения (3) расположено на 5′-конце PEgRNA. Плечо удлинения (3) дополнительно включает в направлении от 3′ к 5′ "участок связывания праймера" (A), "матрицу для редактирования" (B), и "плечо гомологии" (C). Плечо удлинения (3) также может включать необязательный модификаторный участок на 3′ и 5′-концах, который может представлять собой одинаковые или различающиеся последовательности. PEgRNA может также включать последовательность терминатора транскрипции на 3′-конце. Данные элементы последовательности PEgRNA дополнительно описаны и определены здесь.

PE1

[0233] В соответствии с используемым здесь значением "PE1" обозначает комплекс PE, включающий слитый белок, включающий Cas9(H840A) и RT MMLV дикого типа, имеющий следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(дикого типа)] + требуемая PEgRNA, где продукт слияния PE имеет аминокислотную последовательность SEQ ID NO: 123, которая показана следующим образом;

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFD LAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHA ILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK**

VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TL NIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQY PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALL QTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTK PFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSE GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTS TLLIENSSP**SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 123)

ОБОЗНАЧЕНИЯ:
ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133)
CAS9(H840A) (SEQ ID NO: 126)
ЛИНКЕР ДЛИНОЙ 33 АМИНОКИСЛОТЫ (SEQ ID NO: 127)
Обратная транскриптаза M-MLV (SEQ ID NO: 128).

PE2

[0234] В соответствии с используемым здесь значением "PE2" обозначает комплекс PE, включающий слитый белок, включающий Cas9(H840A) и вариант RT MMLV, имеющий следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)] + требуемая PEgRNA, где продукт слияния PE имеет аминокислотную последовательность SEQ ID NO: 134, которая показана следующим образом:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFD LAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHA ILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW**

NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV
TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE
DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV
KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV
ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD
VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW
DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF
LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF
LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV
LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT
LIHQSITGLYETRIDLSQLGGD<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u>*TL*
NIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQY
PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED
IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT
WTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALL
QTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF
LGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTK
PFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG
KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL
PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET
EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTS
EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDT
STLLIENSSP<u>SGGSKRTADGSEFEPKKKRKV</u> (SEQ ID NO: 134)

ОБОЗНАЧЕНИЯ:
<u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS)</u> ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133)
CAS9(H840A) (SEQ ID NO: 137)
<u>ЛИНКЕР ДЛИНОЙ 33 АМИНОКИСЛОТЫ</u> (SEQ ID NO: 127)
Обратная транскриптаза M-MLV (SEQ ID NO: 139).

<u>PE3</u>

[0235] В соответствии с используемым здесь значением термин "PE3" обозначает PE2 плюс никирующую вторую цепь направляющую РНК, которая образует комплекс с PE2 и вносит разрыв в не подвергаемой редактированию цепи ДНК для индуцирования предпочтительно замещения подвергнутой редактированию цепи.

<u>PE3b</u>

[0236] В соответствии с используемым здесь значением термин "PE3b" обозначает PE3, но где никирующая вторую цепь направляющая РНК сконструирована для временно́го контроля, так что разрыв второй цепи не осуществляется до реализации

внесения требуемого изменения. Это достигается за счет конструирования нРНК со спейсерной последовательностью, которая соответствует только подвергаемой редактированию цепи, но не исходному аллелю. За счет использования данной стратегии, обозначаемой далее как PE3b, несовпадения между протоспейсером и не подвергаемым редактированию аллелем должны способствовать отсутствию никирования за счет sgРНК до реализации редактирования на включающей PAM цепи.

PE-short

[0237] В соответствии с используемым здесь значением термин "PE-short" обозначает конструкт PE, который слит с усеченной по С-концу обратной транскриптазой и имеет следующую аминокислотную последовательность:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLG
NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMA
KVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPI
NASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT
EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI
TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS
GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV
VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL
KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD
DSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL
IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL
ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT
STKEVLDATLIHQSITGLYETRIDLSQLGGD**<u>SGGSSGGSSGSETPGTSESATPESS
GGSSGGSS</u>***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPL
KATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQ
DLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR
DPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSEL
DCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPT
PKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLT
APALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRM
VAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF*

*GPVVALNPATLLPLPEEGLQHNCLDNSRLIN*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 765)

ОБОЗНАЧЕНИЯ:
ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133)
CAS9(H840A) (SEQ ID NO: 157)
ЛИНКЕР 1 ДЛИНОЙ 33 АМИНОКИСЛОТЫ (SEQ ID NO: 127)
УСЕЧЕННАЯ ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV
(SEQ ID NO: 766)

Пептидная метка

[0238] Термин "пептидная метка" обозначает аминокислотную последовательность пептида, которая генетически слита с последовательностью белка для придания одной или нескольких функций белкам, которая обеспечивает возможность манипуляции белком для различных целей, например визуализации, очистки, солюбилизации, и отделения, и так далее. Пептидные метки могут включать различные типы меток, разделенных на категории на основе цели или функции, которые могут включать "аффинные метки" (для реализации очистки белка), "метки солюбилизации" (для обеспечения правильного фолдинга белков), "хроматографические метки" (для изменения хроматографических свойств белков), "эпитопные метки" (для связывания с антителами с высокой аффинностью), "флуоресцентные метки" (для обеспечения визуализации белков в клетке или in vitro).

Полимераза

[0239] В соответствии с используемым здесь значением термин "полимераза" обозначает фермент, который синтезирует нуклеотидную цепь и который может применяться совместно с описанными здесь системами праймерных редакторов. Полимераза может являться "зависимой от матрицы" полимеразой (то есть полимеразой, которая синтезирует нуклеотидную цепь на основе порядка нуклеотидных оснований матричной цепи). Полимераза также может являться "независимой от матрицы" полимеразой (то есть полимеразой, которая синтезирует нуклеотидную цепь без необходимости в матричной цепи). Полимераза также может быть дополнительно разделена на категории в виде "ДНК-полимеразы" или "РНК-полимеразы". В различных вариантах осуществления система праймерного редактора включает ДНК-полимеразу. В различных вариантах осуществления ДНК-полимераза может представлять собой "ДНК-зависимую ДНК-полимеразу (где молекула матрицы представляет собой цепь ДНК). В таких случаях молекула матрицы ДНК может

представлять собой PEgRNA, где плечо удлинения включает цепь ДНК. В таких случаях PEgRNA может быть обозначена как химерная или гибридная PEgRNA, которая включает часть в виде РНК (то есть компоненты направляющей РНК, включая спейсер и ядро нРНК) и часть в виде ДНК (например плечо удлинения). В различных других вариантах осуществления ДНК-полимераза может представлять собой "РНК-зависимую ДНК-полимеразу" (где молекула матрицы представляет собой цепь РНК). В таких случаях PEgRNA представляет собой РНК, то есть включает удлинение РНК. Термин "полимераза" также может обозначать фермент, который катализирует полимеразу нуклеотида (то есть полимеразную активность). В целом, данный фермент будет инициировать синтез на 3'-конце праймера, отжигаемого на последовательности полинуклеотидной матрицы (например последовательности праймера, отжигаемой на участке связывания праймера PEgRNA) по направлению к 5'-концу цепи матрицы. "ДНК-полимераза" катализирует полимеризацию дезоксинуклеотидов. В соответствии с используемым здесь значением в отношении ДНК-полимеразы термин "ДНК-полимераза" включает "ее функциональный фрагмент". Термин "ее функциональный фрагмент" обозначает любую часть относящейся к дикому типу или мутантной ДНК-полимеразы, которая охватывает менее чем целую аминокислотную последовательность полимеразы и которая сохраняет способность, по меньшей мере, при одном наборе условий катализировать полимеризацию полинуклеотида. Подобный функциональный фрагмент может присутствовать в виде отдельной сущности или может представлять собой часть более крупного полипептида, например слитого белка.

Праймерное редактирование

[0240] В соответствии с используемым здесь значением термин "праймерное редактирование" обозначает новый подход для редактирования генов с использованием napDNAbp, полимеразы (например обратной транскриптазы), и специализированных направляющих РНК, которые включают матрицу для синтеза ДНК, для кодирования требуемой новой генетической информации (или удаления генетической информации), которая затем встраивается в последовательность-мишень ДНК. Определенные варианты осуществления праймерного редактирования описаны в рамках вариантов осуществления на Фиг.1A-1H и Фиг.72(a)-72(c), среди других фигур.

[0241] Праймерное редактирование представляет собой полностью новую платформу для генетического редактирования, которая является универсальным и точным способом геномного редактирования, в рамках которого осуществляется прямая запись

новой генетической информации в определенный участок ДНК с помощью программируемого нуклеиновой кислотой ДНК-связывающего белка ("napDNAbp"), работающего совместно с полимеразой (то есть в форме слитого белка или иным образом поступающего *in trans* с napDNAbp), при этом система праймерного редактирования программируется направляющей РНК ("PEgRNA") для праймерного редактирования (PE), которая как указывает участок-мишень, так и служит матрицей для синтеза требуемого изменения в форме замещающей цепи ДНК в виде удлинения (ДНК или РНК), встроенного в направляющую РНК (например на 5′ или 3′-конце, или во внутренней части направляющей РНК). Замещающая цепь, содержащая требуемое изменение (например замещение одного азотистого основания), имеет (или является гомологичной) той же последовательности, что и эндогенная цепь (непосредственно ниже участка разрыва) подвергаемого редактированию участка-мишени (за исключением того, что она включает требуемое изменение). С помощью механизма репарации и/или репликации ДНК эндогенная цепь ниже участка разрыва замещается новой синтезированной замещающей цепью, содержащей требуемое изменение. В некоторых случаях праймерное редактирование может быть описано как технология геномного редактирования типа "поиск-и-замена", поскольку праймерные редакторы, как описано здесь, не только обеспечивают поиск и нахождение требуемого участка-мишени для редактирования, но в то же время кодируют замещающую цепь, содержащую требуемое изменение, которая встраивается вместо соответствующей эндогенной цепи ДНК участка-мишени. Праймерные редакторы по настоящему изобретению частично связаны с открытием того, что механизм мишень-праймированной обратной транскрипции (TPRT) или "праймерное редактирование" может быть использован или адаптирован для осуществления точного геномного редактирования на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью (например как показано в рамках различных вариантов осуществления на Фиг.1A-1F). TPRT естественным образом применяется в рамках подвижных элементов ДНК, например ретротранспозонов млекопитающих без длинных концевых повторов (LTR) и бактериальных интронов Группы II[28,29]. Здесь авторы использовали продукты слияния белок Cas-обратная транскриптаза или родственные системы для нацеливания на определенную последовательность ДНК с помощью направляющей РНК, осуществления одноцепочечного разрыва в участке-мишени, и применения никированной ДНК в качестве праймера для обратной транскрипции

сконструированной матрицы в виде обратной транскриптазы, которая встраивается в направляющую РНК. Однако в то время как данная концепция начинается с праймерных редакторов, в рамках которых используется обратная транскриптаза в качестве компонента ДНК-полимеразы, описанные здесь праймерные редакторы не ограничиваются обратными транскриптазами, но могут включать использование практически любой ДНК-полимеразы. Действительно, в то время как по тексту заявки могут указываться праймерные редакторы с "обратными транскриптазами", здесь устанавливается, что обратные транскриптазы являются лишь одним из типов ДНК-полимераз, которые могут работать в рамках праймерного редактирования. Таким образом, при указании в данном описании "обратной транскриптазы" специалист в данной области должен понимать, что вместо обратной транскриптазы может использоваться любая подходящая ДНК-полимераза. Таким образом, в одном аспекте праймерные редакторы могут включать Cas9 (или эквивалентный napDNAbp), который программируется для нацеливания на ДНК-последовательность путем его связывания со специализированной направляющей РНК (то есть PEgRNA), содержащей спейсерную последовательность, которая отжигается на комплементарном протоспейсере в ДНК-мишени. Специализированная направляющая РНК также содержит новую генетическую информацию в форме удлинения, которое кодирует замещающую цепь ДНК, содержащую требуемое генетическое изменение, которая используется для замены соответствующей эндогенной цепи ДНК в участке-мишени. Для переноса информации из PEgRNA в ДНК-мишень механизм праймерного редактирования включает никирование участка-мишени в одной цепи ДНК для высвобождения 3′-гидроксильной группы. Затем свободная 3′-гидроксильная группа может использоваться для праймирования полимеризации ДНК кодирующего изменение удлинения на PEgRNA напрямую в участок-мишень. В различных вариантах осуществления удлинение, которое обеспечивает матрицу для полимеризации замещающей содержащей изменение цепи, может быть образовано из РНК или ДНК. В случае РНК удлинения полимераза праймерного редактора может представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). В случае ДНК удлинения полимераза праймерного редактора может представлять собой ДНК-зависимую ДНК-полимеразу. Новая синтезированная цепь (то есть замещающая цепь ДНК, содержащая требуемое изменение), которая образуется описанными здесь праймерными редакторами, будет являться

гомологичной относительно последовательности геномной мишени (то есть иметь такую же последовательность) за исключением включения требуемого нуклеотидного изменения (например изменения отдельного нуклеотида, делеции, или вставки, или их комбинации). Новая синтезированная (или замещающая) цепь ДНК также может быть обозначена как одноцепочечный ДНК-флэп, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. В определенных вариантах осуществления система может комбинироваться с использованием фермента в виде подверженной ошибкам обратной транскриптазы (например поступающей в виде слитого с Cas9 доменом белка или поступающей с Cas9 доменом *in trans*). Фермент в виде подверженной ошибкам обратной транскриптазы может вносить изменения при синтезе одноцепочечного ДНК-флэпа. Таким образом, в определенных вариантах осуществления подверженная ошибкам обратная транскриптаза может использоваться для внесения нуклеотидных изменений в ДНК-мишень. В зависимости от подверженной ошибкам обратная транскриптазы, которая используется в системе, изменения могут являться случайными или неслучайными. Коррекция гибридизированного промежуточного продукта (включающего одноцепочечный ДНК-флэп, синтезированный обратной транскриптазой, гибридизированной с эндогенной цепью ДНК) может включать удаление образованного вытесненного флэпа эндогенной ДНК (например с помощью эндонуклеазы 5′-концевого ДНК-флэпа, FEN1), лигирование синтезированного одноцепочечного ДНК-флэпа с ДНК-мишенью, и ассимилирование требуемого нуклеотидного изменения в результате процессов репарации и/или репликации клеточной ДНК. Поскольку синтез ДНК на основе матриц обеспечивает точность на уровне отдельных нуклеотидов при модификации любого нуклеотида, включая вставки и делеции, объем данного подхода является очень широким и, как представляется, он может использоваться в рамках множества вариантов применения в фундаментальной науке и терапии.

[0242] В различных вариантах осуществления праймерное редактирование осуществляется путем контактирования молекулы-мишени ДНК (для которое требуется внести изменение в нуклеотидной последовательности) с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) в виде комплекса с направляющей РНК для праймерного редактирования (PEgRNA). Со ссылкой на Фиг.1G, направляющая РНК для праймерного редактирования (PEgRNA) включает

удлинение на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК и кодирует требуемое нуклеотидное изменение (например изменение одного нуклеотида, вставку, или делецию). На этапе (a) комплекс в виде napDNAbp/удлиненной нРНК контактирует с молекулой ДНК, при этом удлиненная нРНК направляет napDNAbp для связывания с локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК, то есть "не являющейся мишенью цепи". Однако разрыв должен быть осуществлен в любой одной из цепей. То есть разрыв должен быть осуществлен в "цепи-мишени" R-петли (то есть цепи, гибридизированной с протоспейсером удлиненной нРНК) или "не являющейся мишенью цепи" (то есть цепи, образующей одноцепочечную часть R-петли, которая является комплементарной цепи-мишени). На этапе (c) 3′-концевая цепь ДНК (образованная разрывом) взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции (то есть "праймированной мишенью RT"). В определенных вариантах осуществления 3′-конец цепи ДНК гибридизируется с определенной праймирующей последовательностью RT на удлиненной части направляющей РНК, то есть "праймирующей последовательностью обратной транскриптазы" или "участком связывания праймера" на PEgRNA. На этапе (d) вносится обратная транскриптаза (или другая подходящая ДНК-полимераза), которая синтезирует одиночную цепь ДНК от 3′-конца праймированного участка к 5′-концу направляющей РНК для праймерного редактирования. ДНК-полимераза (например обратная транскриптаза) может быть слита с napDNAbp или, в качестве альтернативы, быть предоставлена napDNAAbp *in trans*. Это образует одноцепочечный ДНК-флэп, включающий требуемое нуклеотидное изменение (например изменение одного основания, вставку, или делецию, или их комбинацию) и являющийся иным образом гомологичным эндогенной ДНК в участке разрыва или являющейся смежной с ним. На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа, так что требуемое нуклеотидное изменение становится встроенным в локус-мишень. Данный процесс может быть смещен в

сторону образования требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа, который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с последовательностью эндогенной ДНК. Без связи с теорией, процессы репарации и репликации клеточной эндогенной ДНК обеспечивают коррекцию несовпадающей ДНК для встраивания нуклеотидного изменения (изменений) с образованием требуемого измененного продукта. Процесс также может быть смещен в сторону образования продукта с "никированием второй цепи", как показано в качестве примера на Фиг.1F. Данный процесс может обеспечивать внесение, по меньшей мере, одного или нескольких из следующих генетических изменений: трансверсий, транзиций, делеций, и вставок.

[0243] Термин "система праймерного редактора (PE)" или "праймерный редактор (PE)" или "система PE" или "система редактирования PE" обозначает композиции, используемые в рамках способа геномного редактирования с применением описанной здесь мишень-праймированной обратной транскрипции (TPRT), включая, но без ограничения, napDNAbp, обратные транскриптазы, слитые белки (например включающие napDNAbp и обратные транскриптазы), направляющие РНК для праймерного редактирования, и комплексы, включающие слитые белки и направляющие РНК для праймерного редактирования, а также вспомогательные элементы, например компоненты для никирования второй цепи (например sgРНК второй цепи) и эндонуклеазы для удаления 5′ эндогенного ДНК-флэпа (например FEN1) для обеспечения сдвига процесса праймерного редактирования в сторону образования подвергнутого редактированию продукта.

[0244] Несмотря на то, что в описанных до сих пор вариантах осуществления PEgRNA представляет собой одиночную молекулу, включающую направляющую РНК (которая сама по себе включает спейсерную последовательность и ядро или каркас нРНК) и 5′ или 3′ плечо удлинения, включающее участок связывания праймера и матрицу для синтеза ДНК (например см. Фиг.3D), PEgRNA может также принимать форму двух отдельных молекул, состоящих из направляющей РНК и *trans* матрицы РНК праймерного редактора (tPERT), которая по существу включает плечо удлинения (включая, в частности участок связывания праймера и домен синтеза ДНК) и домен рекрутирования РНК-белка (например аптамер или шпильку MS2) в той же молекуле, которая становится солокализованной или рекрутируется к модифицированному комплексу праймерного редактирования, который включает белок рекрутирования

rPERT (например белок MS2cp, который связывается с аптамером MS2). См. Фиг.3G и Фиг.3H в качестве примера tPERT, которая может применяться в рамках праймерного редактирования.

Праймерный редактор

[0245] Термин "праймерный редактор" обозначает описанные здесь слитые конструкты, включающие napDNAbp (например никазу Cas9) и обратную транскриптазу, способные осуществлять праймерное редактирование в нуклеотидной последовательности-мишени в присутствии PEgRNA (или "удлиненной направляющей РНК"). Термин "праймерный редактор" может обозначать слитый белок или слитый белок в виде комплекса с PEgRNA, и/или дополнительно в виде комплекса с никирующей вторую цепь sgРНК. В некоторых вариантах осуществления термин "праймерный редактор" может также обозначать комплекс, включающий слитый белок (обратную транскриптазу, слитую с napDNAbp), PEgRNA, и обычную направляющую РНК, способную направлять этап никирования второго участка не подвергаемой редактированию цепи, как описано здесь. В других вариантах осуществления компонент "праймерного редактора" в виде обратной транскриптазы может поступать *in trans*.

Участок связывания праймера

[0246] Термин "участок связывания праймера" или "PBS" обозначает нуклеотидную последовательность, расположенную на PEgRNA в качестве компонента плеча удлинения (обычно на 3′-конце плеча удлинения) и служащую для связывания с праймерной последовательностью, которая образуется после никирования последовательности-мишени за счет Cas9 с помощью праймерного редактора. Как подробно описано в других разделах, когда компонент праймерного редактора в виде никазы Cas9 осуществляет никирование одной цепи последовательности-мишени ДНК, образуется 3′-концевой оцДНК-флэп, который служит праймерной последовательностью, которая отжигается на участке связывания праймера на PEgRNA для праймирования обратной транскрипции. На Фиг.27 и 28 показаны варианты осуществления участка связывания праймера, расположенного на 3′ и 5′ плече удлинения, соответственно.

Промотор

[0247] Термин "промотор" используется в данной области и обозначает молекулу нуклеиновой кислоты с последовательностью, распознаваемой механизмом клеточной

транскрипции и способной инициировать транскрипцию расположенного ниже гена. Промотор может являться постоянно активным, что означает, что промотор является всегда активным в указанном клеточном контекста, или условно активным, что означает, что промотор является активным лишь при наличии определенного условия. Например, условный промотор может являться активным лишь в присутствии определенного белка, который соединяет белок, ассоциированный с регуляторным элементом в промоторе, с механизмом транскрипции оснований, или лишь при отсутствии ингибиторной молекулы. Подклассом условно активных промоторов являются индуцируемые промоторы, которые требуют наличия низкомолекулярного "индуктора" для активности. Примеры индуцируемых промоторов включают, но без ограничения, индуцируемые арабинозой промоторы, Tet-on промоторы, и индуцируемые тамоксифеном промоторы. Различные постоянные, условные, и индуцируемые промоторы хорошо известны специалисту в данной области, при этом специалист в данной области будет способен определить ряд подобных промоторов, которые могут применяться при реализации настоящего изобретения, которое не является ограниченным в данном отношении.

Протоспейсер

[0248] В соответствии с используемым здесь значением термин "протоспейсер" обозначает последовательность (~20 bp) в ДНК, являющуюся смежной относительно последовательности PAM (смежного с протоспейсером мотива). Протоспейсер имеет ту же последовательность, что и спейсерная последовательность направляющей РНК. Направляющая РНК отжигается на комплементе протоспейсерной последовательности на ДНК-мишени (в частности, одной ее цепи, то есть "цепи-мишени" против "не являющейся мишенью цепи" последовательности-мишени ДНК). Для работы Cas9 ей также требуется определенный смежный с протоспейсером мотив (PAM), который варьируется в зависимости от вида бактерий, к которому относится ген Cas9. Наиболее часто используемая нуклеаза Cas9, относящаяся к *S. pyogenes*, распознает последовательность PAM NGG, которая располагается непосредственно ниже последовательности-мишени в геномной ДНК на не являющейся мишенью цепи. Специалисту в данной области будет известно, что "протоспейсер" в литературе в данной области иногда указывается в виде мишень-специфичной направляющей последовательности длиной ~20 nt на самой направляющей РНК вместо указания его как "спейсера". Таким образом, в некоторых случаях термин "протоспейсер" в

соответствии с используемым здесь значением может использоваться взаимозаменяемым образом с термином "спейсер". Контекст описания, связанный с указанием "протоспейсера" или "спейсера", поможет читателю определить, указывает ли данный термин на нРНК или мишень в виде ДНК.

Смежный с протоспейсером мотив (PAM)

[0249] В соответствии с используемым здесь значением термин "смежная с протоспейсером последовательность" или "PAM" обозначает последовательность ДНК длиной приблизительно 2-6 пар оснований, которая является важным нацеливающим компонентом никазы Cas9. Как правило, последовательность PAM располагается на любой одной из цепей и расположена ниже в направлении от 5′ к 3′ участка разреза Cas9. Стандартная последовательность PAM (то есть последовательность PAM, которая ассоциирована с нуклеазой Cas9 *Streptococcus pyogenes* или SpCas9) представляет собой 5′-NGG-3′, где "N" представляет собой любое азотистое основание с последующими двумя азотистыми основаниями в виде гуанина ("G"). Различные последовательности PAM могут быть ассоциированы с различными нуклеазами Cas9 или эквивалентными белками из различных организмов. Кроме того, любая нуклеаза Cas9, например SpCas9, может быть модифицирована для изменения специфичности нуклеазы в отношении PAM, так что нуклеаза будет распознавать альтернативную последовательность PAM.

[0250] Например, с учетом того, что аминокислотная последовательность стандартной SpCas9 представляет собой SEQ ID NO: 18, последовательность PAM может быть модифицирована путем внесения одной или нескольких мутаций, включая (a) D1135V, R1335Q, и T1337R "вариант VQR", который изменяет специфичность в отношении PAM на NGAN или NGNG, (b) D1135E, R1335Q, и T1337R "вариант EQR", который изменяет специфичность в отношении PAM на NGAG, и (c) D1135V, G1218R, R1335E, и T1337R "вариант VRER", который изменяет специфичность в отношении PAM на NGCG. Кроме того, вариант D1135E стандартной SpCas9 все еще распознает NGG, но является более избирательным по сравнению с белком SpCas9 дикого типа.

[0251] Также будет очевидно, что ферменты Cas9 из бактерий различных видов (то есть ортологов Cas9) могут иметь различные специфичности в отношении PAM. Например, Cas9 из *Staphylococcus aureus* (SaCas9) распознает NGRRT или NGRRN. Кроме того, Cas9 из *Neisseria meningitis* (NmCas) распознает NNNNGATT. В другом примере Cas9 из *Streptococcus thermophilis* (StCas9) распознает NNAGAAW. В другом

примере Cas9 из *Treponema denticola* (TdCas) распознает NAAAAC. Они представляют собой пример, который не является ограничивающим. Также будет очевидно, что отличные от SpCas9 нуклеазы связываются с рядом последовательностей PAM, что делает их полезными при отсутствии подходящей последовательности PAM для SpCas9 в требуемом участке-мишени разрыва. Более того, отличные от SpCas9 нуклеазы могут иметь другие характеристики, которые делают их более полезными, чем SpCas9. Например Cas9 из *Staphylococcus aureus* (SaCas9) имеет на 1 килооснование меньший размер, чем SpCas9, поэтому она может быть упакована в аденоассоциированный вирус (AAV). Дополнительные ссылки могут быть сделаны на Shah et al., "Protospacer recognition motifs: mixed identities and functional diversity," *RNA Biology*, 10(5): 891-899 (который приведен здесь в качестве ссылки).

Рекомбиназа

[0252] Термин "рекомбиназа" в соответствии с используемым здесь значением обозначает участок-специфический фрагмент, который опосредует рекомбинацию ДНК между последовательностями распознавания рекомбиназой, которая приводит к эксцизии, интеграции, инверсии или замене (например транслокации) фрагментов ДНК между последовательностями распознавания рекомбиназой. Рекомбиназы могут быть классифицированы на два отличных друг от друга семейства: сериновые рекомбиназы (например резольвазы и инвертазы) и тирозиновые рекомбиназы (например интегразы). Примеры сериновых рекомбиназ включают, без ограничения, Hin, Gin, Tn3, β-six, CinH, ParA, γδ, Bxb1, ϕC31, TP901, TG1, φBT1, R4, φRV1, φFC1, MR11, A118, U153, и gp29. Примеры тирозиновых рекомбиназ включают, без ограничения, Cre, FLP, R, лямбда, HK101, HK022, и pSAM2. Название сериновых и тирозиновых рекомбиназ основано на консервативном нуклеофильном аминокислотном остатке, который используется рекомбиназой для атаки ДНК и который становится ковалентно связанным с ДНК при замене цепей. Рекомбиназы имеют множества вариантов использования, включая создания генных нокаутов/нокинов и варианты использования в рамках генной терапии. См., например, , Brown *et al.*, "Serine recombinases as tools for genome engineering." *Methods*. 2011;53(4):372-9; Hirano *et al.*, "Site-specific recombinases as tools for heterologous gene integration." *Appl. Microbiol. Biotechnol.* 2011; 92(2):227-39; Chavez and Calos, "Therapeutic applications of the ΦC31 integrase system." *Curr. Gene Ther.* 2011;11(5):375-81; Turan and Bode, "Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications." *FASEB J.* 2011; 25(12):4088-107;

Venken and Bellen, “Genome-wide manipulations of Drosophila melanogaster with transposons, Flp recombinase, and ΦC31 integrase.” *Methods Mol. Biol.* 2012; 859:203-28; Murphy, “Phage recombinases and their applications.” *Adv. Virus Res.* 2012; 83:367-414; Zhang *et al.*, “Conditional gene manipulation: Cre-ating a new biological era.” *J. Zhejiang Univ. Sci. B.* 2012; 13(7):511-24; Karpenshif and Bernstein, “From yeast to mammals: recent advances in genetic control of homologous recombination.” *DNA Repair (Amst)*. 2012; 1;11(10):781-8; полное содержание каждого из которых полностью приведено здесь в качестве ссылки. Указанные здесь рекомбиназы не являются исключающими примерами рекомбиназ, которые могут применяться в рамках вариантов осуществления изобретения. Способы и композиции по изобретению могут быть расширены путем анализа баз данных на предмет новых ортогональных рекомбиназ или конструирования синтетических рекомбиназ с определенными видами специфичности в отношении ДНК (см., например, Groth *et al.*, “Phage integrases: biology and applications.” *J. Mol. Biol*. 2004; 335, 667-678; Gordley *et al.*, “Synthesis of programmable integrases.” *Proc. Natl. Acad. Sci. U S A*. 2009; 106, 5053-5058; полное содержание каждого из которых полностью приведено здесь в качестве ссылки). Другие примеры рекомбиназ, которые могут применяться в рамках описанных здесь способов и композиций известны специалистам в данной области, при этом ожидается, что каждая новая рекомбиназа, которую открывают или получают, может быть использована в рамках различных вариантов осуществления изобретения. В некоторых вариантах осуществления каталитические домены рекомбиназы сливаются с программируемой деактивированной нуклеазой РНК нуклеазой (например dCas9 или ее фрагментом), так что домен рекомбиназы не включает домен связывания нуклеиновой кислоты или не способен связываться с нуклеиновой кислотой-мишенью (например домен рекомбиназы сконструирован таким образом, что он не имеет активности в плане связывания определенной ДНК). Рекомбиназы без активности в виде связывания ДНК и способы их конструирования известны и включают описанные Klippel *et al.*, “Isolation and characterisation of unusual gin mutants.” *EMBO J.* 1988; 7: 3983–3989: Burke *et al.*, “Activating mutations of Tn3 resolvase marking interfaces important in recombination catalysis and its regulation. *Mol Microbiol*. 2004; 51: 937–948; Olorunniji *et al.*, “Synapsis and catalysis by activated Tn3 resolvase mutants.” *Nucleic Acids Res*. 2008; 36: 7181–7191; Rowland *et al.*, “Regulatory mutations in Sin recombinase support a structure-based model of the synaptosome.” *Mol Microbiol*. 2009; 74: 282–298;

Akopian *et al.*, "Chimeric recombinases with designed DNA sequence recognition." *Proc Natl Acad Sci USA.* 2003;100: 8688–8691; Gordley *et al.*, "Evolution of programmable zinc finger-recombinases with activity in human cells. *J Mol Biol.* 2007; 367: 802–813; Gordley *et al.*, "Synthesis of programmable integrases." *Proc Natl Acad Sci USA.* 2009;106: 5053–5058; Arnold *et al.*, "Mutants of Tn*3* resolvase which do not require accessory binding sites for recombination activity." *EMBO J.* 1999;18: 1407–1414; Gaj *et al.*, "Structure-guided reprogramming of serine recombinase DNA sequence specificity." *Proc Natl Acad Sci USA.* 2011;108(2):498-503; и Proudfoot *et al.*, "Zinc finger recombinases with adaptable DNA sequence specificity." *PLoS One.* 2011;6(4):e19537; полное содержание каждого из которого приведено здесь в качестве ссылки. Например, сериновые рекомбиназы группы резольваз-инвертаз, например резольвазы Tn3 и γδ и инвертазы Hin и Gin имеют модульные структуры с автономными каталитическими и связывающими ДНК доменами (см., например, Grindley *et al.*, "Mechanism of site-specific recombination." *Ann Rev Biochem.* 2006; 75: 567–605, полное содержание которого приведено здесь в качестве ссылки). Каталитические домены данных рекомбиназ, таким образом, являются поддающимися рекомбинации с использованием программируемых деактивированной нуклеазой РНК нуклеаз (например dCas9 или ее фрагмента), как описано здесь, например после выделения "активированных" мутантов рекомбиназы, которые не требуют каких-либо вспомогательных факторов (например активности в виде связывания ДНК) (см., например, Klippel *et al.*, "Isolation and characterisation of unusual gin mutants." *EMBO J.* 1988; 7: 3983–3989: Burke *et al.*, "Activating mutations of Tn3 resolvase marking interfaces important in recombination catalysis and its regulation. *Mol Microbiol.* 2004; 51: 937–948; Olorunniji *et al.*, "Synapsis and catalysis by activated Tn3 resolvase mutants." *Nucleic Acids Res.* 2008; 36: 7181–7191; Rowland *et al.*, "Regulatory mutations in Sin recombinase support a structure-based model of the synaptosome." *Mol Microbiol.* 2009; 74: 282–298; Akopian *et al.*, "Chimeric recombinases with designed DNA sequence recognition." *Proc Natl Acad Sci USA.* 2003;100: 8688–8691). Кроме того, известны многие другие природные сериновые рекомбиназы, имеющие N-концевой каталитический домен и С-концевой домен связывания ДНК (например интеграза phiC31, транспозаза TnpX, транспозаза IS607), при этом их каталитические домены могут быть кооптированы для конструирования программируемых специфических в плане участка рекомбиназ, как описано здесь (см., например, Smith *et al.*, "Diversity in the serine recombinases." *Mol Microbiol.* 2002;44: 299–307, полное содержание которого

приведено здесь в качестве ссылки). Аналогичным образом, известны каталитические домены ядра тирозиновых рекомбиназ (например Cre, интегразы λ), при этом они могут быть также кооптированы для конструирования программируемых специфических в плане участка рекомбиназ как описано здесь (см., например, Guo *et al.*, "Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse." *Nature*. 1997; 389:40–46; Hartung *et al.*, "Cre mutants with altered DNA binding properties." *J Biol Chem* 1998; 273:22884–22891; Shaikh *et al.*, "Chimeras of the Flp and Cre recombinases: Tests of the mode of cleavage by Flp and Cre. *J Mol Biol*. 2000; 302:27–48; Rongrong *et al.*, "Effect of deletion mutation on the recombination activity of Cre recombinase." *Acta Biochim Pol*. 2005; 52:541–544; Kilbride *et al.*, "Determinants of product topology in a hybrid Cre-Tn3 resolvase site-specific recombination system." *J Mol Biol*. 2006; 355:185–195; Warren *et al.*, "A chimeric cre recombinase with regulated directionality." *Proc Natl Acad Sci USA*. 2008 105:18278–18283; Van Duyne, "Teaching Cre to follow directions." *Proc Natl Acad Sci USA*. 2009 Jan 6;106(1):4-5; Numrych *et al.*, "A comparison of the effects of single-base and triple-base changes in the integrase arm-type binding sites on the site-specific recombination of bacteriophage λ." *Nucleic Acids Res*. 1990; 18:3953–3959; Tirumalai *et al.*, "The recognition of core-type DNA sites by λ integrase." *J Mol Biol*. 1998; 279:513–527; Aihara *et al.*, "A conformational switch controls the DNA cleavage activity of λ integrase." *Mol Cell*. 2003; 12:187–198; Biswas *et al.*, "A structural basis for allosteric control of DNA recombination by λ integrase." *Nature*. 2005; 435:1059–1066; и Warren *et al.*, "Mutations in the amino-terminal domain of λ-integrase have differential effects on integrative and excisive recombination." *Mol Microbiol*. 2005; 55:1104–1112; полное содержание каждого из которых приведено здесь в качестве ссылки).

Последовательность распознавания рекомбиназой

[0253] В соответствии с используемым здесь значением термин "последовательность распознавания рекомбиназой" или, равным образом, "RRS" или "последовательность-мишень рекомбиназы" обозначает нуклеотидную последовательность-мишень, распознаваемую рекомбиназой, которая подвергается замене цепи с использованием другой молекулы ДНК, имеющей RRS, что приводит к эксцизии, интеграции, инверсии или замене фрагментов ДНК между последовательностями распознавания рекомбиназой.

Рекомбинировать или рекомбинация

[0254] Термин "рекомбинировать" или "рекомбинация" в контексте модификации нуклеиновой кислоты (например геномной модификации) используется для обозначения процесса, с помощью которого две или более молекулы нуклеиновой кислоты или два или более участка одной молекулы нуклеиновой кислоты модифицируются за счет действия белка рекомбиназы (например указанного здесь слитого белка рекомбиназы по изобретению). Рекомбинация может приводить, *inter alia*, к вставке, инверсии, эксцизии или транслокации нуклеиновых кислот, например в рамках или между одной или несколькими молекулами нуклеиновой кислоты. последовательности распознавания рекомбиназой

Обратная транскриптаза

[0255] Термин "обратная транскриптаза" обозначает класс полимераз, характеризующихся как РНК-зависимые ДНК-полимеразы. Все известные обратные транскриптазы требуют праймер для синтеза ДНК транскрипта из РНК матрицы. Исторически обратную транскриптазу использовали в основном для транскрибирования иРНК в кДНК, которая затем может быть клонирована в вектор для дальнейшего манипулирования. Обратная транскриптаза вируса миобластоза птиц (AMV) была первой широко используемой РНК-зависимой ДНК-полимеразой (Verma, Biochim. Biophys. Acta 473:1 (1977)). Данный фермент имеет 5′-3′ РНК-направленную ДНК-полимеразную активность, 5′-3′ ДНК-направленную ДНК-полимеразную активность, и активность РНКазы H. РНКаза H представляет собой процессивную 5′ и 3′ рибонуклеазу, являющуюся специфичной для цепи РНК для РНК-ДНК гибридов (Perbal, A Practical Guide to Molecular Cloning, New York: Wiley & Sons (1984)). Ошибки транскрипции не могут быть скорректированы обратной транскриптазой, поскольку известные вирусные обратные транскриптазы не имеют 3′-5′ экзонуклеазную активность, необходимую для проверки (Saunders and Saunders, Microbial Genetics Applied to Biotechnology, London: Croom Helm (1987)). Подробное исследование активности обратной транскриптазы AMV и связанной с ней активности РНКазы H было представлено Berger et al., Biochemistry 22:2365-2372 (1983). Другая обратная транскриптаза, которая широко применяется в молекулярной биологии, представляет собой обратную транскриптазу из вируса лейкоза мышей Молони (M-MLV). См., например, Gerard, G. R., DNA 5:271-279 (1986) и Kotewicz, M. L., et al., Gene 35:249-258 (1985). Также была описана обратная транскриптаза M-MLV, по существу не имеющая

активности РНКазы H. См., например, патент США 5244797. Изобретение охватывает применение любых подобных обратных транскриптаз или их вариантов или мутантов.

[0256] Кроме того, изобретение охватывает применение обратных транскриптаз, которые подвержены ошибкам, то есть которые могут быть обозначены как подверженные ошибкам обратные транскриптазы или обратные транскриптазы, которые не поддерживают встраивание нуклеотидов с высокой точностью при полимеризации. При синтезе одноцепочечного ДНК-флэпа на основе RT матрицы, встроенной в направляющую ДНК, подверженная ошибкам обратная транскриптаза может вносить один или несколько нуклеотидов, которые не совпадают с последовательностью RT матрицы, тем самым внося изменения в нуклеотидную последовательность за счет ошибочной полимеризации одноцепочечного ДНК-флэпа. Данные ошибки, внесенные при синтезе одноцепочечного ДНК-флэпа, затем становятся встроенными в двуцепочечную молекулу за счет гибридизации с соответствующей эндогенной цепью-мишенью, удаления эндогенной вытесненной цепи, лигирования, и затем за счет еще одного раунда процессов репарации и/или секвенирования эндогенной ДНК.

Обратная транскрипция

[0257] В соответствии с используемым здесь значением термин "обратная транскрипция" обозначает способность фермента синтезировать цепь ДНК (то есть комплементарную ДНК или кДНК) с помощью РНК в качестве матрицы. В некоторых вариантах осуществления обратная транскрипция может представлять собой "подверженную ошибкам обратную транскрипцию", что обозначает свойства определенных ферментов в виде обратной транскриптазы, которые являются подверженными ошибкам в плане их активности, связанной с полимеризацией ДНК.

PACE

[0258] В соответствии с используемым здесь значением термин "непрерывная эволюция с помощью фагов (PACE)" обозначает непрерывную эволюцию, в рамках которой используется фаг в качестве вирусных векторов. Общая концепция технологии PACE была описана, например, в международной PCT заявке PCT/US2009/056194, поданной 8 сентября 2009 г., опубликованной как WO 2010/028347 11 марта 2010 г.; международной PCT заявке PCT/US2011/066747, поданной 22 декабря 2011 г., опубликованной как WO 2012/088381 28 июня 2012 г.; заявке на патент США, патенте США 9023594, выданном 5 мая 2015 г., международной PCT заявке

PCT/US2015/012022, поданной 20 января 2015 г., опубликованной как WO 2015/134121 11 сентября 2015 г., и международной PCT заявке PCT/US2016/027795, поданной 15 апреля 2016 г., опубликованной как WO 2016/168631 20 октября 2016 г., полное содержание каждой из которых приведено здесь в качестве ссылки.

Фаг

[0259] В соответствии с используемым здесь значением термин "фаг", которое используется взаимозаменяемым образом с термином "бактериофаг", обозначает вирус, который инфицирует бактериальные клетки. Как правило, фаги состоят из внешнего белкового капсида, в котором заключен генетический материал. Генетический материал может представлять собой оцРНК, дцРНК, оцДНК, или дцДНК в линейной или циркулярной форме. Фаги и фаговые векторы хорошо известны специалистам в данной области, при этом неограничивающие примеры фагов, которые могут применяться для реализации способов PACE, приведены здесь и представляют собой λ (лизоген), T2, T4, T7, T12, R17, M13, MS2, G4, P1, P2, P4, Phi X174, N4, Φ6, и Φ29. В определенных вариантах осуществления используемый в настоящем изобретении фаг представляет собой M13. Дополнительные подходящие фаги и клетки-хозяева будут очевидны специалистам в данной области, при этом изобретение не является ограниченным в плане данного аспекта. Описание примеров дополнительных подходящих фагов и клеток-хозяев см. в Elizabeth Kutter and Alexander Sulakvelidze: Bacteriophages: Biology and Applications. CRC Press; 1st edition (December 2004), ISBN: 0849313368; Martha R. J. Clokie and Andrew M. Kropinski: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions (Methods in Molecular Biology) Humana Press; 1st edition (December, 2008), ISBN: 1588296822; Martha R. J. Clokie and Andrew M. Kropinski: Bacteriophages: Methods and Protocols, Volume 2: Molecular and Applied Aspects (Methods in Molecular Biology) Humana Press; 1st edition (December 2008), ISBN: 1603275649; все из которых полностью приведены здесь для ссылки в объеме описания подходящих фагов и клеток-хозяев, а также способов и протоколов выделения, культивирования и манипулирования такими фагами).

Белок, пептид, и полипептид

[0260] Термины "белок", "пептид", и "полипептид" используются здесь взаимозаменяемым образом и обозначают полимер из аминокислотных остатков, которые связаны вместе пептидными (амидными) связями. Данные термины

обозначают белок, пептид, или полипептид с любым размером, структурой или функцией. Как правило, белок, пептид, или полипептид будет иметь длину, по меньшей мере, в виде трех аминокислот. Термин "белок", "пептид", или "полипептид" может обозначать отдельный белок или группу белков. Одна или несколько аминокислот в белке, пептиде, или полипептиде могут быть модифицированы, например, путем добавления химической сущности, например углеводородной группы, гидроксильной группы, фосфатной группы, фарнезильной группы, изофарнезильной группы, группы в виде жирной кислоты, линкера для конъюгирования, функционализации или другой модификации, и так далее. Белок, пептид, или пептид также может представлять собой отдельную молекулу или представлять собой мультимолекулярный комплекс. Белок, пептид, или полипептид может представлять собой лишь фрагмент встречающегося в природе белка или пептида. Белок, пептид, или полипептид может являться встречающимся в природе, рекомбинантным, или синтетически, или любой их комбинацией. Любой из указанных здесь белков может быть получен с помощью любого способа, известного в данной области. Например, указанные здесь белки могут быть получены с помощью экспрессии рекомбинантного белка и очистки, которые являются особенно подходящими для слитых белков, включающих пептидный линкер. Способы экспрессии рекомбинантных белков и очистки хорошо известны и включают способы, описанные Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), полное содержание которого приведено здесь в качестве ссылки.

Сплайсинг белков

[0261] В соответствии с используемым здесь значением термин "сплайсинг белков" обозначает процесс, при котором последовательность, интеин (или разделенные интеины, как это может быть) подвергается эксцизии из аминокислотной последовательности, при этом оставшиеся фрагменты аминокислотной последовательности, экстеины, лигируются с помощью амидной связи для образования непрерывной аминокислотной последовательности. Термин "транс-сплайсинг белков" обозначает определенный случай, при котором интеины представляют собой разделенные интеины, при этом они локализованы на различных белках.

Никирование второй цепи

[0262] Коррекция гетеродуплексной ДНК (то есть содержащей одну подвергнутую редактированию и одну не подвергнутую редактированию цепь), образованной в результате праймерного редактирования, определяет долгосрочные результаты редактирования. Простыми словами, целью праймерного редактирования является коррекция гетеродуплексной ДНК (подвергнутой редактированию цепи, спаренной с эндогенной не подвергнутой редактированию цепью), образованной в качестве промежуточного продукта PE за счет перманентного встраивания подвергнутой редактированию цепи в комплементарную эндогенную цепь. Подход в виде "никирования второй цепи" может применяться здесь для смещения коррекции гетеродуплексной ДНК в сторону перманентной интеграции подвергнутой редактированию цепи в молекулу ДНК. В соответствии с используемым здесь значением концепция "никирования второй цепи" обозначает внесение второго разрыва в участке ниже первого разрыва (то есть первоначального участка разрыва, который обеспечивает свободный 3´-конец для использования при праймировании обратной транскриптазы на удлиненной части направляющей РНК), предпочтительно на не подвергаемой редактированию цепи. В определенных вариантах осуществления первый разрыв и второй разрыв располагаются на противоположных цепях. В других вариантах осуществления первый разрыв и второй разрыв располагаются на противоположных цепях. В другом варианте осуществления первый разрыв располагается на не подвергаемой редактированию цепи (то есть цепи, которая образует одноцепочечную часть R-петли), а второй разрыв располагается на цепи-мишени. В других вариантах осуществления первый разрыв располагается на подвергаемой редактированию цепи, а второй разрыв располагается на не подвергаемой редактированию цепи. Второй разрыв может быть расположен, по меньшей мере, на 5 нуклеотидов ниже первого разрыва, или, по меньшей мере, на 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, или 150, или более нуклеотидов ниже первого разрыва. В определенных вариантах осуществления второй разрыв может быть внесен на не подвергаемой редактированию цепи на расстоянии около 5-150 нуклеотидов от участка индуцированного PEgRNA разрыва, или около 5-140, или около 5-130, или около 5-120, или около 5-110, или около 5-100, или около 5-90, или около 5-80, или около 5-70, или около 5-60, или около 5-50, или около 5-40, или около 5-30, или около 5-20, или около 5-10. В одном варианте осуществления второй разрыв вносится на

расстоянии около 14-116 нуклеотидов от индуцированного PEgRNA разрыва. Без связи с теорией, второй разрыв индуцирует процессы репарации и репликации эндогенной ДНК клетки с замещением или редактированием не подвергаемой редактированию цепи, тем самым обеспечивая перманентное внесение отредактированной последовательности в обе цепи и коррекцию гетеродуплекса, который образуется в результате PE. В некоторых вариантах осуществления подвергаемая редактированию цепь представляет собой не являющуюся мишенью цепь, а не подвергаемая редактированию цепь представляет собой цепь-мишень. В других вариантах осуществления подвергаемая редактированию цепь представляет собой цепь-мишень, а не подвергаемая редактированию цепь представляет собой не являющуюся мишенью цепь.

Смысловая цепь

[0263] В генетике "смысловая" цепь представляет собой сегмент в двуцепочечной ДНК, который проходит от 5' к 3' и который является комплементарным антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' к 5'. В случае сегмента ДНК, который кодирует белок, смысловая цепь представляет собой цепь ДНК, которая имеет такую же последовательность, что и иРНК, которая использует антисмысловую цепь в качестве ее матрицы при транскрипции и в итоге подвергается (обычно, но не всегда) трансляции в белок. Антисмысловая цепь, таким образом, отвечает за РНК, которая позже транслируется в белок, в то время как смысловая цепь имеет практически такой же состав, что и иРНК. Следует обратить внимание, что для каждого сегмента дцДНК вероятно будет присутствовать два набора смысловых и антисмысловых цепей в зависимости от направления чтения (поскольку смысл и антисмысл зависят от перспективы). В конечном итоге то, какая цепь одного сегмента дцДНК считается смысловой или антисмысловой, определяется генным продуктом или иРНК.

[0264] В контексте PEgRNA первый этап представляет собой синтез одноцепочечной комплементарной ДНК (то есть 3′ оцДНК-флэпа, который становится встроенным), ориентированной в направлении от 5 ′ к 3′, которая использует в качестве матрицы плечо удлинения PEgRNA. То, следует ли считать 3′ оцДНК-флэп смысловой или антисмысловой цепью, зависит от направления транскрипции, поскольку считается, что обе цепи ДНК могут служить в качестве матрицы для транскрипции (но не одновременно). Таким образом, в некоторых вариантах осуществления 3′ оцДНК-флэп

(который в целом проходит в направлении от 5′ к 3′) будет служить в качестве смысловой цепи, поскольку он представляет собой кодирующую цепь. В других вариантах осуществления 3′ оцДНК-флэп (который в целом проходит в направлении от 5′ к 3′) будет служить в качестве антисмысловой цепи и, таким образом, матрицы для транскрипции.

Спейсерная последовательность

[0265] В соответствии с используемым здесь значением термин "спейсерная последовательность" при использовании в контексте направляющей нРНК или PEgRNA обозначает часть направляющей нРНК или PEgRNA длиной около 20 нуклеотидов, которая содержит нуклеотидную последовательность, которая является комплементарной протоспейсерной последовательности в последовательности-мишени ДНК. Спейсерная последовательность отжигается на протоспейсерной последовательности для образования структуры в виде оцРНК/оцДНК гибрида в участке-мишени и соответствующей структуры в виде оцДНК R-петли эндогенной цепи ДНК, которая комплементарна протоспейсерной последовательности.

Субъект

[0266] В соответствии с используемым здесь значением термин "субъект" обозначает отдельный организм, например отдельное млекопитающее. В некоторых вариантах осуществления субъект представляет собой человека. В некоторых вариантах осуществления субъект представляет собой отличное от человека млекопитающее. В некоторых вариантах осуществления субъект представляет собой отличного от человека примата. В некоторых вариантах осуществления субъект представляет собой грызуна. В некоторых вариантах осуществления субъект представляет собой овцу, козу, крупный рогатый скот, кошку, или собаку. В некоторых вариантах осуществления субъект представляет собой позвоночное, амфибию, рептилию, рыбу, насекомое, муху, или нематоду. В некоторых вариантах осуществления субъект представляет собой лабораторное животное. В некоторых вариантах осуществления субъект является генетически модифицированным, например генетически модифицированным отличным от человека субъектом. Субъект может быть любого пола и на любой стадии развития.

Разделенный интеин

[0267] Несмотря на то, что интеины наиболее часто представлены в виде последовательного домена, некоторые из них существуют в естественной разделенной

форме. В этом случае два фрагмента экспрессируются в виде отдельных полипептидов и должны быть связаны перед осуществлением сплайсинга, так называемого транс-сплайсинга белка.

[0268] Примером разделенного интеина является интеин *Ssp* DnaE, который включает две субъединицы, а именно DnaE-N и DnaE-C. Две различных субъединицы кодируются отдельными генами, а именно *dnaE-n* и *dnaE-c*, которые кодируют субъединицы DnaE-N и DnaE-C, соответственно. DnaE представляет собой встречающийся в природе разделенный интеин в *Synechocytis sp.* PCC6803 и способен направлять транс-сплайсинг двух отдельных белков, каждый из которых включает продукт слияния с DnaE-N или DnaE-C.

[0269] Дополнительные встречающиеся в природе или сконструированные последовательности разделенного интеина известны или могут быть реализованы на основе описанных здесь или доступных в данной области последовательностей цельного интеина. Примеры последовательностей разделенного интеина можно найти в Stevens et al., "A promiscuous split intein with expanded protein engineering applications," PNAS, 2017, Vol.114: 8538-8543; Iwai et al., "Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostc punctiforme, FEBS Lett, 580: 1853-1858, каждый из которых приведен здесь в качестве ссылки. Дополнительные последовательности разделенного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774, и EP2877490, содержание каждой из которых приведено здесь в качестве ссылки.

[0270] Кроме того, транс-сплайсинг белков был описан *in vivo* и *in vitro* (Shingledecker, et al., *Gene* 207:187 (1998), Southworth, et al., *EMBO J.* 17:918 (1998); Mills, et al., *Proc. Natl. Acad. Sci. USA,* 95:3543-3548 (1998); Lew, et al., *J. Biol. Chem.,* 273:15887-15890 (1998); Wu, et al., *Biochim. Biophys. Acta* 35732:1 (1998b), Yamazaki, et al., *J. Am. Chem. Soc.* 120:5591 (1998), Evans, et al., *J. Biol. Chem.* 275:9091 (2000); Otomo, et al., *Biochemistry* 38:16040-16044 (1999); Otomo, et al., *J. Biolmol. NMR* 14:105-114 (1999); Scott, et al., *Proc. Natl. Acad. Sci. USA* 96:13638-13643 (1999)) и предоставляет возможность экспрессировать белок в виде двух неактивных фрагментов, которые далее подвергаются лигирования для образования функционального продукта, например как показано на Фиг.66 и 67 в отношении образования полного слитого белка PE из двух экспрессируемых по отдельности половин.

<u>Участок-мишень</u>

[0271] Термин "участок-мишень" обозначает последовательность в молекуле нуклеиновой кислоты, который подвергается редактированию с помощью описанного здесь праймерного редактора (PE). Термин "участок-мишень" также обозначает последовательность в молекуле нуклеиновой кислоты, с которой связывается комплекс праймерного редактора (PE) и нРНК.

tPERT

[0272] См. определение "РНК-матрица *trans* праймерного редактора (tPERT)".

Временно́й разрыв второй цепи

[0273] В соответствии с используемым здесь значением термин "временно́й разрыв второй цепи" обозначает вариант никирования второй цепи, при котором внесение второго разрыва в не подвергаемую редактированию цепь осуществляется только после внесения требуемого изменения в подвергаемую редактированию цепь. Это обеспечивает возможность избежать одновременных разрывов на обеих цепях, которые могут приводить к двуцепочечным разрывам ДНК. Никирующая вторую цепь направляющая РНК сконструирована для временно́го контроля, так что разрыв второй цепи не осуществляется до реализации внесения требуемого изменения. Это достигается за счет конструирования нРНК со спейсерной последовательностью, которая соответствует только подвергаемой редактированию цепи, но не исходному аллелю. За счет использования данной стратегии несовпадения между протоспейсером и не подвергаемым редактированию аллелем должны способствовать отсутствию никирования за счет sgРНК до реализации редактирования на включающей PAM цепи.

***Trans* праймерное редактирование**

[0274] В соответствии с используемым здесь значением термин "*trans* праймерное редактирование" обозначает модифицированную форму праймерного редактирования, при которой используется разделенная PEgRNA, то есть при которой PEgRNA разделена на две отдельные молекулы: sgРНК и РНК матрицу для *trans* праймерного редактирования (tPERT). sgРНК служит для нацеливания праймерного редактора (или, в более общем смысле, для нацеливания компонента праймерного редактора в виде napDNAbp) на требуемый геномный участок-мишень, в то время как pPERT используется полимеразой (например обратной транскриптазой) для записи новой последовательности ДНК в локус-мишень, когда tPERT рекрутируется *in trans* к праймерному редактору за счет взаимодействия доменов связывания, расположенных на праймерном редакторе и на tPERT. В одном варианте осуществления домены

связывания могут включать фрагменты для рекрутирования РНК-белка, например аптамер MS2, расположенный на tPERT, и белок MS2cp, слитый с праймерным редактором. Преимущество *trans* праймерного редактирования заключается в том, что за счет отделения матрицы для синтеза ДНК от направляющей РНК является потенциально возможным использовать матрицы большей длины.

[0275] Вариант осуществления *trans* праймерного редактирования показан на Фиг.3G и 3H. На Фиг.3G показана композиция комплекса *trans* праймерного редактора слева ("комплекс RP:PE:нРНК"), который включает napDNAbp, слитый с каждой полимеразой (например обратной транскриптазой) и rPERT-рекрутирующим белком (например MS2sc), который образует комплекс с направляющей РНК. На Фиг.3G дополнительно показана отдельная молекула tPERT, которая включает элементы плеча удлинения PEgRNA, включая матрицу для синтеза ДНК и последовательность связывания праймера. Молекула tPERT также включает домен рекрутирования РНК-белок (который, в данном случае, представляет собой структуру типа "стебель-петля" и может являться, например, аптамером MS2). Как показано в рамках процесса, изображенного на Фиг.3H, комплекс RP-PE:нРНК связывает и никирует последовательность-мишень ДНК. Затем рекрутирующий белок (RP) рекрутирует tPERT для солокализации с комплексом праймерного редактора, связанным с участком-мишенью ДНК, тем самым обеспечивая связывание участка связывания праймера с праймерной последовательностью на никированной цепи и далее обеспечивая синтез полимеразой (например RT) одиночной цепи ДНК на матрице для синтеза ДНК на 5′ tPERT.

[0276] В то время как tPERT показана на Фиг.3G и Фиг.3H в качестве включающей PBS и матрицу для синтеза ДНК на 5′-конце домена рекрутирования РНК-белка, tPERT в других конфигурациях может быть сконструирована с PBS и матрицей для синтеза ДНК, расположенными на 3′-конце домена рекрутирования РНК-белка. Однако tPERT с 5’ удлинением имеет преимущество, заключающееся в том, что синтез одиночной цепи ДНК будет естественным образом терминирован на 5'-конце tPERT и, таким образом, не будет связан с риском использования какой-либо части домена рекрутирования РНК-белка в качестве матрицы на этапе праймерного редактирования, связанного с синтезом ДНК.

<u>РНК-матрица *trans* праймерного редактора (tPERT)</u>

[0277] В соответствии с используемым здесь значением термин "РНК-матрица *trans* праймерного редактора (tPERT)" обозначает компонент, используемый при *trans* праймерном редактировании, представляющим собой модифицированную версию праймерного редактирования, которая работает на основе разделения PEgRNA на две отдельные молекулы: направляющую РНК и молекулу tPERT. Молекула tPERT программируется для солокализации с комплексом праймерного редактора в участке-мишени ДНК, направляя участок связывания праймера и матрицу для синтеза ДНК к праймерному редактору *in trans*. Например, см. вариант осуществления *trans* праймерного редактора (tPE) на Фиг.3G в виде двухкомпонентной системы, включающей (1) комплекс RP-PE:нРНК и (2) tPERT, которая включает участок связывания праймера и матрицу для синтеза ДНК, присоединенные к домену рекрутирования РНК-белка, при этом компонент в виде RT (рекрутирующего белка) комплекса RP-PE:нРНК рекрутирует tPERT к подвергаемому редактированию участку-мишени, тем самым связывая PBS и матрицу для синтеза ДНК с праймерным редактором *in trans*. Другими словами, tPERT конструируют для включения (полностью или частично) плеча удлинения PEgRNA, которое включает участок связывания праймера и матрицу для синтеза ДНК.

Транзиции

[0278] В соответствии с используемым здесь значением термин "транзиции" обозначает замену пуриновых азотистых оснований (A ↔ G) или замену пиримидиновых азотистых оснований (C ↔ T). Данный класс замен затрагивает азотистые основания одинаковой формы. Описанные здесь композиции и способы способны обеспечивать индуцирование одной или нескольких транзиций в молекуле-мишени ДНК. Описанные здесь композиции и способы также способны обеспечивать индуцирование как транзиций, так и трансверсии в одной и той же молекуле-мишени ДНК. Данные изменения включают A ↔ G, G ↔ A, C ↔ T, или T ↔ C. В контексте двуцепочечной ДНК с парными азотистыми основаниями Уотсона-Крика трансверсии обозначают следующие замены пар оснований: A:T ↔ G:C, G:G ↔ A:T, C:G ↔ T:A, или T:A↔ C:G. Описанные здесь композиции и способы способны обеспечивать индуцирование одной или нескольких транзиций в молекуле-мишени ДНК. Описанные здесь композиции и способы также способны обеспечивать индуцирование как транзиций, так и трансверсии в одной и той же молекуле-мишени ДНК, а также другие нуклеотидные изменения, включая делеции и вставки.

Трансверсии

[0279] В соответствии с используемым здесь значением термин "трансверсии" обозначает замену пуриновых азотистых оснований пиримидиновыми азотистыми основаниями или наоборот и, таким образом, включает взаимную замену азотистых оснований с различающейся формой. Данные изменения включают T ↔ A, T↔ G, C ↔ G, C ↔ A, A ↔ T, A ↔ C, G ↔ C, и G ↔ T. В контексте двуцепочечной ДНК с парными азотистыми основаниями Уотсона-Крика трансверсии обозначают следующие замены пар оснований: T:A ↔ A:T, T:A ↔ G:C, C:G ↔ G:C, C:G ↔ A:T, A:T ↔ T:A, A:T ↔ C:G, G:C ↔ C:G, и G:C ↔ T:A. Описанные здесь композиции и способы способны обеспечивать индуцирование одной или нескольких трансверсий в молекуле-мишени ДНК. Описанные здесь композиции и способы также способны обеспечивать индуцирование как транзиций, так и трансверсии в одной и той же молекуле-мишени ДНК, а также другие нуклеотидные изменения, включая делеции и вставки.

Лечение

[0280] Термины "лечение", "лечить" и "лечащий" обозначают клиническое вмешательство, направленное на обращение, облегчение, отсрочку развития или ингибирования прогрессирования заболевания или расстройства, одного или нескольких его симптомов, как описано здесь. В соответствии с используемым здесь значением термины "лечение", "лечить" и "лечащий" обозначают клиническое вмешательство, направленное на обращение, облегчение, отсрочку развития или ингибирования прогрессирования заболевания или расстройства, одного или нескольких его симптомов, как описано здесь. В некоторых вариантах осуществления лечение может осуществляться после развития одного или нескольких симптомов и/или после диагностирования заболевания. В других вариантах осуществления лечение может осуществляться при отсутствии симптомов, например для предотвращения или задержки развития симптома или ингибирования развития или прогрессирования заболевания. Например, лечение может осуществляться в отношении уязвимого субъекта до развития симптомов (например в свете истории симптомов и/или в свете генетических или иных факторов предрасположенности). Лечение может также продолжаться после устранения симптомов, например для предотвращения или замедления их повторного возникновения.

Связанное с тринуклеотидными повторами заболевание

[0281] В соответствии с используемым здесь значением термин "связанное с тринуклеотидными повторами заболевание" (или, альтернативно, "связанное экспансией повторов заболевание" или "связанное с повторной экспансией заболевание") обозначает набор генетических заболеваний, которые вызываются "экспансией тринуклеотидных повторов", которая представляет собой разновидность мутации, при которой определенный тринуклеотид повторяется в определенных генах или интронах. Ранее считалось, что тринуклеотидные повторы являются обычными интерациями в рамках генома, однако в 1990-х данные заболевания были выявлены. Данные на первый взгляд "доброкачественные" последовательности ДНК иногда могут подвергаться экспансии и вызывать заболевание. Заболевания, вызываемые экспансией тринуклеотидных повторов, разделяют некоторые определяющие признаки. Во-первых, мутантные повторы демонстрируют нестабильность как соматической, так и зародышевой линии, и они более часто подвергаются удлинению, а не сжатию при последующих передачах. Во-вторых, более ранний возраст развития и повышающаяся тяжесть фенотипа в последующих поколениях (ожидающаяся) в целом коррелируют с большей длиной повторов. Наконец, происхождение связанного с заболеванием аллеля от родителя зачастую может влиять на ожидание вследствие того, что передача от родителя связана с повышающимся риском экспансии для многих из данных заболеваний.

[0282] Считается, что экспансия триплетов вызывается проскальзыванием при репликации ДНК. Вследствие повторяющейся природы последовательности ДНК в данных участках могут образовываться "выступающие в виде петли" структуры при репликации ДНК с одновременным сохранением связывания комплементарных оснований между исходной цепью и дочерней синтезируемой цепью. Если выступающая в виде петли структура образуется из последовательности на дочерней цепи, это приведет к повышению количества повторов. Однако если выступающая в виде петли структура образуется на исходной цепи, то это приводит к снижению количества повторов. Похоже, что экспансия данных повторов является более частой, чем сжатие. В целом чем больше экспансия, тем с большей вероятностью вызывается заболевание или повышается тяжесть заболевания. Это свойство приводит к характеристике в виде ожидания при заболеваниях, связанных с тринуклеотидными повторами. Ожидание описывает тенденцию, связанную с тем, что возраст

возникновения заболевания понижается, а тяжесть симптомов повышается у последующих поколений семьи вследствие экспансии данных повторов.

[0283] Связанные с нуклеотидными повторами заболевания могут включать заболевания, при которых триплетный повтор присутствует в некодирующем участке (то есть некодирующее связанное с тринуклеотидными повторами заболевание) или в кодирующем участке.

[0284] Описанная здесь система праймерного редактора (PE) может использоваться для лечения связанных с нуклеотидными повторами заболеваний, которые могут включать синдром ломкой X-хромосомы (FRAXA), ломкую XE MR (FRAXE), атаксию Фрейдрейха (FRDA), миотоническую дистрофию (DM), спиноцеребеллярную атаксию 8 типа (SCA8) и спиноцеребеллярную атаксию 12 типа (SCA12), среди прочих.

Выше

[0285] В соответствии с используемым здесь значением термины "выше" и "ниже" представляют собой термины для указания относительного положения, которые определяют линейное положение, по меньшей мере, двух элементов, расположенных в молекуле нуклеиновой кислоты (одноцепочечной или двуцепочечной), которая ориентирована в направлении от 5′ к 3′. В частности, первый элемент расположен выше второго элемента в молекуле нуклеиновой кислоты, если первый элемент расположен где-либо 5’ относительно второго элемента. Например, SNP расположен выше Cas9-индуцированного участка разрыва, если SNP расположен на 5’ стороне участка разрыва. С другой стороны, первый элемент расположен ниже второго элемента в молекуле нуклеиновой кислоты, если первый элемент расположен где-либо 3’ относительно второго элемента. Например, SNP расположен ниже Cas9-индуцированного участка разрыва, если SNP расположен на 3’ стороне участка разрыва. Молекула нуклеиновой кислоты может представлять собой ДНК (двуцепочечную или одноцепочечную). РНК (двуцепочечную или одноцепочечную), или гибрид ДНК и РНК. Анализ является таким же для одноцепочечной молекулы нуклеиновой кислоты, что и для двуцепочечной молекулы, поскольку термины "выше" и "ниже" обозначают только одну цепь молекулы нуклеиновой кислоты, за исключением случаев, когда необходимо выбрать указываемую цепь двуцепочечной молекулы. Зачастую цепь двуцепочечной ДНК, которая может применяться для определения относительности положения, по меньшей мере, двух элементов, представляет собой "смысловую" или "кодирующую" цепь. В генетике "смысловая"

цепь представляет собой сегмент в рамках двуцепочечной ДНК, которая проходит от 5' к 3' и которая является комплементарной антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' к 5'. Таким образом, в качестве примера, азотистое основание SNP находится ниже последовательности промотора в геномной ДНК (которая является двуцепочечной), если азотистое основание SNP располагается на 3' стороне промотора на смысловой или кодирующей цепи.

Вариант

[0286] В соответствии с используемым здесь значением термин "вариант" должен пониматься в качестве обозначающего демонстрирование качеств, которые имеют паттерн, являющийся отличным от находящегося в природе, например вариант Cas9 представляет собой Cas9, включающую одно или несколько изменения в аминокислотных остатках по сравнению с аминокислотной последовательностью Cas9 дикого типа. Термин "вариант" охватывает гомологичные белки, имеющие, по меньшей мере, 75%, или, по меньшей мере, 80%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность относительно исходной последовательности и имеющие такую же или по существу такую же функциональную активность или активности, что и исходная последовательность. Термин также охватывает мутанты, усечения, или домены исходной последовательности, и которые демонстрируют такую же или по существу такую же функциональную активность или активности, что и исходная последовательность.

Вектор

[0287] В соответствии с используемым здесь значением термин "вектор обозначает нуклеиновую кислоту, которая может быть модифицирована для кодирования представляющего интерес гена и которая способна проникать в клетку-хозяина, мутировать и реплицироваться в клетке-хозяине, и затем переносить реплицированную форму вектора в другую клетку-хозяин. Примеры подходящих векторов включают вирусные векторы, например ретровирусные векторы или бактериофаги и нитчатый фаг, и конъюгативные плазмиды. Дополнительные подходящие векторы будут очевидны специалистам в данной области на основе настоящего описания.

Дикий тип

[0288] В соответствии с используемым здесь значением термин "дикий тип" представляет собой термин в данной области, который понимается специалистами в данной области и обозначает обычную форму организма, штамма, гена или

характеристику в том виде, в котором она встречается в природе, для различения от мутантных или вариантных форм.

5′ эндогенный ДНК-флэп

[0289] В соответствии с используемым здесь значением термин "5′ эндогенный ДНК-флэп" обозначает цепь ДНК, расположенную непосредственно ниже индуцированного PE участка разрыва в ДНК-мишени. Никирование цепи ДНК-мишени с помощью PE высвобождает 3′ гидроксильную группу на лежащей выше стороне участка разрыва и 5′ гидроксильную группу на лежащей ниже стороне участка разрыва. Эндогенная цепь, оканчивающаяся 3′ гидроксильной группой, используется для праймирования ДНК-полимеразы праймерного редактора (например когда ДНК-полимераза представляет собой обратную транскриптазу). Эндогенная цепь на расположенной ниже стороне участка разрыва, которая начинается со свободной 5′ гидроксильной группы, обозначается как "5′ эндогенный ДНК-флэп" и в конечном итоге удаляется и замещается новой синтезированной замещающей цепью (то есть "3′ замещающим ДНК-флэпом"), кодируемой удлинением PEgRNA.

Удаление 5′ эндогенного ДНК-флэпа

[0290] В соответствии с используемым здесь значением термин "удаление 5′ эндогенного ДНК-флэпа" обозначает удаление 5′ эндогенного ДНК-флэпа, который образуется когда синтезированный RT одноцепочечный ДНК-флэп конкурентно вносится и гибридизируется с эндогенной ДНК, вытесняя при этом эндогенную цепь. Удаление данной эндогенной вытесненной цепи может склонять реакцию к образованию требуемого продукта, включающего требуемое нуклеотидное изменение. Собственные ферменты клетки для репарации ДНК могут катализировать удаление или эксцизию 5′ эндогенного флэпа (например флэп-нуклеаза, такая как EXO1 или FEN1). Кроме того, клетки-хозяева могут быть трансформированы для экспрессии одного или нескольких ферментов, которые катализируют удаление указанных 5′ эндогенных флэпов, тем самым склоняя процесс к образованию продукта (например флэп-эндонуклеазы). Флэп-эндонуклеазы известны в данной области и описаны в Patel et al., “Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends,” *Nucleic Acids Research*, 2012, 40(10): 4507-4519 и Tsutakawa et al., “Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily,” *Cell*, 2011, 145(2): 198-211 (каждый из которых приведен здесь в качестве ссылки).

3′ замещающий ДНК-флэп

[0291] В соответствии с используемым здесь значением термин "3′ замещающий ДНК-флэп" или просто "замещающий ДНК-флэп" обозначает цепь ДНК, которая синтезируется праймерным редактором и которая кодируется плечом удлинения PEgRNA праймерного редактора. Более конкретно, 3′ замещающий ДНК-флэп кодируется полимеразной матрицей PEgRNA. 3′ замещающий ДНК-флэп включает ту же последовательность, что и 5′ эндогенный ДНК-флэп за исключением того, что он также содержит отредактированную последовательность (например изменение одного нуклеотида). 3′ замещающий ДНК-флэп отжигается на ДНК-мишени, вытесняя или замещая 5′ эндогенный ДНК-флэп (который может быть подвергнут эксцизии, например с помощью 5′ флэп-эндонуклеазы, такой как FEN1 или EXO1), и затем лигируется для присоединения 3′-конца 3′ замещающего ДНК-флэпа к высвобожденному 5′ гидроксильному концу эндогенной ДНК (высвобожденному после эксцизии 5′ эндогенного ДНК-флэпа), тем самым восстанавливая фосфодиэфирную связь и обеспечивая внесение 3′ замещающего ДНК-флэпа для образования гетеродуплексной ДНК, содержащей одну подвергнутую редактированию цепь и одну не подвергнутую редактированию цепь. Процессы репарации ДНК корректируют гетеродуплекс за счет копирования информации из подвергнутой редактированию цепи в комплементарную цепь, обеспечивая перманентное внесение изменения в ДНК. Данный процесс корректирования может быть дополнительно сдвинут в сторону завершения путем никирования не подвергнутой редактированию цепи, то есть путем "никирования второй цепи", как описано здесь.

ПОДРОБНОЕ ОПИСАНИЕ НЕКОТОРЫХ ВАРИАНТОВ ОСУЩЕСТВЛЕНИЯ

[0292] Использование системы на основе коротких регулярно расположенных группами палиндромных повторов (CRISPR) для геномного редактирования произвело революцию в биологических науках[1-3]. Несмотря на то, что манипулирование генами с помощью CRISPR теперь является обыденным, точное внесение изменений отдельных нуклеотидов остается трудной задачей, несмотря на то, что оно является необходимым для исследования или корректирования большого количества вызывающих заболевания мутаций. Направленная гомологичная репарация (HDR) способна обеспечивать реализацию подобных изменений, но связана с низкой эффективностью (зачастую <5%), потребностью в донорных матрицах для репарации ДНК, и негативными эффектами образования двуцепочечных разрывов ДНК (DSB). Недавно в

лаборатории профессора David Liu *et al.* было разработано основное редактирование, которое обеспечивает достижение эффективного редактирования отдельных нуклеотидов без DSB. В рамках редакторов оснований (BE) система CRISPR комбинируется с ферментами в виде модифицирующей основания дезаминазы для конвертации пар-мишеней оснований C•G или A•T в A•T или G•C, соответственно[4-6]. Несмотря на то, что они уже широко используются исследователями по всему миру, текущие BE обеспечивают конверсию лишь четырех из двенадцати возможных пар оснований и не способны корректировать малые вставки или делеции. Более того, объем нацеливания при редактировании оснований ограничен редактированием не являющихся мишенями оснований C или A, являющихся смежными с основанием-мишенью ("побочное редактирование"), а также необходимостью наличия последовательности PAM на расстоянии 15±2 bp от основания-мишени. Таким образом, преодоление данных ограничений значительно расширит варианты исследовательского и терапевтического использования редактирования оснований для геномного редактирования.

[0293] В рамках настоящего изобретения предлагается новый подход точного редактирования, который обеспечивает множество преимуществ редактирования оснований, а именно устранение двуцепочечных разрывов и потребности в донорных матрицах для репарации ДНК, в то же время обеспечивая преодоление его основных ограничений. Предложенный описанный здесь подход обеспечивает достижение прямого внесения подвергнутых редактированию цепей ДНК в геномные участки-мишени с помощью мишень-инициированной обратной транскрипции (TPRT). В рамках описанной здесь конструкции направляющая РНК (нРНК) CRISPR будет сконструирована для несения последовательности матрицы в виде обратной транскриптазы (RT), кодирующей одноцепочечную ДНК, включающую требуемое нуклеотидное изменение. Никированный нуклеазой (Cas9) CRISPR участок-мишень ДНК будет служить в качестве праймера для обратной транскрипции последовательности матрицы на модифицированной нРНК, обеспечивая прямое встраивание любого требуемого нуклеотидного изменения.

[0294] Соответственно, настоящее изобретение частично связано с открытием того, что механизм мишень-праймированной обратной транскрипции (TPRT) может быть использован или адаптирован для осуществления точного геномного редактирования на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью

(например как показано в рамках различных вариантов осуществления на Фиг.1A-1F). Авторы предложили здесь использовать продукты слияния белок Cas-обратная транскриптаза для нацеливания на определенную последовательность ДНК с помощью модифицированной направляющей РНК ("удлиненной направляющей РНК"), осуществлять одноцепочечный разрыв в участке-мишени, и применять никированную ДНК в качестве праймера для обратной транскрипции сконструированной матрицы в виде обратной транскриптазы, которая встраивается в удлиненную направляющую РНК. Новая синтезированная цепь будет являться гомологичной относительно последовательности геномной мишени за исключением включения требуемого нуклеотидного изменения (например изменения отдельного нуклеотида, делеции, или вставки, или их комбинации). Новая синтезированная цепь ДНК также может быть обозначена как одноцепочечный ДНК-флэп, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. Корректирования данного гибридизированного промежуточного продукта может включать удаление полученного вытесненного флэпа из эндогенной ДНК (например с помощью эндонуклеазы 5′-концевого ДНК-флэпа, FEN1), лигирование синтезированного одноцепочечного ДНК-флэпа с ДНК-мишенью, и ассимилирование требуемого нуклеотидного изменения в результате процессов репарации и/или репликации клеточной ДНК. Поскольку синтез ДНК на основе матриц обеспечивает точность на уровне отдельных, объем данного подхода является очень широким и, как представляется, он может использоваться в рамках множества вариантов применения в фундаментальной науке и терапии.

[1] napDNAbp

[0295] Описанные здесь праймерные редакторы и *trans* праймерные редакторы могут включать программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp).

[0296] В одном аспекте napDNAbp может быть связан или образовывать комплекс, по меньшей мере, с одной направляющей нуклеиновой кислотой (например направляющей РНК или PEgRNA), которая локализует napDNAbp в последовательности ДНК, которая включает цепь ДНК (то есть цепь-мишень), которая комплементарна направляющей нуклеиновой кислоте или ее части (например спейсеру направляющей РНК, который отжигается на протоспейсере ДНК-мишени). Другими словами, направляющая нуклеиновая кислота "программирует" napDNAbp (например

Cas9 или его эквивалент) для локализации и связывания с комплементарной последовательностью протоспейсера в ДНК.

[0297] В описанных здесь праймерных редакторах может применяться любой подходящий napDNAbp. В различных вариантах осуществления napDNAbp может представлять собой любую систему CRISPR-Cas класса 2, включая любой фермент CRISPR-Cas типа II, типа V, или типа VI. С учетом быстрого развития CRISPR-Cas в качестве инструмента для геномного редактирования происходили постоянные изменения в номенклатуре, используемой для описания и/или идентификации ферментов CRISPR-Cas, например Cas9 и ортологов Cas9. Данная заявка содержит указания ферментов CRISPR-Cas в соответствии с номенклатурой, которая может являться старой и/или новой. Специалист в данной области будет способен идентифицировать конкретный указанный в данной заявке фермент CRISPR-Cas на основе используемой номенклатуры, независимо от того, является ли она старой (то есть "устаревшей") или новой номенклатурой. Номенклатура CRISPR-Cas широко описана в Makarova et al., “Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?,” *The CRISPR Journal*, Vol. 1. No. 5, 2018, полное содержание которого приведено здесь в качестве ссылки. Конкретная номенклатура CRISPR-Cas, используемая в любом случае в рамках данной заявки, не является ограниченной каким-либо образом и специалист в данной области будет способен идентифицировать указываемый фермент CRISPR-Cas.

[0298] Например, следующие ферменты CRISPR-Cas класса 2 типа II, типа V, и типа VI имеют следующие известные в данной области старые (то есть устаревшие) и новые названия. Каждый из данных ферментов и/или их вариантов могут применяться в рамках описанных здесь праймерных редакторов.

Устаревшая номенклатура	**Текущая номенклатура***
Ферменты CRISPR-Cas типа II	
Cas9	Та же
Ферменты CRISPR-Cas типа V	
Cpf1	Cas12a
CasX	Cas12e
C2c1	Cas12b1
Cas12b2	Та же
C2c3	Cas12c
CasY	Cas12d
C2c4	Та же
C2c8	Та же
C2c5	Та же

C2c10	Та же
C2c9	Та же
Ферменты CRISPR-Cas типа VI	
C2c2	Cas13a
Cas13d	Та же
C2c7	Cas13c
C2c6	Cas13b

* См. Makarova et al., *The CRISPR Journal*, Vol. 1, No. 5, 2018

[0299] Без связи с теорией, механизм действия определенных предусмотренных здесь napDNAbp включает этап образования R-петли, при котором napDNAbp индуцирует разматывание двуцепочечной ДНК-мишени, тем самым разделяя цепи в участке, связываемом napDNAbp. Затем спейсер направляющей РНК гибридизируется с "цепью-мишенью" на протоспейсерной последовательности. Это обеспечивает вытеснение "не являющейся мишенью цепи", которая комплементарна цепи-мишени, которая образует одноцепочечный участок R-петли. В некоторых вариантах осуществления napDNAbp включает один или несколько видов нуклеазной активности, которые затем обеспечивают разрезание ДНК, оставляя различные типы разрезов. Например napDNAbp может включать нуклеазную активность, обеспечивающую разрезание не являющейся мишенью цепи в первом месторасположении и/или разрезание цепи-мишени во втором месторасположении. В зависимости от нуклеазной активности, ДНК-мишень может быть разрезана с образованием "двуцепочечного разрыва", при котором обе цепи являются разрезанными. В других вариантах осуществления ДНК-мишень может быть разрезана только в одном участке, то есть ДНК "подвергается никированию" на одной цепи. Пример napDNAbp с различными видами нуклеазной активности включает "никазу Cas9" ("nCas9") и деактивированный Cas9 без нуклеазной активности ("мертвый Cas9" или "dCas9").

[0300] Приведенное ниже описание различных napDNAbp, которые могут применяться совместно с описываемыми здесь праймерными редакторами, не является ограничивающим каким-либо образом. Праймерные редакторы могут включать стандартный SpCas9 или любой ортолог белка Cas9, или любой вариант белка Cas9, включая любой встречающийся в природе вариант, мутант, или иным образом модифицированную версию Cas9, которая известна или которая может быть получена или может быть подвергнута эволюции с помощью связанного с направленной эволюцией или иного мутагенезного процесса. В различных вариантах осуществления Cas9 или варианты Cas9 обладают никазной активностью, то есть расщепляют только

цепь последовательности-мишени ДНК. В других вариантах осуществления Cas9 или варианты Cas9 имеют деактивированные нуклеазы, то есть представляют собой белки в виде "мертвого" Cas9. Другие белки в виде вариантов Cas9, которые могут применяться, представляют собой белки с более низкой молекулярной массой, в отличие от стандартного SpCas9 (например для более легкой доставки) или имеющие модифицированную или перестроенную первичную аминокислотную структуру (например циркулярные пермутантные форматы).

[0301] Описанные здесь праймерные редакторы также могут включать эквиваленты Cas9, включая белки Cas12a (Cpf1) и Cas12b1, которые представляют собой результат конвергентной эволюции. Используемые здесь napDNAbp (например SpCas9, вариант Cas9, или эквиваленты Cas9) могут также содержать различные модификации, которые изменяют/повышают их специфичность в отношении PAM. Наконец, в рамках заявки предусмотрен любой Cas9, вариант Cas9, или эквивалент Cas9, который имеет, по меньшей мере, 70%, по меньшей мере, 75%, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 91%, по меньшей мере, 92%, по меньшей мере, 93%, по меньшей мере, 94%, по меньшей мере, 95%, по меньшей мере, 96%, по меньшей мере, 97%, по меньшей мере, 98%, по меньшей мере, 99%, или, по меньшей мере, 99,9% идентичность последовательности относительно исходной последовательности Cas9, например исходной стандартной последовательности SpCas9 или исходного эквивалента Cas9 (например Cas12a (Cpf1)).

[0302] napDNAbp может представлять собой нуклеазу, ассоциированную с CRISPR (короткими регулярно расположенными группами палиндромными повторами). Как указано выше, CRISPR представляет собой адаптивную иммунную систему, которая обеспечивает защиту против подвижных генетических элементов (вирусов, переносимых элементов, и конъюгативных плазмид). Кластеры CRISPR содержат спейсеры, которые представляют собой последовательности, комплементарные предшествующим подвижным элементам, атакующие мишени нуклеиновые кислоты. Кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В системах CRISPR типа II корректный процессинг пре-crРНК требует транс-кодируемую малую РНК (tracrРНК), эндогенную рибонуклеазу 3 (rnc) и белок Cas9. tracrРНК служит в качестве направляющей для осуществляемого с помощью рибонуклеазы 3 процессинга пре-crРНК. Далее Cas9/crРНК/tracrРНК обеспечивает расщепление эндонуклеазой линейной или циркулярной дцДНК-мишени, являющейся

комплементарной спейсеру. Цепь-мишень, которая не является комплементарной crРНК, сначала разрезается эндонуклеазой, затем подрезается 3′-5′ экзонуклеазой. В природе связывание и расщепление ДНК обычно требует белок и обе РНК. Однако одиночные направляющие РНК ("sgРНК" или просто "нРНК") могут быть сконструированы для включения аспектов как crРНК, так и tracrРНК в виде одиночной РНК. См, например Jinek M. et al., Science 337:816-821(2012), полное содержание которого приведено здесь в качестве ссылки.

[0303] В некоторых вариантах осуществления napDNAbp направляет расщепление одной или обеих цепей в месторасположении последовательности-мишени, например в последовательности-мишени и/или в комплементе последовательности-мишени. В некоторых вариантах осуществления napDNAbp направляет расщепление одной или обеих цепей на расстоянии около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 или более пар оснований от первого или последнего нуклеотида последовательности-мишени. В некоторых вариантах осуществления вектор кодирует napDNAbp, который подвергается мутированию с учетом соответствующего фермента дикого типа, так что мутированный napDNAbp не имеет способности расщеплять одну или обе цепи полинуклеотида-мишени, содержащего последовательность-мишень. Например, замещение аспартата на аланин (D10A) в каталитическом домене RuvC I Cas9 из *S. pyogenes* обеспечивает конвертацию Cas9 из нуклеазы, которая расщепляет обе цепи, в никазу (расщепляет одну цепь). Другие примеры мутаций, которые превращают Cas9 в никазу, включают, без ограничения, H840A, N854A и N863A для стандартной последовательности SpCas9 или для эквивалентных аминокислотных положений в других вариантах Cas9 или эквивалентах Cas9.

[0304] В соответствии с используемым здесь значением термин "белок Cas" обозначает полноразмерный белок Cas, полученный из природы, рекомбинантный белок Cas, имеющий последовательности, которые отличаются от встречающегося в природе белка Cas, или любой фрагмент белка Cas, который тем не менее сохраняет все или значительную часть требуемых основных функций, необходимых для описанных способов, то есть (i) программируемое нуклеиновой кислотой связывание белка Cas с ДНК-мишенью и (ii) способность никировать последовательность-мишень ДНК на одной цепи. Предусмотренные здесь белки Cas охватывают белки Cas9 CRISPR, а также эквиваленты Cas9, варианты (например никазу Cas9 (nCas9) или не имеющий нуклезной активности Cas9 (dCas9)), гомологи, ортологи, или паралоги,

встречающиеся или не встречающиеся в природе (например сконструированные или рекомбинантные), и могут включать эквивалент Cas9 из любой системы CRISPR класса 2 (например типа II, V, VI), включая Cas12a (Cpf1), Cas12e (CasX), Cas12b1 (C2c1), Cas12b2, Cas12c (C2c3), C2c4, C2c8, C2c5, C2c10, C2c9 Cas13a (C2c2), Cas13d, Cas13c (C2c7), Cas13b (C2c6), и Cas13b. Дополнительные эквиваленты Cas описаны в Makarova et al., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," Science 2016; 353(6299) и Makarova et al., "Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?," *The CRISPR Journal*, Vol. 1. No. 5, 2018, полное содержание которых приведено здесь в качестве ссылки.

[0305] Термины "Cas9" или "нуклеаза Cas9" или "фрагмент Cas9" или "домен Cas9" охватывают любой встречающийся в природе Cas9 из любого организма, любой встречающийся в природе эквивалент Cas9 или его функциональный фрагмент, любой гомолог Cas9, ортолог, или паралог из любого организма, и любой мутант или вариант Cas9, являющийся встречающимся в природе или сконструированным. Термин "Cas9" не является конкретно ограничивающим и может рассматриваться как "Cas9 или эквивалент". Примеры белков Cas9 дополнительно описаны здесь и/или описаны в данной области и приведены здесь в качестве ссылки. Настоящее изобретение не ограничено в плане конкретного Cas9, который используется в рамках праймерного редактора (PE) по изобретению.

[0306] Последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной области (см., например, , "Complete genome sequence of an M1 strain of Streptococcus pyogenes." Ferretti et al., J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci.* U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); и "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полное содержание которых приведено здесь в качестве ссылки).

[0307] Примеры Cas9 и эквивалентов Cas9 приведены ниже; однако данные конкретные примеры не являются ограничивающими. В рамках праймерного

редактора по настоящему изобретению может использоваться любой подходящий napDNAbp, включая любой подходящий Cas9 или эквивалент Cas9.

A. Стандартный SpCas9 дикого типа

[0308] В одном варианте осуществления описанный здесь конструкт праймерного редактора может включать "стандартную нуклеазу SpCas9" из *S. pyogenes*, которая широко использовалась в качестве инструмента для геномной инженерии и относится к подгруппе ферментов типа II систем CRISPR-Cas класса 2. Данный белок Cas9 представляет собой крупный имеющий несколько доменов белок, содержащий два отдельных нуклеазных домена. В Cas9 могут быть внесены точечные мутации для устранения одного или обоих видов нуклеазной активности, что приводит к получению никазы Cas9 (nCas9) или мертвого Cas9 (dCas9), соответственно, при этом он все же сохраняет способность связываться с ДНК за счет программирования sgРНК. В принципе, при слиянии с другим белком или доменом Cas9 или его вариант (например nCas9) может нацеливать данный белок практически на любую последовательность ДНК просто за счет совместной экспрессии с подходящей sgРНК. В соответствии с используемым здесь значением термин "стандартный белок SpCas9" обозначает белок дикого типа из *Streptococcus pyogenes*, имеющий следующую аминокислотную последовательность:

Описание	Последовательность	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* M1 SwissProt кат. # Q99ZW2 Дикий тип	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQL PGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV	SEQ ID NO: 18

	ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
SpCas9 Обратная трансляция SwissProt кат. # Q99ZW2 *Streptococcus pyogenes*	ATGGATAAAAAATATAGCATTGGCCTGGATATTGGC ACCAACAGCGTGGGCTGGGCGGTGATTACCGATGAA TATAAAGTGCCGAGCAAAAAATTTAAAGTGCTGGGC AACACCGATCGCCATAGCATTAAAAAAAACCTGATT GGCGCGCTGCTGTTTGATAGCGGCGAAACCGCGGAA GCGACCCGCCTGAAACGCACCGCGCGCCGCCGCTAT ACCCGCCGCAAAAACCGCATTTGCTATCTGCAGGAA ATTTTTAGCAACGAAATGGCGAAAGTGGATGATAGC TTTTTTCATCGCCTGGAAGAAAGCTTTCTGGTGGAAG AAGATAAAAAACATGAACGCCATCCGATTTTTGGCA ACATTGTGGATGAAGTGGCGTATCATGAAAAATATC CGACCATTTATCATCTGCGCAAAAAACTGGTGGATA GCACCGATAAAGCGGATCTGCGCCTGATTTATCTGG CGCTGGCGCATATGATTAAATTTCGCGGCCATTTTCT GATTGAAGGCGATCTGAACCCGGATAACAGCGATGT GGATAAACTGTTTATTCAGCTGGTGCAGACCTATAA CCAGCTGTTTGAAGAAAACCCGATTAACGCGAGCGG CGTGGATGCGAAAGCGATTCTGAGCGCGCGCCTGAG CAAAAGCCGCCGCCTGGAAAACCTGATTGCGCAGCT GCCGGGCGAAAAAAAAAACGGCCTGTTTGGCAACCT GATTGCGCTGAGCCTGGGCCTGACCCCGAACTTTAA AAGCAACTTTGATCTGGCGGAAGATGCGAAACTGCA GCTGAGCAAAGATACCTATGATGATGATCTGGATAA CCTGCTGGCGCAGATTGGCGATCAGTATGCGGATCT GTTTCTGGCGGCGAAAAACCTGAGCGATGCGATTCT GCTGAGCGATATTCTGCGCGTGAACACCGAAATTAC CAAAGCGCCGCTGAGCGCGAGCATGATTAAACGCTA TGATGAACATCATCAGGATCTGACCCTGCTGAAAGC GCTGGTGCGCCAGCAGCTGCCGGAAAAATATAAAG AAATTTTTTTTGATCAGAGCAAAAACGGCTATGCGG GCTATATTGATGGCGGCGCGAGCCAGGAAGAATTTT ATAAATTTATTAAACCGATTCTGGAAAAAATGGATG GCACCGAAGAACTGCTGGTGAAACTGAACCGCGAA GATCTGCTGCGCAAACAGCGCACCTTTGATAACGGC	SEQ ID NO: 19

	AGCATTCCGCATCAGATTCATCTGGGCGAACTGCAT GCGATTCTGCGCCGCCAGGAAGATTTTTATCCGTTTC TGAAAGATAACCGCGAAAAAATTGAAAAAATTCTG ACCTTTCGCATTCCGTATTATGTGGGCCCGCTGGCGC GCGGCAACAGCCGCTTTGCGTGGATGACCCGCAAAA GCGAAGAAACCATTACCCCGTGGAACTTTGAAGAAG TGGTGGATAAAGGCGCGAGCGCGCAGAGCTTTATTG AACGCATGACCAACTTTGATAAAAACCTGCCGAACG AAAAAGTGCTGCCGAAACATAGCCTGCTGTATGAAT ATTTTACCGTGTATAACGAACTGACCAAAGTGAAAT ATGTGACCGAAGGCATGCGCAAACCGGCGTTTCTGA GCGGCGAACAGAAAAAAGCGATTGTGGATCTGCTGT TTAAAACCAACCGCAAAGTGACCGTGAAACAGCTGA AAGAAGATTATTTTAAAAAAATTGAATGCTTTGATA GCGTGGAAATTAGCGGCGTGGAAGATCGCTTTAACG CGAGCCTGGGCACCTATCATGATCTGCTGAAAATTA TTAAAGATAAAGATTTTCTGGATAACGAAGAAAACG AAGATATTCTGGAAGATATTGTGCTGACCCTGACCC TGTTTGAAGATCGCGAAATGATTGAAGAACGCCTGA AAACCTATGCGCATCTGTTTGATGATAAAGTGATGA AACAGCTGAAACGCCGCCGCTATACCGGCTGGGGCC GCCTGAGCCGCAAACTGATTAACGGCATTCGCGATA AACAGAGCGGCAAAACCATTCTGGATTTTCTGAAAA GCGATGGCTTTGCGAACCGCAACTTTATGCAGCTGA TTCATGATGATAGCCTGACCTTTAAAGAAGATATTC AGAAAGCGCAGGTGAGCGGCCAGGGCGATAGCCTG CATGAACATATTGCGAACCTGGCGGGCAGCCCGGCG ATTAAAAAAGGCATTCTGCAGACCGTGAAAGTGGTG GATGAACTGGTGAAAGTGATGGGCCGCCATAAACCG GAAAACATTGTGATTGAAATGGCGCGCGAAAACCA GACCACCCAGAAAGGCCAGAAAAACAGCCGCGAAC GCATGAAACGCATTGAAGAAGGCATTAAAGAACTG GGCAGCCAGATTCTGAAAGAACATCCGGTGGAAAA CACCCAGCTGCAGAACGAAAAACTGTATCTGTATTA TCTGCAGAACGGCCGCGATATGTATGTGGATCAGGA ACTGGATATTAACCGCCTGAGCGATTATGATGTGGA TCATATTGTGCCGCAGAGCTTTCTGAAAGATGATAG CATTGATAACAAAGTGCTGACCCGCAGCGATAAAAA CCGCGGCAAAAGCGATAACGTGCCGAGCGAAGAAG TGGTGAAAAAAATGAAAAACTATTGGCGCCAGCTGC TGAACGCGAAACTGATTACCCAGCGCAAATTTGATA ACCTGACCAAAGCGGAACGCGGCGGCCTGAGCGAA CTGGATAAAGCGGGCTTTATTAAACGCCAGCTGGTG GAAACCCGCCAGATTACCAAACATGTGGCGCAGATT CTGGATAGCCGCATGAACACCAAATATGATGAAAAC GATAAACTGATTCGCGAAGTGAAAGTGATTACCCTG AAAAGCAAACTGGTGAGCGATTTTCGCAAAGATTTT CAGTTTTATAAAGTGCGCGAAATTAACAACTATCAT CATGCGCATGATGCGTATCTGAACGCGGTGGTGGGC	

	ACCGCGCTGATTAAAAAATATCCGAAACTGGAAAGC GAATTTGTGTATGGCGATTATAAAGTGTATGATGTG CGCAAAATGATTGCGAAAAGCGAACAGGAAATTGG CAAAGCGACCGCGAAATATTTTTTTTATAGCAACAT TATGAACTTTTTTAAAACCGAAATTACCCTGGCGAA CGGCGAAATTCGCAAACGCCCGCTGATTGAAACCAA CGGCGAAACCGGCGAAATTGTGTGGGATAAAGGCC GCGATTTTGCGACCGTGCGCAAAGTGCTGAGCATGC CGCAGGTGAACATTGTGAAAAAAACCGAAGTGCAG ACCGGCGGCTTTAGCAAAGAAAGCATTCTGCCGAAA CGCAACAGCGATAAACTGATTGCGCGCAAAAAAGA TTGGGATCCGAAAAAATATGGCGGCTTTGATAGCCC GACCGTGGCGTATAGCGTGCTGGTGGTGGCGAAAGT GGAAAAAGGCAAAAGCAAAAAACTGAAAAGCGTGA AAGAACTGCTGGGCATTACCATTATGGAACGCAGCA GCTTTGAAAAAAACCCGATTGATTTTCTGGAAGCGA AAGGCTATAAAGAAGTGAAAAAAGATCTGATTATTA AACTGCCGAAATATAGCCTGTTTGAACTGGAAAACG GCCGCAAACGCATGCTGGCGAGCGCGGGCGAACTG CAGAAAGGCAACGAACTGGCGCTGCCGAGCAAATA TGTGAACTTTCTGTATCTGGCGAGCCATTATGAAAA ACTGAAAGGCAGCCCGGAAGATAACGAACAGAAAC AGCTGTTTGTGGAACAGCATAAACATTATCTGGATG AAATTATTGAACAGATTAGCGAATTTAGCAAACGCG TGATTCTGGCGGATGCGAACCTGGATAAAGTGCTGA GCGCGTATAACAAACATCGCGATAAACCGATTCGCG AACAGGCGGAAAACATTATTCATCTGTTTACCCTGA CCAACCTGGGCGCGCCGGCGGCGTTTAAATATTTTG ATACCACCATTGATCGCAAACGCTATACCAGCACCA AAGAAGTGCTGGATGCGACCCTGATTCATCAGAGCA TTACCGGCCTGTATGAAACCCGCATTGATCTGAGCC AGCTGGGCGGCGAT	

[0309] Описанные здесь праймерные редакторы могут включать стандартный SpCas9 или любой его вариант, имеющий, по меньшей мере 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно последовательности Cas9 дикого типа, приведенной выше. Данные варианты могут включать варианты SpCas9, содержащие одну или несколько мутаций, включая любую известную мутацию, указанную в записи SwissProt кат. # Q99ZW2 (SEQ ID NO: 18), которая включает:

Мутация SpCas9 (относительно аминокислотной последовательности стандартной последовательности SpCas9, SEQ ID NO: 18)	Функция/Характеристика (как указано) (см. запись UniProtKB – Q99ZW2 (CAS9_STRPT1) – приведена здесь в качестве ссылки)
D10A	Мутант никазы, который расщепляет протоспейсерную цепь (но не расщепляет

	непротоспейсерную цепь)
S15A	Сниженная активность в плане расщепления ДНК
R66A	Сниженная активность в плане расщепления ДНК
R70A	Расщепление ДНК отсутствует
R74A	Сниженное расщепление ДНК
R78A	Сниженное расщепление ДНК
Делеция 97-150	Нуклеазная активность отсутствует
R165A	Сниженное расщепление ДНК
Делеция 175-307	Сниженное приблизительно на 50% расщепление ДНК
Делеция 312-409	Нуклеазная активность отсутствует
E762A	Никаза
H840A	Мутант никазы, который расщепляет непротоспейсерную цепь, но не расщепляет протоспейсерную цепь
N854A	Никаза
N863A	Никаза
H982A	Сниженное расщепление ДНК
D986A	Никаза
Делеция 1099-1368	Нуклеазная активность отсутствует
R1333A	Сниженное связывание ДНК

[0310] Другие последовательности SpCas9 дикого типа, которые могут использоваться в настоящем изобретении, включают:

Описание	Последовательность	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* MGAS1882 дикий тип NC_017053.1	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCA CAAATAGCGTCGGATGGGCGGTGATCACTGATGATTAT AAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATAC AGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCT CTTTTATTTGGCAGTGGAGAGACAGCGGAAGCGACTC GTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCG GAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAA ATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGA CTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGC ATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAA GTTGCTTATCATGAGAAATATCCAACTATCTATCATCT GCGAAAAAAAATTGGCAGATTCTACTGATAAAGCGGAT TTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAA GTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATC CTGATAATAGTGATGTGGACAAACTATTTATCCAGTTG GTACAAATCTACAATCAATTATTTGAAGAAAACCCTAT TAACGCAAGTAGAGTAGATGCTAAAGCGATTCTTTCTG CACGATTGAGTAAATCAAGACGATTAGAAAATCTCATT GCTCAGCTCCCCGGTGAGAAGAGAAATGGCTTGTTTGG GAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTT TAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTAC	SEQ ID NO: 20

	AGCTTTCAAAAGATACTTACGATGATGATTTAGATAAT TTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTT TTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTC AGATATCCTAAGAGTAAATAGTGAAATAACTAAGGCT CCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACA TCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGAC AACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGAT CAATCAAAAAACGGATATGCAGGTTATATTGATGGGG GAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCA ATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGT GAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGG ACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTT GGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGAC TTTTATCCATTTTTAAAAGACAATCGTGAGAAGATTGA AAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCC ATTGGCGCGTGGCAATAGTCGTTTTGCATGGATGACTC GGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGA AGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTA TTGAACGCATGACAAACTTTGATAAAAATCTTCCAAAT GAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGTA TTTTACGGTTTATAACGAATTGACAAAGGTCAAATATG TTACTGAGGGAATGCGAAAACCAGCATTTCTTTCAGGT GAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAAC AAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGAT TATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAAT TTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGCG CCTACCATGATTTGCTAAAAATTATTAAAGATAAAGAT TTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGG ATATTGTTTTAACATTGACCTTATTTGAAGATAGGGGG ATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTT TGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTT ATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAAT GGTATTAGGGATAAGCAATCTGGCAAAACAATATTAG ATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTA TGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAA GATATTCAAAAAGCACAGGTGTCTGGACAAGGCCATA GTTTACATGAACAGATTGCTAACTTAGCTGGCAGTCCT GCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGT TGATGAACTGGTCAAAGTAATGGGGCATAAGCCAGAA AATATCGTTATTGAAATGGCACGTGAAAATCAGACAA CTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAA ACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAG ATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCA AAATGAAAAGCTCTATCTCTATTATCTACAAAATGGAA GAGACATGTATGTGGACCAAGAATTAGATATTAATCGT TTAAGTGATTATGATGTCGATCACATTGTTCCACAAAG TTTCATTAAAGACGATTCAATAGACAATAAGGTACTAA CGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTT CCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATT	

	GGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGT AAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTT GAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAAT TGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACA AATTTTGGATAGTCGCATGAATACTAAATACGATGAAA ATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTA AAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCA ATTCTATAAAGTACGTGAGATTAACAATTACCATCATG CCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCT TTGATTAAGAAATATCCAAAACTTGAATCGGAGTTTGT CTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGA TTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGC AAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAA AACAGAAATTACACTTGCAAATGGAGAGATTCGCAAA CGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAA TTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCGC AAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAA AACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCA ATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCG TAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTT GATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGC TAAGGTGGAAAAAGGGAAATCGAAGAAGTTAAAATCC GTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAA GTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCT AAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTA AACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGT CGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAA AAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAA TTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGG GTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGT GGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGC AAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGAT GCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACA TAGAGACAAACCAATACGTGAACAAGCAGAAAATATT ATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCT GCTTTTAAATATTTTGATACAACAATTGATCGTAAACG ATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTA TCCATCAATCCATCACTGGTCTTTATGAAACACGCATT GATTTGAGTCAGCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* MGAS1882 дикий тип NC_017053.1	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTD RHSIKKNLIGALLFGSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQIYNQLFEE NPINASRVDAKAILSARLSKSRRLENLIAQLPGEKRNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED	SEQ ID NO: 21

	LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGH KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	
SpCas9 *Streptococcus pyogenes* дикий тип SWBC2D7W014	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCAC TAATTCCGTTGGATGGGCTGTCATAACCGATGAATACA AAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACAC AGACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCC TCCTATTCGATAGTGGCGAAACGGCAGAGGCGACTCG CCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGC AAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAA TGAGATGGCCAAAGTTGACGATTCTTTCTTTCACCGTT TGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACA TGAACGGCACCCCATCTTTGGAAACATAGTAGATGAG GTGGCATATCATGAAAAGTACCCAACGATTTATCACCT CAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGAC CTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAAA GTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATC CGGACAACTCGGATGTCGACAAACTGTTCATCCAGTTA GTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTAT AAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGC GCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGAT CGCACAATTACCCGGAGAGAAGAAAAATGGGTTGTTC GGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAA TTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAAT TGCAGCTTAGTAAGGACACGTACGATGACGATCTCGA CAATCTACTGGCACAAATTGGAGATCAGTATGCGGACT TATTTTTGGCTGCCAAAAACCTTAGCGATGCAATCCTC	SEQ ID NO: 22

	CTATCTGACATACTGAGAGTTAATACTGAGATTACCAA GGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATG AACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTC CGTCAGCAACTGCCTGAGAAATATAAGGAAATATTCTT TGATCAGTCGAAAAACGGGTACGCAGGTTATATTGAC GGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCA AACCCATATTAGAGAAGATGGATGGGACGGAAGAGTT GCTTGTAAAACTCAATCGCGAAGATCTACTGCGAAAG CAGCGGACTTTCGACAACGGTAGCATTCCACATCAAAT CCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGG AGGATTTTTATCCGTTCCTCAAAGACAATCGTGAAAAG ATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGT GGGACCCCTGGCCCGAGGGAACTCTCGGTTCGCATGG ATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGA ATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAA TCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTT ACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTT ACGAGTATTTCACAGTGTACAATGAACTCACGAAAGTT AAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCT AAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTA TTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGA AAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCT GTCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTC ACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAG ATAAGGACTTCCTGGATAACGAAGAGAATGAAGATAT CTTAGAAGATATAGTGTTGACTCTTACCCTCTTTGAAG ATCGGGAAATGATTGAGGAAAGACTAAAAACATACGC TCACCTGTTCGACGATAAGGTTATGAAACAGTTAAAGA GGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAA ACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAA ACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAA TAGGAACTTTATGCAGCTGATCCATGATGACTCTTTAA CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGG ACAAGGGGACTCATTGCACGAACATATTGCGAATCTTG CTGGTTCGCCAGCCATCAAAAAGGGCATACTCCAGAC AGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGA CGTCACAAACCGGAAAACATTGTAATCGAGATGGCAC GCGAAAATCAAACGACTCAGAAGGGGCAAAAAAACA GTCGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAA AGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTG GAAAATACCCAATTGCAGAACGAGAAACTTTACCTCT ATTACCTACAAAATGGAAGGGACATGTATGTTGATCA GGAACTGGACATAAACCGTTTATCTGATTACGACGTCG ATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCA ATCGACAATAAAGTGCTTACACGCTCGGATAAGAACC GAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGT AAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAAT GCGAAACTGATAACGCAAAGAAAGTTCGATAACTTAA CTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAA	

	GGCCGGATTTATTAAACGTCAGCTCGTGGAAACCCGCC AAATCACAAAGCATGTTGCACAGATACTAGATTCCCG AATGAATACGAAATACGACGAGAACGATAAGCTGATT CGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATTGG TGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTT AGGGAGATAAATAACTACCACCATGCGCACGACGCTT ATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAAA TACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTA CAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGC GAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCT TTTATTCTAACATTATGAATTTCTTTAAGACGGAAATC ACTCTGGCAAACGGAGAGATACGCAAACGACCTTTAA TTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGA TAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTG TCCATGCCCCAAGTCAACATAGTAAAGAAAACTGAGG TGCAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCCA AAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG ACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCC TACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAAGTTG AGAAGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGA ATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTG AAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTA CAAGGAAGTAAAAAAGGATCTCATAATTAAACTACCA AAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAAC GGATGTTGGCTAGCGCCGGAGAGCTTCAAAAGGGGAA CGAACTCGCACTACCGTCTAAATACGTGAATTTCCTGT ATTTAGCGTCCCATTACGAGAAGTTGAAAGGTTCACCT GAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGC ACAAACATTATCTCGACGAAATCATAGAGCAAATTTCG GAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCT GGACAAAGTATTAAGCGCATACAACAAGCACAGGGAT AAACCCATACGTGAGCAGGCGGAAAATATTATCCATTT GTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCA AGTATTTTGACACAACGATAGATCGCAAACGATACACT TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACC AATCCATCACGGGATTATATGAAACTCGGATAGATTTG TCACAGCTTGGGGGTGACGGATCCCCCAAGAAGAAGA GGAAAGTCTCGAGCGACTACAAAGACCATGACGGTGA TTATAAAGATCATGACATCGATTACAAGGATGACGAT GACAAGGCTGCAGGA	
SpCas9 *Streptococcus pyogenes* дикий тип Кодируемый продукт SWBC2D7W014	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN	SEQ ID NO: 23

	GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGDGSPKKKRKVSSDY KDHDGDYKDHDIDYKDDDDKAAG	
SpCas9 *Streptococcus pyogenes* M1GAS дикий тип NC_002737.2	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCA CAAATAGCGTCGGATGGGCGGTGATCACTGATGAATA TAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATA CAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGC TCTTTTATTTGACAGTGGAGAGACAGCGGAAGCGACTC GTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCG GAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAA ATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGA CTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGC ATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAA GTTGCTTATCATGAGAAATATCCAACTATCTATCATCT GCGAAAAAAAATTGGTAGATTCTACTGATAAAGCGGAT TTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAA GTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATC CTGATAATAGTGATGTGGACAAACTATTTATCCAGTTG GTACAAACCTACAATCAATTATTTGAAGAAAACCCTAT TAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTG CACGATTGAGTAAATCAAGACGATTAGAAAATCTCATT GCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGG GAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTT TAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTAC AGCTTTCAAAAGATACTTACGATGATGATTTAGATAAT	SEQ ID NO: 24

	TTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTT TTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTC AGATATCCTAAGAGTAAATACTGAAATAACTAAGGCT CCCCTATCAGCTTCAATGATTAAACGCTACGATGAACA TCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGAC AACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGAT CAATCAAAAAACGGATATGCAGGTTATATTGATGGGG GAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCA ATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGT GAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGG ACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTT GGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGAC TTTTATCCATTTTTAAAAGACAATCGTGAGAAGATTGA AAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCC ATTGGCGCGTGGCAATAGTCGTTTTGCATGGATGACTC GGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGA AGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTA TTGAACGCATGACAAACTTTGATAAAAATCTTCCAAAT GAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGTA TTTTACGGTTTATAACGAATTGACAAAGGTCAAATATG TTACTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGT GAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAAC AAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGAT TATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAAT TTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTA CCTACCATGATTTGCTAAAAATTATTAAAGATAAAGAT TTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGG ATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAG ATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTT TGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTT ATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAAT GGTATTAGGGATAAGCAATCTGGCAAAACAATATTAG ATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTA TGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAA GACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATA GTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCT GCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGT TGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCA GAAAATATCGTTATTGAAATGGCACGTGAAAATCAGA CAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTAT GAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGT CAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATT GCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATG GAAGAGACATGTATGTGGACCAAGAATTAGATATTAA TCGTTTAAGTGATTATGATGTCGATCACATTGTTCCAC AAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTC TTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAA CGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAAC TATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCA	

	ACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGA GGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACG CCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGG CACAAATTTTGGATAGTCGCATGAATACTAAATACGAT GAAAATGATAAACTTATTCGAGAGGTTAAAGTGATTA CCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGAT TTCCAATTCTATAAAGTACGTGAGATTAACAATTACCA TCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAA CTGCTTTGATTAAGAAATATCCAAAACTTGAATCGGAG TTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAA AATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAGCA ACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTC TTCAAAACAGAAATTACACTTGCAAATGGAGAGATTC GCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGG AGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACA GTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGT CAAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAG GAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTAT TGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGT GGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGT GGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAGTTA AAATCCGTTAAAGAGTTACTAGGGATCACAATTATGG AAAGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTA GAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAA TCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAA AACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAAT TACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATA TGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTT GAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTG TTTGTGGAGCAGCATAAGCATTATTTAGATGAGATTAT TGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAG CAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAAC AAACATAGAGACAAACCAATACGTGAACAAGCAGAAA ATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTC CCGCTGCTTTTAAATATTTTGATACAACAATTGATCGT AAACGATATACGTCTACAAAAGAAGTTTTAGATGCCA CTCTTATCCATCAATCCATCACTGGTCTTTATGAAACA CGCATTGATTTGAGTCAGCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* M1GAS дикий тип Кодируемый продукт NC_002737.2 (100% идентичен стандартном	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR	SEQ ID NO: 25

y Q99ZW2 дикого типа)	EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	

[0311] Описанные здесь праймерные редакторы могут включать любую из указанных выше последовательностей SpCas9 или любой его вариант, имеющий, по меньшей мере 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность относительно них.

B. <u>Ортологи Cas9 дикого типа</u>

[0312] В других вариантах осуществления белок Cas9 может представлять собой ортолог Cas9 дикого типа из другого вида бактерий, отличный от стандартного Cas9 из *S. pyogenes*. Например, с описанными в данном описании конструктами праймерного редактора могут использоваться следующие ортологи Cas9. Кроме того, любой вариант ортологов Cas9, имеющий, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно любого из указанных ниже ортологов, также может использоваться с настоящими праймерными редакторами.

Описание	Последовательность
LfCas9 *Lactobacillus fermentum* дикий тип	MKEYHIGLDIGTSSIGWAVTDSQFKLMRIKGKTAIGVRLFEEGKTAAERR TFRTTRRRLKRRKWRLHYLDEIFAPHLQEVDENFLRRLKQSNIHPEDPTK NQAFIGKLLFPDLLKKNERGYPTLIKMRDELPVEQRAHYPVMNIYKLRE AMINEDRQFDLREVYLAVHHIVKYRGHFLNNASVDKFKVGRIDFDKSFN

Описание	Последовательность
GenBank: SNX31424.1 1	VLNEAYEELQNGEGSFTIEPSKVEKIGQLLLDTKMRKLDRQKAVAKLLE VKVADKEETKRNKQIATAMSKLVLGYKADFATVAMANGNEWKIDLSS ETSEDEIEKFREELSDAQNDILTEITSLFSQIMLNEIVPNGMSISESMMDRY WTHERQLAEVKEYLATQPASARKEFDQVYNKYIGQAPKERGFDLEKGL KKILSKKENWKEIDELLKAGDFLPKQRTSANGVIPHQMHQQELDRIIEKQ AKYYPWLATENPATGERDRHQAKYELDQLVSFRIPYYVGPLVTPEVQK ATSGAKFAWAKRKEDGEITPWNLWDKIDRAESAEAFIKRMTVKDTYLL NEDVLPANSLLYQKYNVLNELNNVRVNGRRLSVGIKQDIYTELFKKKKT VKASDVASLVMAKTRGVNKPSVEGLSDPKKFNSNLATYLDLKSIVGDK VDDNRYQTDLENIIEWRSVFEDGEIFADKLTEVEWLTDEQRSALVKKRY KGWGRLSKKLLTGIVDENGQRIIDLMWNTDQNFKEIVDQPVFKEQIDQL NQKAITNDGMTLRERVESVLDDAYTSPQNKKAIWQVVRVVEDIVKAVG NAPKSISIEFARNEGNKGEITRSRRTQLQKLFEDQAHELVKDTSLTEELEK APDLSDRYYFYFTQGGKDMYTGDPINFDEISTKYDIDHILPQSFVKDNSL DNRVLTSRKENNKKSDQVPAKLYAAKMKPYWNQLLKQGLITQRKFEN LTKDVDQNIKYRSLGFVKRQLVETRQVIKLTANILGSMYQEAGTEIIETR AGLTKQLREEFDLPKVREVNDYHHAVDAYLTTFAGQYLNRRYPKLRSF FVYGEYMKFKHGSDLKLRNFNFFHELMEGDKSQGKVVDQQTGELITTR DEVAKSFDRLLNMKYMLVSKEVHDRSDQLYGATIVTAKESGKLTSPIEI KKNRLVDLYGAYTNGTSAFMTIIKFTGNKPKYKVIGIPTTSAASLKRAGK PGSESYNQELHRIIKSNPKVKKGFEIVVPHVSYGQLIVDGDCKFTLASPTV QHPATQLVLSKKSLETISSGYKILKDKPAIANERLIRVFDEVVGQMNRYF TIFDQRSNRQKVADARDKFLSLPTESKYEGAKKVQVGKTEVITNLLMGL HANATQGDLKVLGLATFGFFQSTTGLSLSEDTMIVYQSPTGLFERRICLK DI (SEQ ID NO: 26)
SaCas9 *Staphylococcus aureus* дикий тип GenBank: AYD60528.1	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQ SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT

Описание	Последовательность
	VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLY LASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 27)
SaCas9 *Staphylococcus aureus*	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS KRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQ KLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEE KYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQ SFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELR SVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKP TLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENA ELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSL KAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSP VVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNR QTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNP FNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISY ETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRY ATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYK HHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGN TLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQ YGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEV NSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIE VNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSK KHPQIIKK (SEQ ID NO: 28)
StCas9 *Streptococcus thermophilus* UniProtKB/Swiss-Prot: G3ECR1.2 Дикий тип	MLFNKCIIISINLDFSNKEKCMTKPYSIGLDIGTNSVGWAVITDNYKVPSK KMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRI LYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKVYH DEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKN NDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLF PGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLL GYIGDDYSDVFLKAKKLYDAILLSGFLTVTDNETEAPLSSAMIKRYNEH KEDLALLKEYIRNISLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLK NLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQA KFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPWNF EDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVR FIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIE LKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIK QRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLI DDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEVVKSLPGSPAI KKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQRLK RLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTG DDLDIDRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDFPSLEVV KKRKTFWYQLLKSKLISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETR QITKHVARLLDEKFNNKKDENNRAVRTVKIITLKSTLVSQFRKDFELYK

Описание	Последовательность
	VREINDFHHAHDAYLNAVIASALLKKYPKLEPEFVYGDYPKYNSFRERK SATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDLA TVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNEN LVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGISI LDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFELSDGSRRMLASILSTN NKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEEL FYYILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKG LFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRID LAKLGEG (SEQ ID NO: 29)
LcCas9 *Lactobacillus crispatus* NCBI Исходная последовательность: WP_133478044.1 Дикий тип	MKIKNYNLALTPSTSAVGHVEVDDDLNILEPVHHQKAIGVAKFGEGETA EARRLARSARRTTKRRANRINHYFNEIMKPEIDKVDPLMFDRIKQAGLSP LDERKEFRTVIFDRPNIASYYHNQFPTIWHLQKYLMITDEKADIRLIYWA LHSLLKHRGHFFNTTPMSQFKPGKLNLKDDMLALDDYNDLEGLSFAVA NSPEIEKVIKDRSMHKKEKIAELKKLIVNDVPDKDLAKRNNKIITQIVNAI MGNSFHLNFIFDMDLDKLTSKAWSFKLDDPELDTKFDAISGSMTDNQIGI FETLQKIYSAISLLDILNGSSNVVDAKNALYDKHKRDLNLYFKFLNTLPD EIAKTLKAGYTLYIGNRKKDLLAARKLLKVNVAKNFSQDDFYKLINKEL KSIDKQGLQTRFSEKVGELVAQNNFLPVQRSSDNVFIPYQLNAITFNKILE NQGKYYDFLVKPNPAKKDRKNAPYELSQLMQFTIPYYVGPLVTPEEQV KSGIPKTSRFAWMVRKDNGAITPWNFYDKVDIEATADKFIKRSIAKDSY LLSELVLPKHSLLYEKYEVFNELSNVSLDGKKLSGGVKQILFNEVFKKTN KVNTSRILKALAKHNIPGSKITGLSNPEEFTSSLQTYNAWKKYFPNQIDNF AYQQDLEKMIEWSTVFEDHKILAKKLDEIEWLDDDQKKFVANTRLRGW GRLSKRLLTGLKDNYGKSIMQRLETTKANFQQIVYKPEFREQIDKISQAA AKNQSLEDILANSYTSPSNRKAIRKTMSVVDEYIKLNHGKEPDKIFLMFQ RSEQEKGKQTEARSKQLNRILSQLKADKSANKLFSKQLADEFSNAIKKS KYKLNDKQYFYFQQLGRDALTGEVIDYDELYKYTVLHIIPRSKLTDDSQ NNKVLTKYKIVDGSVALKFGNSYSDALGMPIKAFWTELNRLKLIPKGKL LNLTTDFSTLNKYQRDGYIARQLVETQQIVKLLATIMQSRFKHTKIIEVR NSQVANIRYQFDYFRIKNLNEYYRGFDAYLAAVVGTYLYKVYPKARRL FVYGQYLKPKKTNQENQDMHLDSEKKSQGFNFLWNLLYGKQDQIFVN GTDVIAFNRKDLITKMNTVYNYKSQKISLAIDYHNGAMFKATLFPRNDR DTAKTRKLIPKKKDYDTDIYGGYTSNVDGYMLLAEIIKRDGNKQYGFYG VPSRLVSELDTLKKTRYTEYEEKLKEIIKPELGVDLKKIKKIKILKNKVPF NQVIIDKGSKFFITSTSYRWNYRQLILSAESQQTLMDLVVDPDFSNHKAR KDARKNADERLIKVYEEILYQVKNYMPMFVELHRCYEKLVDAQKTFKS LKISDKAMVLNQILILLHSNATSPVLEKLGYHTRFTLGKKHNLISENAVL VTQSITGLKENHVSIKQML (SEQ ID NO: 30)
PdCas9 *Pedicoccus damnosus* NCBI Исходная последовательность: WP_062913273.1 Дикий тип	MTNEKYSIGLDIGTSSIGFAVVNDNNRVIRVKGKNAIGVRLFDEGKAAA DRRSFRTTRRSFRTTRRRLSRRRWRLKLLREIFDAYITPVDEAFFIRLKES NLSPKDSKKQYSGDILFNDRSDKDFYEKYPTIYHLRNALMTEHRKFDVR EIYLAIHHIMKFRGHFLNATPANNFKVGRLNLEEKFEELNDIYQRVFPDE SIEFRTDNLEQIKEVLLDNKRSRADRQRTLVSDIYQSSEDKDIEKRNKAV ATEILKASLGNKAKLNVITNVEVDKEAAKEWSITFDSESIDDDLAKIEGQ MTDDGHEIIEVLRSLYSGITLSAIVPENHTLSQSMVAKYDLHKDHLKLFK KLINGMTDTKKAKNLRAAYDGYIDGVKGKVLPQEDFYKQVQVNLDDS AEANEIQTYIDQDIFMPKQRTKANGSIPHQLQQQELDQIIENQKAYYPWL AELNPNPDKKRQQLAKYKLDELVTFRVPYYVGPMITAKDQKNQSGAEF

Описание	Последовательность
	AWMIRKEPGNITPWNFDQKVDRMATANQFIKRMTTTDTYLLGEDVLPA QSLLYQKFEVLNELNKIRIDHKPISIEQKQQIFNDLFKQFKNVTIKHLQDY LVSQGQYSKRPLIEGLADEKRFNSSLSTYSDLCGIFGAKLVEENDRQEDL EKIIEWSTIFEDKKIYRAKLNDLTWLTDDQKEKLATKRYQGWGRLSRKL LVGLKNSEHRNIMDILWITNENFMQIQAEPDFAKLVTDANKGMLEKTDS QDVINDLYTSPQNKKAIRQILLVVHDIQNAMHGQAPAKIHVEFARGEER NPRRSVQRQRQVEAAYEKVSNELVSAKVRQEFKEAINNKRDFKDRLFL YFMQGGIDIYTGKQLNIDQLSSYQIDHILPQAFVKDDSLTNRVLTNENQV KADSVPIDIFGKKMLSVWGRMKDQGLISKGKYRNLTMNPENISAHTENG FINRQLVETRQVIKLAVNILADEYGDSTQIISVKADLSHQMREDFELLKN RDVNDYHHAFDAYLAAFIGNYLLKRYPKLESYFVYGDFKKFTQKETKM RRFNFIYDLKHCDQVVNKETGEILWTKDEDIKYIRHLFAYKKILVSHEVR EKRGALYNQTIYKAKDDKGSGQESKKLIRIKDDKETKIYGGYSGKSLAY MTIVQITKKNKVSYRVIGIPTLALARLNKLENDSTENNGELYKIIKPQFTH YKVDKKNGEIIETTDDFKIVVSKVRFQQLIDDAGQFFMLASDTYKNNAQ QLVISNNALKAINNTNITDCPRDDLERLDNLRLDSAFDEIVKKMDKYFSA YDANNFREKIRNSNLIFYQLPVEDQWENNKITELGKRTVLTRILQGLHAN ATTTDMSIFKIKTPFGQLRQRSGISLSENAQLIYQSPTGLFERRVQLNKIK (SEQ ID NO: 31)
FnCas9 *Fusobacterium nucleatum* NCBI Исходная последовательность: WP_060798984.1	MKKQKFSDYYLGFDIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFEE AKTAAERRVQRNSRRRLKRRKWRLNLLEEIFSNEILKIDSNFFRRLKESSL WLEDKSSKEKFTLFNDDNYKDYDFYKQYPTIFHLRNELIKNPEKKDIRLV YLAIHSIFKSRGHFLFEGQNLKEIKNFETLYNNLIAFLEDNGINKIIDKNNI EKLEKIVCDSKKGLKDKEKEFKEIFNSDKQLVAIFKLSVGSSVSLNDLFD TDEYKKGEVEKEKISFREQIYEDDKPIYYSILGEKIELLDIAKTFYDFMVL NNILADSQYISEAKVKLYEEHKKDLKNLKYIIRKYNKGNYDKLFKDKNE NNYSAYIGLNKEKSKKEVIEKSRLKIDDLIKNIKGYLPKVEEIEEKDKAIF NKILNKIELKTILPKQRISDNGTLPYQIHEAELEKILENQSKYYDFLNYEE NGIITKDKLLMTFKFRIPYYVGPLNSYHKDKGGNSWIVRKEEGKILPWNF EQKVDIEKSAEEFIKRMTNKCTYLNGEDVIPKDTFLYSEYVILNELNKVQ VNDEFLNEENKRKIIDELFKENKKVSEKKFKEYLLVKQIVDGTIELKGVK DSFNSNYISYIRFKDIFGEKLNLDIYKEISEKSILWKCLYGDDKKIFEKKIK NEYGDILTKDEIKKINTFKFNNWGRLSEKLLTGIEFINLETGECYSSVMDA LRRTNYNLMELLSSKFTLQESINNENKEMNEASYRDLIEESYVSPSLKRAI FQTLKIYEEIRKITGRVPKKVFIEMARGGDESMKNKKIPARQEQLKKLYD SCGNDIANFSIDIKEMKNSLISYDNNSLRQKKLYLYYLQFGKCMYTGREI DLDRLLQNNDTYDIDHIYPRSKVIKDDSFDNLVLVLKNENAEKSNEYPV KKEIQEKMKSFWRFLKEKNFISDEKYKRLTGKDDFELRGFMARQLVNV RQTTKEVGKILQQIEPEIKIVYSKAEIASSFREMFDFIKVRELNDTHHAKD AYLNIVAGNVYNTKFTEKPYRYLQEIKENYDVKKIYNYDIKNAWDKEN SLEIVKKNMEKNTVNITRFIKEKKGQLFDLNPIKKGETSNEIISIKPKVYN GKDDKLNEKYGYYKSLNPAYFLYVEHKEKNKRIKSFERVNLVDVNNIK DEKSLVKYLIENKKLVEPRVIKKVYKRQVILINDYPYSIVTLDSNKLMDF ENLKPLFLENKYEKILKNVIKFLEDNQGKSEENYKFIYLKKKDRYEKNET LESVKDRYNLEFNEMYDKFLEKLDSKDYKNYMNNKKYQELLDVKEKFI KLNLFDKAFTLKSFLDLFNRKTMADFSKVGLTKYLGKIQKISSNVLSKNE LYLLEESVTGLFVKKIKL (SEQ ID NO: 32)
EcCas9	RRKQRIQILQELLGEEVLKTDPGFFHRMKESRYVVEDKRTLDGKQVELP

Описание	Последовательность
Enterococcus cecorum NCBI Исходная последовательность: WP_047338501.1 Дикий тип	YALFVDKDYTDKEYYKQFPTINHLIVYLMTTSDTPDIRLVYLALHYYMK NRGNFLHSGDINNVKDINDILEQLDNVLETFLDGWNLKLKSYVEDIKNIY NRDLGRGERKKAFVNTLGAKTKAEKAFCSLISGGSTNLAELFDDSSLKEI ETPKIEFASSSLEDKIDGIQEALEDRFAVIEAAKRLYDWKTLTDILGDSSS LAEARVNSYQMHHEQLLELKSLVKEYLDRKVFQEVFVSLNVANNYPAY IGHTKINGKKKELEVKRTKRNDFYSYVKKQVIEPIKKKVSDEAVLTKLSE IESLIEVDKYLPLQVNSDNGVIPYQVKLNELTRIFDNLENRIPVLRENRDK IIKTFKFRIPYYVGSLNGVVKNGKCTNWMVRKEEGKIYPWNFEDKVDLE ASAEQFIRRMTNKCTYLVNEDVLPKYSLLYSKYLVLSELNNLRIDGRPLD VKIKQDIYENVFKKNRKVTLKKIKKYLLKEGIITDDDELSGLADDVKSSL TAYRDFKEKLGHLDLSEAQMENIILNITLFGDDKKLLKKRLAALYPFIDD KSLNRIATLNYRDWGRLSERFLSGITSVDQETGELRTIIQCMYETQANLM QLLAEPYHFVEAIEKENPKVDLESISYRIVNDLYVSPAVKRQIWQTLLVIK DIKQVMKHDPERIFIEMARENQESKKTKSRKQVLSEVYKKAKEYEHLFE KLNSLTEEQLRSKKIYLYFTQLGKCMYSGEPIDFENLVSANSNYDIDHIYP QSKTIDDSFNNIVLVKKSLNAYKSNHYPIDKNIRDNEKVKTLWNTLVSK GLITKEKYERLIRSTPFSDEELAGFIARQLVETRQSTKAVAEILSNWFPESE IVYSKAKNVSNFRQDFEILKVRELNDCHHAHDAYLNIVVGNAYHTKFTN SPYRFIKNKANQEYNLRKLLQKVNKIESNGVVAWVGQSENNPGTIATVK KVIRRNTVLISRMVKEVDGQLFDLTLMKKGKGQVPIKSSDERLTDISKY GGYNKATGAYFTFVKSKKRGKVVRSFEYVPLHLSKQFENNNELLKEYIE KDRGLTDVEILIPKVLINSLFRYNGSLVRITGRGDTRLLLVHEQPLYVSNS FVQQLKSVSSYKLKKSENDNAKLTKTATEKLSNIDELYDGLLRKLDLPIY SYWFSSIKEYLVESRTKYIKLSIEEKALVIFEILHLFQSDAQVPNLKILGLS TKPSRIRIQKNLKDTDKMSIIHQSPSGIFEHEIELTSL (SEQ ID NO: 33)
AhCas9 *Anaerostipes hadrus* NCBI Исходная последовательность: WP_044924278.1 Дикий тип	MQNGFLGITVSSEQVGWAVTNPKYELERASRKDLWGVRLFDKAETAED RRMFRTNRRLNQRKKNRIHYLRDIFHEEVNQKDPNFFQQLDESNFCEDD RTVEFNFDTNLYKNQFPTVYHLRKYLMETKDKPDIRLVYLAFSKFMKN RGHFLYKGNLGEVMDFENSMKGFCESLEKFNIDFPTLSDEQVKEVRDIL CDHKIAKTVKKKNIITITKVKSKTAKAWIGLFCGCSVPVKVLFQDIDEEIV TDPEKISFEDASYDDYIANIEKGVGIYYEAIVSAKMLFDWSILNEILGDHQ LLSDAMIAEYNKHHDDLKRLQKIIKGTGSRELYQDIFINDVSGNYVCYV GHAKTMSSADQKQFYTFLKNRLKNVNGISSEDAEWIDTEIKNGTLLPKQ TKRDNSVIPHQLQLREFELILDNMQEMYPFLKENREKLLKIFNFVIPYYV GPLKGVVRKGESTNWMVPKKDGVIHPWNFDEMVDKEASAECFISRMT GNCSYLFNEKVLPKNSLLYETFEVLNELNPLKINGEPISVELKQRIYEQLF LTGKKVTKKSLTKYLIKNGYDKDIELSGIDNEFHSNLKSHIDFEDYDNLS DEEVEQIILRITVFEDKQLLKDYLNREFVKLSEDERKQICSLSYKGWGNL SEMLLNGITVTDSNGVEVSVMDMLWNTNLNLMQILSKKYGYKAEIEHY NKEHEKTIYNREDLMDYLNIPPAQRRKVNQLITIVKSLKKTYGVPNKIFF KISREHQDDPKRTSSRKEQLKYLYKSLKSEDEKHLMKELDELNDHELSN DKVYLYFLQKGRCIYSGKKLNLSRLRKSNYQNDIDYIYPLSAVNDRSMN NKVLTGIQENRADKYTYFPVDSEIQKKMKGFWMELVLQGFMTKEKYFR LSRENDFSKSELVSFIEREISDNQQSGRMIASVLQYYFPESKIVFVKEKLIS SFKRDFHLISSYGHNHLQAAKDAYITIVVGNVYHTKFTMDPAIYFKNHK RKDYDLNRLFLENISRDGQIAWESGPYGSIQTVRKEYAQNHIAVTKRVV EVKGGLFKQMPLKKGHGEYPLKTNDPRFGNIAQYGGYTNVTGSYFVLV ESMEKGKKRISLEYVPVYLHERLEDDPGHKLLKEYLVDHRKLNHPKILL

Описание	Последовательность
	AKVRKNSLLKIDGFYYRLNGRSGNALILTNAVELIMDDWQTKTANKISG YMKRRAIDKKARVYQNEFHIQELEQLYDFYLDKLKNGVYKNRKNNQA ELIHNEKEQFMELKTEDQCVLLTEIKKLFVCSPMQADLTLIGGSKHTGMI AMSSNVTKADFAVIAEDPLGLRNKVIYSHKGEK (SEQ ID NO: 34)
KvCas9 *Kandleria vitulina* NCBI Исходная последовательность: WP_031589969.1 Дикий тип	MSQNNNKIYNIGLDIGDASVGWAVVDEHYNLLKRHGKHMWGSRLFTQ ANTAVERRSSRSTRRRYNKRRERIRLLREIMEDMVLDVDPTFFIRLANVS FLDQEDKKDYLKENYHSNYNLFIDKDFNDKTYYDKYPTIYHLRKHLCES KEKEDPRLIYLALHHIVKYRGNFLYEGQKFSMDVSNIEDKMIDVLRQFN EINLFEYVEDRKKIDEVLNVLKEPLSKKHKAEKAFALFDTTKDNKAAYK ELCAALAGNKFNVTKMLKEAELHDEDEKDISFKFSDATFDDAFVEKQPL LGDCVEFIDLLHDIYSWVELQNILGSAHTSEPSISAAMIQRYEDHKNDLK LLKDVIRKYLPKKYFEVFRDEKSKKNNYCNYINHPSKTPVDEFYKYIKK LIEKIDDPDVKTILNKIELESFMLKQNSRTNGAVPYQMQLDELNKILENQ SVYYSDLKDNEDKIRSILTFRIPYYFGPLNITKDRQFDWIIKKEGKENERIL PWNANEIVDVDKTADEFIKRMRNFCTYFPDEPVMAKNSLTVSKYEVLN EINKLRINDHLIKRDMKDKMLHTLFMDHKSISANAMKKWLVKNQYFSN TDDIKIEGFQKENACSTSLTPWIDFTKIFGKINESNYDFIEKIIYDVTVFED KKILRRRLKKEYDLDEEKIKKILKLKYSGWSRLSKKLLSGIKTKYKDSTR TPETVLEVMERTNMNLMQVINDEKLGFKKTIDDANSTSVSGKFSYAEVQ ELAGSPAIKRGIWQALLIVDEIKKIMKHEPAHVYIEFARNEDEKERKDSF VNQMLKLYKDYDFEDETEKEANKHLKGEDAKSKIRSERLKLYYTQMG KCMYTGKSLDIDRLDTYQVDHIVPQSLLKDDSIDNKVLVLSSENQRKLD DLVIPSSIRNKMYGFWEKLFNNKIISPKKFYSLIKTEFNEKDQERFINRQIV ETRQITKHVAQIIDNHYENTKVVTVRADLSHQFRERYHIYKNRDINDFHH AHDAYIATILGTYIGHRFESLDAKYIYGEYKRIFRNQKNKGKEMKKNND GFILNSMRNIYADKDTGEIVWDPNYIDRIKKCFYYKDCFVTKKLEENNG TFFNVTVLPNDTNSDKDNTLATVPVNKYRSNVNKYGGFSGVNSFIVAIK GKKKKGKKVIEVNKLTGIPLMYKNADEEIKINYLKQAEDLEEVQIGKEIL KNQLIEKDGGLYYIVAPTEIINAKQLILNESQTKLVCEIYKAMKYKNYDN LDSEKIIDLYRLLINKMELYYPEYRKQLVKKFEDRYEQLKVISIEEKCNII KQILATLHCNSSIGKIMYSDFKISTTIGRLNGRTISLDDISFIAESPTGMYSK KYKL (SEQ ID NO: 35)
EfCas9 *Enterococcus faecalis* NCBI Исходная последовательность: WP_016631044.1 Дикий тип	MRLFEEGHTAEDRRLKRTARRRISRRRNRLRYLQAFFEEAMTDLDENFF ARLQESFLVPEDKKWHRHPIFAKLEDEVAYHETYPTIYHLRKKLADSSE QADLRLIYLALAHIVKYRGHFLIEGKLSTENTSVKDQFQQFMVIYNQTFV NGESRLVSAPLPESVLIEEELTEKASRTKKSEKVLQQFPQEKANGLFGQF LKLMVGNKADFKKVFGLEEEAKITYASESYEEDLEGILAKVGDEYSDVF LAAKNVYDAVELSTILADSDKKSHAKLSSSMIVRFTEHQEDLKKFKRFIR ENCPDEYDNLFKNEQKDGYAGYIAHAGKVSQLKFYQYVKKIIQDIAGAE YFLEKIAQENFLRKQRTFDNGVIPHQIHLAELQAIIHRQAAYYPFLKENQE KIEQLVTFRIPYYVGPLSKGDASTFAWLKRQSEEPIRPWNLQETVDLDQS ATAFIERMTNFDTYLPSEKVLPKHSLLYEKFMVFNELTKISYTDDRGIKA NFSGKEKEKIFDYLFKTRRKVKKKDIIQFYRNEYNTEIVTLSGLEEDQFN ASFSTYQDLLKCGLTRAELDHPDNAEKLEDIIKILTIFEDRQRIRTQLSTFK GQFSAEVLKKLERKHYTGWGRLSKKLINGIYDKESGKTILDYLVKDDGV SKHYNRNFMQLINDSQLSFKNAIQKAQSSEHEETLSETVNELAGSPAIKK GIYQSLKIVDELVAIMGYAPKRIVVEMARENQTTSTGKRRSIQRLKIVEK

Описание	Последовательность
	AMAEIGSNLLKEQPTTNEQLRDTRLFLYYMQNGKDMYTGDELSLHRLS HYDIDHIIPQSFMKDDSLDNLVLVGSTENRGKSDDVPSKEVVKDMKAY WEKLYAAGLISQRKFQRLTKGEQGGLTLEDKAHFIQRQLVETRQITKNV AGILDQRYNAKSKEKKVQIITLKASLTSQFRSIFGLYKVREVNDYHHGQD AYLNCVVATTLLKVYPNLAPEFVYGEYPKFQTFKENKATAKAIIYTNLL RFFTEDEPRFTKDGEILWSNSYLKTIKKELNYHQMNIVKKVEVQKGGFS KESIKPKGPSNKLIPVKNGLDPQKYGGFDSPVVAYTVLFTHEKGKKPLIK QEILGITIMEKTRFEQNPILFLEEKGFLRPRVLMKLPKYTLYEFPEGRRRL LASAKEAQKGNQMVLPEHLLTLLYHAKQCLLPNQSESLAYVEQHQPEF QEILERVVDFAEVHTLAKSKVQQIVKLFEANQTADVKEIAASFIQLMQFN AMGAPSTFKFFQKDIERARYTSIKEIFDATIIYQSPTGLYETRRKVVD (SEQ ID NO: 36)
Staphylococcus aureus Cas9	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKL SEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKY VAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFI DTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSV KYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTL KQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAEL LDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAI NLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVK RSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTN ERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNY EVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETF KKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATR GLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHA EDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYK EIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDE KNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYP NSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKC YEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMI DITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG (SEQ ID NO: 37)
Geobacillus thermodenitrificans Cas9	MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRL ARSARRRLRRRKHRLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRV EALDRKLNNDELARILLHLAKRRGFRSNRKSERTNKENSTMLKHIEENQ SILSSYRTVAEMVVKDPKFSLHKRNKEDNYTNTVARDDLEREIKLIFAKQ REYGNIVCTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFEPKEKRAPK ATYTFQSFTVWEHINKLRLVSPGGIRALTDDERRLIYKQAFHKNKITFHD VRTLLNLPDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVY GKGAAKSFRPIDFDTFGYALTMFKDDTDIRSYLRNEYEQNGKRMENLA DKVYDEELIEELLNLSFSKFGHLSLKALRNILPYMEQGEVYSTACERAGY TFTGPKKKQKTVLLPNIPPIANPVVMRALTQARKVVNAIIKKYGSPVSIHI ELARELSQSFDERRKMQKEQEGNRKKNETAIRQLVEYGLTLNPTGLDIV KFKLWSEQNGKCAYSLQPIEIERLLEPGYTEVDHVIPYSRSLDDSYTNKV LVLTKENREKGNRTPAEYLGLGSERWQQFETFVLTNKQFSKKKRDRLLR

Описание	Последовательность
	LHYDENEENEFKNRNLNDTRYISRFLANFIREHLKFADSDDKQKVYTVN GRITAHLRSRWNFNKNREESNLHHAVDAAIVACTTPSDIARVTAFYQRR EQNKELSKKTDPQFPQPWPHFADELQARLSKNPKESIKALNLGNYDNEK LESLQPVFVSRMPKRSITGAAHQETLRRYIGIDERSGKIQTVVKKKLSEIQ LDKTGHFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGE LGPIIRTIKIIDTTNQVIPLNDGKTVAYNSNIVRVDVFEKDGKYYCVPIYTI DMMKGILPNKAIEPNKPYSEWKEMTEDYTFRFSLYPNDLIRIEFPREKTIK TAVGEEIKIKDLFAYYQTIDSSNGGLSLVSHDNNFSLRSIGSRTLKRFEKY QVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 38)
ScCas9 *S. canis* 1375 AA 159.2 кДа	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLM GALLFDSGETAEATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSF FQRLEESFLVEEDKKNERHPIFGNLADEVAYHRNYPTIYHLRKKLADSPE KADLRLIYLALAHIIKFRGHFLIEGKLNAENSDVAKLFYQLIQTYNQLFEE SPLDEIEVDAKGILSARLSKSKRLEKLIAVFPNEKKNGLFGNIIALALGLTP NFKSNFDLTEDAKLQLSKDTYDDDLDELLGQIGDQYADLFSAAKNLSDA ILLSDILRSNSEVTKAPLSASMVKRYDEHHQDLALLKTLVRQQFPEKYAE IFKDDTKNGYAGYVGIGIKHRKRTTKLATQEEFYKFIKPILEKMDGAEEL LAKLNRDDLLRKQRTFDNGSIPHQIHLKELHAILRRQEEFYPFLKENREKI EKILTFRIPYYVGPLARGNSRFAWLTRKSEEAITPWNFEEVVDKGASAQS FIERMTNFDEQLPNKKVLPKHSLLYEYFTVYNELTKVKYVTERMRKPEF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIIGVEDRFNA SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRHYTGWGRLSRKMINGIRDKQSGKTILDFLKSDGF SNRNFMQLIHDDSLTFKEEIEKAQVSGQGDSLHEQIADLAGSPAIKKGIL QTVKIVDELVKVMGHKPENIVIEMARENQTTTKGLQQSRERKKRIEEGIK ELESQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFIKDDSIDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSEADKAGFIKRQLVETRQITKHVARILDS RMNTKRDKNDKPIREVKVITLKSKLVSDFRKDFQLYKVRDINNYHHAH DAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKAT AKRFFYSNIMNFFKTEVKLANGEIRKRPLIETNGETGEVVWNKEKDFAT VRKVLAMPQVNIVKKTEVQTGGFSKESILSKRESAKLIPRKKGWDTRKY GGFGSPTVAYSILVVAKVEKGKAKKLKSVKVLVGITIMEKGSYEKDPIGF LEAKGYKDIKKELIFKLPKYSLFELENGRRRMLASATELQKANELVLPQ HLVRLLYYTQNISATTGSNNLGYIEQHREEFKEIFEKIIDFSEKYILKNKV NSNLKSSFDEQFAVSDSILLSNSFVSLLKYTSFGASGGFTFLDLDVKQGRL RYQTVTEVLDATLIYQSITGLYETRTDLSQLGGD (SEQ ID NO: 39)

[0313] Описанные здесь праймерные редакторы могут включать любую из указанных выше последовательностей ортолога Cas9 или любой его вариант, имеющий, по меньшей мере 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность относительно них.

[0314] napDNAbp может включать любые подходящие гомологи и/или ортологи или встречающиеся в природе ферменты, например Cas9. Гомологи и/или ортологи Cas9 были описаны у различных видов, включая, но без ограничения, *S. pyogenes* и *S.*

thermophilus. Предпочтительно, фрагмент Cas имеет конфигурацию (например мутагенизирован, рекомбинантно сконструирован, или в ином случае получен из природы) в виде никазы, то есть способен расщеплять только одну цепь мишени. Дополнительные подходящие нуклеазы и последовательности Cas9 будут очевидны специалисту в данной области на основании данного описания, при этом подобные нуклеазы и последовательности Cas9 включают последовательности Cas9 из организмов и локусов, описанные в Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) RNA Biology 10:5, 726-737, полное содержание которого приведено здесь в качестве ссылки. В некоторых вариантах осуществления нуклеаза Cas9 имеет неактивный (например деактивированный) домен расщепления ДНК, то есть Cas9 представляет собой никазу. В некоторых вариантах осуществления белок Cas9 включает любую аминокислотную последовательность, которая, по меньшей мере, на 80% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 3. В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 92%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична аминокислотной последовательности белка Cas9 любого одного из ортологов Cas9 в приведенных выше таблицах.

C. Вариант в виде мертвого Cas9

[0315] В определенных вариантах осуществления описанные здесь праймерные редакторы могут включать мертвый Cas9, например мертвый SpCas9, который не имеет нуклеазной активности вследствие одной или нескольких мутаций, которые деактивируют оба нуклеазных домена Cas9, а именно домен RuvC (который расщепляет непротоспейсерную цепь ДНК) и домен HNH (который расщепляет протоспейсерную цепь ДНК). Нуклеазная деактивация может быть вызвана одной или несколькими мутациями, которые приводят к одному или нескольким замещениям и/или делециям в аминокислотной последовательности кодируемого белка или любого его варианта, имеющего, по меньшей мере 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере 95%, или, по меньшей мере, 99% идентичность последовательности относительно нее.

[0316] В соответствии с используемым здесь значением термин "dCas9" обозначает Cas9 с деактивированной нуклеазной активностью Cas9 или не имеющий нуклеазной активности Cas9 или его функциональный фрагмент и охватывает любой встречающийся в природе dCas9 из любого организма, любой эквивалент встречающегося в природе dCas9 или его функциональный фрагмент, любой гомолог dCas9, ортолог, или паралог из любого организма и любой мутант или вариант dCas9, являющийся встречающимся в природе или сконструированным. Термин "dCas9" не является конкретно ограничивающим и может рассматриваться как "dCas9 или эквивалент". Примеры белков dCas9 и способ получения белков dCas9 дополнительно описаны здесь и/или описаны в данной области и приведены здесь в качестве ссылки.
[0317] В других вариантах осуществления dCas9 соответствует или включает частично или полностью аминокислотную последовательность Cas9, имеющую одну или несколько мутаций, которые деактивируют нуклеазную активность Cas9. В других вариантах осуществления предусмотрены варианты Cas9, имеющие мутации, отличные от D10A и H840A, которые могут приводить к полной или частичной деактивации нуклеазной активности эндогенного Cas9 (например nCas9 или dCas9, соответственно). В качестве примера такие мутации включают другие аминокислотные замещения в D10 и H820 или другие замещения в нуклеазных доменах Cas9 (например замещения в нуклеазном субдомене HNH и/или субдомене RuvC1) со ссылкой на последовательность дикого типа, например Cas9 из *Streptococcus pyogenes* (кат. # NCBI: NC_017053.1). В некоторых вариантах осуществления предусмотрены варианты или гомологи Cas9 (например варианты Cas9 из *Streptococcus pyogenes* (кат. # NCBI: NC_017053.1 (SEQ ID NO: 20))), которые, по меньшей мере, на 70% идентичны, по меньшей мере, на 80% идентичны, по меньшей мере, на 90% идентичны, по меньшей мере, на 95% идентичны, по меньшей мере, на 98% идентичны, по меньшей мере, на 99% идентичны, по меньшей мере, на 99,5% идентичны, или, по меньшей мере, на 99,9% идентичны последовательности NCBI NC_017053.1. В некоторых вариантах осуществления предусмотрены варианты dCas9 (например варианты кат. # NCBI NC_017053.1 (SEQ ID NO: 20)), имеющие аминокислотные последовательности, которые являются более короткими или более длинными, чем NC_017053.1 (SEQ ID NO: 20) приблизительно на 5 аминокислот, приблизительно на 10 аминокислот, приблизительно на 15 аминокислот, приблизительно на 20 аминокислот, приблизительно на 25 аминокислот, приблизительно на 30 аминокислот,

приблизительно на 40 аминокислот, приблизительно на 50 аминокислот, приблизительно на 75 аминокислот, приблизительно на 100 аминокислот или более.

[0318] В одном варианте осуществления мертвый Cas9 может быть основан на стандартной последовательности SpCas9 Q99ZW2 и может иметь следующую последовательность, которая включает D10X и H810X, где X может представлять собой любую аминокислоту, замещения (подчеркнуты или указаны жирным шрифтом), или являться вариантом SEQ ID NO: 40, имеющим, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно нее.

[0319] В одном варианте осуществления мертвый Cas9 может быть основан на стандартной последовательности SpCas9 Q99ZW2 и может иметь следующую последовательность, которая включает замещения D10A и H810A (подчеркнуты или указаны жирным шрифтом), или являться вариантом SEQ ID NO: 41, имеющим, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно нее.

Описание	Последовательность	SEQ ID NO:
мертвый Cas9 или dCas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**X** и H810**X** Где "X" представляет собой любую аминокислоту	MDKKYSIGL**X**IGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVD**X**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND	SEQ ID NO: 40

Описание	Последовательность	SEQ ID NO:
	KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
мертвый Cas9 или dCas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**A** и H810**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 41

D. <u>Вариант в виде никазы Cas9</u>

[0320] В одном варианте осуществления описанные здесь праймерные редакторы включают никазу Cas9. Термин "никаза Cas9" или "nCas9" обозначает вариант Cas9, который способен вносить одноцепочечный разрыв в двуцепочечную молекулу-мишень ДНК. В некоторых вариантах осуществления никаза Cas9 включает только один функционирующий нуклеазный домен. Cas9 дикого типа (например стандартный SpCas9) включает два отдельных нуклеазных домена, а именно домен RuvC (который расщепляет непротоспейсерную цепь ДНК) и домен HNH (который расщепляет протоспейсерную цепь ДНК). В одном варианте осуществления никаза Cas9 включает мутацию в домене RuvC, которая деактивирует нуклеазную активность RuvC. Например, сообщалось о мутациях в аспартате (D) 10, гистидине (H) 983, аспартате (D) 986, или глутамате (E) 762 как о связанных с потерей функции мутациях нуклеазного домена RuvC и создании функциональной никазы Cas9 (например Nishimasu et al., “Crystal structure of Cas9 in complex with guide RNA and target DNA,” *Cell* 156(5), 935–949, который приведен здесь в качестве ссылки). Таким образом, никазные мутации в домене RuvC могут включать D10X, H983X, D986X, или E762X, где X представляет собой любую аминокислоту, отличную от аминокислоты дикого типа. В определенных вариантах осуществления никаза может представлять собой D10A, или H983A, или D986A, или E762A, или их комбинацию.

[0321] В различных вариантах осуществления никаза Cas9 может иметь мутацию в нуклеазном домене RuvC и иметь одну из следующих аминокислотных последовательностей или являться ее вариантом, имеющим аминокислотную последовательность, которая имеет, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно нее.

Описание	Последовательность	SEQ ID NO:
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**X**, где X представляет собой любую альтернативную	MDKKYSIGL**X**IGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF	SEQ ID NO: 42

Описание	Последовательность	SEQ ID NO:
аминокислоту	EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с E762X, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVI**X**MARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD	SEQ ID NO: 43

Описание	Последовательность	SEQ ID NO:
	AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с H983X, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH**X**AHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 44
Никаза Cas9 *Streptococcus*	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF	SEQ ID NO: 45

Описание	Последовательность	SEQ ID NO:
pyogenes Q99ZW2 Cas9 с D986X, где X представляет собой любую альтернативную аминокислоту	GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH**X** AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL	SEQ ID NO: 46

Описание	Последовательность	SEQ ID NO:
	KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c E762A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVI**A**MARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS	SEQ ID NO: 47

Описание	Последовательность	SEQ ID NO:
	VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c H983A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH**A**AHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 48
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c D986A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN	SEQ ID NO: 49

Описание	Последовательность	SEQ ID NO:
	LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH**A** AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	

[0322] В другом варианте осуществления никаза Cas9 включает мутацию в домене HNH, которая деактивирует активность нуклеазы HNH. Например, сообщалось, что мутации в гистидине (H) 840 или аспарагине (R) 863 являются связанными с потерей функции мутациями домена нуклеазы HNH и созданием функциональной никазы Cas9 (например Nishimasu et al., “Crystal structure of Cas9 in complex with guide RNA and target DNA,” *Cell* 156(5), 935–949, который приведен здесь в качестве ссылки). Таким образом, никазные мутации в домене HNH могут включать H840X и R863X, где X представляет собой любую аминокислоту, отличную от аминокислоты дикого типа. В определенных вариантах осуществления никаза может представлять собой H840A, или R863A, или их комбинацию.

[0323] В различных вариантах осуществления никаза Cas9 может иметь мутацию в нуклеазном домене HNH и иметь одну из следующих аминокислотных последовательностей или представлять собой их вариант, имеющий аминокислотную

последовательность, которая имеет, по меньшей мере, 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, или, по меньшей мере, 99% идентичность последовательности относительно них.

Описание	Последовательность	SEQ ID NO:
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**X**, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVD**X**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 50
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**A**	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS	SEQ ID NO: 51

Описание	Последовательность	SEQ ID NO:
	ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKN**R**GKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с R863X, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN	SEQ ID NO: 52

Описание	Последовательность	SEQ ID NO:
	RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**X**GKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c R863**A**	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**A**GKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE	SEQ ID NO: 53

Описание	Последовательность	SEQ ID NO:
	VLDATLIHQSITGLYETRIDLSQLGGD	

[0324] В некоторых вариантах осуществления N-концевой метионин удаляется из никазы Cas9 или из любого варианта Cas9, ортолога, или эквивалента, описанного или подразумеваемого здесь. Например, никазы Cas9 без метионина включают следующие последовательности или их варианты, имеющие аминокислотную последовательность, которая имеет, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно них.

Описание	Последовательность
Никаза Cas9 (без Met) *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**X**, где X представляет собой любую альтернативную аминокислоту	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVD**X**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGD (SEQ ID NO: 54)
Никаза Cas9 (без Met) *Streptococcus pyogenes*	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL

Q99ZW2 Cas9 c H840**A**	GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKN<u>R</u>GKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGD (SEQ ID NO: 55)
Никаза Cas9 (без Met) *Streptococcus pyogenes* Q99ZW2 Cas9 c R863X, где X представляет собой любую альтернативную аминокислоту	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**X**GKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM

	LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGD(SEQ ID NO: 56)
Никаза Cas9 (без Met) *Streptococcus pyogenes* Q99ZW2 Cas9 с R863**A**	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**A**GKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGD (SEQ ID NO: 57)

E. Другие варианты Cas9

[0325] Помимо вариантов в виде мертвого Cas9 и никазы Cas9, используемые здесь белки Cas9 могут также включать другие "варианты Cas9", являющиеся, по меньшей мере, приблизительно на 70% идентичными, по меньшей мере, приблизительно на 80% идентичными, по меньшей мере, приблизительно на 90% идентичными, по меньшей мере, приблизительно на 95% идентичными, по меньшей мере, приблизительно на 96% идентичными, по меньшей мере, приблизительно на 97% идентичными, по меньшей мере, приблизительно на 98% идентичными, по меньшей мере, приблизительно на 99% идентичными, по меньшей мере, приблизительно на 99,5% идентичными, или, по меньшей мере, приблизительно на 99,9% идентичными любому исходному белку Cas9, включая любой Cas9 дикого типа или мутант Cas9 (например мертвый Cas9 или никазу

Cas9), или фрагмент Cas9, или циркулярный пермутант Cas9, или другой описанный здесь или известный в данной области вариант Cas9. В некоторых вариантах осуществления вариант Cas9 может иметь 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 или более аминокислотных изменений по сравнению с исходным Cas9. В некоторых вариантах осуществления вариант Cas9 включает фрагмент исходного (например связывающий нРНК домен или домен расщепления ДНК), так что фрагмент является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным соответствующему фрагменту Cas9 дикого типа. В некоторых вариантах осуществления фрагмент имеет длину, соответствующую, по меньшей мере, на 30%, по меньшей мере, на 35%, по меньшей мере, на 40%, по меньшей мере, на 45%, по меньшей мере, на 50%, по меньшей мере, на 55%, по меньшей мере, на 60%, по меньшей мере, на 65%, по меньшей мере, на 70%, по меньшей мере, на 75%, по меньшей мере, на 80%, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% длине аминокислоты соответствующего Cas9 дикого типа (например SEQ ID NO: 18).

[0326] В некоторых вариантах осуществления в изобретении также могут использоваться фрагменты Cas9, которые сохраняют свою функциональность и которые представляют собой фрагменты любого описанного здесь белка Cas9. В некоторых вариантах осуществления фрагмент Cas9 имеет длину, по меньшей мере, 100 аминокислот. В некоторых вариантах осуществления фрагмент имеет длину, по меньшей мере, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, или, по меньшей мере, 1300 аминокислот.

[0327] В различных вариантах осуществления описанные здесь праймерные редакторы могут включать один из вариантов Cas9, описанных далее, или его вариант, который

является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным любым исходным вариантам Cas9.

F. Варианты Cas9 малого размера

[0328] В некоторых вариантах осуществления указанные здесь праймерные редакторы могут включать белок Cas9, который имеет меньшую молекулярную массу, чем стандартная последовательность SpCas9. В некоторых вариантах осуществления варианты Cas9 малого размера могут обеспечивать доставку в клетки, например с помощью вектора экспрессии, наночастицы, или других средств доставки. В определенных вариантах осуществления варианты Cas9 малого размера могут включать ферменты, относящиеся к ферментам типа II систем CRISPR-Cas Класса 2. В некоторых вариантах осуществления варианты Cas9 малого размера могут включать ферменты, относящиеся к ферментам типа V систем CRISPR-Cas Класса 2. В других вариантах осуществления варианты Cas9 малого размера могут включать ферменты, относящиеся к ферментам типа VI систем CRISPR-Cas Класса 2.

[0329] Стандартный белок SpCas9 имеет длину 1368 аминокислот и имеет предполагаемую молекулярную массу 158 кДа. В соответствии с используемым здесь значением термин "вариант Cas9 малого размера" обозначает любой вариант Cas9 — встречающийся в природе, сконструированный или полученный иным образом — длиной менее чем приблизительно 1300 аминокислот, или менее чем 1290 аминокислот, или менее чем 1280 аминокислот, или менее чем 1270 аминокислот, или менее чем 1260 аминокислот, или менее чем 1250 аминокислот, или менее чем 1240 аминокислот, или менее чем 1230 аминокислот, или менее чем 1220 аминокислот, или менее чем 1210 аминокислот, или менее чем 1200 аминокислот, или менее чем 1190 аминокислот, или менее чем 1180 аминокислот, или менее чем 1170 аминокислот, или менее чем 1160 аминокислот, или менее чем 1150 аминокислот, или менее чем 1140 аминокислот, или менее чем 1130 аминокислот, или менее чем 1120 аминокислот, или менее чем 1110 аминокислот, или менее чем 1100 аминокислот, или менее чем 1050

аминокислот, или менее чем 1000 аминокислот, или менее чем 950 аминокислот, или менее чем 900 аминокислот, или менее чем 850 аминокислот, или менее чем 800 аминокислот, или менее чем 750 аминокислот, или менее чем 700 аминокислот, или менее чем 650 аминокислот, или менее чем 600 аминокислот, или менее чем 550 аминокислот, или менее чем 500 аминокислот, но, по меньшей мере, более чем приблизительно 400 аминокислот, сохраняющий требуемые функции белка Cas9. Варианты Cas9 могут включать варианты, относящиеся к ферментам типа II, типа V, или типа VI системы CRISPR-Cas Класса 2.

[0330] В различных вариантах осуществления описанные здесь праймерные редакторы могут включать один из вариантов Cas9 малого размера, описанных далее, или вариант Cas9, который является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным любому исходному белку Cas9 малого размера.

Описание	Последовательность	SEQ ID NO:
SaCas9 *Staphylococcus aureus* 1053 AA 123 кДа	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEA NVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLT DHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLER LKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQ SFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLM GHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENE KLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYR VTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILT IYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSL KAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIP TTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAR EKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKY LIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIP RSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKIS YETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKD FINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGG FTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEW KKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITP HQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDK GNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDP	SEQ ID NO: 58

Описание	Последовательность	SEQ ID NO:
	QTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKD NGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKP YRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEE AKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLL NRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKY STDILGNLYEVKSKKHPQIIKK	
NmeCas9 *N. meningitidis* 1083 AA 124,5 кДа	MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDL GVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRL LRTRRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAAL DRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGA LLKGVAGNAHALQTGDFRTPAELALNKFEKESGHIRNQR SDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIE TLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTA ERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSK LTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKA YHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDE DITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLM EQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNP VVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKD RKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILK LRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDAALPFSRT WDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREW QEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRY VNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFW GLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYK EMNAFDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRV FGKPDGKPEFEEADTLEKLRTLLAEKLSSRPEAVHEYVTP LFVSRAPNRKMSGQGHMETVKSAKRLDEGVSVLRVPLT QLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAF AEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNG IADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRA VVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMF GYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVKTALSF QKYQIDELGKEIRPCRLKKRPPVR	SEQ ID NO: 59
CjCas9 *C. jejuni* 984 AA 114,9 кДа	MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTG ESLALPRRLARSARKRLARRKARLNHLKHLIANEFKLNY EDYQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFAR VILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQ SVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCIAQSFL KDELKLIFKKQREFGFSFSKKFEEEVLSVAFYKRALKDFS HLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKN TEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDY EFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDIT LIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKA LKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFN ETYYKDEVTNPVVLRAIKEYRKVLNALLKKYGKVHKINI ELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLG LKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEI	SEQ ID NO: 60

Описание	Последовательность	SEQ ID NO:
	DHIYPYSRSFDDSYMNKVLVFTKQNQEKLNQTPFEAFGN DSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKD RNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDTQK GSKVHVEAKSGMLTSALRHTWGFSAKDRNNHLHHAIDA VIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKN KRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETF RKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDM FRVDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKK GEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFVYY NAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIA KSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK	
GeoCas9 *G. stearothermophilus* 1087 AA 127 кДа	MRYKIGLDIGITSVGWAVMNLDIPRIEDLGVRIFDRAENP QTGESLALPRRLARSARRRLRRRKHRLERIRRLVIREGILT KEELDKLFEEKHEIDVWQLRVEALDRKLNNDELARVLLH LAKRRGFKSNRKSERSNKENSTMLKHIEENRAILSSYRTV GEMIVKDPKFALHKRNKGENYTNTIARDDLEREIRLIFSK QREFGNMSCTEEFENEYITIWASQRPVASKDDIEKKVGFC TFEPKEKRAPKATYTFQSFIAWEHINKLRLISPSGARGLTD EERRLLYEQAFQKNKITYHDIRTLLHLPDDTYFKGIVYDR GESRKQNENIRFLELDAYHQIRKAVDKVYGKGKSSSFLPI DFDTFGYALTLFKDDADIHSYLRNEYEQNGKRMPNLAN KVYDNELIEELLNLSFTKFGHLSLKALRSILPYMEQGEVY SSACERAGYTFTGPKKKQKTMLLPNIPPIANPVVMRALT QARKVVNAIIKKYGSPVSIHIELARDLSQTFDERRKTKKE QDENRKKNETAIRQLMEYGLTLNPTGHDIVKFKLWSEQN GRCAYSLQPIEIERLLEPGYVEVDHVIPYSRSLDDSYTNK VLVLTRENREKGNRIPAEYLGVGTERWQQFETFVLTNKQ FSKKKRDRLLRLHYDENEETEFKNRNLNDTRYISRFFANF IREHLKFAESDDKQKVYTVNGRVTAHLRSRWEFNKNRE ESDLHHAVDAVIVACTTPSDIAKVTAFYQRREQNKELAK KTEPHFPQPWPHFADELRARLSKHPKESIKALNLGNYDD QKLESLQPVFVSRMPKRSVTGAAHQETLRRYVGIDERSG KIQTVVKTKLSEIKLDASGHFPMYGKESDPRTYEAIRQRL LEHNNDPKKAFQEPLYKPKKNGEPGPVIRTVKIIDTKNQV IPLNDGKTVAYNSNIVRVDVFEKDGKYYCVPVYTMDIM KGILPNKAIEPNKPYSEWKEMTEDYTFRFSLYPNDLIRIEL PREKTVKTAAGEEINVKDVFVYYKTIDSANGGLELISHDH RFSLRGVGSRTLKRFEKYQVDVLGNIYKVRGEKRVGLAS SAHSKPGKTIRPLQSTRD	SEQ ID NO: 61
LbaCas12a *L. bacterium* 1228 AA 143,9 кДа	MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVED EKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISL FRKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFK KDIIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENM FSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEV QEIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGG FVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSD RESLSFYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKL FKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIRDKWNA	SEQ ID NO: 62

Описание	Последовательность	SEQ ID NO:
	EYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEY ADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLE KSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNR DESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKF KLYFQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMD KKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVF FSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKL IDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEE QGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSH GTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKK EELVVHPANSPIANKNPDNPKKTTTLSYDVYKDKRFSED QYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIDR GERNLLYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHS LLDKKEKERFEARQNWTSIENIKELKAGYISQVVHKICEL VEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLID KLNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQN GFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFD RIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNR IRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDI RALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFL ISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIARK VLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKH	
BhCas12b *B. hisashii* 1108 AA 130,4 кДа	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILK LIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSN KFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEE EKKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPI VKEIKWMEKSRNQSVRRLDKDMFIQALERFLSWESWNL KVKEEYEKVEKEYKTLEERIKEDIQALKALEQYEKERQE QLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSE KYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHP EYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRF EERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTES GGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKD ESIKFPLKGTLGGARVQFDRDHLRRYPHKVESGNVGRIYF NMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIK DSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVD QKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREV LRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVT KWISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQ LHKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDR TRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKE DRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILF EDLSNYNPYEERSRFENSKLMKWSRREIPRQVALQGEIY GLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRF FKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRK CVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDG QTVYIPESKDQKQKIIEEFGEGYFILKDGVYEWVNAGKLK	SEQ ID NO: 63

Описание	Последовательность	SEQ ID NO:
	IKKGSSKQSSSELVDSDILKDSFDLASELKGEKLMLYRDP SGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDS SKQSM	

G. Эквиваленты Cas9

[0331] В некоторых вариантах осуществления описанные здесь праймерные редакторы могут включать любой эквивалент Cas9. В соответствии с используемым здесь значением термин "эквивалент Cas9" является широким термином, который охватывает любой белок napDNAbp, который служит той же функции, что и Cas9 в настоящих праймерных редакторах, несмотря на то, что его первичная аминокислотная последовательность и/или его трехмерная структура может отличаться и/или не являться родственной с эволюционной точки зрения. Таким образом, в то время как эквиваленты Cas9 включают любой ортолог, гомолог, мутант или вариант Cas9, описанный или охватываемый здесь, который является эволюционно родственным, эквиваленты Cas9 также охватывают белки, которые могли эволюционировать за счет конвергентных эволюционных процессов, имеющие такую же или аналогичную функцию, что и Cas9, но которые не обязательно имеют какое-либо сходство в плане аминокислотной последовательности и/или трехмерной структуры. Описанные здесь праймерные редакторы охватывают любой эквивалент Cas9, обеспечивающий предоставление такой же или аналогичной функции, что и Cas9, несмотря на то, что эквивалент Cas9 может быть основан на белке, который образовался за счет конвергентной эволюции. Например, если Cas9 обозначает фермент типа II системы CRISPR-Cas, эквивалент Cas9 может обозначать фермент типа V или типа VI системы CRISPR-Cas.

[0332] Например Cas12e (CasX) представляет собой эквивалент Cas9, который, как сообщается, имеет такую же функцию, что и Cas9, но который эволюционировал за счет конвергентной эволюции. Таким образом, предполагается, что белок Cas12e (CasX), описанный в Liu et al., “CasX enzymes comprises a distinct family of RNA-guided genome editors,” *Nature*, 2019, Vol.566: 218-223, будет использоваться с описанными здесь праймерными редакторами. Кроме того, в рамках объема настоящего изобретения может рассматриваться любой вариант или модификация Cas12e (CasX).

[0333] Cas9 представляет собой бактериальный фермент, который эволюционировал у широкого ряда видов. Однако рассматриваемые здесь эквиваленты Cas9 также могут

быть получены из архей, которые составляют домен и царство одноклеточных прокариотических микробов, отличных от бактерий.

[0334] В некоторых вариантах осуществления эквиваленты Cas9 могут обозначать Cas12e (CasX) или Cas12d (CasY), которые были описаны, например, в Burstein et al., "New CRISPR–Cas systems from uncultivated microbes." *Cell Res*. 2017 Feb 21. doi: 10.1038/cr.2017.21, полное содержание которого приведено здесь в качестве ссылки. С помощью метагеномики с геномным корректированием был идентифицирован ряд систем CRISPR–Cas, включая первый обнаруженный Cas9 в домене жизни архей. Данный отличающийся белок Cas9 был обнаружен у малоизученных наноархей в виде части активной системы CRISPR–Cas. В бактериях были открыты две ранее неизвестные системы, CRISPR– Cas12e и CRISPR– Cas12d, которые являются одними из наиболее компактных систем, открытых на сегодняшний день. В некоторых вариантах осуществления Cas9 обозначает Cas12e или вариант Cas12e. В некоторых вариантах осуществления Cas9 обозначает Cas12d или вариант Cas12d. Следует понимать, что в качестве программируемого нуклеиновой кислотой ДНК-связывающего белка могут использоваться другие РНК-направляемые связывающие ДНК белки, при этом они находятся в рамках объема данного изобретения. Также см. Liu et al., "CasX enzymes comprises a distinct family of RNA-guided genome editors," *Nature*, 2019, Vol.566: 218-223. Включает любой из данных эквивалентов Cas9.

[0335] В некоторых вариантах осуществления эквивалент Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 91%, по меньшей мере, на 92%, по меньшей мере, на 93%, по меньшей мере, на 94%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична встречающемуся в природе белку Cas12e (CasX) или Cas12d (CasY). В некоторых вариантах осуществления napDNAbp представляет собой встречающийся в природе белок Cas12e (CasX) или Cas12d (CasY). В некоторых вариантах осуществления napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 91%, по меньшей мере, на 92%, по меньшей мере, на 93%, по меньшей мере, на 94%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по

меньшей мере, на 99,5% идентична фрагменту Cas дикого типа или любому указанному здесь фрагменту Cas.

[0336] В различных вариантах осуществления программируемые нуклеиновой кислотой ДНК-связывающие белки включают, без ограничения, Cas9 (например dCas9 и nCas9), Cas12e (CasX), Cas12d (CasY), Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), Cas12c (C2c3), Argonaute, и Cas12b1. Один пример программируемого нуклеиновой кислотой ДНК-связывающего белка, который имеет другую специфичность в отношении PAM, чем Cas9, представляет собой короткие регулярно расположенные группами палиндромные повторы из Prevotella и Francisella 1 (то есть Cas12a (Cpf1)). Аналогично Cas9, Cas12a (Cpf1) также представляет собой эффектор CRISPR Класса 2, но он является членом подгруппы ферментов типа V, а не подгруппы типа II. Было показано, что Cas12a (Cpf1) опосредует устойчивую интерференцию ДНК со свойствами, отличающимися от Cas9. Cas12a (Cpf1) представляет собой направляемую одной РНК эндонуклеазу без tracrРНК и использует смежный с протоспейсером мотив с высоким содержанием T (TTN, TTTN, или YTN). Более того, Cpf1 расщепляет ДНК за счет зигзагообразного двуцепочечного разрыва ДНК. Показано, что из 16 белков семейства Cpf1 два фермента из *Acidaminococcus* и *Lachnospiraceae* имеют эффективную связанную с редактированием генома активность в клетках человека. Белки Cpf1 известны в данной области и были ранее описаны, например, в Yamano *et al.*, “Crystal structure of Cpf1 in complex with guide RNA and target DNA.” *Cell* (165) 2016, p. 949-962; полное содержание которого приведено здесь в качестве ссылки.

[0337] В других вариантах осуществления белок Cas может включать любой ассоциированный с CRISPR белок, включая, но без ограничения, Cas12a, Cas12b1, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, их гомологи или модифицированные версии, и, предпочтительно, включающие никазную мутацию (например мутацию, соответствующую мутации D10A полипептида Cas9 дикого типа согласно SEQ ID NO: 18).

[0338] В различных других вариантах осуществления napDNAbp может представлять собой любой из следующих белков: Cas9, Cas12(Cpf1), Cas12e (CasX), Cas12d (CasY), Cas12b1 (C2c1), Cas13(C2c2), Cas12c (C2c3), GeoCas9, CjCas9, Cas12g, Cas12h, Cas12i,

Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, циркулярно пермутированный Cas9, домен Argonaute (Ago), или их вариант.

[0339] Примеры последовательностей белка-эквивалента Cas9 могут включать следующие:

Описание	Последовательность
AsCas12a (ранее известный как Cpf1) *Acidaminococcus sp. (штамм BV3L6)* UniProtKB U2UMQ6	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKEL KPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQA TYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTV TTTEHENALLRSFDKFTTYFSGFYENRKNVFSAEDISTAIPHRIVQDNFPK FKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQT QIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPL FKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNEL NSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEK VQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLK KQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMEP SLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILF VKNGLYYLGIMPKQKGRYKALSFEPTEKTSEGFDKMYYDYFPDAAKMI PKCSTQLKAVTAHFQTHTTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQT AYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKD LGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHG KPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGE KMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKE VSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPII GIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAARQ AWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFGFKSKRTG IAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAK MGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFL HYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTP FIAGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLE NDDSHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQ NPEWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYI QELRN (SEQ ID NO: 64)
Никаза AsCas12a (например R1226A)	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKEL KPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQA TYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTV TTTEHENALLRSFDKFTTYFSGFYENRKNVFSAEDISTAIPHRIVQDNFPK FKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQT QIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPL FKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNEL NSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEK VQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLK KQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMEP SLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILF VKNGLYYLGIMPKQKGRYKALSFEPTEKTSEGFDKMYYDYFPDAAKMI PKCSTQLKAVTAHFQTHTTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQT AYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKD LGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHG

	KPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGE KMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKE VSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPII GIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAARQ AWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFGFKSKRTG IAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAK MGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFL HYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTP FIAGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLE NDDSHAIDTMVALIRSVLQMANSNAATGEDYINSPVRDLNGVCFDSRFQ NPEWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYI QELRN (SEQ ID NO: 65)
LbCas12a (ранее известный как Cpf1) *Lachnospiraceae bacterium GAM79* Исх. послед. WP_119623382.1	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQ QELKEIMDDYYRTFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKI QNEKRKEICCYFTSDKRFKDLFNAKLITDILPNFIKDNKEYTEEEKAEKEQ TRVLFQRFATAFTNYFNQRRNNFSEDNISTAISFRIVNENSEIHLQNMRAF QRIEQQYPEEVCGMEEEYKDMLQEWQMKHIYSVDFYDRELTQPGIEYY NGICGKINEHMNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFES DQEVYDALNEFIKTMKKKEIIRRCVHLGQECDDYDLGKIYISSNKYEQIS NALYGSWDTIRKCIKEEYMDALPGKGEKKEEKAEAAAKKEEYRSIADID KIISLYGSEMDRTISAKKCITEICDMAGQISIDPLVCNSDIKLLQNKEKTTE IKTILDSFLHVYQWGQTFIVSDIIEKDSYFYSELEDVLEDFEGITTLYNHVR SYVTQKPYSTVKFKLHFGSPTLANGWSQSKEYDNNAILLMRDQKFYLGI FNVRNKPDKQIIKGHEKEEKGDYKKMIYNLLPGPSKMLPKVFITSRSGQE TYKPSKHILDGYNEKRHIKSSPKFDLGYCWDLIDYYKECIHKHPDWKNY DFHFSDTKDYEDISGFYREVEMQGYQIKWTYISADEIQKLDEKGQIFLFQI YNKDFSVHSTGKDNLHTMYLKNLFSEENLKDIVLKLNGEAELFFRKASI KTPIVHKKGSVLVNRSYTQTVGNKEIRVSIPEEYYTEIYNYLNHIGKGKL SSEAQRYLDEGKIKSFTATKDIVKNYRYCCDHYFLHLPITINFKAKSDVA VNERTLAYIAKKEDIHIIGIDRGERNLLYISVVDVHGNIREQRSFNIVNGY DYQQKLKDREKSRDAARKNWEEIEKIKELKEGYLSMVIHYIAQLVVKY NAVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLHYLVFKDREVCE EGGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFVNLFSF KNLTNRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTILASTKW KVYTNGTRLKRIVVNGKYTSQSMEVELTDAMEKMLQRAGIEYHDGKD LKGQIVEKGIEAEIIDIFRLTVQMRNSRSESEDREYDRLISPVLNDKGEFFD TATADKTLPQDADANGAYCIALKGLYEVKQIKENWKENEQFPRNKLVQ DNKTWFDFMQKKRYL (SEQ ID NO: 66)
PcCas12a – ранее известный как Cpf1 *Prevotella copri* Исх. послед. WP_11922	MAKNFEDFKRLYSLSKTLRFEAKPIGATLDNIVKSGLLDEDEHRAASYV KVKKLIDEYHKVFIDRVLDDGCLPLENKGNNNSLAEYYESYVSRAQDED AKKKFKEIQQNLRSVIAKKLTEDKAYANLFGNKLIESYKDKEDKKKIIDS DLIQFINTAESTQLDSMSQDEAKELVKEFWGFVTYFYGFFDNRKNMYTA EEKSTGIAYRLVNENLPKFIDNIEAFNRAITRPEIQENMGVLYSDFSEYLN VESIQEMFQLDYYNMLLTQKQIDVYNAIIGGKTDDEHDVKIKGINEYINL YNQQHKDDKLPKLKALFKQILSDRNAISWLPEEFNSDQEVLNAIKDCYE RLAENVLGDKVLKSLLGSLADYSLDGIFIRNDLQLTDISQKMFGNWGVI QNAIMQNIKRVAPARKHKESEEDYEKRIAGIFKKADSFSISYINDCLNEA DPNNAYFVENYFATFGAVNTPTMQRENLFALVQNAYTEVAALLHSDYP TVKHLAQDKANVSKIKALLDAIKSLQHFVKPLLGKGDESDKDERFYGEL

7726.1	ASLWAELDTVTPLYNMIRNYMTRKPYSQKKIKLNFENPQLLGGWDANK EKDYATIILRRNGLYYLAIMDKDSRKLLGKAMPSDGECYEKMVYKFFK DVTTMIPKCSTQLKDVQAYFKVNTDDYVLNSKAFNKPLTITKEVFDLNN VLYGKYKKFQKGYLTATGDNVGYTHAVNVWIKFCMDFLNSYDSTCIY DFSSLKPESYLSLDAFYQDANLLLYKLSFARASVSYINQLVEEGKMYLFQ IYNKDFSEYSKGTPNMHTLYWKALFDERNLADVVYKLNGQAEMFYRK KSIENTHPTHPANHPILNKNKDNKKKESLFDYDLIKDRRYTVDKFMFHV PITMNFKSVGSENINQDVKAYLRHADDMHIIGIDRGERHLLYLVVIDLQG NIKEQYSLNEIVNEYNGNTYHTNYHDLLDVREEERLKARQSWQTIENIK ELKEGYLSQVIHKITQLMVRYHAIVVLEDLSKGFMRSRQKVEKQVYQKF EKMLIDKLNYLVDKKTDVSTPGGLLNAYQLTCKSDSSQKLGKQSGFLFY IPAWNTSKIDPVTGFVNLLDTHSLNSKEKIKAFFSKFDAIRYNKDKKWFE FNLDYDKFGKKAEDTRTKWTLCTRGMRIDTFRNKEKNSQWDNQEVDL TTEMKSLLEHYYIDIHGNLKDAISAQTDKAFFTGLLHILKLTLQMRNSIT GTETDYLVSPVADENGIFYDSRSCGNQLPENADANGAYNIARKGLMLIE QIKNAEDLNNVKFDISNKAWLNFAQQKPYKNG (SEQ ID NO: 67)
ErCas12a – ранее известный как Cpf1 *Eubacterium rectale* Исх. послед. WP_119223642.1	MFSAKLISDILPEFVIHNNNYSASEKEEKTQVIKLFSRFATSFKDYFKNRA NCFSANDISSSSCHRIVNDNAEIFFSNALVYRRIVKNLSNDDINKISGDMK DSLKEMSLEEIYSYEKYGEFITQEGISFYNDICGKVNLFMNLYCQKNKEN KNLYKLRKLHKQILCIADTSYEVPYKFESDEEVYQSVNGFLDNISSKHIV ERLRKIGENYNGYNLDKIYIVSKFYESVSQKTYRDWETINTALEIHYNNI LPGNGKSKADKVKKAVKNDLQKSITEINELVSNYKLCPDDNIKAETYIH EISHILNNFEAQELKYNPEIHLVESELKASELKNVLDVIMNAFHWCSVFM TEELVDKDNNFYAELEEIYDEIYPVISLYNLVRNYVTQKPYSTKKIKLNF GIPTLADGWSKSKEYSNNAIILMRDNLYYLGIFNAKNKPDKKIIEGNTSE NKGDYKKMIYNLLPGPNKMIPKVFLSSKTGVETYKPSAYILEGYKQNKH LKSSKDFDITFCHDLIDYFKNCIAIHPEWKNFGFDFSDTSTYEDISGFYRE VELQGYKIDWTYISEKDIDLLQEKGQLYLFQIYNKDFSKKSSGNDNLHT MYLKNLFSEENLKDIVLKLNGEAEIFFRKSSIKNPIIHKKGSILVNRTYEAE EKDQFGNIQIVRKTIPENIYQELYKYFNDKSDKELSDEAAKLKNVVGHH EAATNIVKDYRYTYDKYFLHMPITINFKANKTSFINDRILQYIAKEKDLH VIGIDRGERNLIYVSVIDTCGNIVEQKSFNIVNGYDYQIKLKQQEGARQIA RKEWKEIGKIKEIKEGYLSLVIHEISKMVIKYNAIIAMEDLSYGFKKGRFK VERQVYQKFETMLINKLNYLVFKDISITENGGLLKGYQLTYIPDKLKNV GHQCGCIFYVPAAYTSKIDPTTGFVNIFKFKDLTVDAKREFIKKFDSIRYD SDKNLFCFTFDYNNFITQNTVMSKSSWSVYTYGVRIKRRFVNGRFSNES DTIDITKDMEKTLEMTDINWRDGHDLRQDIIDYEIVQHIFEIFKLTVQMR NSLSELEDRDYDRLISPVLNENNIFYDSAKAGDALPKDADANGAYCIAL KGLYEIKQITENWKEDGKFSRDKLKISNKDWFDFIQNKRYL (SEQ ID NO: 68)
CsCas12a – ранее известный как Cpf1 *Clostridium sp. AF34-10BH*	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQ QELKEIMDDYYRAFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKI QNEKRKEICCYFTSDKRFKDLFNAKLITDILPNFIKDNKEYTEEEKAEKEQ TRVLFQRFATAFTNYFNQRRNNFSEDNISTAISFRIVNENSEIHLQNMRAF QRIEQQYPEEVCGMEEEYKDMLQEWQMKHIYLVDFYDRVLTQPGIEYY NGICGKINEHMNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFES DQEVYDALNEFIKTMKEKEIICRCVHLGQKCDDYDLGKIYISSNKYEQIS NALYGSWDTIRKCIKEEYMDALPGKGEKKEEKAEAAAKKEEYRSIADID KIISLYGSEMDRTISAKKCITEICDMAGQISTDPLVCNSDIKLLQNKEKTTE

Исх. послед. WP_118538418.1	IKTILDSFLHVYQWGQTFIVSDIIEKDSYFYSELEDVLEDFEGITTLYNHVR SYVTQKPYSTVKFKLHFGSPTLANGWSQSKEYDNNAILLMRDQKFYLGI FNVRNKPDKQIIKGHEKEEKGDYKKMIYNLLPGPSKMLPKVFITSRSGQE TYKPSKHILDGYNEKRHIKSSPKFDLGYCWDLIDYYKECIHKHPDWKNY DFHFSDTKDYEDISGFYREVEMQGYQIKWTYISADEIQKLDEKGQIFLFQI YNKDFSVHSTGKDNLHTMYLKNLFSEENLKDIVLKLNGEAELFFRKASI KTPVVHKKGSVLVNRSYTQTVGDKEIRVSIPEEYYTEIYNYLNHIGRGKL STEAQRYLEERKIKSFTATKDIVKNYRYCCDHYFLHLPITINFKAKSDIAV NERTLAYIAKKEDIHIIGIDRGERNLLYISVVDVHGNIREQRSFNIVNGYD YQQKLKDREKSRDAARKNWEEIEKIKELKEGYLSMVIHYIAQLVVKYN AVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLHYLVFKDREVCEE GGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFVNLFSFK NLTNRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTMLASTKW KVYTNGTRLKRIVVNGKYTSQSMEVELTDAMEKMLQRAGIEYHDGKD LKGQIVEKGIEAEIIDIFRLTVQMRNSRSESEDREYDRLISPVLNDKGEFFD TATADKTLPQDADANGAYCIALKGLYEVKQIKENWKENEQFPRNKLVQ DNKTWFDFMQKKRYL (SEQ ID NO: 69)
BhCas12b *Bacillus hisashii* Исх. послед. WP_095142515.1	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYEHH EQDPKNPKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRELY EELVPSSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLK IAGDPSWEEEKKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIV KEIKWMEKSRNQSVRRLDKDMFIQALERFLSWESWNLKVKEEYEKVEK EYKTLEERIKEDIQALKALEQYEKERQEQLLRDTLNTNEYRLSKRGLRG WREIIQKWLKMDENEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKE NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFE ERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDI VLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDH LRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFKP KELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQK PDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKAREDNLKL MNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENSDVPLVYQDELIQI RELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLYGIS LKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKED RLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILFEDLSNYNPYEE RSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAKTGS PGIRCSVVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEK FISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQT VYIPESKDQKQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSE LVDSDILKDSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKL ERILISKLTNQYSISTIEDDSSKQSM (SEQ ID NO: 70)
ThCas12b *Thermomonas hydrothermalis* Исх. послед.	MSEKTTQRAYTLRLNRASGECAVCQNNSCDCWHDALWATHKAVNRG AKAFGDWLLTLRGGLCHTLVEMEVPAKGNNPPQRPTDQERRDRRVLLA LSWLSVEDEHGAPKEFIVATGRDSADDRAKKVEEKLREILEKRDFQEHEI DAWLQDCGPSLKAHIREDAVWVNRRALFDAAVERIKTLTWEEAWDFLE PFFGTQYFAGIGDGKDKDDAEGPARQGEKAKDLVQKAGQWLSARFGIG TGADFMSMAEAYEKIAKWASQAQNGDNGKATIEKLACALRPSEPPTLD TVLKCISGPGHKSATREYLKTLDKKSTVTQEDLNQLRKLADEDARNCRK KVGKKGKKPWADEVLKDVENSCELTYLQDNSPARHREFSVMLDHAAR RVSMAHSWIKKAEQRRRQFESDAQKLKNLQERAPSAVEWLDRFCESRS

WP_072754838	MTTGANTGSGYRIRKRAIEGWSYVVQAWAEASCDTEDKRIAAARKVQA DPEIEKFGDIQLFEALAADEAICVWRDQEGTQNPSILIDYVTGKTAEHNQ KRFKVPAYRHPDELRHPVFCDFGNSRWSIQFAIHKEIRDRDKGAKQDTR QLQNRHGLKMRLWNGRSMTDVNLHWSSKRLTADLALDQNPNPNPTEV TRADRLGRAASSAFDHVKIKNVFNEKEWNGRLQAPRAELDRIAKLEEQ GKTEQAEKLRKRLRWYVSFSPCLSPSGPFIVYAGQHNIQPKRSGQYAPH AQANKGRARLAQLILSRLPDLRILSVDLGHRFAAACAVWETLSSDAFRR EIQGLNVLAGGSGEGDLFLHVEMTGDDGKRRTVVYRRIGPDQLLDNTP HPAPWARLDRQFLIKLQGEDEGVREASNEELWTVHKLEVEVGRTVPLID RMVRSGFGKTEKQKERLKKLRELGWISAMPNEPSAETDEKEGEIRSISRS VDELMSSALGTLRLALKRHGNRARIAFAMTADYKPMPGGQKYYFHEA KEASKNDDETKRRDNQIEFLQDALSLWHDLFSSPDWEDNEAKKLWQNH IATLPNYQTPEEISAELKRVERNKKRKENRDKLRTAAKALAENDQLRQH LHDTWKERWESDDQQWKERLRSLKDWIFPRGKAEDNPSIRHVGGLSITR INTISGLYQILKAFKMRPEPDDLRKNIPQKGDDELENFNRRLLEARDRLR EQRVKQLASRIIEAALGVGRIKIPKNGKLPKRPRTTVDTPCHAVVIESLKT YRPDDLRTRRENRQLMQWSSAKVRKYLKEGCELYGLHFLEVPANYTSR QCSRTGLPGIRCDDVPTGDFLKAPWWRRAINTAREKNGGDAKDRFLVD LYDHLNNLQSKGEALPATVRVPRQGGNLFIAGAQLDDTNKERRAIQADL NAAANIGLRALLDPDWRGRWWYVPCKDGTSEPALDRIEGSTAFNDVRS LPTGDNSSRRAPREIENLWRDPSGDSLESGTWSPTRAYWDTVQSRVIELL RRHAGLPTS (SEQ ID NO: 71)
LsCas12b *Laceyella sacchari* WP_132221894.1	MSIRSFKLKLKTKSGVNAEQLRRGLWRTHQLINDGIAYYMNWLVLLRQ EDLFIRNKETNEIEKRSKEEIQAVLLERVHKQQQRNQWSGEVDEQTLLQ ALRQLYEEIVPSVIGKSGNASLKARFFLGPLVDPNNKTTKDVSKSGPTPK WKKMKDAGDPNWVQEYEKYMAERQTLVRLEEMGLIPLFPMYTDEVG DIHWLPQASGYTRTWDRDMFQQAIERLLSWESWNRRVRERRAQFEKKT HDFASRFSESDVQWMNKLREYEAQQEKSLEENAFAPNEPYALTKKALR GWERVYHSWMRLDSAASEEAYWQEVATCQTAMRGEFGDPAIYQFLAQ KENHDIWRGYPERVIDFAELNHLQRELRRAKEDATFTLPDSVDHPLWVR YEAPGGTNIHGYDLVQDTKRNLTLILDKFILPDENGSWHEVKKVPFSLA KSKQFHRQVWLQEEQKQKKREVVFYDYSTNLPHLGTLAGAKLQWDRN FLNKRTQQQIEETGEIGKVFFNISVDVRPAVEVKNGRLQNGLGKALTVLT HPDGTKIVTGWKAEQLEKWVGESGRVSSLGLDSLSEGLRVMSIDLGQRT SATVSVFEITKEAPDNPYKFFYQLEGTEMFAVHQRSFLLALPGENPPQKI KQMREIRWKERNRIKQQVDQLSAILRLHKKVNEDERIQAIDKLLQKVAS WQLNEEIATAWNQALSQLYSKAKENDLQWNQAIKNAHHQLEPVVGKQ ISLWRKDLSTGRQGIAGLSLWSIEELEATKKLLTRWSKRSREPGVVKRIE RFETFAKQIQHHINQVKENRLKQLANLIVMTALGYKYDQEQKKWIEVYP ACQVVLFENLRSYRFSFERSRRENKKLMEWSHRSIPKLVQMQGELFGLQ VADVYAAYSSRYHGRTGAPGIRCHALTEADLRNETNIIHELIEAGFIKEE HRPYLQQGDLVPWSGGELFATLQKPYDNPRILTLHADINAAQNIQKRFW HPSMWFRVNCESVMEGEIVTYVPKNKTVHKKQGKTFRFVKVEGSDVYE WAKWSKNRNKNTFSSITERKPPSSMILFRDPSGTFFKEQEWVEQKTFWG KVQSMIQAYMKKTIVQRMEE (SEQ ID NO: 72)
DtCas12b *Dsulfonatronum*	MVLGRKDDTAELRRALWTTHEHVNLAVAEVERVLLRCRGRSYWTLDR RGDPVHVPESQVAEDALAMAREAQRRNGWPVVGEDEEILLALRYLYEQ IVPSCLLDDLGKPLKGDAQKIGTNYAGPLFDSDTCRRDEGKDVACCGPF HEVAGKYLGALPEWATPISKQEFDGKDASHLRFKATGGDDAFFRVSIEK

thiodismutans *WP_031386437*	ANAWYEDPANQDALKNKAYNKDDWKKEKDKGISSWAVKYIQKQLQL GQDPRTEVRRKLWLELGLLPLFIPVFDKTMVGNLWNRLAVRLALAHLL SWESWNHRAVQDQALARAKRDELAALFLGMEDGFAGLREYELRRNESI KQHAFEPVDRPYVVSGRALRSWTRVREEWLRHGDTQESRKNICNRLQD RLRGKFGDPDVFHWLAEDGQEALWKERDCVTSFSLLNDADGLLEKRKG YALMTFADARLHPRWAMYEAPGGSNLRTYQIRKTENGLWADVVLLSP RNESAAVEEKTFNVRLAPSGQLSNVSFDQIQKGSKMVGRCRYQSANQQ FEGLLGGAEILFDRKRIANEQHGATDLASKPGHVWFKLTLDVRPQAPQG WLDGKGRPALPPEAKHFKTALSNKSKFADQVRPGLRVLSVDLGVRSFA ACSVFELVRGGPDQGTYFPAADGRTVDDPEKLWAKHERSFKITLPGENP SRKEEIARRAAMEELRSLNGDIRRLKAILRLSVLQEDDPRTEHLRLFMEAI VDDPAKSALNAELFKGFGDDRFRSTPDLWKQHCHFFHDKAEKVVAERF SRWRTETRPKSSSWQDWRERRGYAGGKSYWAVTYLEAVRGLILRWNM RGRTYGEVNRQDKKQFGTVASALLHHINQLKEDRIKTGADMIIQAARGF VPRKNGAGWVQVHEPCRLILFEDLARYRFRTDRSRRENSRLMRWSHREI VNEVGMQGELYGLHVDTTEAGFSSRYLASSGAPGVRCRHLVEEDFHDG LPGMHLVGELDWLLPKDKDRTANEARRLLGGMVRPGMLVPWDGGELF ATLNAASQLHVIHADINAAQNLQRRFWGRCGEAIRIVCNQLSVDGSTRY EMAKAPKARLLGALQQLKNGDAPFHLTSIPNSQKPENSYVMTPTNAGK KYRAGPGEKSSGEEDELALDIVEQAEELAQGRKTFFRDPSGVFFAPDRW LPSEIYWSRIRRRIWQVTLERNSSGRQERAEMDEMPY (SEQ ID NO: 73)

[0340] Описанные здесь праймерные редакторы также могут включать варианты Cas12a (Cpf1) (dCpf1), которые могут применяться в качестве домена программируемого направляющей нуклеотидной последовательностью ДНК-связывающего белка. Белок Cas12a (Cpf1) имеет RuvC-подобный эндонуклеазный домен, который аналогичен домену RuvC Cas9, но не имеет эндонуклеазного домена HNH, при этом N-конец Cas12a (Cpf1) не имеет доли распознавания альфа-цепи Cas9. В Zetsche *et al.*, *Cell*, 163, 759–771, 2015 (который включает здесь в качестве ссылки) было показано, что RuvC-подобный домен Cas12a (Cpf1) отвечает за расщепление обеих цепей ДНК, а деактивация RuvC-подобного домена обеспечивает деактивацию нуклеазной активности Cas12a (Cpf1).

[0341] В некоторых вариантах осуществления napDNAbp представляет собой одиночный эффектор микробной системы CRISPR-Cas. Одиночные эффекторы микробных систем CRISPR-Cas включают, без ограничения, Cas9, Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), и Cas12c (C2c3). Как правило, микробные системы CRISPR-Cas разделяются на системы Класса 1 и Класса 2. Системы Класса 1 имеют мультисубъединичные эффекторные комплексы, в то время как системы Класса 2 имеют одиночный белковый эффектор. Например, Cas9 и Cas12a (Cpf1) представляют собой эффекторы Класса 2. Помимо Cas9 и Cas12a (Cpf1), в Shmakov et al., “Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems”, *Mol. Cell*, 2015

Nov 5; 60(3): 385–397, полное содержание которого приведено здесь в качестве ссылки, были описаны три различных системы CRISPR-Cas Класса 2 (Cas12b1, Cas13a, и Cas12c).

[0342] Эффекторы двух систем, Cas12b1 и Cas12c, содержат RuvC-подобные эндонуклеазные домены, являющиеся родственными Cas12a. Третья система, Cas13a, содержит эффектор с двумя предсказанными доменами HEPN РНКазы. Продуцирование зрелой РНК CRISPR не зависит от tracrРНК, в отличие от продуцирования РНК CRISPR за счет Cas12b1. Cas12b1 зависит как от РНК CRISPR, так и tracrРНК для расщепления ДНК. Было показано, что бактериальный Cas13a обладает уникальной РНКазной активностью в плане созревания РНК CRISPR, отличной от его РНК-активируемой связанной с расщеплением одноцепочечной РНК активностью. Данные РНКазные функции отличаются друг от друга и от связанного с процессингом РНК CRISPR поведения Cas12a. См., например, East-Seletsky, et al., "Two distinct RNase activities of CRISPR-Cas13a enable guide-RNA processing and RNA detection", *Nature,* 2016 Oct 13;538(7624):270-273, полное содержание которого приведено здесь в качестве ссылки. *In vitro* биохимический анализ Cas13a в *Leptotrichia shahii* показал, что Cas13a направляется одиночной РНК CRISPR и может быть запрограммирован для расщепления мишеней в виде оцРНК, несущих комплементарные протоспейсеры. Каталитические остатки в двух консервативных доменах HEPN опосредуют расщепление. Мутации в каталитических остатках продуцируют каталитически деактивированные РНК-связывающие белки. См., например, Abudayyeh et al., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector", *Science,* 2016 Aug 5; 353(6299), полное содержание которого приведено здесь в качестве ссылки.

[0343] Сообщалось о кристаллической структуре Cas12b1 (AacC2c1) *Alicyclobaccillus acidoterrastris* в комплексе с химерной направляющей РНК в виде одиночной молекулы (sgРНК). См., например, Liu et al., "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell,* 2017 Jan 19;65(2):310-322, полное содержание которого приведено здесь в качестве ссылки. Также сообщалось о кристаллической структуре в C2c1 *Alicyclobacillus acidoterrestris*, связанном с ДНК-мишенями в виде третичных комплексов. См., например,Yang et al., "PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease", *Cell,* 2016 Dec 15;167(7):1814-1828, полное содержание которого приведено здесь в качестве

ссылки. Каталитически компетентные конформации AacC2c1 как с являющимися, так и не являющимися мишенью цепями ДНК были обнаружены в виде независимо расположенных в одном каталитическом кармане RuvC с C2c1-опосредованным расщеплением, приводящим к зигзагообразному семинуклеотидному разрыву ДНК-мишени. Структурные сравнения третичных комплексов C2c1 и ранее идентифицированных Cas9 и Cpf1 демонстрируют разнообразие механизмов, используемых в рамках систем CRISPR-Cas9.

[0344] В некоторых вариантах осуществления napDNAbp может представлять собой белок C2c1, 2c2, или C2c3. В некоторых вариантах осуществления napDNAbp представляет собой белок C2c1. В некоторых вариантах осуществления napDNAbp представляет собой белок Cas13a. В некоторых вариантах осуществления napDNAbp представляет собой белок Cas12c. В некоторых вариантах осуществления napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 91%, по меньшей мере, на 92%, по меньшей мере, на 93%, по меньшей мере, на 94%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична встречающемуся в природе белку Cas12b1 (C2c1), Cas13a (C2c2), или Cas12c (C2c3). В некоторых вариантах осуществления napDNAbp представляет собой встречающийся в природе белок Cas12b1 (C2c1), Cas13a (C2c2), или Cas12c (C2c3).

H. Циркулярные пермутанты Cas9

[0345] В различных вариантах осуществления описанные здесь праймерные редакторы могут включать циркулярный пермутант Cas9.

[0346] Термин "циркулярно пермутированный Cas9" или "циркулярный пермутант" Cas9 или "CP-Cas9" обозначает любой белок Cas9 или его вариант, который возникает или был сконструирован в виде циркулярного пермутантного варианта, что означает, что N-конец и C-конец белка Cas9 (например белка Cas9 дикого типа) были локально перенесены. Подобные циркулярно пермутированные белки Cas9 или их варианты сохраняют способность связывать ДНК при образовании комплексов с направляющей РНК (нРНК). См. Oakes et al., “Protein Engineering of Cas9 for enhanced function,” *Methods Enzymol*, 2014, 546: 491–511 and Oakes et al., “CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification,” *Cell*, January 10, 2019, 176: 254-267, каждый из которых приведен здесь в качестве ссылки. Настоящее изобретение

охватывает любые ранее известные CP-Cas9 или предусматривает использование нового CP-Cas9 при условии, если полученный циркулярно пермутированный белок сохраняет способность связывать ДНК при образовании комплексов с направляющей РНК (нРНК).

[0347] Любой из описанных здесь белков Cas9, включая любой вариант, ортолог, или встречающийся в природе Cas9 или его эквивалент может быть подвергнут реконфигурированию в виде циркулярного пермутантного варианта.

[0348] В различных вариантах осуществления циркулярные пермутанты Cas9 могут иметь следующую структуру:

N-конец-[исходный C-конец] – [необязательны линкер] – [исходных N-конец]-C-конец.

[0349] В качестве примера, настоящее изобретение включает следующие циркулярные пермутанты стандартного Cas9 *S. pyogenes* (1368 аминокислот UniProtKB - Q99ZW2 (CAS9_STRP1) (нумерация основана на аминокислотном положении в SEQ ID NO: 18)):

N-конец-[1268-1368]-[необязательный линкер]-[1-1267]-C-конец;

N-конец-[1168-1368]-[необязательный линкер]-[1-1167]-C-конец;

N-конец-[1068-1368]-[необязательный линкер]-[1-1067]-C-конец;

N-конец-[968-1368]-[необязательный линкер]-[1-967]-C-конец;

N-конец-[868-1368]-[необязательный линкер]-[1-867]-C-конец;

N-конец-[768-1368]-[необязательный линкер]-[1-767]-C-конец;

N-конец-[668-1368]-[необязательный линкер]-[1-667]-C-конец;

N-конец-[568-1368]-[необязательный линкер]-[1-567]-C-конец;

N-конец-[468-1368]-[необязательный линкер]-[1-467]-C-конец;

N-конец-[368-1368]-[необязательный линкер]-[1-367]-C-конец;

N-конец-[268-1368]-[необязательный линкер]-[1-267]-C-конец;

N-конец-[168-1368]-[необязательный линкер]-[1-167]-C-конец;

N-конец-[68-1368]-[необязательный линкер]-[1-67]-C-конец; или

N-конец-[10-1368]-[необязательный линкер]-[1-9]-C-конец, или соответствующие циркулярные пермутанты других белков Cas9 (включая другие ортологи Cas9, варианты, и так далее).

[0350] В определенных вариантах осуществления циркулярный пермутант Cas9 имеет следующую структуру (на основе Cas9 *S. pyogenes* (1368 аминокислот UniProtKB -

Q99ZW2 (CAS9_STRP1) (нумерация основана на аминокислотном положении в SEQ ID NO: 18):

N-конец-[102-1368]-[необязательный линкер]-[1-101]-C-конец;

N-конец-[1028-1368]-[необязательный линкер]-[1-1027]-C-конец;

N-конец-[1041-1368]-[необязательный линкер]-[1-1043]-C-конец;

N-конец-[1249-1368]-[необязательный линкер]-[1-1248]-C-конец; или

N-конец-[1300-1368]-[необязательный линкер]-[1-1299]-C-конец, или соответствующие циркулярные пермутанты других белков Cas9 (включая другие ортологи Cas9, варианты, и так далее).

[0351] В других вариантах осуществления циркулярный пермутант Cas9 имеет следующую структуру (на основе Cas9 *S. pyogenes* (1368 аминокислот UniProtKB - Q99ZW2 (CAS9_STRP1) (нумерация основана на аминокислотном положении в SEQ ID NO: 18):

N-конец-[103-1368]-[необязательный линкер]-[1-102]-C-конец;

N-конец-[1029-1368]-[необязательный линкер]-[1-1028]-C-конец;

N-конец-[1042-1368]-[необязательный линкер]-[1-1041]-C-конец;

N-конец-[1250-1368]-[необязательный линкер]-[1-1249]-C-конец; или

N-конец-[1301-1368]-[необязательный линкер]-[1-1300]-C-конец, или соответствующие циркулярные пермутанты других белков Cas9 (включая другие ортологи Cas9, варианты, и так далее).

[0352] В некоторых вариантах осуществления циркулярный пермутант может быть образован путем связывания C-концевого фрагмент Cas9 с N-концевым фрагментом Cas9, либо напрямую, или с помощью линкера, например аминокислотного линкера. В некоторых вариантах осуществления C-концевой фрагмент может соответствовать 95% или более C-концевым аминокислотам Cas9 (например приблизительно 1300-1368 аминокислотам) или 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, или 5% или более C-концевым аминокислотам Cas9 (например любой одной из SEQ ID NO: 77-86). N-концевая часть может соответствовать 95% или более N-концевым аминокислотам Cas9 (например приблизительно 1-1300 аминокислотам) или 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, или 5% или более N-концевым аминокислотам Cas9 (например SEQ ID NO: 18).

[0353] В некоторых вариантах осуществления циркулярный пермутант может быть образован путем связывания С-концевого фрагмент Cas9 с N-концевым фрагментом Cas9, либо напрямую, или с помощью линкера, например аминокислотного линкера. В некоторых вариантах осуществления С-концевой фрагмент, который переносится на N-конец, включает или соответствует 30% или менее С-концевым аминокислотам на Cas9 (например аминокислотам 1012-1368 SEQ ID NO: 18). В некоторых вариантах осуществления С-концевой фрагмент, который переносится на N-конец, включает или соответствует 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, или 1% С-концевым аминокислотам Cas9 (например Cas9 согласно SEQ ID NO: 18). В некоторых вариантах осуществления С-концевой фрагмент, который переносится на N-конец, включает или соответствует 410 или менее С-концевым остаткам Cas9 (например Cas9 SEQ ID NO: 18). В некоторых вариантах осуществления С-концевая часть, которая переносится на N-конец, включает или соответствует 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20, или 10 С-концевым остаткам Cas9 (например Cas9 согласно SEQ ID NO: 18). В некоторых вариантах осуществления С-концевая часть, которая переносится на N-конец, включает или соответствует 357, 341, 328, 120, или 69 С-концевым остаткам Cas9 (например Cas9 SEQ ID NO: 18).

[0354] В других вариантах осуществления циркулярные пермутантные варианты Cas9 могут быть определены как топологическое перестроение первичной структуры Cas9 на основе следующего способа, который основан на Cas9 *S. pyogenes* согласно SEQ ID NO: 18: (a) выборе циркулярного пермутантного (CP) участка, соответствующего внутреннему аминокислотному остатку первичной структуры Cas9, который рассекает исходный белок на две половины: N-концевой участок и С-концевой участок; (b) модифицировании последовательности белка Cas9 (например с помощью способов генной инженерии) путем переноса исходного С-концевого участка (включающего аминокислоту CP участка) для расположения перед исходным N-концевым участком, тем самым образуя новый N-конец белка Cas9, который теперь начинается с аминокислотного остатка CP участка. CP участок может быть расположен в любом домене белка Cas9, включая, например, домен спирали II, домен RuvCIII, или домен CTD. Например, CP участок может быть расположен (относительно Cas9 *S. pyogenes*

согласно SEQ ID NO: 18) в исходном аминокислотном остатке 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249, или 1282. Таким образом, после переноса на N-конец исходная аминокислота 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249, или 1282 становится новой N-концевой аминокислотой. Номенклатура данных белков CP-Cas9 может быть обозначена как Cas9-CP^{181}, Cas9-CP^{199}, Cas9-CP^{230}, Cas9-CP^{270}, Cas9-CP^{310}, Cas9-CP^{1010}, Cas9-CP^{1016}, Cas9-CP^{1023}, Cas9-CP^{1029}, Cas9-CP^{1041}, Cas9-CP^{1247}, Cas9-CP^{1249}, и Cas9-CP^{1282}, соответственно. Данное описание не ограничивается получением вариантов CP на основе SEQ ID NO: 18, но может быть реализовано для получения вариантов CP в любой последовательности Cas9, либо в CP участках, которые соответствуют данным положениям, или в полностью других CP участках. Данное описание не является ограничивающим конкретные CP участки каким-либо образом. Практически любой CP участок может быть использован для образования варианта CP-Cas9.

[0355] Ниже приведены примеры аминокислотных последовательностей CP-Cas9 на основе Cas9 согласно SEQ ID NO: 18, в рамках которых последовательности линкера отмечены подчеркиванием, а необязательные остатки метионина (M) указаны жирным шрифтом. Следует понимать, что в рамках изобретения предусмотрены последовательности CP-Cas9, которые не включают последовательность линкера или которые включают различные последовательности линкера. Следует понимать, что последовательности CP-Cas9 могут быть основаны на последовательностях Cas9, отличных от SEQ ID NO: 18, при этом любые представленные здесь примеры не являются ограничивающими. Примеры последовательностей CP-Cas9 являются следующими:

Название CP	Последовательность	SEQ ID NO:
CP1012	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTE ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLAS HYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGDKKYSIG LAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEI	SEQ ID NO: 77

Название CP	Последовательность	SEQ ID NO:
	FSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPI NASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAP LSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSE ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIH DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKY PKLESEFVYG	
CP1028	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGG FSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGG SGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVIT DEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYH LRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN PDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNF KSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV	SEQ ID NO: 78

Название CP	Последовательность	SEQ ID NO:
	VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRI EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEF VYGDYKVYDVRKMIAKSEQ	
CP1041	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSG GSGGDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVL GNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDK KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIA QLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLAR GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVT EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY FKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV	SEQ ID NO: 79

Название CP	Последовательность	SEQ ID NO:
	PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYS	
CP1249	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFK YFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTN SVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGIL QTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKD DSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSK LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALI KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW DKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGS	
CP1300	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTST KEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGS GGSGGSGGSGGDKKYSIGLAIGTNSVGWAVITDEYKV PSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEE	SEQ ID NO: 81

Название CP	Последовательность	SEQ ID NO:
	SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE EFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIP YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGA SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNEL TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNR GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGD YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEI TLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLAS HYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRD	

[0356] Здесь описаны циркулярные пермутанты Cas9, которые могут применяться в рамках конструктов для праймерного редактирования. Примеры С-концевых фрагментов Cas9 на основе Cas9 согласно SEQ ID NO: 18, которые могут быть перенесены на N-конец Cas9, приведены ниже. Следует понимать, то подобные С-концевые фрагменты Cas9 представляют собой примеры и не являются ограничивающими. Данные примеры фрагментов CP-Cas9 имеют следующие последовательности:

Название CP	Последовательность	SEQ ID NO:
С-концевой фрагмент	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL	SEQ ID NO: 82

Название CP	Последовательность	SEQ ID NO:
CP1012	SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGL YETRIDLSQLGGD	
C-концевой фрагмент CP1028	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGG FSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 83
C-концевой фрагмент CP1041	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 84
C-концевой фрагмент CP1249	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFK YFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD	SEQ ID NO: 85
C-концевой фрагмент CP1300	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTST KEVLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 86

I. Варианты Cas9 с модифицированной специфичностью в отношении PAM

[0357] Праймерные редакторы по настоящему изобретению также могут включать варианты Cas9 с модифицированной специфичностью в отношении PAM. Некоторые аспекты данного изобретения предусматривают белки Cas9, которые демонстрируют активность на последовательности-мишени, которая не включает стандартный PAM (5′-NGG-3′, где N представляет собой A, C, G, или T) на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NGG-3′ на ее

3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NNG-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NNA-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NNC-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NNT-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NGT-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NGA-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NGC-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NAA-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NAC-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NAT-3′ на ее 3′-конце. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, включающей последовательность PAM 5′-NAG-3′на ее 3′-конце.

[0358] Следует понимать, что любые описанные здесь аминокислотные мутации (например A262T) из первого аминокислотного остатка (например A) во второй аминокислотный остаток (например T) также могут включать мутации из первого аминокислотного остатка в аминокислотный остаток, который аналогичен второму аминокислотному остатку (например является консервативным). Например, мутация аминокислоты с гидрофобной боковой цепью (например аланина, валина, изолейцина, лейцина, метионина, фенилаланина, тирозина, или триптофана) может представлять собой мутацию второй аминокислоты с другой гидрофобной боковой цепью (например

аланина, валина, изолейцина, лейцина, метионина, фенилаланина, тирозина, или триптофана). Например, мутация аланина на треонин (например мутация A262T) также может представлять собой мутацию из аланина в аминокислоту, которая имеет размер и химические свойства, аналогичные треонину, например серину. В качестве другого примера мутация аминокислоты с положительно заряженной боковой цепью (например аргинина, гистидина или лизина) может представлять собой мутацию во вторую аминокислоту с другой положительно заряженной боковой цепью (например аргинин, гистидин, или лизин). В качестве другого примера мутация аминокислоты с полярной боковой цепью (например серина. треонина, аспарагина, или глутамина) может представлять собой мутацию во вторую аминокислоту с другой полярной боковой цепью (например серин, треонин, аспарагин, или глутамин). Другие аналогичные аминокислотные пары включают, но без ограничения, следующие: фенилаланин и тирозин; аспарагин и глутамин; метионин и цистеин; аспарагиновую кислоту и глутаминовую кислоту; и аргинин и лизин. Специалист в данной области определит, что подобные консервативные аминокислотные замещения скорее всего будут иметь минимальное влияние не структуру белка и скорее всего будут хорошо переноситься без ухудшения функции. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в треонин могут представлять собой аминокислотную мутацию в серин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в аргинин могут представлять собой аминокислотную мутацию в лизин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в изолейцин могут представлять собой аминокислотную мутацию в аланин, валин, метионин, или лейцин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в лизин могут представлять собой аминокислотную мутацию в аргинин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в аспарагиновую кислоту могут представлять собой аминокислотную мутацию в глутаминовую кислоту или аспарагин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в валин могут представлять собой аминокислотную мутацию в аланин, изолейцин, метионин, или лейцин. В некоторых вариантах осуществления любые указанные здесь аминокислотные мутации из одной аминокислоты в глицин могут

представлять собой аминокислотную мутацию в аланин. Следует понимать, однако, что дополнительные консервативные аминокислотные остатки будут очевидны специалисту в данной области и любые аминокислотные мутации в другие консервативные аминокислотные остатки также входят в объем данного изобретения.

[0359] В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций, которые демонстрируют активность на последовательности-мишени, включающей последовательность PAM 5′-NAA-3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 1. В некоторых вариантах осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 1. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 1.

[0360] Таблица 1: Клоны PAM NAA

Мутации из SpCas9 дикого типа (например SEQ ID NO: 18)
D177N, K218R, D614N, D1135N, P1137S, E1219V, A1320V, A1323D, R1333K
D177N, K218R, D614N, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, G715C, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A367T, K710E, R1114G, D1135N, P1137S, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D861N, D1135N, K1188R, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, V743I, R753G, E762G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, S1274R, A1320V, R1333K
A10T, I322V, S409I, E427G, A589S, R753G, D1135N, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, E757K, G865G, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, R753G, E757K, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, K599R, M631A, R654L, K673E, V743I, R753G, N758H, E762G, D1135N, D1180G, E1219V, Q1221H, Q1256R, H1264Y, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N869S, N1054D, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, L727I, V743I, R753G, E762G, R859S, N946D, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, N1317T, A1320V, A1323D,

R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, K1151E, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1080S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, E630K, R654L, K673E, V743I, R753G, E762G, Q768H, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, Q768H, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, G1223S, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1801S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, M673I, N803S, N869S, G1077D, R1114G, D1135N, V1139A, D1180G, E1219V, Q1221H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, R1114G, D1135N, E1219V, Q1221H, A1320V, R1333K

[0361] В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 80% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 1. В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 92%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 1.

[0362] В некоторых вариантах осуществления белок Cas9 демонстрирует повышенную активность на последовательности-мишени, которая не включает стандартный PAM (5′-NGG-3′) на ее 3′-конце в отличие от Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, имеющей 3′-конец, которая не является

непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 5-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 10-кратно, по меньшей мере, 50-кратно, по меньшей мере, 100-кратно, по меньшей мере, 500-кратно, по меньшей мере, 1000-кратно, по меньшей мере, 5000-кратно, по меньшей мере, 10000-кратно, по меньшей мере, 50000-кратно, по меньшей мере, 100000-кратно, по меньшей мере, 500000-кратно, или, по меньшей мере, 1000000-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 2 на той же последовательности-мишени. В некоторых вариантах осуществления 3′-конец последовательности-мишени является непосредственно смежным с последовательностью AAA, GAA, CAA, или TAA. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций, которые демонстрируют активность на последовательности-мишени, включающей последовательность PAM 5′-NAC-3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 2. В некоторых вариантах осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 2. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 2.

[0363] Таблица 2: Клоны PAM NAC

МУТАЦИИ ИЗ SPCAS9 ДИКОГО ТИПА (НАПРИМЕР SEQ ID NO: 18)
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
D1135N, E1219V, D1332N, R1335Q, T1337N
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
T472I, R753G, Q771H, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, K1156E, E1219V, D1332N, R1335Q, T1337N

E627K, T638P, V647I, R753G, N803S, K959N, G1030R, I1055E, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, E630G, T638P, V647A, G687R, N767D, N803S, K959N, R1114G, D1135N, E1219V, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, I1057T, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, M631I, T638P, R753G, N803S, K959N, Y1036H, R1114G, D1135N, E1219V, D1251G, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, V875I, K959N, Y1016C, R1114G, D1135N, E1219V, D1251G, D1332G, R1335Q, T1337N, I1348V
K608R, E627K, T638P, V647I, R654L, R753G, N803S, T804A, K848N, V922A, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R753G, N803S, V922A, K959N, K1014N, V1015A, R1114G, D1135N, K1156N, E1219V, N1252D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, A711T, R753G, K775R, K789E, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
I670S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, K797N, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, G752R, R753G, K797N, N803S, K948E, K959N, V1015A, Y1016S, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
I570T, A589V, K608R, E627K, T638P, V647I, R654L, Q716R, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1207G, E1219V, N1234D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, R654L, Q740R, R753G, N803S, K959N, N990S, T995S, V1015A, Y1036D, R1114G, D1135N, E1207G, E1219V, N1234D, N1266H, D1332N, R1335Q, T1337N
I562F, V565D, I570T, K608R, L625S, E627K, T638P, V647I, R654I, G752R, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, N1177S, N1234D, D1332N, R1335Q, T1337N
I562F, I570T, K608R, E627K, T638P, V647I, R753G, E790A, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, D1180E, A1184T, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1015A, R1114G, D1127A, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, L625S, E627K, T638P, V647I, R654I, T703P, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1016A, R1114G, D1135N, E1219V, K1246E, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R654L, K673E, R753G, E790A, N803S, K948E, K959N, R1114G, D1127G, D1135N, D1180E, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, L625S, E627K, T638P, V647I, R654I, I670T, R753G, N803S, N808D, K959N,

M1021L, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
E627K, M631V, T638P, V647I, K710E, R753G, N803S, N808D, K948E, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N, S1338T, H1349R

[0364] В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 80% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 2. В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 92%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 2.

[0365] В некоторых вариантах осуществления белок Cas9 демонстрирует повышенную активность на последовательности-мишени, которая не включает стандартный PAM (5′-NGG-3′) на ее 3′-конце в отличие от Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, имеющей 3′-конец, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 5-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 10-кратно, по меньшей мере, 50-кратно, по меньшей мере, 100-кратно, по меньшей мере, 500-кратно, по меньшей мере, 1000-кратно, по меньшей мере, 5000-кратно, по меньшей мере, 10000-кратно, по меньшей мере, 50000-кратно, по меньшей мере, 100000-кратно, по меньшей мере, 500000-кратно, или, по меньшей мере, 1000000-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления 3′-конец последовательности-мишени является непосредственно смежным с последовательностью AAC, GAC, CAC, или TAC.

[0366] В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций, которые демонстрируют активность на последовательности-мишени,

включающей последовательность PAM 5′-NAT-3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 3. В некоторых вариантах осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 3. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 3.

[0367] Таблица 3: Клоны PAM NAT

МУТАЦИИ ИЗ SPCAS9 ДИКОГО ТИПА (НАПРИМЕР SEQ ID NO: 18)
K961E, H985Y, D1135N, K1191N, E1219V, Q1221H, A1320A, P1321S, R1335L
D1135N, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
V743I, R753G, E790A, D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, V748I, V743I, R753G, D853E, V922A, R1114G D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1286K, P1321S, D1322G, R1335L
M631L, R654L, R753G, K797E, D853E, V922A, D1012A, R1114G D1135N, G1218S, E1219V, Q1221H, P1249S, N1317K, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, D596Y, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, Q1256R, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, K710E, V750A, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, K649R, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, K1156E, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, I1057G, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1308D, P1321S, D1322G, R1335L
M631L, R654L, R753G, D853E, V922A, R1114G, Y1131C, D1135N, E1150V, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, D853E, I1057V, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, I1057V, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L

[0368] Приведенное выше описание различных napDNAbp, которые могут применяться совместно с описываемыми здесь праймерными редакторами, не является ограничивающим каким-либо образом. Праймерные редакторы могут включать стандартный SpCas9 или любой ортолог белка Cas9, или любой вариант белка Cas9,

включая любой встречающийся в природе вариант, мутант, или иным образом модифицированную версию Cas9, которая известна или которая может быть получена или может быть подвергнута эволюции с помощью связанного с направленной эволюцией или иного мутагенезного процесса. В различных вариантах осуществления Cas9 или варианты Cas9 обладают никазной активностью, то есть только расщепляют цепь последовательности-мишени ДНК. В других вариантах осуществления Cas9 или варианты Cas9 имеют деактивированные нуклеазы, то есть представляют собой белки в виде "мертвого" Cas9. Другие белки в виде вариантов Cas9, которые могут применяться, представляют собой белки с более низкой молекулярной массой, в отличие от стандартного SpCas9 (например для более легкой доставки) или имеющие модифицированную или перестроенную первичную аминокислотную структуру (например циркулярные пермутантные форматы). Описанные здесь праймерные редакторы также могут включать эквиваленты Cas9, включая белки Cas12a/Cpf1 и Cas12b, которые представляют собой результат конвергентной эволюции. Используемые здесь napDNAbp (например SpCas9, вариант Cas9, или эквиваленты Cas9) могут также содержать различные модификации, которые изменяют/повышают их специфичность в отношении PAM. Наконец, изобретение охватывает любой Cas9, вариант Cas9, или эквивалент Cas9, который имеет, по меньшей мере, 70%, по меньшей мере, 75%, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 91%, по меньшей мере, 92%, по меньшей мере, 93%, по меньшей мере, 94%, по меньшей мере, 95%, по меньшей мере, 96%, по меньшей мере, 97%, по меньшей мере, 98%, по меньшей мере, 99%, или, по меньшей мере, 99,9% идентичность последовательности относительно исходной последовательности Cas9, например исходной стандартной последовательности SpCas9 или исходного эквивалента Cas9 (например Cas12a (Cpf1).

[0369] В определенном варианте осуществления вариант Cas9, имеющий расширенные возможности в отношении PAM, представляет собой SpCas9 (H840A) VRQR (SEQ ID NO: 87), который имеет следующую аминокислотную последовательность (при этом замещения V, R, Q, R, относящиеся к SpCas9 (H840A) согласно SEQ ID NO: 51, отмечены жирным шрифтом и подчеркнуты). Кроме того, остаток метионина в SpCas9 (H840A) был удален для SpCas9 (H840A) VRQR:

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI

FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL
AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT
KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISG
VEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH
LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK
PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL
YYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDN
VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR
QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS
NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT
EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVVAKVEKGKSK
KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIE
QISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTI
DRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 87)

[0370] В другом определенном варианте осуществления вариант Cas9, имеющий расширенные возможности в отношении PAM, представляет собой SpCas9 (H840A) VRER, который имеет следующую аминокислотную последовательность (при этом замещения V, R, E, R, относящиеся к SpCas9 (H840A) согласно SEQ ID NO: 51, отмечены жирным шрифтом и подчеркнуты). Кроме того, остаток метионина в SpCas9 (H840A) был удален для SpCas9 (H840A) VRER:

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL
AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR
KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT
KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISG
VEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH
LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK
PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL
YYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDN
VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR
QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHH
AHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS

NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT
EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVVAKVEKGKSK
KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML
ASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIE
QISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTI
DRK**E**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 88)

[0371] В некоторых вариантах осуществления napDNAbp, который функционирует с нестандартной последовательностью PAM, представляет собой белок Argonaute. Один пример такого программируемого нуклеиновой кислотой ДНК-связывающего белка представляет собой белок Argonaute из *Natronobacterium gregoryi* (NgAgo). NgAgo представляет собой оцДНК-направляемую эндонуклеазу. NgAgo связывает 5' фосфорилированную оцДНК из ~24 нуклеотидов (нДНК) для ее направления к ее участку-мишени и осуществляет двуцепочечные разрывы ДНК в участке нДНК. В отличие от Cas9, система NgAgo–нДНК не требует смежного с протоспейсером мотива (PAM). Использование NgAgo без нуклеазной активности (dNgAgo) может значительно расширить перечень оснований, на которое может осуществляться нацеливание. Характеристики и применение NgAgo было описано в Gao *et al.*, *Nat Biotechnol.*, 2016 Jul;34(7):768-73. PubMed PMID: 27136078; Swarts *et al.*, *Nature*. 507(7491) (2014):258-61; и Swarts *et al.*, *Nucleic Acids Res.* 43(10) (2015):5120-9, каждый из которых приведен здесь в качестве ссылки.

[0372] В некоторых вариантах осуществления napDNAbp представляет собой прокариотический гомолог белка Argonaute. Прокариотические гомологи белков Argonaute известны и были описаны, например, в Makarova K., et al., "Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements", *Biol Direct.* 2009 Aug 25;4:29. doi: 10.1186/1745-6150-4-29, полное содержание которого приведено здесь в качестве ссылки. В некоторых вариантах осуществления napDNAbp представляет собой белок Argunaute Marinitoga piezophila (MpAgo). Ассоциированный с CRISPR белок Argunaute Marinitoga piezophila (MpAgo) расщепляет одноцепочечные последовательности-мишени с помощью 5′-фосфорилированных направляющих. 5′ направляющие используются всеми известными Argonaute. Кристаллическая структура комплекса MpAgo-РНК демонстрирует участок связывания направляющей цепи, включающий остатки, которые блокируют 5′ фосфатные взаимодействия. Эти данные свидетельствуют об эволюции подкласса Argonaute с нестандартной специфичностью в отношении 5′-гидроксилированной направляющей. См., например, Kaya et al., "A

bacterial Argonaute with noncanonical guide RNA specificity", *Proc Natl Acad Sci U S A*. 2016 Apr 12;113(15):4057-62, полное содержание которого приведено здесь в качестве ссылки). Является очевидным, что могут использоваться другие белки Argonaute, при этом они находятся в рамках объема данного изобретения

[0373] Некоторые аспекты изобретения предусматривают домены Cas9, которые имеют различную специфичность в отношении PAM. Как правило, белки Cas9, такие как Cas9 из *S. pyogenes* (spCas9), требуют стандартную последовательность PAM NGG для связывания с определенным участком нуклеиновой кислоты. Это может ограничивать возможность редактировать требуемые основания в рамках генома. В некоторых вариантах осуществления может требоваться расположение предусмотренных здесь слитых белков для редактирования оснований в точном месторасположении, например где основание-мишень располагается в рамках участка длиной 4 основания (например "окна редактирования"), что приблизительно соответствует 15 основаниям выше PAM. См. Komor, A.C., *et al.*, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016), полное содержание которого приведено здесь в качестве ссылки. Соответственно, в некоторых вариантах осуществления любой из предусмотренных здесь слитых белков может включать домен Cas9, который способен связывать нуклеотидную последовательность, которая не содержит стандартной (например NGG) последовательности PAM. Домены Cas9, которые связываются с нестандартными последовательностями PAM, были описаны в данной области и будут очевидны специалисту в данной области. Например, домены Cas9, которые связываются с нестандартными последовательностями PAM, были описаны в Kleinstiver, B. P., *et al.*, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" *Nature* 523, 481-485 (2015); и Kleinstiver, B. P., *et al.*, "Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition" *Nature Biotechnology* 33, 1293-1298 (2015); содержание каждого из которых приведено здесь в качестве ссылки.

[0374] Например, домен napDNAbp с измененной специфичностью в отношении PAM, например домен, имеющий, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно Cpf1 *Francisella novicida* дикого типа (D917, E1006, and D1255) (SEQ ID NO: 74), который имеет следующую аминокислотную последовательность:

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVEN
QGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNL
QDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTE
DKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGN
IIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHE
IAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT
GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQ
EFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNW
DTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSK
TGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQ
EGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 74)

[0375] Дополнительный домен napDNAbp с измененной специфичностью в отношении PAM, например домен, имеющий, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно Cas9 *Geobacillus thermodenitrificans* дикого типа (SEQ ID NO: 75), который имеет следующую аминокислотную последовательность:

MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLARSARRRL
RRRKHRLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEALDRKLNNDELARIL
LHLAKRRGFRSNRKSERTNKENSTMLKHIEENQSILSSYRTVAEMVVKDPKFSLHKR
NKEDNYTNTVARDDLEREIKLIFAKQREYGNIVCTEAFEHEYISIWASQRPFASKDDIE
KKVGFCTFEPKEKRAPKATYTFQSFTVWEHINKLRLVSPGGIRALTDDERRLIYKQAF
HKNKITFHDVRTLLNLPDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVY
GKGAAKSFRPIDFDTFGYALTMFKDDTDIRSYLRNEYEQNGKRMENLADKVYDEELI
EELLNLSFSKFGHLSLKALRNILPYMEQGEVYSTACERAGYTFTGPKKKQKTVLLPNI
PPIANPVVMRALTQARKVVNAIIKKYGSPVSIHIELARELSQSFDERRKMQKEQEGNR
KKNETAIRQLVEYGLTLNPTGLDIVKFKLWSEQNGKCAYSLQPIEIERLLEPGYTEVDH
VIPYSRSLDDSYTNKVLVLTKENREKGNRTPAEYLGLGSERWQQFETFVLTNKQFSKK
KRDRLLRLHYDENEENEFKNRNLNDTRYISRFLANFIREHLKFADSDDKQKVYTVNG
RITAHLRSRWNFNKNREESNLHHAVDAAIVACTTPSDIARVTAFYQRREQNKELSKKT
DPQFPQPWPHFADELQARLSKNPKESIKALNLGNYDNEKLESLQPVFVSRMPKRSITG
AAHQETLRRYIGIDERSGKIQTVVKKKLSEIQLDKTGHFPMYGKESDPRTYEAIRQRL
LEHNNDPKKAFQEPLYKPKKNGELGPIIRTIKIIDTTNQVIPLNDGKTVAYNSNIVRVD
VFEKDGKYYCVPIYTIDMMKGILPNKAIEPNKPYSEWKEMTEDYTFRFSLYPNDLIRI
EFPREKTIKTAVGEEIKIKDLFAYYQTIDSSNGGLSLVSHDNNFSLRSIGSRTLKRFEKY

QVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 75)

[0376] В некоторых вариантах осуществления программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) представляет собой программируемый нуклеиновой кислотой ДНК-связывающий белок, который не требует стандартной (NGG) последовательности PAM. В некоторых вариантах осуществления napDNAbp представляет собой белок Argonaute. Один пример такого программируемого нуклеиновой кислотой ДНК-связывающего белка представляет собой белок Argonaute из *Natronobacterium gregoryi* (NgAgo). NgAgo представляет собой оцДНК-направляемую эндонуклеазу. NgAgo связывает 5' фосфорилированную оцДНК из ~24 нуклеотидов (нДНК) для ее направления к ее участку-мишени и осуществляет двуцепочечные разрывы ДНК в участке нДНК. В отличие от Cas9, система NgAgo–нДНК не требует смежного с протоспейсером мотива (PAM). Использование NgAgo без нуклеазной активности (dNgAgo) может значительно расширить перечень оснований, на которое может осуществляться нацеливание. Характеристики и применение NgAgo было описано в Gao *et al.*, *Nat Biotechnol.*, 34(7): 768-73 (2016), PubMed PMID: 27136078; Swarts *et al.*, *Nature*, 507(7491): 258-61 (2014); и Swarts *et al.*, *Nucleic Acids Res.* 43(10) (2015): 5120-9, каждый из которых приведен здесь в качестве ссылки. Последовательность Argonaute *Natronobacterium gregoryi* приведена в SEQ ID NO: 76.

[0377] Описанные слиты белки могут включать домен napDNAbp, имеющий, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, или, по меньшей мере, 99% идентичность последовательности относительно Argonaute *Natronobacterium gregoryi* дикого типа (SEQ ID NO: 76), который имеет следующую аминокислотную последовательность:

MTVIDLDSTTTADELTSGHTYDISVTLTGVYDNTDEQHPRMSLAFEQDNGERRYITLW
KNTTPKDVFTYDYATGSTYIFTNIDYEVKDGYENLTATYQTTVENATAQEVGTTDEDE
TFAGGEPLDHHLDDALNETPDDAETESDSGHVMTSFASRDQLPEWTLHTYTLTATDG
AKTDTEYARRTLAYTVRQELYTDHDAAPVATDGLMLLTPEPLGETPLDLDCGVRVEA
DETRTLDYTTAKDRLLARELVEEGLKRSLWDDYLVRGIDEVLSKEPVLTCDEFDLHER
YDLSVEVGHSGRAYLHINFRHRFVPKLTLADIDDDNIYPGLRVKTTYRPRRGHIVWG
LRDECATDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVETRRQGHGDDAV
SFPQELLAVEPNTHQIKQFASDGFHQQARSKTRLSASRCSEKAQAFAERLDPVRLNGS
TVEFSSEFFTGNNEQQLRLLYENGESVLTFRDGARGAHPDETFSKGIVNPPESFEVAVV
LPEQQADTCKAQWDTMADLLNQAGAPPTRSETVQYDAFSSPESISLNVAGAIDPSEV
DAAFVVLPPDQEGFADLASPTETYDELKKALANMGIYSQMAYFDRFRDAKIFYTRNV
ALGLLAAAGGVAFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINIAATATAVYKDG
TILGHSSTRPQLGEKLQSTDVRDIMKNAILGYQQVTGESPTHIVIHRDGFMNEDLDPA

TEFLNEQGVEYDIVEIRKQPQTRLLAVSDVQYDTPVKSIAAINQNEPRATVATFGAPEY LATRDGGGLPRPIQIERVAGETDIETLTRQVYLLSQSHIQVHNSTARLPITTAYADQAST HATKGYLVQTGAFESNVGFL (SEQ ID NO: 76)

[0378] твии с используемым здесь значением термин "мутация" обозначает замещение остатка в последовательности, например нуклеиновой кислоты или аминокислотной последовательности, другим остатком, или делецию или вставку одного или нескольких остатков в последовательности. Мутации обычно описаны здесь на основе указания исходного остатка и далее положения остатка в последовательности и на основе указания нового замещенного остатка. Различные способы осуществления указанных здесь аминокислотных замещений (мутаций) хорошо известны в данной области и описаны, например, в Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)). Мутации могут включать различные категории, например полиморфизмы одиночных оснований, участки микродупликации, индели и инверсии, и не являются ограниченными каким-либо образом. Мутации могут включать мутации типа "потеря функции", что является нормальным результатом мутации, которая снижает или устраняет активность белка. Большинство мутаций потери функции являются рецессивными, поскольку в гетерозиготе вторая копия хромосомы несет не подверженную мутации версию гена, кодирующего полностью функциональный пептид, чье присутствие компенсирует эффект мутации.Мутации также охватывают мутации типа "приобретение функции", которые придают нехарактерную активность белку или клетки, которая иным образом не присутствует при нормальных условиях. Многие мутации, связанные с приобретением функции, располагаются в регуляторных последовательностях, а не в кодирующих участках, и, таким образом, имеют ряд последствий.Например, мутация может приводить к тому, что один или несколько генов экспрессируются в неправильных тканях, при этом эти ткани приобретают функции, которые они обычно не имеют.Вследствие их природы, мутации, связанные с приобретением функции, обычно являются доминантными.

[0379] Мутации могут быть внесены в исходный белок Cas9 с помощью сайт-направленного мутагенеза. Старые известные в данной области способы сайт-направленного мутагенеза основаны на субклонировании подвергаемой мутированию последовательности в вектор, например вектор на основе бактериофага M13, который обеспечивает изолирование одноцепочечной матрицы ДНК. В рамках данных способов осуществляется отжиг мутагенного праймера (то есть праймера, способного

отжигаться на подвергаемом мутированию участке, но несущего один или несколько несовпадающих нуклеотидов в подвергаемом мутированию участке) на одноцепочечной матрице и затем полимеризация комплемента матрицы начиная с 3′-конца мутагенного праймера. Получаемые в результате этого дуплексы затем переносятся в бактерии-хозяева и осуществляется скрининг бляшек на предмет требуемой мутации. Недавно в рамках сайт-направленного мутагенеза стали использоваться способы ПЦР, которые имеют преимущество, заключающееся в отсутствии потребности в одноцепочечной матрице. Кроме того, были разработаны способы, не требующие субклонирования. При осуществлении сайт-направленного мутагенеза на основе ПЦР необходимо учитывать несколько нюансов. Во-первых, в рамках этих способов является желательным снижать количество циклов ПЦР для предотвращения экспансии нежелательных мутаций, вносимых полимеразой. Во-вторых, должен осуществляться отбор для снижения количества не подвергнутых мутированию исходных молекул, остающихся в реакции. В-третьих, ПЦР способ с повышенной длиной является предпочтительным для обеспечения возможности использования одного набора праймеров ПЦР. И в-четвертых, вследствие активности не зависящего от матрицы концевого удлинения некоторых термостабильных полимераз зачастую является необходимым включать конечный финишный этап в процедуру перед лигированием тупого конца полученного с помощью ПЦР мутантного продукта.

[0380] Мутации также могут быть внесены с помощью процессов направленной эволюции, например непрерывной эволюции с помощью фагов (PACE) или прерывистой эволюции с помощью фагов (PANCE). В соответствии с используемым здесь значением термин "непрерывная эволюция с помощью фагов (PACE)" обозначает непрерывную эволюцию, в рамках которой используется фаг в качестве вирусных векторов. Общая концепция технологии PACE была описана, например, в международной PCT заявке PCT/US2009/056194, поданной 8 сентября 2009 г., опубликованной как WO 2010/028347 11 марта 2010 г.; международной PCT заявке PCT/US2011/066747, поданной 22 декабря 2011 г., опубликованной как WO 2012/088381 28 июня 2012 г.; заявке на патент США, патенте США 9023594, выданном 5 мая 2015 г., международной PCT заявке PCT/US2015/012022, поданной 20 января 2015 г., опубликованной как WO 2015/134121 11 сентября 2015 г., и международной PCT заявке PCT/US2016/027795, поданной 15 апреля 2016 г., опубликованной как WO

2016/168631 20 октября 2016 г., полное содержание каждой из которых приведено здесь в качестве ссылки. Варианты Cas9 также могут быть получены с помощью "прерывистой эволюции с помощью фагов (PANCE)", которая в соответствии с используемым здесь значением обозначает прерывистую эволюцию, в рамках которой используется фаг в качестве вирусных векторов. PANCE представляет собой упрощенный способ быстрой *in vivo* направленной эволюции с применением последовательных переносов во флаконах подвергаемого эволюции "фага отбора" (SP), который содержит представляющий интерес подвергаемый эволюции ген, между свежими клетками-хозяевами *E. coli*, тем самым сохраняя гены внутри *E. coli* неизменными, в то время как гены, содержащиеся в SP, непрерывно подвергаются эволюции. Последовательные переносы во флаконах долго служили в качестве широко доступного подхода для лабораторной эволюции микробов, при этом в недавнее время были разработаны аналогичные подходы для эволюции бактериофагов. Системе PANCE присуща более низкая строгость, чем системе PACE.

[0381] Любые из приведенных выше ссылок, которые относятся к Cas9 или эквивалентам Cas9, полностью приведены здесь в качестве ссылки, если не указано иное.

J. Разделенные домены napDNAbp для доставки разделенного PE

[0382] В различных вариантах осуществления описанные здесь праймерные редакторы могут быть доставлены в клетки в виде двух или более фрагментов, которые становятся собранными внутри клетки (либо за счет пассивной сборки, или за счет активной сборки, например использования последовательностей разделенного интеина) с образованием восстановленного праймерного редактора. В некоторых случаях самосборка может являться пассивной, когда два или более фрагмента праймерного редактора ковалентно или нековалентно связываются внутри клетки для восстановления праймерного редактора. В других случаях самосборка может катализироваться доменами димеризации, располагаемыми на каждом из фрагментов. Примеры доменов димеризации описаны здесь. В других случаях самосборка может катализироваться последовательностями разделенного интеина, располагаемыми на каждом из фрагментов праймерного редактора.

[0383] Доставка разделенного PE может быть полезной для преодоления различных связанных с размером ограничений разных подходов к доставке. Например, подходы к доставке могут включать способы доставки на основе вирусов, способы доставки на

основе информационной РНК, или доставку на основе RNP (доставку на основе рибонуклеопротеина). При этом каждый из данных способов доставки может быть более эффективным и/или действенным за счет разделения праймерного редактора на части более малого размера. При попадании внутрь клетки части малого размера могут собираться в функциональный праймерный редактор. В зависимости от способа разделения разделенные фрагменты праймерного редактора могут быть собраны обратно нековалентным способом или ковалентным способом для восстановления праймерного редактора. В одном варианте осуществления праймерный редактор может быть разделен в одном или нескольких участках разделения на два или более фрагмента. Фрагменты могут являться немодифицированными (не считая того, что они являются разделенными). После доставки фрагментов в клетку (например с помощью прямой доставки рибонуклеопротеинового комплекса или с помощью ядерной доставки – например доставки с помощью иРНК или доставки на основе вирусного вектора) фрагменты могут ковалентно или нековалентно связываться для восстановления праймерного редактора. В другом варианте осуществления праймерный редактор может быть разделен в одном или нескольких участках разделения на два или более фрагмента. Каждый из фрагментом может быть модифицирован для включения домена димеризации, при этом каждый образованный фрагмент соединяется с доменом димеризации. При доставке или экспрессии в клетке домены димеризации различных фрагментов ассоциируются и связываются друг с другом, объединяя различные фрагменты праймерного редактора вместе для восстановления функционального праймерного редактора. В другом варианте осуществления фрагмент праймерного редактора может быть модифицирован для включения разделенного интеина. При доставке или экспрессии в клетке домены разделенного интеина различных фрагментов ассоциируются и связываются друг с другом и затем подвергаются транс-сплайсингу, который приводит к эксцизии доменов разделенного интеина из каждого из фрагментов и сопутствующему образованию пептидной связи между фрагментами, тем самым обеспечивая восстановление праймерного редактора.

[0384] В одном варианте осуществления праймерный редактор может быть доставлен с применением похода на основе разделенного интеина.

[0385] Участок разделения может быть расположен между любой одной или несколькими парами остатков в праймерном редакторе и в любом из доменов в нем,

включая домен napDNAbp, полимеразный домен (например домен RT), линкерный домен, который соединяет домен napDNAbp и полимеразный домен.

[0386] В одном варианте осуществления, изображенном на Фиг.66, праймерный редактор (PE) разделен в участке разделения в napDNAbp.

[0387] В определенных вариантах осуществления napDNAbp представляет собой стандартный полипептид SpCas9 согласно SEQ ID NO: 18, а именно:

SpCas9 *Streptococcus pyogenes* M1 SwissProt кат. # Q99ZW2 Дикий тип 1368 AA	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQL PGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 18

[0388] В определенных вариантах осуществления SpCas9 разделен на два фрагмента в участке разделения, расположенном между остатками 1 и 2, или 2 и 3, или 3 и 4, или 4 и 5, или 5 и 6, или 6 и 7, или 7 и 8, или 8 и 9, или 9 и 10, или между любыми двумя

парами остатков, расположенными между остатками 1-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 стандартного SpCas9 согласно SEQ ID NO: 18.

[0389] В определенных вариантах осуществления napDNAbp разделен на два фрагмента в участке разделения, расположенном в паре остатков, которые соответствуют любым двум парам остатков, расположенным где-либо между положениями 1-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 стандартного SpCas9 согласно SEQ ID NO: 18.

[0390] В определенных вариантах осуществления SpCas9 разделен на два фрагмента в участке разделения, расположенном между остатками 1 и 2, или 2 и 3, или 3 и 4, или 4 и 5, или 5 и 6, или 6 и 7, или 7 и 8, или 8 и 9, или 9 и 10, или между любыми двумя парами остатков, расположенными между остатками 1-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 стандартного SpCas9 согласно SEQ ID NO: 18. В определенных вариантах осуществления участок разделения расположен в одном или нескольких участках полипептидных связей (то есть "участок разделения или участок разделения разделенного интеина), слит с разделенным интеином и затем доставляется в клетки в виде отдельно кодируемых слитых белков. При экспрессии слитых белков в виде разделенного интеина (то есть половин белка) в клетке белки подвергаются транс-сплайсингу с образованием полного или цельного PE с сопутствующим удалением последовательностей объединенного разделенного интеина.

[0391] Например, как показано на Фиг.66, N-концевой экстеин может быть слит с первым разделенным интеином (например N интеином), а C-концевой экстеин может быть слит со вторым разделенным интеином (например C интеином). N-концевой экстеин становится слитым с C-концевым экстеином для восстановления цельного слитого белка праймерного редактора, включающего домен napDNAbp и полимеразный домен (например домен RT), при самоассоциации N интеина и C интеина внутри клетки и далее их самоэксцизии с сопутствующим образованием пептидной связи между частями N-концевого экстеина и C-концевого экстеина цельного праймерного редактора (PE).

[0392] Для использования преимущества, связанного со стратегией доставки разделенного PE с использованием разделенных интеинов, праймерный редактор должен быть разделен в одном или нескольких участках разделения для создания, по меньшей мере, двух отдельных половин праймерного редактора, каждая из которых может быть соединена внутри клетки, если каждая половина слита с последовательностью разделенного интеина.

[0393] В определенных вариантах осуществления праймерный редактор разделен в одном участке разделения. В определенных других вариантах осуществления праймерный редактор разделен в двух участках разделения, или трех участках разделения, или четырех участках разделения, или более.

[0394] В предпочтительном варианте осуществления праймерный редактор разделен в одном участке разделения для создания двух отдельных половин праймерного редактора, каждая из которых может быть слита с последовательностью разделенного интеина.

[0395] Примерном разделенного интеина является интеин *Ssp* DnaE, который включает две субъединицы, а именно DnaE-N и DnaE-C. Две различных субъединицы кодируются отдельными генами, а именно *dnaE-n* и *dnaE-c*, которые кодируют субъединицы DnaE-N и DnaE-C, соответственно. DnaE представляет собой встречающийся в природе разделенный интеин в *Synechocytis sp.* PCC6803 и способен направлять транс-сплайсинг двух отдельных белков, каждый из которых включает продукт слияния с DnaE-N или DnaE-C.

[0396] Дополнительные встречающиеся в природе или сконструированные последовательности разделенного интеина известны или могут быть реализованы на основе описанных здесь или доступных в данной области последовательностей цельного интеина. Примеры последовательностей разделенного интеина можно найти в Stevens et al., "A promiscuous split intein with expanded protein engineering applications," PNAS, 2017, Vol.114: 8538-8543; Iwai et al., "Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostc punctiforme, FEBS Lett, 580: 1853-1858, каждый из которых приведен здесь в качестве ссылки. Дополнительные последовательности разделенного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774, и EP2877490, содержание каждой из которых приведено здесь в качестве ссылки.

[0397] Кроме того, сплайсинг белка *in trans* был описан *in vivo* и *in vitro* (Shingledecker, et al., [0076] *Gene* 207:187 (1998), Southworth, et al., *EMBO J.* 17:918 (1998); Mills, et al., *Proc. Natl. Acad. Sci. USA,* 95:3543-3548 (1998); Lew, et al., *J. Biol. Chem.,* 273:15887-15890 (1998); Wu, et al., *Biochim. Biophys. Acta* 35732:1 (1998b), Yamazaki, et al., *J. Am. Chem. Soc.* 120:5591 (1998), Evans, et al., *J. Biol. Chem.* 275:9091 (2000); Otomo, et al., *Biochemistry* 38:16040-16044 (1999); Otomo, et al., *J. Biolmol. NMR* 14:105-114 (1999); Scott, et al., *Proc. Natl. Acad. Sci. USA* 96:13638-13643 (1999)), при этом он обеспечивает возможность экспрессировать белок в виде двух неактивных фрагментов, которые далее подвергаются лигированию для образования функционального продукта, например как показано на Фиг.66 и 67 в отношении образования полного слитого белка PE из двух отдельно экспрессируемых половин.

[0398] В различных описанных здесь вариантах осуществления могут использоваться способы непрерывной эволюции (например PACE) для подвергания эволюции первой части редактора оснований. Первая часть может включать один компонент или домен, например домен Cas9, домен дезаминазы, или домен UGI. Отдельно подвергнутый эволюции компонент или домен затем может быть слит с остальными частями редактора оснований в клетке путем отдельной экспрессии как подвергнутой эволюции части, так и остальных не подвергнутых эволюции частей с полипептидными доменами разделенного интеина. Первая часть может более широким образом включать любую первую аминокислотную часть редактора оснований, которую требуется подвергнуть эволюции с использованием описанного здесь способа непрерывной эволюции. Вторая часть в рамках данного варианта осуществления будет обозначать оставшуюся аминокислотную часть редактора оснований, которая не подвергается эволюции с помощью описанных здесь способов. Подвергнутая эволюции первая часть и вторая часть редактора оснований могут экспрессироваться с полипептидными доменами разделенного интеина в клетке. Природные механизмы сплайсинга белка клетки будут обеспечивать сборку подвергнутой эволюции первой части и не подвергнутой эволюции второй части для образования подвергнутого эволюции редактора оснований в виде единого слитого белка. Подвергнутая эволюции первая часть может включать N- или C-концевую часть единого слитого белка. Аналогичным образом применение второй ортогональной подвергаемой транс-сплайсингу пары интеина может обеспечивать включение подвергнутой эволюции первой частью внутренней части единого слитого белка.

[0399] Таким образом, любые из подвергнутых или не подвернутых эволюции компонентов описанных здесь редакторов оснований могут экспрессироваться с метками в виде разделенного интеина для реализации образования полного редактора оснований, включающего подвергнутый и не подвергнутый эволюции компонент в клетке.

[0400] Механизм процесса сплайсинга белка был подробно изучен (Chong, et al., J. Biol. Chem. 1996, 271, 22159-22168; Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153), при этом консервативные аминокислоты были обнаружены в точках сплайсинга интеина и экстеина (Xu, et al., EMBO Journal, 1994, 13 5517-522). Описанные здесь конструкты содержат последовательность интеина, слитую с 5′-концом первого гена (например подвергнутой эволюции частью редактора оснований). Подходящие последовательности интеина могут быть выбраны из любых белков, в отношении которых известно, что они содержат элементы сплайсинга белка. Базу данных, содержащую все известные интеины, можно найти в интернете (Perler, F. B. Nucleic Acids Research, 1999, 27, 346-347). Последовательность интеина слита на 3′ конце с 5′ концом второго гена. Для нацеливания данного гена на определенную органеллу пептидный сигнал может быть слит с кодирующей последовательностью гена. После второго гена последовательность гена интеина может быть повторена необходимое количество раз для экспрессии нескольких белков в одной клетке. Для содержащих несколько интеинов конструктов может являться полезным использовать элементы в виде интеина из разных источников. После последовательности последнего подвергаемого экспрессии гена должна быть вставлена последовательности терминации транскрипции. В одном варианте осуществления сконструирована единица для сплайсинга в виде модифицированного интеина, так что она может как катализировать эксцизию экстеинов из интеинов, так и предотвращать лигирование экстеинов. Было обнаружено, что мутагенез С-концевого соединения экстеина в ДНК-полимеразе GB-D вида Pyrococcus продуцирует измененный элемент сплайсинга, который индуцирует расщепление экстеинов и интеинов, но предотвращает последующее лигирование экстеинов (Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153). Мутация серина 538 в аланин или глицин индуцировала расщепление, но предотвращала лигирование. Мутация эквивалентных остатков в других единицах для сплайсинга в виде интеина также должна предотвращать лигирование экстеина вследствие консервации аминокислот в С-концевом соединении экстеина с интеином.

Предпочтительный интеин, не содержащий эндонуклеазный домен, представляет собой белок GyrA Mycobacterium xenopi (Telenti, et al. J. Bacteriol. 1997, 179, 6378-6382). Другие из них были обнаружены в природе или были искусственно созданы путем удаления эндонуклеазных доменов из содержащих эндонуклеазу интеинов (Chong, et al. J. Biol. Chem. 1997, 272, 15587-15590). В предпочтительном варианте осуществления интеин выбран таким образом, что он состоит из минимального количества аминокислот, необходимых для реализации функции сплайсинга, например интеин из белка GyrA Mycobacterium xenopi (Telenti, A., et al., J. Bacteriol. 1997, 179, 6378-6382). В альтернативном варианте осуществления выбирают интеин без эндонуклеазной активности, например интеин из белка GyrA Mycobacterium xenopi или интеин VMA Saccharaomyces cerevisiae, который был модифицирован для удаления эндонуклеазных доменов.Дополнительная модификация единицы для сплайсинга в виде интеина может обеспечивать изменение скорости реакции расщепления, обеспечивающее управление дозировкой белка путем простого модифицирования генной последовательности единицы для сплайсинга.

[0401] Интеины также могут существовать в виде двух фрагментов, кодируемых двумя отдельно транскрибируемыми и транслируемыми генами. Данные так называемые разделенные интеины самоассоциируются и катализируют активность сплайсинга белка *in trans*. Разделенные интеины были идентифицированы у многих цианобактерий и архей (Caspi et al, Mol Microbiol. 50: 1569-1577 (2003); Choi J. et al, J Mol Biol. 556: 1093-1106 (2006.); Dassa B. et al, Biochemistry. 46:322-330 (2007.); Liu X. and Yang J., J Biol Chem. 275:26315-26318 (2003); Wu H. et al.

[0402] Proc Natl Acad Sci USA. £5:9226-9231 (1998.); и Zettler J. et al, FEBS Letters. 553:909-914 (2009)), но до сих пор не были обнаружены у эукариотов. Недавно биоинформационный анализ природных метагеномных данных выявил 26 различных локусов с новым геномным строением. В каждом локусе консервативный кодирующий фермент участок прерывается разделенным интеином со свободным эндонуклеазным геном, вставленным между участками, кодирующими субдомены интеина. Среди них пять локусов были полностью собраны: ДНК хеликазы (gp41-l, gp41-8); инозин-5 '-монофосфатдегидрогеназа (IMPDH-1); и каталитические субъединицы рибонуклеотид-редуктазы (NrdA-2 и NrdJ-1). Похоже, что данная фрагментированная организация генов присутствует в основном у фагов (Dassa et al, Nucleic Acids Research. 57:2560-2573 (2009)).

[0403] Разделенный интеин Npu DnaE был описан как имеющий наиболее высокую отмеченную скорость реакции транс-сплайсинга белка. Кроме того, реакция сплайсинга белка Npu DnaE считается устойчивой и связанной с высоким выходом в плане различных последовательностей экстеина, температурами от 6 до 37°C, и наличием до 6M мочевины (Zettler J. et al, FEBS Letters. 553:909-914 (2009); Iwai I. et al, FEBS Letters 550: 1853-1858 (2006)). Как ожидается, при внесении мутации Cysl Ala в N-домене данных интеинов первоначальный N-C-ацильный сдвиг и, таким образом, сплайсинг белка блокируется. К сожалению, реакция C-концевого расщепления также практически полностью ингибировалась. Зависимость циклизации аспарагина в C-концевом сплайсинговом соединении от ацильного сдвига в N-концевой неустойчивой пептидной связи вероятно является уникальным свойством, присущим аллелям естественным образом разделенного интеина DnaE (Zettler J. et al. FEBS Letters. 555:909-914 (2009).

[0404] Механизм сплайсинга белка обычно имеет четыре этапа [29-30]" 1) N-S или N-O ацильный сдвиг на N-конце интеина, который разрывает расположенную выше пептидную связь и образует эфирную связь между N-экстеином и боковой цепью первой аминокислоты интеина (Cys или Ser); 2) трансэстерификацию с переносом N-экстеина на C-конец интеина, образующую новую эфирную связь, присоединяющую N-экстеин к боковой цепи первой аминокислоты C-экстеина (Cys, Ser, или Thr); 4) циклизацию Asn, разрывающую пептидную связь между интеином и C-экстеином; и 4) S-N или O-N ацильный сдвиг, который замещает эфирную связь пептидной связью между N-экстеином и C-экстеином.

[0405] Транс-сплайсинг белка, катализируемый разделенными интеинами, обеспечивает полностью ферментативный способ лигирования белка [31]. Разделенный интеин представляет собой по существу непрерывный интеин (например мини-интеин), разделенный на две части, называемые N-интеином и C-интеином, соответственно. N-интеин и C-интеин разделенного интеина могут нековалентно связываться для образования активного интеина и катализировать реакцию сплайсинга по существу так же, как и непрерывный интеин. Разделенные интеины были обнаружены в природе, а также сконструированы в лабораториях [31-35]. В соответствии с используемым здесь значением термин "разделенный интеин" обозначает любой интеин, в котором присутствует один или несколько разрывов пептидных связей между N-концевыми и C-концевыми последовательностями

аминокислот, так что N-концевые и С-концевые последовательности становятся отдельными молекулами, которые могут нековалентно связываться или восстанавливаться с образованием интеина, который является функциональным для реакций транс-сплайсинга белка. Для получения разделенного интеина для использования в рамках способов по изобретению может применяться любой каталитически активный интеин или его фрагмент. Например, в одном аспекте разделенный интеин может быть получен из эукариотического интеина. В другом аспекте разделенный интеин может быть получен из бактериального интеина. В другом аспекте разделенный интеин может быть получен из архейного интеина. Предпочтительно, полученный таким образом разделенный интеин будет иметь только аминокислотные последовательности, необходимые для катализирования реакций транс-сплайсинга.

[0406] В соответствии с используемым здесь значением термин "N-концевой разделенный интеин (In)" обозначает любую последовательность интеина, которая включает N-концевую аминокислотную последовательность, которая является функциональной для реакций транс-сплайсинга. Таким образом, In также включает последовательность, которая устраняется при осуществлении транс-сплайсинга. In может включать последовательность, которая представляет собой модификацию N-концевой части встречающейся в природе последовательности интеина. Например, In может включать дополнительные аминокислотные остатки и/или мутированные остатки, если включение подобных дополнительных и/или мутированных остатков не делает In нефункциональным в плане транс-сплайсинга. Предпочтительно, включение дополнительных и/или мутированных остатков повышает или стимулирует активность In, связанную с транс-сплайсингом.

[0407] В соответствии с используемым здесь значением термин "С-концевой разделенный интеин (Ic)" обозначает любую последовательность интеина, которая включает С-концевую аминокислотную последовательность, которая является функциональной для реакций транс-сплайсинга. В одном аспекте Ic включает от 4 до 7 последовательных аминокислотных остатков, по меньшей мере, 4 аминокислоты из которых происходят из последней β -цепи интеина, из которого он был получен. Таким образом, Ic также включает последовательность, которая устраняется при осуществлении транс-сплайсинга. Ic может включать последовательность, которая представляет собой модификацию С-концевой части встречающейся в природе

последовательности интеина. Например, Ic может включать дополнительные аминокислотные остатки и/или мутированные остатки, если включение подобных дополнительных и/или мутированных остатков не делает Ic нефункциональным в плане транс-сплайсинга. Предпочтительно, включение дополнительных и/или мутированных остатков повышает или стимулирует активность Ic, связанную с транс-сплайсингом.

[0408] В некоторых вариантах осуществления изобретения пептид, связанный с Ic или In, может включать дополнительный химический фрагмент, включая, помимо прочено, флуоресцентные группы, биотин, полиэтиленгликоль (ПЭГ), аналоги аминокислот, искусственные аминокислоты, фосфатные группы, гликозильные группы, радиоизотопные метки, и фармацевтические молекулы. В других вариантах осуществления пептид, связанный с Ic, может включать одну или несколько химически реактивных групп, включая, среди прочих, кетон, альдегид, остатки Cys и остатки Lys. N-интеин и C-интеин разделенного интеина могут нековалентно связываться для образования активного интеина и катализировать реакцию сплайсинга при присутствии "полипептида сплайсинга интеина (ISP)". В соответствии с используемым здесь значением термин "полипептид сплайсинга интеина (ISP)" обозначает часть аминокислотной последовательности разделенного интеина, которая сохраняется, когда Ic, In, или они оба удаляются из разделенного интеина. В определенных вариантах осуществления In включает ISP. В другом варианте осуществления Ic включает ISP. В другом варианте осуществления ISP представляет собой отдельный пептид, который не является ковалентно связанным ни с In, ни с Ic.

[0409] Разделенные интеины могут быть созданы из непрерывных интеинов путем конструирования одного или нескольких участков разделения в неструктурированной петле или внесения аминокислотной последовательности между 12 консервативными бета-цепями, находящимися в структуре мини-интеинов [25-28]. Может присутствовать некоторая гибкость в плане расположения участка разрыва в участках между бета-цепями при условии, что создание разрыва не будет нарушать структуру интеина, в частности структурированных бета-цепей, до значительной степени, при которой теряется активность, связанная со сплайсингом белка.

[0410] При транс-сплайсинге белка один белок-предшественник состоит из части в виде N-экстеина и далее N-интеина, другой белок-предшественник состоит из C-интеина и далее части в виде C-экстеина, при этом реакция транс-сплайсинга

(катализируемая совместно N- и C-интеинами) обеспечивает эксцизию данных двух последовательностей интеина и соединение данных двух последовательностей экстеина с помощью пептидной связи. Транс-сплайсинг белка, являясь ферментативной реакцией, может работать при очень низких (например микромолярных) концентрациях белков и может осуществляться при физиологических условиях.

[2] Другие программируемые нуклеазы

[0411] В различных описанных здесь вариантах осуществления праймерные редакторы включают napDNAbp, например белок Cas9. Данные белки "программируются" за счет образования комплексов с направляющей РНК (или PEgRNA, как может быть в этом случае), которая направляет белок Cas9 к участку-мишени на ДНК, которая обладает последовательностью, являющейся комплементарной спейсерной части нНРК (или PEgRNA), и которая также обладает требуемой последовательностью PAM. Однако в определенных предполагаемых здесь вариантах осуществления napDNAbp может быть замещен другим типом программируемого белка, например нуклеазой цинкового пальца или подобной активатору транскрипции эффекторной нуклеазой (TALEN).

[0412] На **Фиг.1J** показан вариант предполагаемого здесь праймерного редактирования, в рамках которого осуществляется замещение napDNAbp (например никазы SpCas9) любым доменом программируемой нуклеазы, например нуклеазами цинкового пальца (ZFN) или подобными активатору транскрипции эффекторными нуклеазами (TALEN). Таким образом, предполагается, что подходящие нуклеазы не обязательно должны "программироваться" нацеленной на нуклеиновую кислоту молекулой (например направляющей РНК), но вместо этого могут программироваться путем задания специфичности ДНК-связывающего домена, например (и в частности) нуклеазы. Как и при праймерном редактировании с использованием фрагментов napDNAbp, является предпочтительным, чтобы такие альтернативные программируемые нуклеазы были модифицированы таким образом, что только одна цепь ДНК-мишени подвергается разрыву. Другими словами, программируемые нуклеазы предпочтительно должны функционировать в качестве никаз. После выбора программируемой нуклеазы (например ZFN или TALEN) в систему могут быть добавлены дополнительные виды функциональности для обеспечения ее работы в соответствии с подобным праймерному редактированию механизмом. Например, программируемые нуклеазы могут быть модифицированы путем присоединения к ним

(например через химический линкер) РНК или ДНК плеча удлинения, при этом плечо удлинения включает участок связывания праймера (PBS) и матрицу для синтеза ДНК. Программируемая нуклеаза также может быть присоединена (например через химический или аминокислотный линкер) к полимеразе, природа которой будет зависеть от того, является ли плечо удлинения ДНК или РНК. В случае плеча удлинения в виде РНК полимераза может представлять собой РНК-зависимую ДНК-полимеразу (например обратную транскриптазу). В случае плеча удлинения в виде ДНК полимераза может представлять собой ДНК-зависимую ДНК-полимеразу (например прокариотическую полимеразу, включая Pol I, Pol II, или Pol III, или эукариотическую полимеразу, включая Pol a, Pol b, Pol g, Pol d, Pol e, или Pol z). Система может также включать другие виды функциональности, добавленные в виде продуктов слияния с программируемыми нуклеазами или добавленные *in trans* для реализации реакции в целом (например (a) хеликазу для разматывания ДНК в участке разрыва, чтобы сделать цепь с разрывом с 3'-концом доступной в качестве праймера, (b) FEN1 для обеспечения удаления эндогенной цепи на цепи с разрывом для смещения реакции в сторону замещения эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:нРНК для создания второго участка разрыва на противоположной цепи, что может обеспечивать интеграцию синтезированной коррекции за счет предпочтительной клеточной репарации не подвергнутой редактированию цепи). Аналогично праймерному редактированию с использованием napDNAbp, подобный комплекс с иным образом программируемой нуклеазой может использоваться для синтеза и последующего встраивания новой синтезированной замещающей цепи ДНК, несущей представляющее интерес изменение, перманентно в участок-мишень ДНК.

[0413] Подходящие альтернативно программируемые нуклеазы хорошо известны в данной области и могут использоваться вместо комплекса napDNAbp:нРНК для конструирования альтернативной системы праймерного редактирования, которая может быть запрограммирована для избирательного связывания с участком-мишенью ДНК и которая может быть дополнительно модифицирована описанным выше образом для солокализации полимеразы и РНК или ДНК плеча удлинения, включающего участок связывания праймера и матрицу для синтеза ДНК, в определенном участке разрыва. В качестве примера и как показано на Фиг.1J подобные активатору транскрипции эффекторные нуклеазы (TALEN) могут применяться в качестве программируемой нуклеазы в рамках описанных здесь способов праймерного

редактирования и соответствующих композиций. TALEN представляют собой искусственные ферменты рестрикции, полученные путем слиянии эффекторного связывающего ДНК домена TAL с доменом расщепления ДНК. Данные реагенты обеспечивают эффективное, программируемое, и направленное расщепление ДНК и представляют собой мощные инструменты для геномного редактирования *in situ*. Подобные активатору транскрипции эффекторы (TALE) могут быть быстро сконструированы для связывания практически с любой последовательностью ДНК. В соответствии с используемым здесь значением термин "TALEN" является широким и включает мономерный TALEN, который может расщеплять двуцепочечную ДНК без помощи со стороны другого TALEN. Термин "TALEN" также используется для обозначения одного или обоих членов пары TALEN, которые конструируются совместной работы для расщепления ДНК в одном и том же участке. TALEN, которые работают совместно, могут быть обозначены как левый TALEN и правый TALEN, что обозначает хиральность ДНК. См. заявку на патент США 12/965,590; заявку на патент США 13/426,991 (патент США. 8450471); заявку на патент США 13/427,040 (патент США 8,440,431); заявку на патент США 13/427,137 (патент США 8,440,432); и заявку на патент США 13/738,381, все из которых полностью приведены здесь в качестве ссылки. Кроме того, TALEN описаны в WO 2015/027134, патенте США 9181535, Boch et al., "Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors", Science, vol. 326, pp. 1509-1512 (2009), Bogdanove et al., TAL Effectors: Customizable Proteins for DNA Targeting, Science, vol. 333, pp. 1843-1846 (2011), Cade et al., "Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs", Nucleic Acids Research, vol. 40, pp. 8001-8010 (2012), и Cermak et al., "Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting", Nucleic Acids Research, vol. 39, No. 17, e82 (2011), каждый из которых приведен здесь в качестве ссылки.

[0414] Как показано на Фиг.1J, нуклеазы цинкового пальца также могут использоваться в качестве альтернативных программируемых нуклеаз в рамках праймерного редактирования вместо napDNAb, например никаз Cas9. Подобно TALEN, белки ZFN могут быть модифицированы таким образом, что они функционируют в качестве никаз, то есть за счет конструирования ZFN, так что он расщепляет только одну цепь ДНК-мишени таким же образом, что и napDNAbp, используемый с описанными здесь праймерными редакторами. Белки ZFN были

обширно описаны в данной области, например в Carroll et al., “Genome Engineering with Zinc-Finger Nucleases,” *Genetics*, Aug 2011, Vol. 188: 773-782; Durai et al., “Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells,” *Nucleic Acids Res*, 2005, Vol. 33: 5978-90; и Gaj et al., “ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering,” *Trends Biotechnol*. 2013, Vol.31: 397-405, каждый из которых полностью приведен здесь в качестве ссылки.

[3] Полимеразы (например обратные транскриптазы)

[0415] В различных вариантах осуществления описанная здесь система праймерного редактора (PE) включает полимеразу (например ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, например обратную транскриптазу) или ее вариант, который может поступать в виде слитого белка с napDNAbp или другой программируемой нуклеазой или поступать *in trans*.

[0416] В рамках описанных здесь праймерных редакторов может использоваться любая полимераза. Полимеразы могут представлять собой полимеразы дикого типа, функциональные фрагменты, мутанты, варианты, или усеченные варианты, и им подобные. Полимеразы могут включать полимеразы дикого типа из эукариотических, прокариотических, архейных, или вирусных организмов, и/или полимеразы могут быть модифицированы с помощью генной инженерии, мутагенеза, процессов на основе направленной эволюции. Полимеразы могут включать ДНК-полимеразу T7, ДНК-полимеразу T5, ДНК-полимеразу T4, фрагмент ДНК-полимеразы Klenow, ДНК-полимеразу III, и им подобные. Полимеразы также могут являться термостабильными и могут включать Taq, Tne, Tma, Pfu, Tfl, Tth, фрагмент Stoffel, ДНК-полимеразы VENT® и DEEPVENT®, KOD, Tgo, JDF3, и их мутанты, варианты, и производные (см. патент США 5436149; патент США 4889818; патент США 4965185; патент США 5079352; патент США 5614365; патент США 5374553; патент США 5270179; патент США 5047342; патент США 5512462; WO 92/06188; WO 92/06200; WO 96/10640; Barnes, W. M., Gene 112:29-35 (1992); Lawyer, F. C., et al., PCR Meth. Appl. 2:275-287 (1993); Flaman, J.-M, et al., Nuc. Acids Res. 22(15):3259-3260 (1994), каждый из которых приведен здесь в качестве ссылки). Для синтеза более длинных молекул нуклеиновой кислоты (например молекул нуклеиновой кислоты длиной более чем около 3-5 kb) могут использоваться, по меньшей мере, две ДНК-полимеразы. В определенных вариантах осуществления одна из полимераз может по существу не иметь 3' экзонуклеазной активности, а другая может иметь 3' экзонуклеазную активность.

Подобные пары могут включать полимеразы, которые являются одинаковыми или различающимися. Примеры ДНК-полимераз, которые по существу не имеют 3' экзонуклеазной активности, включают, но без ограничения, Taq, Tne(exo-), Tma(exo-), Pfu(exo-), Pwo(exo-), exo-KOD и Tth ДНК-полимеразы, и их мутанты, варианты и производные.

[0417] Предпочтительно, полимеразы, которые могут использоваться в рамках описанных здесь праймерных редакторов, представляют собой "зависимую от матрицы" полимеразу (поскольку полимеразы должны полагаться на матрицу для синтеза ДНК для указания последовательности цепи ДНК при синтезе во время праймерного редактирования). В соответствии с используемым здесь значением термин "молекула матричной ДНК" обозначает цепь нуклеиновой кислоты, из которой синтезируется комплементарная цепь нуклеиновой кислоты ДНК-полимеразой, например при реакции праймерного удлинения матрицы для синтеза ДНК PEgRNA.

[0418] В соответствии с используемым здесь значением термин "зависимым от матрицы образом" обозначает процесс, который включает зависимое от матрицы удлинение праймерной молекулы (например синтез ДНК ДНК-полимеразой). Термин "зависимым от матрицы образом" обозначает полинуклеотидный синтез РНК или ДНК, при котором последовательность новой синтезированной цепи полинуклеотида определяется хорошо известными правилами связывания комплементарных оснований (см., например, Watson, J. D. et al., In: Molecular Biology of the Gene, 4th Ed., W. A. Benjamin, Inc., Menlo Park, Calif. (1987)). Термин "комплементарный" обозначает широкую концепцию комплементарности последовательностей между участками двух полинуклеотидных цепей или между двумя нуклеотидами за счет спаривания оснований. Известно, что нуклеотид аденин способен образовывать определенные водородные связи ("спаривание оснований") с нуклеотидом, который представляет собой тимин или урацил. Аналогичным образом, известно, что нуклеотид цитозин способен обеспечивать спаривание оснований с нуклеотидом гуанином. Таким образом, в случае праймерного редактирования может быть сказано, что одна цепь ДНК, синтезированная полимеразой праймерного редактора на основе матрицы для синтеза ДНК, является "комплементарной" последовательности матрицы для синтеза ДНК.

A. Примеры полимераз

[0419] В различных вариантах осуществления описанные здесь праймерные редакторы включают полимеразу. Изобретение охватывает любую полимеразу дикого типа, полученную из любого встречающегося в природе организма или вируса, или полученную из коммерческого или некоммерческого источника. Кроме того, полимеразы, которые могут применяться в рамках праймерных редакторов по изобретению, могут включать любую встречающуюся в природе мутантную полимеразу, сконструированную мутантную полимеразу, или другую представляющую собой вариант полимеразу, включая усеченные варианты, которые сохраняют функцию. Полимеразы, которые могут использоваться здесь, также могут быть сконструированы для включения определенных аминокислотных замещений, например тех, которые конкретно указаны здесь. В определенных предпочтительных вариантах осуществления полимеразы, которые могут использоваться в рамках праймерных редакторов по изобретению, представляют собой полимеразы на основе матрицы, то есть они синтезируют нуклеотидные последовательности зависимым от матрицы образом.

[0420] Полимераза представляет собой фермент, который синтезирует нуклеотидную цепь и который может применяться совместно с описанными здесь системами праймерных редакторов. Полимеразы предпочтительно являются "зависимыми от матрицы" полимеразами (то есть полимеразой, которая синтезирует нуклеотидную цепь на основе порядка нуклеотидных оснований матричной цепи). В определенных конфигурациях полимеразы также могут являться "независимыми от матрицы" (то есть полимеразой, которая синтезирует нуклеотидную цепь без необходимости в матричной цепи). Полимераза также может быть дополнительно разделена на категории в виде "ДНК-полимеразы" или "РНК-полимеразы". В различных вариантах осуществления система праймерного редактора включает ДНК-полимеразу. В различных вариантах осуществления ДНК-полимераза может представлять собой "ДНК-зависимую ДНК-полимеразу (где молекула матрицы представляет собой цепь ДНК). В таких случаях молекула матрицы ДНК может представлять собой PEgRNA, где плечо удлинения включает цепь ДНК. В таких случаях PEgRNA может быть обозначена как химерная или гибридная PEgRNA, которая включает часть в виде РНК (то есть компоненты направляющей РНК, включая спейсер и ядро нРНК) и часть в виде ДНК (например плечо удлинения). В различных других вариантах осуществления ДНК-полимераза может представлять собой "РНК-зависимую ДНК-полимеразу" (где молекула матрицы

представляет собой цепь РНК). В таких случаях PEgRNA представляет собой РНК, то есть включает удлинение РНК. Термин "полимераза" также может обозначать фермент, который катализирует полимеразу нуклеотида (то есть полимеразную активность). В целом, данный фермент будет инициировать синтез на 3'-конце праймера, отжигаемого на последовательности полинуклеотидной матрицы (например последовательности праймера, отжигаемой на участке связывания праймера PEgRNA) по направлению к 5'-концу цепи матрицы. "ДНК-полимераза" катализирует полимеризацию дезоксинуклеотидов. В соответствии с используемым здесь значением в отношении ДНК-полимеразы термин "ДНК-полимераза" включает "ее функциональный фрагмент". Термин "ее функциональный фрагмент" обозначает любую часть относящейся к дикому типу или мутантной ДНК-полимеразы, которая охватывает менее чем целую аминокислотную последовательность полимеразы и которая сохраняет способность, по меньшей мере, при одном наборе условий катализировать полимеризацию полинуклеотида. Подобный функциональный фрагмент может присутствовать в виде отдельной сущности или может представлять собой часть более крупного полипептида, например слитого белка.

[0421] В некоторых вариантах осуществления полимеразы могут быть получены из бактериофага. ДНК-полимеразы бактериофагов в целом лишены экзонуклеазной активности от 5' к 3', поскольку данная активность кодируется отдельным полипептидом. Примеры подходящих ДНК-полимераз представляют собой ДНК-полимеразу T4, T7, и phi29. Коммерчески доступные ферменты представляют собой: T4 (доступный из многих источников, например Epicentre) и T4 (доступный из многих источников, например Epicentre для немодифицированного и USB для ДНК-полимеразы T7 exo от 3' к 5' "Sequenase").

[0422] В других вариантах осуществления полимеразы представляют собой архейные полимеразы. Существует 2 различных класса ДНК-полимераз, которые были идентифицированы у архей: 1. Семейство B/pol типа I (гомологи Pfu из *Pyrococcus furiosus*) и 2. pol типа II (гомологи *P. furiosus* DP1/2-субъединичной полимеразы DP2). Было показано, что ДНК-полимеразы обоих классов естественным образом не имеют ассоциированной экзонуклеазной активности от 5' к 3' и обладают экзонуклеазной (проверочной) активностью от 3' к 5'. Подходящие ДНК-полимеразы (pol I или pol II) могут быть получены из архей с оптимальными температурами роста, которые аналогичны требуемым температурам в рамках анализов.

[0423] Термостабильные архейные ДНК-полимеразы выделяют из вида *Pyrococcus* (furiosus, вид GB-D, w*oesii, abysii, horikoshii*), вида *Thermococcus* (kodakaraensis KOD1, litoralis, вид 9 degrees North-7, вид JDF-3, gorgonarius), *Pyrodictium occultum*, и *Archaeoglobus fulgidus*.

[0424] Полимеразы также могут быть получены из эубактериальных видов. Существует 3 класса эубактериальных ДНК-полимераз: pol I, II, и III. Ферменты семейства ДНК-полимераз pol I обладают экзонуклеазной активностью от 5' к 3', при этом определенные участники также демонстрируют экзонуклеазную активность от 5' к 3'. ДНК-полимеразы pol II естественным образом не имеют экзонуклеазную активность от 5' к 3', но демонстрируют экзонуклеазную активность от 3' к 5'. ДНК-полимеразы pol III представляют собой крупную репликативную ДНК-полимеразу клетки и состоят из нескольких субъединиц. Каталитическая субъединица pol III не имеет экзонуклеазную активность от 5' к 3', но в некоторых случаях экзонуклеазная активность от 3' к 5' располагается в том же полипептиде.

[0425] Существует множество коммерчески доступных ДНК-полимераз pol I, некоторые из которых были модифицированы для снижения или устранения экзонуклеазной активности от 5' к 3'.

[0426] Подходящие термостабильные ДНК-полимеразы pol I могут быть выделены из множества термофильных эубактерий, включая виды *Thermus* и *Thermotoga maritime*, например *Thermus aquaticus* (Taq), *Thermus thermophilus* (Tth) и *Thermotoga maritima* (Tma UlTma).

[0427] Дополнительные эубактерии, являющиеся родственными тем, которые перечислены выше, описаны в Thermophilic Bacteria (Kristjansson, J. K., ed.) CRC Press, Inc., Boca Raton, Fla., 1992.

[0428] Изобретение дополнительно включает химерные и нехимерные ДНК-полимеразы, которые химически модифицированы в соответствии со способами, описанными в патентах США 5677152, 6479264 и 6183998, содержание которых полностью приведено здесь в качестве ссылки.

[0429] Дополнительные архейные ДНК-полимеразы, являющиеся родственными тем, которые перечислены выше, описаны в следующих ссылках: Archaea: A Laboratory Manual (Robb, F. T. and Place, A. R., eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1995 и Thermophilic Bacteria (Kristjansson, J. K., ed.) CRC Press, Inc., Boca Raton, Fla., 1992.

B. Примеры обратных транскриптаз

[0430] В различных вариантах осуществления описанные здесь праймерные редакторы включают обратную транскриптазу в качестве полимеразы. Изобретение охватывает любую обратную транскриптазу дикого типа, полученную из любого встречающегося в природе организма или вируса, или полученную из коммерческого или некоммерческого источника. Кроме того, обратные транскриптазы, которые могут применяться в рамках праймерных редакторов по изобретению, могут включать любую встречающуюся в природе мутантную RT, сконструированную мутантную RT, или другой вариант RT, включая усеченные варианты, которые сохраняют функцию. RT, которые могут использоваться здесь, также могут быть сконструированы для включения определенных аминокислотных замещений, например тех, которые конкретно указаны здесь.

[0431] Обратные транскриптазы представляют собой многофункциональные ферменты обычно с тремя видами ферментативной активности, включая активность в виде РНК и ДНК-зависимой полимеризации ДНК, и активность РНКазы Н, которая катализирует расщепление РНК в РНК-ДНК гибридах. Некоторые мутанты обратной транскриптазы вызвали устранение способности фрагмента в виде РНКазы Н предотвращать ненамеренное повреждение иРНК. Данные ферменты, которые синтезируют комплементарную ДНК (кДНК) с использованием иРНК в качестве матрицы были сначала идентифицированы в РНК вирусах. После этого обратные транскриптазы были выделены и очищены напрямую из вирусных частиц, клеток, или тканей. (см., например, Kacian et al., 1971, Biochim. Biophys. Acta 46: 365-83; Yang et al., 1972, Biochem. Biophys. Res. Comm. 47: 505-11; Gerard et al., 1975, J. Virol. 15: 785-97; Liu et al., 1977, Arch. Virol. 55 187-200; Kato et al., 1984, J. Virol. Methods 9: 325-39; Luke et al., 1990, Biochem. 29: 1764-69 и Le Grice et al., 1991, J. Virol. 65: 7004-07, каждый из которых приведен здесь в качестве ссылки). Недавно были созданы мутанты и слитые белки в попытке улучшить свойства, такие как термостабильность, точность и активность. Здесь охватываются любые относящиеся к дикому типу, вариантные, и/или мутантные формы обратной транскриптазы, которые известны в данной области или которые могут быть созданы с помощью известных в данной области способов.

[0432] Ген обратной транскриптазы (RT) (или содержащаяся в нем генетическая информация) может быть получены из ряда различных источников. Например ген может быть получен из эукариотических клеток, которые инфицированы

ретровирусом, или из ряда плазмид, которые содержат либо часть, либо полный геном ретровируса. Кроме того, из ретровирусов может быть получена РНК, подобная информационной РНК, которая содержит ген RT. Примеры источников RT включают, но без ограничения, вирус лейкоза мышей Молони (M-MLV или MLVRT); вирус Т-клеточного лейкоза человека типа 1 (HTLV-1); вирус лейкоза крупного рогатого скота (BLV); вирус саркомы Рауса (RSV); вирус иммунодефицита человека (ВИЧ); дрожжи, включая Saccharomyces, Neurospora, Drosophila; приматов; и грызунов. См., например, Weiss, et al., патент США 4663290 (1987); Gerard, G. R., DNA:271-79 (1986); Kotewicz, M. L., et al., Gene 35:249-58 (1985); Tanese, N., et al., Proc. Natl. Acad. Sci. (USA):4944-48 (1985); Roth, M. J., at al., J. Biol. Chem. 260:9326-35 (1985); Michel, F., et al., Nature 316:641-43 (1985); Akins, R. A., et al., Cell 47:505-16 (1986), EMBO J. 4:1267-75 (1985); и Fawcett, D. F., Cell 47:1007-15 (1986) (каждый из которых полностью приведен здесь в качестве ссылки).

RT дикого типа

[0433] Примеры ферментов для применения в рамках описанных здесь праймерных редакторов могут включать, но без ограничения, обратную транскриптазу M-MLV и обратную транскриптазу RSV. Ферменты, имеющие активность обратной транскриптазы, являются коммерчески доступными. В определенных вариантах осуществления обратная транскриптаза предоставляется *in trans* другим компонентам системы праймерного редактора (PE). То есть обратная транскриптаза экспрессируется или иным образом предоставляется в виде отдельного компонента, то есть не в виде слитого белка с napDNAbp.

[0434] Среднему специалисту в данной области будет очевидно, что обратные транскриптазы дикого типа, включая, но без ограничения, обратную транскриптазу вируса лейкемии мышей Молони (M-MLV), обратную транскриптазу вируса иммунодефицита человека (ВИЧ) и обратную транскриптазу вируса саркомы-лейкоза птиц (ASLV), которая включает, но без ограничения, обратную транскриптазу вируса саркомы Рауса (RSV), обратную транскриптазу вируса миелобластоза птиц (AMV), обратную транскриптазу хелперного вируса MCAV вируса эритробластоза птиц (AEV), обратную транскриптазу хелперного вируса MCAV вируса миелоцитоматоза птиц MC29, обратную транскриптазу хелперного вируса REV-A вируса ретикулоэндотелиоза птиц (REV-T), обратную транскриптазу хелперного вируса UR2AV вируса саркомы птиц UR2, обратную транскриптазу хелперного вируса YAV

вируса саркомы птиц Y73, обратную транскриптазу Раус-ассоциированного вируса (RAV) и обратную транскриптазу миелобластоз-ассоциированного вируса (MAV), могут подходящим образом использоваться в рамках настоящих описанных здесь способов и композиций.

[0435] Примеры ферментов RT дикого типа приведены ниже:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ОБРАТНАЯ ТРАНСКРИПТАЗА (M-MLV RT) ДИКОГО ТИПА ВИРУС ЛЕЙКЕМИИ МЫШЕЙ МОЛОНИ ИСПОЛЬЗУЕТСЯ В PE1 (ОПИСАННЫЙ ЗДЕСЬ СЛИТЫЙ БЕЛОК ПРАЙМЕРНОГО РЕДАКТОРА 1)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLA VRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGIL VPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTV PNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEW RDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHP DLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAK KAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGK LTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDT DRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLT DQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKA LPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQ KGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP(SEQ ID NO: 89)
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУС ЛЕЙКЕМИИ МЫШЕЙ МОЛОНИ # ПОСЛ. AAA66622.1	AFPLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVS MSFIWQSAPDIGRKLGRLEDLKSKTLGDLVREAEKIFNKRETP EEREERIRRETEEKEERRRTVDEQKEKERDRRRHREMSKLLAT VVIGQEQDRQEGERKRPQLDKDQCAYCKEKGHWAKDCPKKP RGPRGPRPQTSLLTLGDXGGQGQDPPPEPRITLKVGGQPVTFL VDTGAQHSVLTQNPGPLSDKSAWVQGATGGKRYRWTTDRKV HLATGKVTHSFLHVPDCPYPLLGRDLLTKLKAQIHFEGSGAQV VGPMGQPLQVLTLNIEDEYRLHETSKEPDVSLGFTWLSDFPQA WAESGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKP HIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREV NKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRL HPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEAL HRDLADFR(SEQ ID NO: 90)
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУС ЛЕЙКЕМИИ КОШЕК # ПОСЛ. NP955579.1	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAH CQAPVLIQLKATATPISIRQYPMPHEAYQGIKPHIRRMLDQGIL KPCQSPWNTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHPTV PNPYNLLSTLPPSHPWYTVLDLKDAFFCLRLHSESQLLFAFEW RDPEIGLSGQLTWTRLPQGFKNSPTLFDEALHSDLADFRVRYP ALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGYRASAK KAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQV REFLGTAGYCRLWIPGFAELAAPLYPLTRPGTLFQWGTEQQLA FEDIKKALLSSPALGLPDITKPFELFIDENSGFAKGVLVQKLGP

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	WKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKDAGKLTLG QPLTILTSHPVEALVRQPPNKWLSNARMTHYQAMLLDAERVH FGPTVSLNPATLLPLPSGGNHHDCLQILAETHGTRPDLTDQPLP DADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSA QRAELIALTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRR GLLTSEGKEIKNKNEILALLEALFLPKRLSIIHCPGHQKGDSPQ AKGNRLADDTAKKAATETHSSLTVL SEQ ID NO: 91)
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИЧ-1, ЦЕПЬ A # ПОСЛ. ITL3-A	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKE GKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDF WEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTA FTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQ NPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTP DKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQ KLVGKLNWASQIYPGIKVRQLXKLLRGTKALTEVIPLTEEAEL ELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQ EPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGK TPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLW YQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVV TLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPD QSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSA GIRKV(SEQ ID NO: 92) CM. MARTINELLI ET AL., VIROLOGY, 1990, 174(1): 135-144, КОТОРАЯ ПРИВЕДЕНА ЗДЕСЬ В КАЧЕСТВЕ ССЫЛКИ
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИЧ-1, ЦЕПЬ B # ПОСЛ. ITL3-B	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKE GKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDF WEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTA FTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQ NPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTP DKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQ KLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELE LAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEP FKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTP KFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQ LEKEPIVGAETF(SEQ ID NO: 93) CM. STAMMERS ET AL., J. MOL. BIOL., 1994, 242(4): 586-588, КОТОРАЯ ПРИВЕДЕНА ЗДЕСЬ В КАЧЕСТВЕ ССЫЛКИ
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИРУСА САРКОМЫ РАУСА # ПОСЛ. ACL14945	TVALHLAIPLKWKPNHTPVWIDQWPLPEGKLVALTQLVEKEL QLGHIEPSLSCWNTPVFVIRKASGSYRLLHDLRAVNAKLVPFG AVQQGAPVLSALPRGWPLMVLDLKDCFFSIPLAEQDREAFAF TLPSVNNQAPARRFQWKVLPQGMTCSPTICQLIVGQILEPLRL KHPSLRMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTIS PDKVQKEPGVQYLGYKLGSTYAAPVGLVAEPRIATLWDVQKL VGSLQWLRPALGIPPRLRGPFYEQLRGSDPNEAREWNLDMK MAWREIVQLSTTAALERWDPALPLEGAVARCEQGAIGVLGQG LSTHPRPCLWLFSTQPTKAFTAWLEVLTLLITKLRASAVRTFGK EVDILLLPACFRDELPLPEGILLALRGFAGKIRSSDTPSIFDIARP

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	LHVSLKVRVTDHPVPGPTVFTDASSSTHKGVVVWREGPRWEI KEIADLGASVQQLEARAVAMALLLWPTTPTNVVTDSAFVAKM LLKMGQEGVPSTAAAFILEDALSQRSAMAAVLHVRSHSEVPG FFTEGNDVADSQATFQAYPLREAKDLHTALHIGPRALSKACNI SMQQAREVVQTCPHCNSAPALEAGVNPRGLGPLQIWQTDFTL EPRMAPRSWLAVTVDTASSAIVVTQHGRVTSVAAQHHWATVI AVLGRPKAIKTDNGSCFTSKSTREWLARWGIAHTTGIPGNSQG QAMVERANRLLKDKIRVLAEGDGFMKRIPTSKQGELLAKAM YALNHFERGENTKTPIQKHWRPTVLTEGPPVKIRIETGEWEKG WNVLVWGRGYAAVKNRDTDKVIWVPSRKVKPDIAQKDEVT KKDEASPLFA(SEQ ID NO: 94) CM. YASUKAWA ET AL., J. BIOCHEM. 2009, 145(3): 315-324, КОТОРАЯ ПРИВЕДЕНА ЗДЕСЬ В КАЧЕСТВЕ ССЫЛКИ
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИРУСА МОЗАИКИ ЦВЕТНОЙ КАПУСТЫ # ПОСЛ. AGT42196	MMDHLLQKTQIQNQTEQVMNITNPNSIYIKGRLYFKGYKKIEL HCFVDTGASLCIASKFVIPEEHWINAERPIMVKIADGSSITINKV CRDIDLIIAGEIFHIPTVYQQESGIDFIIGNNFCQLYEPFIQFTDRV IFTKDRTYPVHIAKLTRAVRVGTEGFLESMKKRSKTQQPEPVNI STNKIAILSEGRRLSEEKLFITQQRMQKIEELLEKVCSENPLDP NKTKQWMKASIKLSDPSKAIKVKPMKYSPMDREEFDKQIKEL LDLKVIKPSKSPHMAPAFLVNNEAEKRRGKKRMVVNYKAMN KATVGDAYNLPNKDELLTLIRGKKIFSSFDCKSGFWQVLLDQD SRPLTAFTCPQGHYEWNVVPFGLKQAPSIFQRHMDEAFRVFR KFCCVYVDDILVFSNNEEDHLLHVAMILQKCNQHGIILSKKKA QLFKKKINFLGLEIDEGTHKPQGHILEHINKFPDTLEDKKQLQR FLGILTYASDYIPKLAQIRKPLQAKLKENVPWKWTKEDTLYM QKVKKNLQGFPPLHHPLPEEKLIIETDASDDYWGGMLKAIKIN EGTNTELICRYASGSFKAAEKNYHSNDKETLAVINTIKKFSIYL TPVHFLIRTDNTHFKSFVNLNYKGDSKLGRNIRWQAWLSHYS FDVEHIKGTDNHFADFLSREFNRVNS(SEQ ID NO: 95) CM. FARZADFAR ET AL., VIRUS GENES, 2013, 47(2): 347-356, КОТОРАЯ ПРИВЕДЕНА ЗДЕСЬ В КАЧЕСТВЕ ССЫЛКИ
ОБРАТНАЯ ТРАНСКРИПТАЗА КЛЕБСИЕЛЛА ПНЕВМОНИИ # ПОСЛ. RFF81513.1	MKEKISKIDKNFYTDIFIKTSFQNEFEAGGVIPPIAKNQVSTISN KNKTFYSLAHSSPHYSIQTRIEKFLLKNIPLSASSFAFRKERSYL HYLEPHTQNVKYCHLDIVSFFHSIDVNIVRDTFSVYFSDEFLVK EKQSLLDAFMASVTLTAELDGVEKTFIPMGFKSSPSISNIIFRKI DILIQKFCDKNKITYTRYADDLLFSTKKENNILSSTFFINEISSIL SINKFKLNKSKYLYKEGTISLGGYVIENILKDNSSGNIRLSSSKL NPLYKALYEIKKGSSSKHICIKVFNLKLKRFIYKKNKEKFEAKF YSSQLKNKLLGYRSYLLSFVIFHKKYKCINPIFLEKCVFLISEIE SIMNRKF(SEQ ID NO: 96)
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ESCERICHIA COLI	MKITSNNVTAVINGKGWHSINWKKCHQHVKTIQTRIAKAACN QQWRTVGRLQRLLVRSFSARALAVKRVTENSGRKTPGVDGQI WSTPESKWEAIFKLRRKGYKPLPLKRVFIPKSNGKKRPLGIPV MLDRAMQALHLLGLEPVSETNADHNSYGFRPARCTADAIQQ VCNMYSSRNASKWVLEGDIKGCFEHISHEWLLENIPMDKQIL RNWLKAGIIEKSIFSKTLSGTPQGGIISPVLANMALDGLERLLQ

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
# ПОСЛ. TGH57013	NRFGRNRLI(SEQ ID NO: 97)
ОБРАТНАЯ ТРАНСКРИПТАЗА RT BACILLUS SUBTILIS # ПОСЛ. QBJ66766	MSKIKINYEKYHIKPFPHFDQRIKVNKKVKENLQNPFYIAAHS FYPFIHYKKISYKFKNGTLSSPKERDIFYSGHMDGYIYKHYGEI LNHKYNNTCIGKGIDHVSLAYRNNKMGKSNIHFAAEVINFISE QQQAFIFVSDFSSYFDSLDHAILKEKLIEVLEEQDKLSKDWWN VFKHITRYNWVEKEEVISDLECTKEKIARDKKSRERYYTPAEF REFRKRVNIKSNDTGVGIPQGTAISAVLANVYAIDLDQKLNQY ALKYGGIYRRYSDDIIMVLPMTSDGQDPSNDHVSFIKSVVKRN KVTMGDSKTSVLYYANNNIYEDYQRKRESKMDYLGFSFDGM TVKIREKSLFKYYHRTYKKINSINWASVKKEKKVGRKKLYLL YSHLGRNYKGHGNFISYCKKAHAVFEGNKKIESLINQQIKRH WKKIQKRLVDV(SEQ ID NO: 98)
RT ИНТРОНА ГРУППЫ II EUBACTERIUM RECTALE	DTSNLMEQILSSDNLNRAYLQVVRNKGAEGVDGMKYTELKE HLAKNGETIKGQLRTRKYKPQPARRVEIPKPDGGVRNLGVPT VTDRFIQQAIAQVLTPIYEEQFHDHSYGFRPNRCAQQAILTALN IMNDGNDWIVDIDLEKFFDTVNHDKLMTLIGRTIKDGDVISIV RKYLVSGIMIDDEYEDSIVGTPQGGNLSPLLANIMLNELDKEM EKRGLNFVRYADDCIIMVGSEMSANRVMRNISRFIEEKLGLKV NMTKSKVDRPSGLKYLGFGFYFDPRAHQFKAKPHAKSVAKF KKRMKELTCRSWGVSNSYKVEKLNQLIRGWINYFKIGSMKTL CKELDSRIRYRLRMCIWKQWKTPQNQEKNLVKLGIDRNTARR VAYTGKRIAYVCNKGAVNVAISNKRLASFGLISMLDYYIEKCV TC(SEQ ID NO: 99)
RT ИНТРОНА ГРУППЫ II GEOBACILLUS STEAROTHERMOPHILUS	ALLERILARDNLITALKRVEANQGAPGIDGVSTDQLRDYIRAH WSTIHAQLLAGTYRPAPVRRVEIPKPGGGTRQLGIPTVVDRLIQ QAILQELTPIFDPDFSSSSFGFRPGRNAHDAVRQAQGYIQEGYR YVVDMDLEKFFDRVNHDILMSRVARKVKDKRVLKLIRAYLQA GVMIEGVKVQTEEGTPQGGPLSPLLANILLDDLDKELEKRGL KFCRYADDCNIYVKSLRAGQRVKQSIQRFLEKTLKLKVNEEKS AVDRPWKRAFLGFSFTPERKARIRLAPRSIQRLKQRIRQLTNPN WSISMPERIHRVNQYVMGWIGYFRLVETPSVLQTIEGWIRRRL RLCQWLQWKRVRTRIRELRALGLKETAVMEIANTRKGAWRTT KTPQLHQALGKTYWTAQGLKSLTQR(SEQ ID NO: 100)

Вариантные и подверженные ошибкам RT

[0436] Обратные транскриптазы являются необходимыми для синтеза цепей комплементарных ДНК (кДНК) из РНК-матриц. Обратные транскриптазы представляют собой ферменты, состоящие из отдельных доменов, которые демонстрируют различные виды биохимической активности. Данные ферменты катализируют синтез ДНК из РНК-матрицы следующим образом: В присутствии подвергнутого отжигу праймера обратная транскриптаза связывается с РНК-матрицей и инициирует реакцию полимеризации. Активность РНК-зависимой ДНК-полимеразы обеспечивает синтез цепи комплементарной ДНК (кДНК), включающей dNTP. Активность РНКазы H обеспечивает расщепление РНК-матрицы комплекса ДНК:РНК.

Таким образом, обратные транскриптазы включают (a) активность связывания, которая обеспечивает распознавание и связывание с РНК/ДНК гибридом, (b) активность РНК-зависимой ДНК-полимеразы, и (c) активность РНКазы H. Кроме того, в целом считается, что обратные транскриптазы имеют различные атрибуты, включая их термостабильность, процессирование (скорость встраивания dNTP), и точность (или величину ошибок). Предусмотренные здесь варианты обратной транскриптазы могут включать любые мутации обратной транскриптазы, которые воздействуют или изменяют любой один или несколько из данных видов ферментативной активности (например активность РНК-зависимой ДНК-полимеразы, активность РНКазы H, или активность связывания ДНК/РНК гибрида) или ферментативные свойства (например термостабильность, процессирование, или точность). Такие варианты могут являться доступными в данной области публично, доступными коммерчески, или могут быть получены с помощью известных способов мутагенеза, включая процессы направленной эволюции (например PACE или PANCE).

[0437] В различных вариантах осуществления обратная транскриптаза может представлять собой вариант обратной транскриптазы. В соответствии с используемым здесь значением термин "вариант обратной транскриптазы" включает любой встречающийся в природе или генетически сконструированный вариант, включающий одну или несколько мутаций (включая единичные мутации, инверсии, делеции, вставки, и перестроения) относительно исходных последовательностей (например исходной последовательности дикого типа). RT естественным образом имеет несколько видов активности, включая активность РНК-зависимой ДНК-полимеразы, активность рибонуклеазы H, и активность ДНК-зависимой ДНК-полимеразы. Данные виды активности совместно обеспечивают возможность фермента конвертировать одноцепочечную РНК в двуцепочечную кДНК. В ретровирусах и ретротранспозонах данная кДНК затем может встраиваться в геном хозяина, из которого могут быть получены новые копии РНК за счет транскрипции в клетке-хозяине. Варианты RT могут включать мутацию, которая влияет на один или несколько данных видов активности (которая снижает или повышает данные виды активности или которая полностью устраняет данные виды активности). Кроме того, варианты RT могут включать одну или несколько мутаций, которые делают RT более или менее стабильной, более склонной к агрегации, и обеспечивают очистку и/или детектирование, и/или другое модифицирование свойств или характеристик.

[0438] Среднему специалисту в данной области будет очевидно, что варианты обратных транскриптаз, являющиеся производными из других обратных транскриптаз, включая, но без ограничения, обратную транскриптазу вируса лейкемии мышей Молони (M-MLV), обратную транскриптазу вируса иммунодефицита человека (ВИЧ) и обратную транскриптазу вируса саркомы-лейкоза птиц (ASLV), которая включает, но без ограничения, обратную транскриптазу вируса саркомы Рауса (RSV), обратную транскриптазу вируса миелобластоза птиц (AMV), обратную транскриптазу хелперного вируса MCAV вируса эритробластоза птиц (AEV), обратную транскриптазу хелперного вируса MCAV вируса миелоцитоматоза птиц MC29, обратную транскриптазу хелперного вируса REV-A вируса ретикулоэндотелиоза птиц (REV-T), обратную транскриптазу хелперного вируса UR2AV вируса саркомы птиц UR2, обратную транскриптазу хелперного вируса YAV вируса саркомы птиц Y73, обратную транскриптазу Раус-ассоциированного вируса (RAV) и обратную транскриптазу миелобластоз-ассоциированного вируса (MAV), могут подходящим образом использоваться в рамках настоящих описанных здесь способов и композиций.

[0439] Один способ получения вариантов RT заключается в генетическом модифицировании (например путем модифицирования последовательности ДНК обратной транскриптазы дикого типа). В данной области известен ряд способов, которые обеспечивают случайное, а также направленное мутирование последовательностей ДНК (см., например, Ausubel et. al. *Short Protocols in Molecular Biology* (1995) 3.sup.rd Ed. John Wiley & Sons, Inc.). Кроме того, существует ряд коммерчески доступных наборов для сайт-направленного мутагенеза, включая как стандартные, так и основанные на ПЦР способы. Примеры включают наборы для сайт-направленного мутагенеза QuikChange (AGILENT®), набор для сайт-направленного мутагенеза Q5® (NEW ENGLAND BIOLABS®), и систему для сайт направленного мутагенеза GeneArt™ (THERMOFISHER SCIENTIFIC®).

[0440] Кроме того, мутантные обратные транскриптазы могут быть получены с помощью мутации в виде вставки или усечения (N-концевых, внутренних, или C-концевых вставок или усечений) в соответствии со способами, известными специалисту в данной области. В соответствии с используемым здесь значением термин "мутация" обозначает замещение остатка в последовательности, например нуклеиновой кислоты или аминокислотной последовательности, другим остатком, или делецию или вставку одного или нескольких остатков в последовательности. Мутации

обычно описаны здесь на основе указания исходного остатка и далее положения остатка в последовательности и на основе указания нового замещенного остатка. Различные способы осуществления указанных здесь аминокислотных замещений (мутаций) хорошо известны в данной области и приведены, например, в Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)). Мутации могут включать различные категории, например полиморфизмы одиночных оснований, участки микродупликации, индели и инверсии, и не являются ограниченными каким-либо образом. Мутации могут включать мутации типа "потеря функции", что является нормальным результатом мутации, которая снижает или устраняет активность белка. Большинство мутаций потери функции являются рецессивными, поскольку в гетерозиготе вторая копия хромосомы несет не подверженную мутации версию гена, кодирующего полностью функциональный пептид, чье присутствие компенсирует эффект мутации. Мутации также охватывают мутации типа "приобретение функции", которые придают нехарактерную активность белку или клетки, которая иным образом не присутствует при нормальных условиях. Многие мутации, связанные с приобретением функции, располагаются в регуляторных последовательностях, а не в кодирующих участках, и, таким образом, имеют ряд последствий. Например, мутация может приводить к тому, что один или несколько генов экспрессируются в неправильных тканях, при этом эти ткани приобретают функции, которые они обычно не имеют. Вследствие их природы, мутации, связанные с приобретением функции, обычно являются доминантными.

[0441] Старые известные в данной области способы сайт-направленного мутагенеза основаны на субклонировании подвергаемой мутированию последовательности в вектор, например вектор на основе бактериофага M13, который обеспечивает изолирование одноцепочечной матрицы ДНК. В рамках данных способов осуществляется отжиг мутагенного праймера (то есть праймера, способного отжигаться на подвергаемом мутированию участке, но несущего один или несколько несовпадающих нуклеотидов в подвергаемом мутированию участке) на одноцепочечной матрице и затем полимеризация комплемента матрицы начиная с 3′-конца мутагенного праймера. Получаемые в результате этого дуплексы затем переносятся в бактерии-хозяева и осуществляется скрининг бляшек на предмет требуемой мутации.

[0442] Недавно в рамках сайт-направленного мутагенеза стали использоваться способы ПЦР, которые имеют преимущество, заключающееся в отсутствии потребности в одноцепочечной матрице. Кроме того, были разработаны способы, не требующие субклонирования. При осуществлении сайт-направленного мутагенеза на основе ПЦР необходимо учитывать несколько проблем. Во-первых, в рамках этих способов является желательным снижать количество циклов ПЦР для предотвращения экспансии нежелательных мутаций, вносимых полимеразой. Во-вторых, должен осуществляться отбор для снижения количества не подвергнутых мутированию исходных молекул, остающихся в реакции. В-третьих, ПЦР способ с повышенной длиной является предпочтительным для обеспечения возможности использования одного набора праймеров ПЦР. И в-четвертых, вследствие активности не зависящего от матрицы концевого удлинения некоторых термостабильных полимераз зачастую является необходимым включать конечный финишный этап в процедуру перед лигированием тупого конца полученного с помощью ПЦР мутантного продукта.

[0443] В данной области существуют способы случайного мутагенеза, которые будут приводить к получению группы мутантов, несущих одну или несколько случайно расположенных мутаций. Такая группа мутантов затем может быть подвергнута скринингу на предмет тех из них, которые демонстрируют требуемые свойства, например повышенную стабильность, по сравнению с обратной транскриптазой дикого типа.

[0444] Пример способа случайного мутагенеза представляет собой так называемый "способ на основе подверженной ошибкам ПЦР". Как следует из названия, способ обеспечивает амплификацию заданной последовательности в условиях, при которых ДНК-полимераза не поддерживает встраивание с высокой точностью. Несмотря на то, что условия, стимулирующие подверженное ошибкам встраивание, варьируются для различных ДНК-полимераз, специалист в данной области может определить такие условия для заданного фермента. Ключевая переменная для многих ДНК-полимераз в плане точности амплификации представляет собой, например, тип и концентрацию двухвалентного иона металла в буфере. Применение иона марганца и/или варьирование концентрации иона магния или марганца может, таким образом, применяться для управления величиной ошибок полимеразы.

[0445] В рамках различных аспектов RT праймерных редакторов может представлять собой вариант "подверженной ошибкам" обратной транскриптазы. Могут применяться

подверженные ошибкам обратные транскриптазы, которые известны и/или доступны в данной области. Будет являться очевидным, что обратные транскриптазы естественным образом не имеют какой-либо проверочной функции; таким образом, величина ошибок обратной транскриптазы в целом является более высокой, чем у ДНК-полимераз, включающих проверочную активность. Величина ошибок какой-либо конкретной обратной транскриптазы представляет собой свойство "точности" фермента, которое отражает точность направляемой матрицей полимеризации ДНК в соответствии с ее РНК-матрицей. RT с высокой точностью имеет низкую величину ошибок. Напротив, RT с низкой точностью имеет высокую величину ошибок. Сообщается, что точность обратных транскриптаз на основе M-MLV соответствует величине ошибок в рамках диапазоне в виде одной ошибки на 15000-27000 синтезированных нуклеотидов. См. Boutabout et al. (2001) “DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1,” *Nucleic Acids Res* 29(11):2217–2222, который полностью приведен здесь в качестве ссылки. Таким образом, для целей настоящего изобретения обратные транскриптазы, считающиеся "подверженными ошибкам" или считающиеся имеющими "подверженную ошибкам точность", соответствуют тем, которые имеют величину ошибок в виде менее чем одной ошибки на 15000 синтезированных нуклеотидов.

[0446] Подверженная ошибкам обратная транскриптаза также может быть создана с помощью мутагенеза исходного фермента RT (например RT M-MLV дикого типа). Способ мутагенеза не является ограниченным и может включать процессы направленной эволюции, например непрерывную эволюцию с помощью фагов (PACE) или прерывистую эволюцию с помощью фагов (PANCE). В соответствии с используемым здесь значением термин "непрерывная эволюция с помощью фагов (PACE)" обозначает непрерывную эволюцию, в рамках которой используется фаг в качестве вирусных векторов. Общая концепция технологии PACE была описана, например, в международной PCT заявке PCT/US2009/056194, поданной 8 сентября 2009 г., опубликованной как WO 2010/028347 11 марта 2010 г.; международной PCT заявке PCT/US2011/066747, поданной 22 декабря 2011 г., опубликованной как WO 2012/088381 28 июня 2012 г.; заявке на патент США, патенте США 9023594, выданном 5 мая 2015 г., международной PCT заявке PCT/US2015/012022, поданной 20 января 2015 г., опубликованной как WO 2015/134121 11 сентября 2015 г., и международной PCT заявке PCT/US2016/027795, поданной 15 апреля 2016 г., опубликованной как WO

2016/168631 20 октября 2016 г., полное содержание каждой из которых приведено здесь в качестве ссылки.

[0447] Подверженные ошибкам обратные транскриптазы также могут быть получены с помощью "прерывистой эволюции с помощью фагов (PANCE)", которая в соответствии с используемым здесь значением обозначает прерывистую эволюцию, в рамках которой используется фаг в качестве вирусных векторов. PANCE представляет собой упрощенный способ быстрой in vivo направленной эволюции с применением последовательных переносов во флаконе подвергаемого эволюции "фага отбора" (SP), который содержит представляющий интерес подвергаемый эволюции ген, между свежими клетками-хозяевами *E. coli*, тем самым сохраняя гены внутри *E. coli* неизменными, в то время как гены, содержащиеся в SP, непрерывно подвергаются эволюции. Последовательные переносы во флаконах долго служили в качестве широко доступного подхода для лабораторной эволюции микробов, при этом в недавнее время были разработаны аналогичные подходы для эволюции бактериофагов. Системе PANCE присуща более низкая строгость, чем системе PACE.

[0448] В литературе были описаны другие подверженные ошибкам обратные транскриптазы, каждая из которых предусмотрена для применения в рамках описанных здесь способов и композиций. Например, подверженные ошибкам обратные транскриптазы были описаны в Bebenek et al., “Error-prone Polymerization by HIV-1 Reverse Transcriptase,” *J Biol Chem*, 1993, Vol. 268: 10324-10334 и Sebastian-Martin et al., “Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases,” *Scientific Reports*, 2018, Vol. 8: 627, каждый из которых приведен здесь в качестве ссылки. Кроме того, обратные транскриптазы, включая подверженные ошибкам обратные транскриптазы, могут быть получены от коммерческого поставщика, включая обратную транскриптазу ProtoScript® (II), обратную транскриптазу AMV, обратную транскриптазу WarmStart®, и обратную транскриптазу M-MuLV, все от NEW ENGLAND BIOLABS®, или обратную транскриптазу AMV XL, обратную транскриптазу SMARTScribe, ультра-чистую обратную транскриптазу MMLV GPR, все от TAKARA BIO USA, INC. (бывш. CLONTECH).

[0449] Данное изобретение также охватывает обратные транскриптазы, имеющие мутации в домене РНКазы. Как указано выше, одно из характерных свойств обратных транскриптаз представляет собой активность РНКазы H, которая расщепляет РНК-

матрицу РНК:кДНК гибрида одновременно с полимеризацией. Активность РНКазы Н может являться нежелательной для синтеза длинных кДНК, поскольку РНК-матрица может расщепляться до завершения полноразмерной обратной транскрипции. Активность РНКазы Н может также снижать эффективность обратной транскрипции, вероятно вследствие ее конкурирования с полимеразной активностью фермента. Таким образом, настоящее изобретение охватывает любые варианты обратной транскриптазы, которые включают модифицированную активность РНКазы Н.

[0450] Данное изобретение также охватывает обратные транскриптазы, имеющие мутации в домене РНК-зависимой ДНК-полимеразы. Как указано выше, одно из характерных свойств обратных транскриптаз представляет собой активность РНК-зависимой ДНК-полимеразы, которая обеспечивает встраивание азотистых оснований в образующуюся цепь кДНК, кодируемую цепью матричной РНК РНК:кДНК гибрида. Активность РНК-зависимой ДНК-полимеразы может быть повышена или понижена (то есть в терминах ее скорости встраивания) для повышения или понижения процессирования фермента. Таким образом, настоящее изобретение охватывает любые варианты обратной транскриптазы, которые включают модифицированную активность РНК-зависимой ДНК-полимеразы, например повышенное или пониженное процессирование фермента по сравнению с немодифицированной версией.

[0451] Также здесь предусмотрены варианта обратной транскриптазы, которые имеют измененные характеристики в плане термостабильности. Способность обратной транскриптазы выдерживать высокие температуры является важным аспектом синтеза кДНК. Повышенные температуры реакции способствуют денатурации РНК с сильными вторичными структурами и/или высоким содержанием GC, обеспечивая чтение последовательности обратными транскриптазами. В результате этого обратная транскрипция при более высоких температурах обеспечивает синтез полноразмерных кДНК и более высокие выходы, что может приводить к повышенному продуцированию 3′ оцДНК-флэпа в результате процесса праймерного редактирования. Обратная транскриптаза M-MLV дикого типа обычно имеет оптимальную температуру в диапазоне 37-48°C; однако могут быть внесены мутации, которые обеспечивают активность в виде обратной транскрипции при более высоких температурах, превышающих 48°C, включая 49°C, 50°C, 51°C, 52°C, 53°C, 54°C, 55°C, 56°C, 57°C, 58°C, 59°C, 60°C, 61°C, 62°C, 63°C ,64°C ,65°C ,66°C, и выше.

[0452] Предусмотренные здесь варианты обратных транскриптаз, включая подверженные ошибкам RT, термостабильные RT, RT с повышенным процессированием, могут быть сконструированы с помощью различных стандартных стратегий, включая мутагенез или эволюционные процессы. В некоторых случаях варианты могут быть получены путем внесения одиночной мутации. В других случаях варианты могут требовать более чем одной мутации. Для мутантов, которые включают более чем одну мутацию, эффект определенной мутации может быть проанализирован путем внесения идентифицированной мутации в ген дикого типа с помощью сайт-направленного мутагенеза с изолированием от других мутаций, которые несет конкретный мутант. Затем скрининговые анализы одного полученного таким образом мутанта обеспечат определение эффекта данной мутации в отдельном виде.

[0453] Используемые здесь ферменты в виде вариантов RT могут также включать другие "варианты RT", являющиеся, по меньшей мере, приблизительно на 70% идентичными, по меньшей мере, приблизительно на 80% идентичными, по меньшей мере, приблизительно на 90% идентичными, по меньшей мере, приблизительно на 95% идентичными, по меньшей мере, приблизительно на 96% идентичными, по меньшей мере, приблизительно на 97% идентичными, по меньшей мере, приблизительно на 98% идентичными, по меньшей мере, приблизительно на 99% идентичными, по меньшей мере, приблизительно на 99,5% идентичными, или, по меньшей мере, приблизительно на 99,9% идентичными любому исходному белку RT, включая RT дикого типа, или мутантную RT, или фрагмент RT, или другой вариант RT, описанный или охватываемый здесь или известный в данной области.

[0454] В некоторых вариантах осуществления вариант RT может иметь 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, или до 100, или до 200, или до 300, или до 400, или до 500 или более аминокислотных изменений по сравнению с исходной RT. В некоторых вариантах осуществления вариант RT включает фрагмент исходной RT, так что фрагмент является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90% идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей

мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, по меньшей мере, приблизительно на 99,9% идентичным соответствующему фрагменту исходной RT. В некоторых вариантах осуществления фрагмент имеет длину, соответствующую, по меньшей мере, на 30%, по меньшей мере, на 35%, по меньшей мере, на 40%, по меньшей мере, на 45%, по меньшей мере, на 50%, по меньшей мере, на 55%, по меньшей мере, на 60%, по меньшей мере, на 65%, по меньшей мере, на 70%, по меньшей мере, на 75%, по меньшей мере, на 80%, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% длине аминокислоты соответствующей RT дикого типа (обратной транскриптазы M-MLV) (например SEQ ID NO: 89) или любой из обратных транскриптаз согласно SEQ ID NO: 90-100.

[0455] В некоторых вариантах осуществления в изобретении также могут использоваться фрагменты RT, которые сохраняют свою функциональность и которые представляют собой фрагменты любых описанных здесь белков RT. В некоторых вариантах осуществления фрагмент RT имеет длину, по меньшей мере, 100 аминокислот. В некоторых вариантах осуществления фрагмент имеет длину, по меньшей мере, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, или до 600 или более аминокислот.

[0456] В других вариантах осуществления в изобретении также могут использоваться варианты RT, которые являются усеченными на N-конце или C-конце или на них обоих на определенное количество аминокислот, что приводит к образованию усеченного варианта, который все же сохраняет достаточную полимеразную функцию. В некоторых вариантах осуществления усеченный вариант RT имеет усечение длиной, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22, по меньшей мере, 23, по меньшей мере, 24, по меньшей мере, 25, по меньшей мере, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, или 250 аминокислот на N-конце белка.

В других вариантах осуществления усеченный вариант RT имеет усечение длиной, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22, по меньшей мере, 23, по меньшей мере, 24, по меньшей мере, 25, по меньшей мере, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, или 250 аминокислот на С-конце белка. В других вариантах осуществления усеченный вариант RT имеет усечения на N-конце и С-конце, которые имеют одинаковую или различающуюся длину.

[0457] Например, описанные здесь праймерные редакторы могут включать усеченную версию обратной транскриптазы M-MLV. В данном варианте осуществления обратная транскриптаза содержит 4 мутации (D200N, T306K, W313F, T330P; следует отметить, что мутация L603W, присутствующая в PE2, является отсутствующей вследствие усечения). Последовательность ДНК, кодирующая данный усеченный редактор, на 522 bp короче, чем PE2, и, таким образом, является потенциально полезной для вариантов использования, в рамках которых доставка последовательности ДНК является сложной вследствие ее размера (то есть при доставке на основе адено-ассоциированного вируса и лентивируса). Данный вариант осуществления обозначается как MMLV-RT(trunc) и имеет следующую аминокислотную последовательность:

MMLV-RT(TRUNC)	*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQ APLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGL PPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAA TSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEG QRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMA APLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLR MVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDNSRLI N* (SEQ ID NO: 766)

[0458] В различных вариантах осуществления описанные здесь праймерные редакторы могут включать один из вариантов RT, описанных здесь, или ее вариант, который является, по меньшей мере, приблизительно на 70% идентичным, по меньшей мере, приблизительно на 80% идентичным, по меньшей мере, приблизительно на 90%

идентичным, по меньшей мере, приблизительно на 95% идентичным, по меньшей мере, приблизительно на 96% идентичным, по меньшей мере, приблизительно на 97% идентичным, по меньшей мере, приблизительно на 98% идентичным, по меньшей мере, приблизительно на 99% идентичным, по меньшей мере, приблизительно на 99,5% идентичным, или, по меньшей мере, приблизительно на 99,9% идентичным любым исходным вариантам RT.

[0459] В других вариантах осуществления в настоящих способах и композициях может применяться ДНК-полимераза, которая была подвергнута эволюции в обратную транскриптазу, как описано в Effefson et al., "Synthetic evolutionary origin of a proofreading reverse transcriptase," *Science*, June 24, 2016, Vol. 352: 1590-1593, содержание которого приведено здесь в качестве ссылки.

[0460] В определенных других вариантах осуществления обратная транскриптаза предусмотрена в виде компонента слитого белка, который также включает napDNAbp. Другими словами, в некоторых вариантах осуществления обратная транскриптаза слита с napDNAbp в виде слитого белка.

[0461] В различных вариантах осуществления варианты обратных транскриптаз могут быть сконструированы из обратной транскриптазы M-MLV дикого типа согласно SEQ ID NO: 89.

[0462] В различных вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий одну или несколько из следующих мутаций: P51L, S67K, E69K, L139P, T197A, D200N, H204R, F209N, E302K, E302R, T306K, F309N, W313F, T330P, L345G, L435G, N454K, D524G, E562Q, D583N, H594Q, L603W, E607K, или D653N в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа.

[0463] Некоторые примеры обратных транскриптаз, которые могут быть слиты с белками napDNAbp или поступать в виде отдельных белков в соответствии с различными вариантами осуществления данного изобретения, приведены ниже. Примеры обратных транскриптаз включают варианты, по меньшей мере, с 80%, по меньшей мере, с 85%, по меньшей мере, с 90%, по меньшей мере, с 95%, по меньшей мере, с 99% идентичностью последовательности относительно следующих дикого типа ферментов или частичных ферментов:

Описание	Последовательность (вариантные замещения относительно дикого типа)
Обратная транскриптаза (RT M-MLV) дикого типа вирус лейкоза мышей Молони Используется в PE1 (описанный здесь слитый белок праймерного редактора 1)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLK ALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL IENSSP (SEQ ID NO: 89)
RT M-MLV D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLK ALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL IENSSP (SEQ ID NO: 701)
RT M-MLV D200N T330P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL

Описание	Последовательность (вариантные замещения относительно дикого типа)
	LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLK ALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL IENSSP (SEQ ID NO: 702)
RT M-MLV D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 740)
RT M-MLV D200N T330P L603W E69K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQ**K**ARLGIKPHIQRLLDQGILVPCQSPWNTP LLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQW YTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFK NSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRA LLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETV MGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMG QPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVV ALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTD GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILA LLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDT STLLIENSSP (SEQ ID NO: 703)
RT M-MLV D200N T330P L603W E302R	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLR**R**FLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG

Описание	Последовательность (вариантные замещения относительно дикого типа)
	PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP(SEQ ID NO: 704)
RT M-MLV D200N T330P L603W E607K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 705)
RT M-MLV D200N T330P L603W L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 706)
RT M-MLV D200N T330P L603W L435G	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP

Описание	Последовательность (вариантные замещения относительно дикого типа)
	LVI**G**APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 707)
RT M-MLV D200N T330P L603W N454K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 708)
RT M-MLV D200N T330P L603W T306K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLG**K**AGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 709)
RT M-MLV D200N T330P L603W W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGP DQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS

Описание	Последовательность (вариантные замещения относительно дикого типа)
	SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 710)
RT M-MLV D200N T330P L603W D524G E562Q D583N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRA**Q**LIALTQALKMA EGKKLNVYT**N**SRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 711)
RT M-MLV D200N T330P L603W E302R W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLR**R**FLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 712)
RT M-MLV D200N T330P L603W E607K L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALL

Описание	Последовательность (вариантные замещения относительно дикого типа)
	KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 713)
RT M-MLV P51L S67K T197A H204R E302K F309N W313F T330P L435G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LII**L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTP LLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQW YTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFK NSP**A**LFDEAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTR ALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET VMGQPTPKTPRQLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFN WGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLT QKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTM GQPLVI**G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPV VALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT **G**GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILA LLKALFLPKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDT STLLIENSSP (SEQ ID NO: 714)
RT M-MLV D200N P51L S67K T197A H204R E302K F309N W313F T330P L345G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LII**L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTP LLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQW YTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFK NSP**A**LF**N**EAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTR ALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET VMGQPTPKTPRQLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFN WGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLT QKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTM GQPLVI**G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPV VALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT **G**GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILA LLKALFLPKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDT STLLIENSSP (SEQ ID NO: 715)
RT M-MLV D200N T330P L603W T306K W313F в PE2	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLG**K**AGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWG PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKL GPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP LVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGS SLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTST LLIENSSP (SEQ ID NO: 716)

[0464] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий одну или несколько из следующих мутаций: P51X, S67X, E69X, L139X, T197X, D200X, H204X, F209X, E302X, T306X, F309X, W313X, T330X, L345X, L435X, N454X, D524X, E562X, D583X, H594X, L603X, E607X, или D653X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту.

[0465] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию P51X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой L.

[0466] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию S67X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0467] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию E69X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0468] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию L139X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой

последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой P.

[0469] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию T197X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой A.

[0470] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию D200X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой N.

[0471] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию H204X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой R.

[0472] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию F209X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой N.

[0473] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать

вариант RT, включающий мутацию E302X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0474] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию E302X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой R.

[0475] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию T306X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0476] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию F309X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой N.

[0477] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию W313X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой F.

[0478] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию T330X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой P.

[0479] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию L345X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой G.

[0480] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию L435X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой G.

[0481] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию N454X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0482] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию D524X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой

любую аминокислоту. В определенных вариантах осуществления X представляет собой G.

[0483] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию E562X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой Q.

[0484] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию D583X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой N.

[0485] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию H594X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой Q.

[0486] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию L603X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой W.

[0487] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию E607X в RT M-MLV дикого типа согласно SEQ ID

NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой K.

[0488] В различных других вариантах осуществления описанные здесь праймерные редакторы (с RT, поступающей в виде слитого партнера или *in trans*) могут включать вариант RT, включающий мутацию D653X в RT M-MLV дикого типа согласно SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой последовательности полипептида RT дикого типа, где "X" может представлять собой любую аминокислоту. В определенных вариантах осуществления X представляет собой N.

[0489] Некоторые примеры обратных транскриптаз, которые могут быть слиты с белками napDNAbp или поступать в виде отдельных белков в соответствии с различными вариантами осуществления данного изобретения, приведены ниже. Примеры обратных транскриптаз включают варианты, по меньшей мере, с 80%, по меньшей мере, с 85%, по меньшей мере, с 90%, по меньшей мере, с 95%, или, по меньшей мере, с 99% идентичностью последовательности относительно следующих ферментов дикого типа или частичных ферментов:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
ОБРАТНАЯ ТРАНСКРИПТАЗА (RT M-MLV) ДИКОГО ТИПА ВИРУС ЛЕЙКОЗА МЫШЕЙ МОЛОНИ ИСПОЛЬЗУЕТСЯ В PE1 (ОПИСАННЫЙ ЗДЕСЬ СЛИТЫЙ БЕЛОК ПРАЙМЕРНОГО	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 89)

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
РЕДАКТОРА 1)	
RT M-MLV D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 106)
RT M-MLV D200N T330P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 107)
RT M-MLV D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 108)

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
RT M-MLV E69K D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQ**K**ARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 109)
RT M-MLV D200N T330P L603W E302R	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLR**R**FLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP(S EQ ID NO: 110)
RT M-MLV D200N T330P L603W E607K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALLKALFLPK RLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 111)
RT M-MLV D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
T330P L603W L139P	PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 112)
RT M-MLV D200N T330P L603W L435G	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI**G**A PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 113)
RT M-MLV D200N T330P L603W N454K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 114)
RT M-MLV D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
T306K	TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLG**K**AGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPAT LLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQ EGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKL NVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLP KRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSS P (SEQ ID NO: 115)
RT M-MLV D200N T330P L603W W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 116)
RT M-MLV D200N T330P L603W D524G E562Q D583N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRA**Q**LIALTQALKMAEGKKLN VYT**N**SRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 117)
RT M-MLV D200N T330P L603W E302R W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	PTPKTPRQLR**R**FLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 118)
RT M-MLV D200N T330P L603W E607K L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALLKALFLPK RLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 119)
RT M-MLV P51L S67K T197A H204R E302K F309N W313F T330P L435G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL II**L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SP**A**LFDEAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGP DQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI **G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPA TLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLL QEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK KLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDTSTLLIENS SP (SEQ ID NO: 120)
RT M-MLV P51L S67K T197A D200N H204R E302K F309N W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL II**L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPL LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN SP**A**LF**N**EAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGP DQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ВАРИАНТНЫЕ ЗАМЕЩЕНИЯ ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
T330P L345G N454K D524G D583N H594Q D653N	PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI **G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPA TLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLL QEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK KLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDTSTLLIENS SP (SEQ ID NO: 121)
RT M-MLV D200N T330P L603W T306K W313F B PE2	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLG**K**AGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 122)

[0490] Описанная здесь система праймерного редактора (PE) охватывает любую публично доступную обратную транскриптазу, описанную или раскрытую в любом из следующих патентов США (каждый из которых полностью приведен здесь в качестве ссылки): 10202658; 10189831; 10150955; 9932567; 9783791; 9580698; 9534201; и 9458484, и любой ее вариант, который может быть получен с помощью известных способов внесения мутаций или известных способов эволюции белков. В следующих ссылках описаны обратные транскриптазы в рамках данной области. Содержание каждой из них полностью приведено здесь в качестве ссылки.

[0491] Herzig, E., Voronin, N., Kucherenko, N. & Hizi, A. A Novel Leu92 Mutant of HIV-1 Reverse Transcriptase with a Selective Deficiency in Strand Transfer Causes a Loss of Viral Replication. *J. Virol.* **89**, 8119–8129 (2015).

[0492] Mohr, G. *et al.* A Reverse Transcriptase-Cas1 Fusion Protein Contains a Cas6 Domain Required for Both CRISPR RNA Biogenesis and RNA Spacer Acquisition. *Mol. Cell* **72**, 700-714.e8 (2018).

[0493] Zhao, C., Liu, F. & Pyle, A. M. An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA* **24**, 183–195 (2018).

[0494] Zimmerly, S. & Wu, L. An Unexplored Diversity of Reverse Transcriptases in Bacteria. *Microbiol Spectr* **3**, MDNA3-0058–2014 (2015).

[0495] Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. *Annual Review of Genetics* **35**, 501–538 (2001).

[0496] Perach, M. & Hizi, A. Catalytic Features of the Recombinant Reverse Transcriptase of Bovine Leukemia Virus Expressed in Bacteria. *Virology* **259**, 176–189 (1999).

[0497] Lim, D. *et al.* Crystal structure of the moloney murine leukemia virus RNase H domain. *J. Virol.* **80**, 8379–8389 (2006).

[0498] Zhao, C. & Pyle, A. M. Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nature Structural & Molecular Biology* **23**, 558–565 (2016).

[0499] Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biol.* **2**, REVIEWS1017 (2001).

[0500] Baranauskas, A. *et al.* Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Eng Des Sel* **25**, 657–668 (2012).

[0501] Zimmerly, S., Guo, H., Perlman, P. S. & Lambowltz, A. M. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**, 545–554 (1995).

[0502] Feng, Q., Moran, J. V., Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).

[0503] Berkhout, B., Jebbink, M. & Zsíros, J. Identification of an Active Reverse Transcriptase Enzyme Encoded by a Human Endogenous HERV-K Retrovirus. *Journal of Virology* **73**, 2365–2375 (1999).

[0504] Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M. & Gerard, G. F. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res* **16**, 265–277 (1988).

[0505] Arezi, B. & Hogrefe, H. Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res* **37**, 473–481 (2009).

[0506] Blain, S. W. & Goff, S. P. Nuclease activities of Moloney murine leukemia virus reverse transcriptase. Mutants with altered substrate specificities. *J. Biol. Chem.* **268**, 23585–23592 (1993).

[0507] Xiong, Y. & Eickbush, T. H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* **9**, 3353–3362 (1990).

[0508] Herschhorn, A. & Hizi, A. Retroviral reverse transcriptases. *Cell. Mol. Life Sci.* **67**, 2717–2747 (2010).

[0509] Taube, R., Loya, S., Avidan, O., Perach, M. & Hizi, A. Reverse transcriptase of mouse mammary tumour virus: expression in bacteria, purification and biochemical characterization. *Biochem. J.* **329 (Pt 3)**, 579–587 (1998).

[0510] Liu, M. *et al.* Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science* **295**, 2091–2094 (2002).

[0511] Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).

[0512] Nottingham, R. M. *et al.* RNA-seq of human reference RNA samples using a thermostable group II intron reverse transcriptase. *RNA* **22**, 597–613 (2016).

[0513] Telesnitsky, A. & Goff, S. P. RNase H domain mutations affect the interaction between Moloney murine leukemia virus reverse transcriptase and its primer-template. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1276–1280 (1993).

[0514] Halvas, E. K., Svarovskaia, E. S. & Pathak, V. K. Role of Murine Leukemia Virus Reverse Transcriptase Deoxyribonucleoside Triphosphate-Binding Site in Retroviral Replication and In Vivo Fidelity. *Journal of Virology* **74**, 10349–10358 (2000).

[0515] Nowak, E. *et al.* Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res* **41**, 3874–3887 (2013).

[0516] Stamos, J. L., Lentzsch, A. M. & Lambowitz, A. M. Structure of a Thermostable Group II Intron Reverse Transcriptase with Template-Primer and Its Functional and Evolutionary Implications. *Molecular Cell* **68**, 926-939.e4 (2017).

[0517] Das, D. & Georgiadis, M. M. The Crystal Structure of the Monomeric Reverse Transcriptase from Moloney Murine Leukemia Virus. *Structure* **12**, 819–829 (2004).

[0518] Avidan, O., Meer, M. E., Oz, I. & Hizi, A. The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus. *European Journal of Biochemistry* **269**, 859–867 (2002).

[0519] Gerard, G. F. *et al.* The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res* **30**, 3118–3129 (2002).

[0520] Monot, C. *et al.* The Specificity and Flexibility of L1 Reverse Transcription Priming at Imperfect T-Tracts. *PLOS Genetics* **9**, e1003499 (2013).

[0521] Mohr, S. *et al.* Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA* **19**, 958–970 (2013).

Любая из приведенных выше ссылок, связанная с обратными транскриптазами, полностью приведена здесь в качестве ссылки, если иное еще не было указано.

[4] Слитые белки PE

[0523] Описанная здесь система праймерного редактора (PE) охватывает слитые белки, включающие napDNAbp и полимеразу (например ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, например обратную транскриптазу), и необязательно соединенные линкером. Изобретение охватывает любой подходящий napDNAbp и полимеразу (например ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, например обратную транскриптазу) для комбинирования в виде единого слитого белка. Примеры napDNAbp и полимераз (например ДНК-зависимой ДНК-полимеразы или РНК-зависимой ДНК-полимеразы, например обратной транскриптазы) приведены здесь. Поскольку полимеразы хорошо известны в данной области и аминокислотные последовательности являются легкодоступными, данное изобретение не является каким-либо образом ограниченным данными конкретными указанными здесь полимеразами.

[0524] В различных вариантах осуществления слитые белки могут включать любую подходящую структурную конфигурацию. Например, слитый белок может включать в направлении от N-конца к C-концу napDNAbp, слитый с полимеразой (например ДНК-зависимой ДНК-полимеразой или РНК-зависимой ДНК-полимеразой, например обратной транскриптазой). В других вариантах осуществления слитый белок может включать по направлению от N-конца к C-концу полимеразу (например обратную транскриптазу), слитую с napDNAbp. Слитый домен может быть необязательно соединен линкером, например аминокислотной последовательностью. В других вариантах осуществления слитые белки могут включать структуру NH_2-[napDNAbp]-[полимераза]-COOH; или NH_2-[полимераза]-[napDNAbp]-COOH, где каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера. В вариантах осуществления, где полимераза представляет собой обратную транскриптазу, слитые белки могут включать структуру NH_2-[napDNAbp]-[RT]-COOH;

или NH_2-[RT]-[napDNAbp]-COOH, где каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера.

[0525] Пример слитого белка изображен на Фиг.14, на которой показан слитый белок, включающий обратную транскриптазу MLV ("MLV-RT"), слитую с никазой Cas9 ("Cas9(H840A)") с помощью последовательности линкера. Данный пример не является ограничивающим объем слитых белков, которые могут использоваться для описанной здесь системы праймерного редактора (PE).

[0526] В различных вариантах осуществления слитый белок праймерного редактора может иметь следующую аминокислотную последовательность (обозначенную здесь как "PE1"), которая включает вариант Cas9, включающий мутацию H840A (то есть никазу Cas9) и RT M-MLV дикого типа, а также N-концевую последовательность NLS (19 аминокислот) и аминокислотный линкер (32 аминокислоты), который присоединяет C-конец домена никазы Cas9 к N-концу домена RT. Слитый белок PE1 имеет следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(дикого типа)]. Аминокислотная последовательность PE1 и его отдельные компоненты являются следующими:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
СЛИТЫЙ БЕЛОК PE1 CAS9(H840A)-MMLV_RT(ДИКОГО ТИПА)	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR KQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFR IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV**

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u>*TLNIEDEYRLH ETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQD LREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHP DLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQV KYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPG FAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSS P**<u>**SGGSKRTADGSEFEPKKKRKV**</u> (SEQ ID NO: 123) ОБОЗНАЧЕНИЯ: <u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS)</u> ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 126) **<u>ЛИНКЕР ДЛИНОЙ 33 АМИНОКИСЛОТЫ</u>** (SEQ ID NO: 127) *ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV* (SEQ ID NO: 128)
PE1 – N-КОНЦЕВАЯ NLS	MKRTADGSEFESPKKKRKV (SEQ ID NO: 124)
PE1 – CAS9 (H840A) (БЕЗ MET)	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDK ADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFE ENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	RLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIM ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAG ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GD (SEQ ID NO: 130)
PE1 – ЛИНКЕР МЕЖДУ ДОМЕНОМ CAS9 И ДОМЕНОМ RT (33 АМИНОКИСЛОТЫ)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127)
PE1 – RT M-MLV	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 132)
PE1 – C-КОНЦЕВАЯ NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 133)

[0527] В другом варианте осуществления слитый белок праймерного редактора может иметь следующую аминокислотную последовательность (обозначенную здесь как "PE2"), которая включает вариант Cas9, включающий мутацию H840A (то есть никазу Cas9) и RT M-MLV, включающую мутации D200N, T330P, L603W, T306K, и W313F, а также N-концевую последовательность NLS (19 аминокислот) и аминокислотный линкер (33 аминокислоты), который присоединяет C-конец домена никазы Cas9 к N-

концу домена RT. Слитый белок PE2 имеет следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)].

Аминокислотная последовательность PE2 является следующей:

СЛИТЫЙ БЕЛОК PE2 CAS9(H840A)-MMLV_RT D200N T330P L603W T306K W313F	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR KQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFR IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD**<u>SGGSSGGSSGSETPGTSESATPESSGGSSGGSS</u>*TLNIEDEYRLH ETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQD LREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHP DLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQV KYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPG FAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENS*

	*SP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 134) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 137) **ЛИНКЕР ДЛИНОЙ 33 АМИНОКИСЛОТЫ** (SEQ ID NO: 127) *ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV* (SEQ ID NO: 139)
PE2 – N-КОНЦЕВАЯ NLS	MKRTADGSEFESPKKKRKV (SEQ ID NO: 124)
PE2 – CAS9 (H840A) (БЕЗ MET)	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDK ADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFE ENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIM ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAG ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GD (SEQ ID NO: 141)
PE2 – ЛИНКЕР МЕЖДУ ДОМЕНОМ CAS9 И ДОМЕНОМ RT (33 АМИНОКИСЛОТЫ)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127)

PE2 – MMLV_RT D200N T330P L603W T306K W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPL IIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLL PVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP TLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQ PTPKTPRQLREFLG**K**AGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 143)
PE2 – C-КОНЦЕВАЯ NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 133)

[0528] В других вариантах осуществления слитый белок праймерного редактора может иметь следующие аминокислотные последовательности:

СЛИТЫЙ БЕЛОК PE MMLV_RT(ДИКОГО ТИПА)-32AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQ AWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLD QGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYN LLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELD CQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARK ETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPH AVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEE GLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAA VTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAH IHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP**SGGSSGGSSGSETPGTSES ATPESSGGSSGGSS**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC YLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL SKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER

	LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIL DFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALI KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 145) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 147) **ЛИНКЕР ДЛИНОЙ 33 АМИНОКИСЛОТЫ** (SEQ ID NO: 127) *ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV* (SEQ ID NO: 149)
СЛИТЫЙ БЕЛОК PE MMLV_RT(ДИКОГО ТИПА)-60AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQ AWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLD QGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYN LLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELD CQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARK ETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPH AVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEE GLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAA VTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAH IHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPS**GGSSGGSSGSETPGTSES ATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS**DKKYS IGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS

	EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVT VKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD YDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 150) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)**(SEQ ID NO: 153) **АМИНОКИСЛОТНЫЙ ЛИНКЕР** *ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV*
СЛИТЫЙ БЕЛОК РЕ CAS9(H840A)-FEN1-MMLV_RT D200N T330P L603W T306K W313F	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR KQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFR IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK**

	LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u>GIQGLAKLIA DVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQNEEGETT SHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAE KQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDA PSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLP IQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHK SIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPNE EELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLDDFFKVTGSLSS AKRKEPEPKGSTKKKAKTGAAGKFKRGK**<u>**SGGSSGGSSGSETPGTSES ATPESSGGSSGGSS**</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETG GMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPC QSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPP SHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQG FKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRA LLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQP TPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQK AYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVA YLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALV KQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN CLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEI YRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGN RMADQAARKAAITETPDTSTLLIENSSP**<u>**SGGSKRTADGSEFEPKKKRKV**</u> (SEQ ID NO: 154) ОБОЗНАЧЕНИЯ: <u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS)</u> ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 157) **<u>ЛИНКЕР 1 ДЛИНОЙ 33 АМИНОКИСЛОТЫ</u>** (SEQ ID NO: 127) *ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV* (SEQ ID NO: 159) **<u>ЛИНКЕР 2 ДЛИНОЙ 33 АМИНОКИСЛОТЫ</u>** (SEQ ID NO: 127) FEN1 (SEQ ID NO: 161)

[0529] В других вариантах осуществления слитые белки праймерного редактора могут быть основаны на никазах SaCas9 или SpCas9 с измененной специфичностью в отношении PAM, таких как следующие примеры последовательностей:

ПРАЙМЕРНЫЙ РЕДАКТОР	MKRTADGSEFESPKKKRKVGKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK

SACAS9-RT M-MLV	LLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLA KRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLE RLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFID TYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPE ELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYH DIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKL VPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYG LPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGK ENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHII PRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETF KKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVD TRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFK KERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMF EEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKK PNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLIN KSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGN YLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE EAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIE VNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNL YEVKSKKHPQIIKKGSGGSSGGSSGSETPGTSESATPESSGGSSG GSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGL AVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGIL VPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR DPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDL ILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQ ICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRR PVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVV ALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAEL IALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSE GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFEPKKKR KV (SEQ ID NO: 162)
ПРАЙМЕРНЫЙ РЕДАКТОР SPCAS9(H840A)-VRQR-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI

	LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 163)
ПРАЙМЕРНЫЙ РЕДАКТОР SPCAS9(H840A)-VRER-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI

	LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 164)

[0530] В других вариантах осуществления предусмотренные здесь слитые белки праймерного редактора могут включать никазу Cas9 (например Cas9 (H840A)), слитую с усеченной версией обратной транскриптазы M-MLV. В данном варианте осуществления обратная транскриптаза также содержит 4 мутации (D200N, T306K, W313F, T330P; следует отметить, что мутация L603W, присутствующая в PE2, является отсутствующей вследствие усечения). Последовательность ДНК, кодирующая

данный усеченный редактор, на 522 bp короче, чем PE2, и, таким образом, является потенциально полезной для вариантов использования, в рамках которых доставка последовательности ДНК является сложной вследствие ее размера (то есть при доставке на основе адено-ассоциированного вируса и лентивируса). Данный вариант осуществления обозначается как Cas9(H840A)-MMLV-RT(trunc) или "PE2-short" или "PE2-trunc" и имеет следующую аминокислотную последовательность:

CAS9(H840A)-RT MMLV(TRUNC) ИЛИ PE2-SHORT	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDE YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDK ADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREK IEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE KLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIV KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKN PIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAG ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFV EQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSITGLYETRIDLSQLGGD**<u>SGGSSGGSSGSETPGTSE SATPESSGGSSGGSS</u>***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQ AWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVE DIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFA FEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQH PDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQI

	*CQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGK AGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQAL LTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEA LVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLP EEGLQHNCLDNSRLIN*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 765) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 157) **ЛИНКЕР 1 ДЛИНОЙ 33 АМИНОКИСЛОТЫ** (SEQ ID NO: 127) *УСЕЧЕННАЯ ОБРАТНАЯ ТРАНСКРИПТАЗА M-MLV* (SEQ ID NO: 766)

[0531] См. Фиг.75, на которой приведена гистограмма для сравнения эффективности (то есть "% общих считываний секвенирования с указанным изменением или инделями") PE2, PE2-trunc, PE3, и PE3-trunc в различных участках-мишенях в различных линиях клеток. Данные показывают, что праймерные редакторы, включающие усеченные варианты RT, были примерно настолько же эффективны, что и праймерные редакторы, включающие неусеченные RT белки.

[0532] В различных вариантах осуществления указанные здесь слитые белки праймерного редактора также могут включать любые варианты описанных выше последовательностей, имеющие аминокислотную последовательность, которая, по меньшей мере, приблизительно на 70% идентична, по меньшей мере, приблизительно на 80% идентична, по меньшей мере, приблизительно на 90% идентична, по меньшей мере, приблизительно на 95% идентична, по меньшей мере, приблизительно на 96% идентична, по меньшей мере, приблизительно на 97% идентична, по меньшей мере, приблизительно на 98% идентична, по меньшей мере, приблизительно на 99% идентична, по меньшей мере, приблизительно на 99,5% идентична, или, по меньшей мере, приблизительно на 99,9% идентична PE1, PE2 или любой из указанных выше последовательностей слитого праймерного редактора.

[0533] В определенных вариантах осуществления могут использоваться линкеры для связывания любых пептидов или пептидных доменов или фрагментов по изобретению (например napDNAbp, связанный или слитый с обратной транскриптазой).

[5] Линкеры и другие домены

[0534] Слитые белки PE могут включать различные другие домены помимо napDNAbp (например домен Cas9) и полимеразный домен (например домен RT). Например, в случае, если napDNAbp представляет собой Cas9, а полимераза представляет собой RT, слитые белки PE могут включать один или несколько линкеров, которые соединяют домен Cas9 с доменом RT. Линкеры могут также соединять другие функциональные домены, например последовательности ядерной локализации (NLS) или FEN1 (или другую флэп-эндонуклеазу) со слитыми белками PE или его доменом.

[0535] Кроме того, в вариантах осуществления, включающих *trans* праймерное редактирование, линкеры могут использоваться для связывания белка рекрутирования tPERT с праймерным редактором, например между белком рекрутирования tPERT и napDNAbp. См, например, пример схемы *trans* праймерного редактора (tPE), который включает линкеры для отдельного связывания полимеразного домена и домена рекрутирования белка с napDNAbp на Фиг.3G.

A. Линкеры

[0536] В соответствии с используемым здесь значением термин "линкер" обозначает химическую группу или молекулу, связывающую две молекулы или фрагмента, например домен связывания и домен расщепления нуклеазы. В некоторых вариантах осуществления линкер соединяет домен связывания нРНК РНК-программируемой нуклеазы и каталитический домен полимеразы (например обратной транскриптазы). В некоторых вариантах осуществления линкер соединяет dCas9 и обратную транскриптазу. Как правило, линкер располагается между или фланкируется двумя группами, молекулами, или иными фрагментами и присоединяется к каждому из них через ковалентную связь, тем самым связывая их. В некоторых вариантах осуществления линкер представляет собой аминокислоту или несколько аминокислот (например пептид или белок). В некоторых вариантах осуществления линкер представляет собой органическую молекулу, группу, полимер, или химический фрагмент. В некоторых вариантах осуществления линкер имеет длину 5-100 аминокислот, например длину 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, или 150-200 аминокислот. Также охватываются более длинные или короткие линкеры.

[0537] Линкер может быть представлять собой просто ковалентную цепь или он может представлять собой полимерный линкер, имеющий длину, соответствующую многим

атомам. В определенных вариантах осуществления линкер представляет собой полипептид или основан на аминокислотах. В других вариантах осуществления линкер не является подобным пептиду. В определенных вариантах осуществления линкер представляет собой ковалентную связь (например связь углерод-углерод, дисульфидную связь, связь углерод-гетероатом, и так далее). В определенных вариантах осуществления линкер представляет собой связь углерод-азот в рамках амидной связи. В определенных вариантах осуществления линкер является циклическим или ациклическим, замещенным или незамещенным, разветвленным или неразветвленным, алифатическим или гетероалифатическим линкером. В определенных вариантах осуществления линкер является полимерным (например полиэтиленовым, полиэтиленгликолевым, полиамидным, полиэфирным, и так далее). В определенных вариантах осуществления линкер включает мономер, димер, или полимер аминоалкановой кислоты. В определенных вариантах осуществления линкер включает аминоалкановую кислоту (например глицин, этановую кислоту, аланин, бета-аланин, 3-аминопропановую кислоту, 4-аминобутановую кислоту, 5-пентановую кислоту, и так далее). В определенных вариантах осуществления линкер включает мономер, димер, или полимер аминогексановой кислоты (Ahx). В определенных вариантах осуществления линкер основан на карбоциклическом фрагменте (например циклопентане, циклогексане). В других вариантах осуществления линкер включает полиэтиленгликолевый фрагмент (ПЭГ). В других вариантах осуществления линкер включает аминокислоты. В определенных вариантах осуществления линкер включает пептид. В определенных вариантах осуществления линкер включает ариловый или гетероариловый фрагмент. В определенных вариантах осуществления линкер основан на фенильном кольце. Линкер может включать функционализированные фрагменты для обеспечения присоединения нуклеофила (например тиола, амино) из пептида к линкеру. В качестве части линкера может использоваться любой электрофил. Примеры электрофилов включают, но без ограничения, активированные сложные эфиры, активированные амиды, акцепторы Майкла, алкильные галиды, арильные галиды, ацильные галиды, и изотиоцианаты.

[0538] В некоторых других вариантах осуществления линкер включает аминокислотную последовательность $(GGGGS)_n$ (SEQ ID NO: 165), $(G)_n$ (SEQ ID NO: 166), $(EAAAK)_n$ (SEQ ID NO: 167), $(GGS)_n$ (SEQ ID NO: 168), $(SGGS)_n$ (SEQ ID NO: 169), $(XP)_n$ (SEQ ID NO: 170), или любую их комбинацию, при этом n независимо

представляет собой целое число от 1 до 30, и при этом X представляет собой любую аминокислоту. В некоторых вариантах осуществления линкер включает аминокислотную последовательность (GGS)N (SEQ ID NO: 176), при этом n представляет собой 1, 3, или 7. В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGS (SEQ ID NO: 174). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS (SEQ ID NO: 175, 60AA).

[0539] В определенных вариантах осуществления могут использоваться линкеры для связывания любых пептидов или пептидных доменов или фрагментов по изобретению (например napDNAbp, связанный или слитый с обратной транскриптазой).

[0540] В соответствии с используемым здесь значением термин "линкер" обозначает химическую группу или молекулу, связывающую две молекулы или фрагмента, например домен связывания и домен расщепления нуклеазы. В некоторых вариантах осуществления линкер соединяет домен связывания нРНК РНК-программируемой нуклеазы и каталитический домен рекомбиназы. В некоторых вариантах осуществления линкер соединяет dCas9 и обратную транскриптазу. Как правило, линкер располагается между или фланкируется двумя группами, молекулами, или иными фрагментами и присоединяется к каждому из них через ковалентную связь, тем самым связывая их. В некоторых вариантах осуществления линкер представляет собой аминокислоту или несколько аминокислот (например пептид или белок). В некоторых вариантах осуществления линкер представляет собой органическую молекулу, группу, полимер, или химический фрагмент. В некоторых вариантах осуществления линкер имеет длину 5-100 аминокислот, например длину 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, или 150-200 аминокислот. Также охватываются более длинные или короткие линкеры.

[0541] Линкер может быть представлять собой просто ковалентную цепь или он может представлять собой полимерный линкер, имеющий длину, соответствующую многим атомам. В определенных вариантах осуществления линкер представляет собой полипептид или основан на аминокислотах. В других вариантах осуществления линкер не является подобным пептиду. В определенных вариантах осуществления линкер представляет собой ковалентную связь (например связь углерод-углерод, дисульфидную связь, связь углерод-гетероатом, и так далее). В определенных вариантах осуществления линкер представляет собой связь углерод-азот в рамках амидной связи. В определенных вариантах осуществления линкер является циклическим или ациклическим, замещенным или незамещенным, разветвленным или неразветвленным, алифатическим или гетероалифатическим линкером. В определенных вариантах осуществления линкер является полимерным (например полиэтиленовым, полиэтиленгликолевым, полиамидным, полиэфирным, и так далее). В определенных вариантах осуществления линкер включает мономер, димер, или полимер аминоалкановой кислоты. В определенных вариантах осуществления линкер включает аминоалкановую кислоту (например глицин, этановую кислоту, аланин, бета-аланин, 3-аминопропановую кислоту, 4-аминобутановую кислоту, 5-пентановую кислоту, и так далее). В определенных вариантах осуществления линкер включает мономер, димер, или полимер аминогексановой кислоты (Ahx). В определенных вариантах осуществления линкер основан на карбоциклическом фрагменте (например циклопентане, циклогексане). В других вариантах осуществления линкер включает полиэтиленгликолевый фрагмент (ПЭГ). В других вариантах осуществления линкер включает аминокислоты. В определенных вариантах осуществления линкер включает пептид. В определенных вариантах осуществления линкер включает ариловый или гетероариловый фрагмент. В определенных вариантах осуществления линкер основан на фенильном кольце. Линкер может включать функционализированные фрагменты для обеспечения присоединения нуклеофила (например тиола, амино) из пептида к линкеру. В качестве части линкера может использоваться любой электрофил. Примеры электрофилов включают, но без ограничения, активированные сложные эфиры, активированные амиды, акцепторы Майкла, алкильные галиды, арильные галиды, ацильные галиды, и изотиоцианаты.

[0542] В некоторых других вариантах осуществления линкер включает аминокислотную последовательность (GGGGS)n (SEQ ID NO: 165), (G)n (SEQ ID NO:

166), $(EAAAK)_n$ (SEQ ID NO: 167), $(GGS)_n$ (SEQ ID NO: 168), (SGGS)n (SEQ ID NO: 169), (XP)n (SEQ ID NO: 170), или любую их комбинацию, при этом n независимо представляет собой целое число от 1 до 30, и при этом X представляет собой любую аминокислоту. В некоторых вариантах осуществления линкер включает аминокислотную последовательность (GGS)N (SEQ ID NO: 176), при этом n представляет собой 1, 3, или 7. В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGS (SEQ ID NO: 174).

[0543] В частности, в различных вариантах осуществления могут использоваться следующие линкеры для соединения доменов праймерного редактора друг с другом:
GGS (SEQ ID NO: 767);
GGSGGS (SEQ ID NO: 768);
GGSGGSGGS (SEQ ID NO: 769);
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127);
SGSETPGTSESATPES (SEQ ID NO: 171);
SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGG S (SEQ ID NO: 175).

B. Последовательность ядерной локализации (NLS)

[0544] В различных вариантах осуществления слитые белки PE могут включать одну или несколько последовательностей ядерной локализации (NLS), которые помогают промотировать транслокацию белка в ядро клетки. Такие последовательности хорошо известны в данной области и могут включать следующие примеры:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ	SEQ ID NO:
NLS БОЛЬШОГО Т-АНТИГЕНА SV40	PKKKRKV	SEQ ID NO: 16
NLS	MKRTADGSEFESPKKKRKV	SEQ ID NO: 124

NLS	MDSLLMNRRKFLYQFKNVRWAKGRRETYLC	SEQ ID NO: 17
NLS НУКЛЕОПЛАЗМИНА	AVKRPAATKKAGQAKKKKLD	SEQ ID NO: 190
NLS EGL-13	MSRRRKANPTKLSENAKKLAKEVEN	SEQ ID NO: 191
NLS C-MYC	PAAKRVKLD	SEQ ID NO: 192
NLS TUS-БЕЛКА	KLKIKRPVK	SEQ ID NO: 193
NLS БОЛЬШОГО Т-АНТИГЕНА ПОЛИОМАВИРУСА	VSRKRPRP	SEQ ID NO: 194
NLS АНТИГЕНА ВИРУСА ГЕПАТИТА D	EGAPPAKRAR	SEQ ID NO: 195
NLS P53 МЫШЕЙ	PPQPKKKPLDGE	SEQ ID NO: 196
NLS PE1 И PE2	SGGSKRTADGSEFEPKKKRKV	SEQ ID NO: 133

[0545] Приведенные выше примеры NLS не являются ограничивающими. Слитые белки PE могут включать любую последовательность NLS, включая любую из описанных в Cokol et al., "Finding nuclear localization signals," *EMBO Rep.*, 2000, 1(5): 411-415 и Freitas et al., "Mechanisms and Signals for the Nuclear Import of Proteins," *Current Genomics*, 2009, 10(8): 550-7, каждая из которых приведена здесь в качестве ссылки.

[0546] В различных вариантах осуществления описанные здесь праймерные редакторы и конструкты, кодирующие праймерные редакторы, дополнительно включают один или несколько, предпочтительно, по меньшей мере, два сигнала ядерной локализации. В определенных вариантах осуществления праймерные редакторы включают, по меньшей мере, две NLS. В вариантах осуществления, включающих, по меньшей мере, две NLS, они могут являться одинаковыми NLS или они могут являться различающимися NLS. Кроме того, NLS могут быть экспрессированы в виде части слитого белка с остальными частями праймерных редакторов. В некоторых вариантах осуществления одна или несколько NLS представляют собой состоящие из двух частей

NLS ("bpNLS"). В определенных вариантах осуществления описанные слитые белки включают две состоящие из двух частей NLS. В некоторых вариантах осуществления описанные слитые белки включают более чем две состоящие из двух частей NLS.

[0547] Слитая NLS может располагаться на N-конце, C-конце, или внутри последовательности праймерного редактора (например будучи вставленной между кодируемым компонентом в виде napDNAbp (например Cas9) и полимеразным доменом (Например доменом обратной транскпритазы).

[0548] NLS могут представлять собой любую известную в данной области последовательность NLS. NLS также могут представлять собой любые открытые в будущем NLS для ядерной локализации. NLS также могут представлять собой любую встречающуюся в природе NLS или любую не встречающуюся в природе NLS (например NLS с одной или несколькими требуемыми мутациями).

[0549] Термин "последовательность ядерной локализации" или "NLS" обозначает аминокислотную последовательность, которая промотирует импорт белка в ядро клетки, например за счет ядерного транспорта. Последовательности ядерной локализации известны в данной области и будут очевидны специалисту в данной области. Например, последовательности NLS описаны у Plank et al., в международной PCT заявке PCT/EP2000/011690, поданной 23 ноября 23 2000 г., опубликованной как WO/2001/038547 31 мая 2001 г., содержание которой приведено здесь в качестве ссылки. В некоторых вариантах осуществления NLS включает аминокислотную последовательность PKKKRKV (SEQ ID NO: 16), MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 17), KRTADGSEFESPKKKRKV (SEQ ID NO: 3864), или KRTADGSEFEPKKKRKV (SEQ ID NO: 13). В других вариантах осуществления NLS включает аминокислотные последовательности NLSKRPAAIKKAGQAKKKK (SEQ ID NO: 3865), PAAKRVKLD (SEQ ID NO: 192), RQRRNELKRSF (SEQ ID NO: 3866), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 3867).

[0550] В одном аспекте изобретения праймерный редактор может быть модифицирован с использованием одного или нескольких сигналов ядерной локализации (NLS), предпочтительно, по меньшей мере, двух NLS. В определенных вариантах осуществления праймерные редакторы модифицированы с использованием двух или более NLS. Изобретение охватывает использование любого сигнала ядерной локализации, известного в данной области на момент изобретения, или любого сигнала

ядерной локализации, который идентифицирован или иным образом сделан доступным в данной области после подачи настоящей заявки. Соответствующий сигнал ядерной локализации представляет собой пептидную последовательность, которая направляет белок к ядру клетки, в которой экспрессируется последовательность. Сигнал ядерной локализации является преимущественно основ́ным, может быть расположен практически где угодно в рамках аминокислотной последовательности белка, в целом включает короткую последовательность от четырех аминокислот до восьми аминокислот (Autieri & Agrawal, (1998) *J. Biol. Chem*. 273: 14731-37, incorporated herein by reference), и обычно имеет высокое содержание остатков лизина и аргинина (Magin *et al*., (2000) *Virology* 274: 11-16, приведен здесь в качестве ссылки). Сигналы ядерной локализации зачастую включают остатки пролина. Было идентифицировано множество сигналов ядерной локализации, при этом они применялись для обеспечения транспорта биологических молекул из цитоплазмы в ядро клетки. См., например, Tinland *et al*., (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:7442-46; Moede *et al*., (1999) *FEBS Lett.* 461:229-34, который приведен здесь в качестве ссылки. В настоящее время считается, что в транслокации принимают участие белки ядерных пор.

[0551] Большинство NLS могут быть отнесены к трем основным группам: (i) состоящий из одной части NLS, примером которого является NLS большого Т-антигена SV40 PKKKRKV (SEQ ID NO: 16)); (ii) состоящий из двух частей мотив, состоящий из двух основных доменов, разделенных переменным количеством спейсерных аминокислот, примером которого является NLS *Xenopus nucleoplasmin* (KRXXXXXXXXXXKKKL (SEQ ID NO: 3868)); и (iii) нестандартные последовательности, такие как M9 белка hnRNP Al, NLS нуклеопротеина вируса гриппа, и NLS белка Gal4 дрожжей (Dingwall and Laskey 1991).

[0552] Сигналы ядерной локализации располагаются в различных точках аминокислотных последовательностей белков. NLS были идентифицированы на N-конце, С-конце, и в центральной части белков. Таким образом, изобретение предусматривает праймерные редакторы, которые могут быть модифицированы с использованием одного или нескольких NLS на С-конце, N-конце, а также во внутренней части праймерного редактора. Остатки более длинной последовательности, которые не функционирует в качестве остатков компонента в виде NLS, должны быть выбраны таким образом, чтобы не нарушать, например тонически или стерически, сам сигнал ядерной локализации. Таким образом, несмотря

на отсутствие строгих ограничений в плане композиции включающей NLS последовательности, на практике такая последовательность может быть функционально ограничена в плане длины и композиции.

[0553] Настоящее изобретение охватывает любые подходящие средства для модификации праймерного редактора для включения одного или нескольких NLS. В одном аспекте праймерные редакторы могут быть сконструированы для экспрессии белка праймерного редактора, который трансляционным образом слит на его N-конце или С-конце (или их обоих) с одним или несколькими NLS, то есть для образования слитого конструкта праймерный редактор-NLS. В других вариантах осуществления кодирующая праймерный редактор нуклеотидная последовательность может быть генетически модифицирована для включения рамки считывания, которая кодирует один или несколько NLS во внутреннем участке кодируемого праймерного редактора. Кроме того, NLS могут включать различные аминокислотные линкеры или спейсерные участки, кодируемые между праймерным редактором и присоединенной на N-конце, С-конце, или в середине аминокислотной последовательностью NLS, например в центральной части белков. Таким образом, настоящее изобретение также предусматривает нуклеотидные конструкты, векторы и клетки-хозяева для экспрессии слитых белков, которые включают праймерный редактор и один или несколько NLS.

[0554] Описанные здесь праймерные редакторы также могут включать сигналы ядерной локализации, которые связаны с праймерным редактором с помощью одного или нескольких линкеров, например элемента в виде полимерного, аминокислотного, основанного на нуклеиновой кислоте, полисахаридного, или химического линкера. В рамках охватываемого объема изобретения линкеры не имеют каких-либо ограничений и могут представлять собой любой подходящий тип молекулы (например полимер, аминокислоту, полисахарид, нуклеиновую кислоту, липид, или любой синтетический химический линкерный домен) и быть присоединенными к праймерному редактору с помощью любой подходящей стратегии, которая обеспечивает образование связи (например ковалентной связи, водородной связи) между праймерным редактором и одним или несколькими NLS.

C. Флэп-эндонуклеазы (например FEN1)

[0555] В различных вариантах осуществления слитые белки PE могут включать одну или несколько флэп-эндонуклеаз (например FEN1), которые обозначают фермент, который катализирует удаление 5′ одноцепочечных ДНК-флэпов. Они представляют

собой встречающиеся в природе ферменты, которые процессируют удаление 5′ флэпов, образованных в рамках клеточных процессов, включая репликацию ДНК. В рамках описанных здесь способов праймерного редактирования могут применяться эндогенно или *in trans* поступающие флэп-эндонуклеазы для удаления 5′ флэпа эндогенной ДНК, образованного в участке-мишени при праймерном редактировании. Флэп-эндонуклеазы известны в данной области и описаны в Patel et al., “Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends,” *Nucleic Acids Research*, 2012, 40(10): 4507-4519 и Tsutakawa et al., “Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily,” *Cell*, 2011, 145(2): 198-211 (каждый из которых приведен здесь в качестве ссылки). Пример флэп-эндонуклеазы представляет собой FEN1, которая может быть описана в виде следующей аминокислотной последовательности:

Описание	Последовательность	SEQ ID NO:
FEN1 дикого типа (wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSI YQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMME NGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQ AQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGI PYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGS PVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVD LCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDP NKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPN EEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLD DFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	SEQ ID NO: 198

[0556] Флэп-эндонуклеазы также могут включать любой вариант FEN1, мутант, или другой ортолог, гомолог, или вариант флэп-эндонуклеазы. Неограничивающие примеры вариантов FEN1 являются следующими:

Описание	Последовательность	SEQ ID NO:
FEN1 K168R (относительно FEN1 дикого типа)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSI YQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMME NGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQ AQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGI PYLDAPSEAEASCAALV**R**AGKVYAAATEDMDCLTFGS PVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVD LCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDP NKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPN EEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLD DFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	SEQ ID NO: 199
FEN1	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSI	SEQ ID NO:

S187A (относительно FEN1 дикого типа)	YQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMME NGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQ AQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGI PYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFG **A**PVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFV DLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLD PNKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEP NEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRL DDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFK RGK	200
FEN1 K354R (относительно FEN1 дикого типа)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSI YQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMME NGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQ AQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGI PYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGS PVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVD LCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDP NKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPN EEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLD DFFKVTGSLSSA**R**RKEPEPKGSTKKKAKTGAAGKFKR GK	SEQ ID NO: 201
GEN1	MGVNDLWQILEPVKQHIPLRNLGGKTIAVDLSLWVCE AQTVKKMMGSVMKPHLRNLFFRISYLTQMDVKLVFV MEGEPPKLKADVISKRNQSRYGSSGKSWSQKTGRSHF KSVLRECLHMLECLGIPWVQAAGEAEAMCAYLNAGG HVDGCLTNDGDTFLYGAQTVYRNFTMNTKDPHVDCY TMSSIKSKLGLDRDALVGLAILLGCDYLPKGVPGVGKE QALKLIQILKGQSLLQRFNRWNETSCNSSPQLLVTKKL AHCSVCSHPGSPKDHERNGCRLCKSDKYCEPHDYEYC CPCEWHRTEHDRQLSEVENNIKKKACCCEGFPFHEVIQ EFLLNKDKLVKVIRYQRPDLLLFQRFTLEKMEWPNHY ACEKLLVLLTHYDMIERKLGSRNSNQLQPIRIVKTRIRN GVHCFEIEWEKPEHYAMEDKQHGEFALLTIEEESLFEA AYPEIVAVYQKQKLEIKGKKQKRIKPKENNLPEPDEVM SFQSHMTLKPTCEIFHKQNSKLNSGISPDPTLPQESISAS LNSLLLPKNTPCLNAQEQFMSSLRPLAIQQIKAVSKSLI SESSQPNTSSHNISVIADLHLSTIDWEGTSFSNSPAIQRN TFSHDLKSEVESELSAIPDGFENIPEQLSCESERYTANIK KVLDEDSDGISPEEHLLSGITDLCLQDLPLKERIFTKLSY PQDNLQPDVNLKTLSILSVKESCIANSGSDCTSHLSKDL PGIPLQNESRDSKILKGDQLLQEDYKVNTSVPYSVSNT VVKTCNVRPPNTALDHSRKVDMQTTRKILMKKSVCLD RHSSDEQSAPVFGKAKYTTQRMKHSSQKHNSSHFKES GHNKLSSPKIHIKETEQCVRSYETAENEESCFPDSTKSS LSSLQCHKKENNSGTCLDSPLPLRQRLKLRFQST	SEQ ID NO: 202
ERCC5	MGVQGLWKLLECSGRQVSPEALEGKILAVDISIWLNQ ALKGVRDRHGNSIENPHLLTLFHRLCKLLFFRIRPIFVF DGDAPLLKKQTLVKRRQRKDLASSDSRKTTEKLLKTF	SEQ ID NO: 203

	LKRQAIKTAFRSKRDEALPSLTQVRRENDLYVLPPLQE EEKHSSEEEDEKEWQERMNQKQALQEEFFHNPQAIDIE SEDFSSLPPEVKHEILTDMKEFTKRRRTLFEAMPEESDD FSQYQLKGLLKKNYLNQHIEHVQKEMNQQHSGHIRRQ YEDEGGFLKEVESRRVVSEDTSHYILIKGIQAKTVAEV DSESLPSSSKMHGMSFDVKSSPCEKLKTEKEPDATPPSP RTLLAMQAALLGSSSEEELESENRRQARGRNAPAAVD EGSISPRTLSAIKRALDDDEDVKVCAGDDVQTGGPGAE EMRINSSTENSDEGLKVRDGKGIPFTATLASSSVNSAEE HVASTNEGREPTDSVPKEQMSLVHVGTEAFPISDESMI KDRKDRLPLESAVVRHSDAPGLPNGRELTPASPTCTNS VSKNETHAEVLEQQNELCPYESKFDSSLLSSDDETKCK PNSASEVIGPVSLQETSSIVSVPSEAVDNVENVVSFNAK EHENFLETIQEQQTTESAGQDLISIPKAVEPMEIDSEESE SDGSFIEVQSVISDEELQAEFPETSKPPSEQGEEELVGTR EGEAPAESESLLRDNSERDDVDGEPQEAEKDAEDSLHE WQDINLEELETLESNLLAQQNSLKAQKQQQERIAATVT GQMFLESQELLRLFGIPYIQAPMEAEAQCAILDLTDQTS GTITDDSDIWLFGARHVYRNFFNKNKFVEYYQYVDFH NQLGLDRNKLINLAYLLGSDYTEGIPTVGCVTAMEILN EFPGHGLEPLLKFSEWWHEAQKNPKIRPNPHDTKVKK KLRTLQLTPGFPNPAVAEAYLKPVVDDSKGSFLWGKP DLDKIREFCQRYFGWNRTKTDESLFPVLKQLDAQQTQ LRIDSFFRLAQQEKEDAKRIKSQRLNRAVTCMLRKEKE AAASEIEAVSVAMEKEFELLDKAKRKTQKRGITNTLEE SSSLKRKRLSDSKRKNTCGGFLGETCLSESSDGSSSEDA ESSSLMNVQRRTAAKEPKTSASDSQNSVKEAPVKNGG ATTSSSSDSDDDGGKEKMVLVTARSVFGKKRRKLRRA RGRKRKT	

[0557] В различных вариантах осуществления указанные здесь слитые белки праймерного редактора также могут включать любой вариант флэп-эндонуклеазы описанных выше последовательностей, имеющий аминокислотную последовательность, которая, по меньшей мере, приблизительно на 70% идентична, по меньшей мере, приблизительно на 80% идентична, по меньшей мере, приблизительно на 90% идентична, по меньшей мере, приблизительно на 95% идентична, по меньшей мере, приблизительно на 96% идентична, по меньшей мере, приблизительно на 97% идентична, по меньшей мере, приблизительно на 98% идентична, по меньшей мере, приблизительно на 99% идентична, по меньшей мере, приблизительно на 99,5% идентична, или, по меньшей мере, приблизительно на 99,9% идентична любой из указанных выше последовательностей.

[0558] Другие эндонуклеазы, которые могут использоваться в настоящих способах для обеспечения удаления 5'-концевого одноцепочечного ДНК-флэпа, включают, но без

ограничения, эндонуклеазу (1) trex 2, (2) exo1 (например Keijzers et al., *Biosci Rep.* 2015, 35(3): e00206)

Trex 2

[0559] 3’ праймерная репарационная экзонуклеаза 2 (TREX2) - человек

[0560] Кат. # NM_080701

MSEAPRAETFVFLDLEATGLPSVEPEIAELSLFAVHRSSLENPEHDESGALVLPRVLD KLTLCMCPERPFTAKASEITGLSSEGLARCRKAGFDGAVVRTLQAFLSRQAGPICLVA HNGFDYDFPLLCAELRRLGARLPRDTVCLDTLPALRGLDRAHSHGTRARGRQGYSL GSLFHRYFRAEPSAAHSAEGDVHTLLLIFLHRAAELLAWADEQARGWAHIEPMYLPP DDPSLEA (SEQ ID NO: 3865).

[0561] 3’ праймерная репарационная экзонуклеаза 2 (TREX2) - мышь

[0562] Кат. # NM_011907

[0563] MSEPPRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSLVLP RVLDKLTLCMCPERPFTAKASEITGLSSESLMHCGKAGFNGAVVRTLQGFLSRQEGPI CLVAHNGFDYDFPLLCTELQRLGAHLPQDTVCLDTLPALRGLDRAHSHGTRAQGRK SYSLASLFHRYFQAEPSAAHSAEGDVHTLLLIFLHRAPELLAWADEQARSWAHIEPM YVPPDGPSLEA (SEQ ID NO: 3866).

[0564] 3’ праймерная репарационная экзонуклеаза 2 (TREX2) - крыса

[0565] Кат. # NM_001107580

[0566] MSEPLRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSLVLP RVLDKLTLCMCPERPFTAKASEITGLSSEGLMNCRKAAFNDAVVRTLQGFLSRQEGP ICLVAHNGFDYDFPLLCTELQRLGAHLPRDTVCLDTLPALRGLDRVHSHGTRAQGRK SYSLASLFHRYFQAEPSAAHSAEGDVNTLLLIFLHRAPELLAWADEQARSWAHIEPM YVPPDGPSLEA (SEQ ID NO: 3867).

ExoI

[0567] Экзонуклеаза 1 человека (EXO1) принимает участие во многих различных метаболических процессах ДНК, включая репарацию несовпадений ДНК (MMR), опосредованное микрогомологией соединение концов, гомологичную рекомбинацию (HR), и репликацию. EXO1 человека принадлежит к семейству эукариотических нуклеаз, Rad2/XPG, которые также включают FEN1 и GEN1. Семейство Rad2/XPG является консервативным в рамках нуклеазного домена у видов от фага до человека. Продукт гена EXO1 демонстрирует активность как 5' экзонуклеазы, так и 5' флэпа. Кроме того, EXO1 содержит собственную активность 5' РНКазы H. EXO1 человека

имеет высокую аффинность в плане процессинга двуцепочечной ДНК (дцДНК), разрывов, гэпов, псевдо-Y структур, и может корректировать Соединения Холлидея за счет собственной флэп-активности. EXO1 человека принимает участие в MMR и содержит консервативные домены связывания, взаимодействующие напрямую с MLH1 и MSH2. Нуклеолитическая активность положительно стимулируется PCNA, MutSα (комплексом MSH2/MSH6), 14-3-3, MRN и комплексом 9-1-1.

[0568] экзонуклеаза 1 (EXO1) Кат. # NM_003686 (экзонуклеаза 1 Homo sapiens (EXO1), транскриптный вариант 3) – изоформа A

MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDRYV GFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVS EARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQAIITE DSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILSGCDY LSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLY QLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYN PDTAMPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVG VERVISTKGLNLPRKSSIVKRPRSAELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSFSE VFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDSTDC VSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDH IPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQ FRRKSDSPTSLPENNMSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASK LSQCSSKDSDSEESDCNIKLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSSADS LSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLNELWKNFGFKKF (SEQ ID NO: 3868).

[0569] экзонуклеаза 1 (EXO1) Кат. # NM_006027 (экзонуклеаза 1 Homo sapiens (EXO1), транскриптный вариант 3) – изоформа B

[0570] MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPT DRYVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLRE GKVSEARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQ AIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILS GCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANN TFLYQLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQID DYNPDTAMPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPS TVGVERVISTKGLNLPRKSSIVKRPRSAELSEDDLLSQYSLSFTKKTKKNSSEGNKSLS

FSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDS
TDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIP
GDHIPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTA
LQQFRRKSDSPTSLPENNMSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSN
ASKLSQCSSKDSDSEESDCNIKLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSS
ADSLSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLNELWKNFGFKKD
SEKLPPCKKPLSPVRDNIQLTPEAEEDIFNKPECGRVQRAIFQ (SEQ ID NO: 3869).

[0571] экзонуклеаза 1 (EXO1) Кат. # NM_001319224 (экзонуклеаза 1 Homo sapiens (EXO1), транскриптный вариант 4) – изоформа C

[0572] MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPT
DRYVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLRE
GKVSEARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQ
AIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILS
GCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANN
TFLYQLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQID
DYNPDTAMPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPS
TVGVERVISTKGLNLPRKSSIVKRPRSELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSF
SEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDST
DCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPG
DHIPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTAL
QQFRRKSDSPTSLPENNMSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSN
ASKLSQCSSKDSDSEESDCNIKLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSS
ADSLSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLNELWKNFGFKKD
SEKLPPCKKPLSPVRDNIQLTPEAEEDIFNKPECGRVQRAIFQ (SEQ ID NO: 3870).

D. <u>Интеины и разделенные интеины</u>

[0573] Будет являться очевидным, что в некоторых вариантах осуществления (например при доставке праймерного редактора *in vivo* с использованием частиц AAV) может являться полезным разделять полипептид (например дезаминазу или napDNAbp) или слитый белок (например праймерный редактор) на N-концевую половину и C-концевую половину и доставлять их по отдельности с последующим восстановлением полного белка (или слитого белка в соответствующем случае) за счет их солокализации в клетке. Каждая из отдельных половин белка или слитого белка может включать метку в виде разделенного интеина для осуществления

восстановления полного белка или слитого белка за счет механизма транс-сплайсинга белка.

[0574] Транс-сплайсинг белка, катализируемый разделенными интеинами, представляет собой полностью ферментативный способ лигирования белка. Разделенный интеин представляет собой по существу непрерывный интеин (например мини-интеин), разделенный на две части, называемые N-интеином и C-интеином, соответственно. N-интеин и C-интеин разделенного интеина могут нековалентно связываться для образования активного интеина и катализировать реакцию сплайсинга по существу так же, как и непрерывный интеин. Разделенные интеины были обнаружены в природе, а также сконструированы в лабораториях. В соответствии с используемым здесь значением термин "разделенный интеин" обозначает любой интеин, в котором присутствует один или несколько разрывов пептидных связей между N-концевыми и C-концевыми последовательностями аминокислот, так что N-концевые и C-концевые последовательности становятся отдельными молекулами, которые могут нековалентно связываться или восстанавливаться с образованием интеина, который является функциональным для реакций транс-сплайсинга белка. Для получения разделенного интеина для использования в рамках способов по изобретению может применяться любой каталитически активный интеин или его фрагмент. Например, в одном аспекте разделенный интеин может быть получен из эукариотического интеина. В другом аспекте разделенный интеин может быть получен из бактериального интеина. В другом аспекте разделенный интеин может быть получен из архейного интеина. Предпочтительно, полученный таким образом разделенный интеин будет иметь только аминокислотные последовательности, необходимые для катализирования реакций транс-сплайсинга.

[0575] В соответствии с используемым здесь значением термин "N-концевой разделенный интеин (In)" обозначает любую последовательность интеина, которая включает N-концевую аминокислотную последовательность, которая является функциональной для реакций транс-сплайсинга. Таким образом, In также включает последовательность, которая устраняется при осуществлении транс-сплайсинга. In может включать последовательность, которая представляет собой модификацию N-концевой части встречающейся в природе последовательности интеина. Например, In может включать дополнительные аминокислотные остатки и/или мутированные остатки, если включение подобных дополнительных и/или мутированных остатков не

делает In нефункциональным в плане транс-сплайсинга. Предпочтительно, включение дополнительных и/или мутированных остатков повышает или стимулирует активность In, связанную с транс-сплайсингом.

[0576] В соответствии с используемым здесь значением термин "С-концевой разделенный интеин (Ic)" обозначает любую последовательность интеина, которая включает С-концевую аминокислотную последовательность, которая является функциональной для реакций транс-сплайсинга. В одном аспекте Ic включает от 4 до 7 последовательных аминокислотных остатков, по меньшей мере, 4 аминокислоты из которых происходят из последней β -цепи интеина, из которого он был получен. Таким образом, Ic также включает последовательность, которая устраняется при осуществлении транс-сплайсинга. Ic может включать последовательность, которая представляет собой модификацию С-концевой части встречающейся в природе последовательности интеина. Например, Ic может включать дополнительные аминокислотные остатки и/или мутированные остатки, если включение подобных дополнительных и/или мутированных остатков не делает Ic нефункциональным в плане транс-сплайсинга. Предпочтительно, включение дополнительных и/или мутированных остатков повышает или стимулирует активность Ic, связанную с транс-сплайсингом.

[0577] В некоторых вариантах осуществления изобретения пептид, связанный с Ic или In, может включать дополнительный химический фрагмент, включая, помимо прочено, флуоресцентные группы, биотин, полиэтиленгликоль (ПЭГ), аналоги аминокислот, искусственные аминокислоты, фосфатные группы, гликозильные группы, радиоизотопные метки, и фармацевтические молекулы. В других вариантах осуществления пептид, связанный с Ic, может включать одну или несколько химически реактивных групп, включая, среди прочих, кетон, альдегид, остатки Cys и остатки Lys. N-интеин и С-интеин разделенного интеина могут нековалентно связываться для образования активного интеина и катализировать реакцию сплайсинга при присутствии "полипептида сплайсинга интеина (ISP)". В соответствии с используемым здесь значением термин "полипептид сплайсинга интеина (ISP)" обозначает часть аминокислотной последовательности разделенного интеина, которая сохраняется, когда Ic, In, или они оба удаляются из разделенного интеина. В определенных вариантах осуществления In включает ISP. В другом варианте осуществления Ic

включает ISP. В другом варианте осуществления ISP представляет собой отдельный пептид, который не является ковалентно связанным ни с In, ни с Ic.

[0578] Разделенные интеины могут быть созданы из непрерывных интеинов путем конструирования одного или нескольких участков разделения в неструктурированной петле или внесения аминокислотной последовательности между 12 консервативными бета-цепями, находящимися в структуре мини-интеинов. Может присутствовать некоторая гибкость в плане расположения участка разрыва в участках между бета-цепями при условии, что создание разрыва не будет нарушать структуру интеина, в частности структурированных бета-цепей, до значительной степени, при которой теряется активность, связанная со сплайсингом белка.

[0579] При транс-сплайсинге белка один белок-предшественник состоит из части в виде N-экстеина и далее N-интеина, другой белок-предшественник состоит из C-интеина и далее части в виде C-экстеина, при этом реакция транс-сплайсинга (катализируемая совместно N- и C-интеинами) обеспечивает эксцизию данных двух последовательностей интеина и соединение данных двух последовательностей экстеина с помощью пептидной связи. Транс-сплайсинг белка, являясь ферментативной реакцией, может работать при очень низких (например микромолярных) концентрациях белков и может осуществляться при физиологических условиях.

[0580] Примеры последовательностей являются следующими:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
ИНТЕИН 2-4	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 8)
ИНТЕИН 3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDR VAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASM MGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEIL MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATS SRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRAL DKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHL YSMKYTNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDDK

	FLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHN C (SEQ ID NO: 9)
ИНТЕИН 30R3-1	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 10)
ИНТЕИН 30R3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 11)
ИНТЕИН 30R3-3	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 12)
ИНТЕИН 37R3-1	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC ((SEQ ID NO: 13)
ИНТЕИН 37R3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH

	LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 14)
ИНТЕИН 37R3-3	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 15)

[0581] Несмотря на то, что интеины наиболее часто представлены в виде последовательного домена, некоторые из них существуют в естественной разделенной форме. В этом случае два фрагмента экспрессируются в виде отдельных полипептидов и должны быть связаны перед осуществлением сплайсинга, так называемого транс-сплайсинга белка.

[0582] Примерном разделенного интеина является интеин *Ssp* DnaE, который включает две субъединицы, а именно DnaE-N и DnaE-C. Две различных субъединицы кодируются отдельными генами, а именно *dnaE-n* и *dnaE-c*, которые кодируют субъединицы DnaE-N и DnaE-C, соответственно. DnaE представляет собой встречающийся в природе разделенный интеин в *Synechocytis sp.* PCC6803 и способен направлять транс-сплайсинг двух отдельных белков, каждый из которых включает продукт слияния с DnaE-N или DnaE-C.

[0583] Дополнительные встречающиеся в природе или сконструированные последовательности разделенного интеина известны или могут быть реализованы на основе описанных здесь или доступных в данной области последовательностей цельного интеина. Примеры последовательностей разделенного интеина можно найти в Stevens et al., “A promiscuous split intein with expanded protein engineering applications,” PNAS, 2017, Vol.114: 8538-8543; Iwai et al., “Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostc punctiforme, FEBS Lett, 580: 1853-1858, каждый из которых приведен здесь в качестве ссылки. Дополнительные последовательности разделенного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774, и EP2877490, содержание каждой из которых приведено здесь в качестве ссылки.

[0584] Кроме того, сплайсинг белка in trans был описан *in vivo* и *in vitro* (Shingledecker, et al., [0076] *Gene* 207:187 (1998), Southworth, et al., *EMBO J.* 17:918 (1998); Mills, et al., *Proc. Natl. Acad. Sci. USA,* 95:3543-3548 (1998); Lew, et al., *J. Biol. Chem.,* 273:15887-15890 (1998); Wu, et al., *Biochim. Biophys. Acta* 35732:1 (1998b), Yamazaki, et al., *J. Am. Chem. Soc.* 120:5591 (1998), Evans, et al., *J. Biol. Chem.* 275:9091 (2000); Otomo, et al., *Biochemistry* 38:16040-16044 (1999); Otomo, et al., *J. Biolmol. NMR* 14:105-114 (1999); Scott, et al., *Proc. Natl. Acad. Sci. USA* 96:13638-13643 (1999)), при этом он обеспечивает возможность экспрессировать белок в виде двух неактивных фрагментов, которые далее подвергаются лигирования для образования функционального продукта, например как показано на Фиг.66 и 67в плане образования полного слитого белка PE из двух отдельно экспрессируемых половин.

E. Домен взаимодействия РНК-белок

[0585] В различных вариантах осуществления два отдельных домена белка (например домен Cas9 и домен полимеразы) могут быть солокализованы друг с другом для образования функционального комплекса (аналогично функции слитого белка, включающего два данных отдельных домена белка) путем использования "системы рекрутирования РНК-белок", например "способа мечения MS2". Такие системы в целом обеспечивают мечение одного домена белка "доменом взаимодействия РНК-белок" (также известным как "домен рекрутирования РНК-белка"), а другого — "РНК-связывающим белком", который специфично распознает и связывается с доменом взаимодействия РНК-белок, например определенной структурой в виде шпильки. Данные типы систем могут быть использованы для солокализации доменов праймерного редактора, а также рекрутирования дополнительных видов функциональности к праймерному редактору, например домена UGI. В одном варианте осуществления система мечения MS2 основана на естественном взаимодействии белка оболочки бактериофага MS2 ("MCP" или "MS2cp") со структурой в виде шпильки, присутствующей в геноме фага, то есть "шпилькой MS2". В случае шпильки MS2 она распознается и связывается белком оболочки бактериофага MS2 (MCP). Таким образом, в одном примере сценария продукт слияния в виде дезаминазы-MS2 может рекрутировать продукт слияния Cas9-MCP.

[0586] Другие модулярные домены взаимодействия РНК-белок описаны в данной области, например в Johansson et al. , “RNA recognition by the MS2 phage coat protein,” *Sem Virol.*, 1997, Vol. 8(3): 176-185; Delebecque et al., “Organization of intracellular

reactions with rationally designed RNA assemblies," *Science*, 2011, Vol. 333: 470-474; Mali et al., "Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," *Nat. Biotechnol.*, 2013, Vol.31: 833-838; и Zalatan et al., "Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds," *Cell*, 2015, Vol.160: 339-350, каждый из которых полностью приведен здесь в качестве примера. Другие системы включают шпильку PP7, которая специфически рекрутирует белок PCP, и шпильку "com", которая специфически рекрутирует белок Com. См. Zalatan et al.

[0587] Нуклеотидная последовательность шпильки MS2 (или эквивалентным образом обозначаемой как "аптамер MS2") является следующей:
GCCAACATGAGGATCACCCATGTCTGCAGGGCC (SEQ ID NO: 3871).

[0588] Аминокислотная последовательность MCP или MS2cp является следующей:
GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
NRKYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIVKAMQGL
LKDGNPIPSAIAANSGIY (SEQ ID NO: 3872).

F. Домен UGI

[0589] В других вариантах осуществления описанные здесь праймерные редакторы могут включать один или несколько доменов ингибитора урацилгликозилазы. Термин "ингибитор урацилгликозилазы (UGI)" или "домен UGI" в соответствии с используемым здесь значением обозначает белок, который способен ингибировать фермент эксцизионной репарации оснований урацил-ДНК-гликозилазы. В некоторых вариантах осуществления домен UGI включает UGI дикого типа или UGI согласно SEQ ID NO: 3873. В некоторых вариантах осуществления предусмотренные здесь белки UGI включают фрагменты UGI и белки, являющиеся гомологичными UGI или фрагменту UGI. Например, в некоторых вариантах осуществления домен UGI включает фрагмент аминокислотной последовательности, указанной в SEQ ID NO: 3873. В некоторых вариантах осуществления фрагмент UGI включает аминокислотную последовательность, которая включает, по меньшей мере, 60%, по меньшей мере, 65%, по меньшей мере, 70%, по меньшей мере, 75%, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 96%, по меньшей мере, 97%, по меньшей мере, 98%, по меньшей мере, 99%, или, по меньшей мере, 99,5% аминокислотной последовательности, указанной в SEQ ID NO: 3873. В некоторых вариантах осуществления UGI включает

аминокислотную последовательность, являющуюся гомологичной аминокислотной последовательности, указанной в SEQ ID NO: 3873, или аминокислотную последовательность, являющуюся гомологичной фрагменту аминокислотной последовательности, указанной в SEQ ID NO: 3873. В некоторых вариантах осуществления белки, включающие UGI, или фрагменты UGI, или гомологи UGI, или фрагменты UGI, обозначаются как "варианты UGI". Вариант UGI имеют такую же гомологию, что и UGI или его фрагмент. Например, вариант UGI является, по меньшей мере, на 70% идентичным, по меньшей мере, на 75% идентичным, по меньшей мере, на 80% идентичным, по меньшей мере, на 85% идентичным, по меньшей мере, на 90% идентичным, по меньшей мере, на 95% идентичным, по меньшей мере, на 96% идентичным, по меньшей мере, на 97% идентичным, по меньшей мере, на 98% идентичным, по меньшей мере, на 99% идентичным, по меньшей мере, на 99,5% идентичным, или, по меньшей мере, на 99,9% идентичными UGI дикого типа или UGI согласно SEQ ID NO: 3873. В некоторых вариантах осуществления вариант UGI включает фрагмент UGI, так что фрагмент является, по меньшей мере, на 70% идентичным, по меньшей мере, на 75% идентичным, по меньшей мере, на 80% идентичным, по меньшей мере, на 85% идентичным, по меньшей мере, на 90% идентичным, по меньшей мере, на 95% идентичным, по меньшей мере, на 96% идентичным, по меньшей мере, на 97% идентичным, по меньшей мере, на 98% идентичным, по меньшей мере, на 99% идентичным, по меньшей мере, на 99,5% идентичным, или, по меньшей мере, на 99,9% идентичными соответствующему фрагменту UGI дикого типа или UGI согласно SEQ ID NO: 3873. В некоторых вариантах осуществления UGI включает следующую аминокислотную последовательность:

[0590] Ингибитор урацил-ДНК-гликозилазы:

[0591] >sp|P14739|UNGI_BPPB2 MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLT SDAPEYKPWALVIQDSNGENKIKML (SEQ ID NO: 3873).

[0592] Описанные здесь праймерные редакторы могут включать более чем один домен UGI, которые могут быть разделены одним или несколькими линкерами, как описано здесь.

G. Дополнительные элементы PE

[0593] В определенных вариантах осуществления описанные здесь праймерные редакторы могут включать ингибитор репарации оснований. Термин "ингибитор репарации оснований" или "IBR" обозначает белок, который способен ингибировать активность фермента репарации нуклеиновой кислоты, например фермента эксцизионной репарации оснований. В некоторых вариантах осуществления IBR представляет собой ингибитор эксцизионной репарации основания OGG. В некоторых вариантах осуществления IBR представляет собой ингибитор эксцизионной репарации основания ("iBER"). Примеры ингибиторов эксцизионной репарации оснований включают ингибиторы APE1, Endo III, Endo IV, Endo V, Endo VIII, Fpg, hOGG1, hNEIL1, T7 EndoI, T4PDG, UDG, hSMUG1, и hAAG. В некоторых вариантах осуществления IBR представляет собой ингибитор Endo V или hAAG. В некоторых вариантах осуществления IBR представляет собой iBER, который может представлять собой каталитическую неактивную гликозилазу или каталитически неактивную деоксигеназу или низкомолекулярный или пептидный ингибитор оксидазы или его варианты. В некоторых вариантах осуществления IBR представляет собой iBER, который может представлять собой ингибитор TDG, ингибитор MBD4, или ингибитор фермента AlkBH. В некоторых вариантах осуществления IBR представляет собой iBER, который включает каталитически неактивный TDG или каталитически неактивный MBD4. Пример каталитически неактивного TDG представляет собой мутант N140A согласно SEQ ID NO: 3872 (TDG человека).

[0594] Некоторые примеры гликозилаз приведены ниже. Каталитически неактивные варианты любого из данных доменов гликозилазы представляют собой iBER, которые могут быть слиты с napDNAbp или доменом полимеразы праймерных редакторов, предусмотренных в рамках данного изобретения.

[0595] OGG (человек)

[0596] MPARALLPRRMGHRTLASTPALWASIPCPRSELRLDLVLPSGQSFRWREQSPA HWSGVLADQVWTLTQTEEQLHCTVYRGDKSQASRPTPDELEAVRKYFQLDVTLAQ LYHHWGSVDSHFQEVAQKFQGVRLLRQDPIECLFSFICSSNNNIARITGMVERLCQAF GPRLIQLDDVTYHGFPSLQALAGPEVEAHLRKLGLGYRARYVSASARAILEEQGGLA WLQQLRESSYEEAHKALCILPGVGTKVADCICLMALDKPQAVPVDVHMWHIAQRD YSWHPTTSQAKGPSPQTNKELGNFFRSLWGPYAGWAQAVLFSADLRQSRHAQEPPA KRRKGSKGPEG (SEQ ID NO: 3869)

[0597] MPG (человек)

[0598] MVTPALQMKKPKQFCRRMGQKKQRPARAGQPHSSSDAAQAPAEQPHSSSDA AQAPCPRERCLGPPTTPGPYRSIYFSSPKGHLTRLGLEFFDQPAVPLARAFLGQVLVR RLPNGTELRGRIVETEAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYF CMNISSQGDGACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQ ALAINKSFDQRDLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFY VRGSPWVSVVDRVAEQDTQA (SEQ ID NO: 3870)

[0599] MBD4 (человек)

[0600] MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQ MMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDV YFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSH LQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDV NFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQ KSQLDRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSA KDSEHNEKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGE KIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHD PWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLR AKTIVKFSDEYLTKQWKYPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHD WLWENHEKLSLS (SEQ ID NO: 3871)

[0601] TDG (человек)

[0602] MEAENAGSYSLQQAQAFYTFPFQQLMAEAPNMAVVNEQQMPEEVPAPAPA QEPVQEAPKGRKRKPRTTEPKQPVEPKKPVESKKSGKSAKSKEKQEKITDTFKVKRK VDRFNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFM SGLSEVQLNHMDDHTLPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKY QPRIAVFNGKCIYEIFSKEVFGVKVKNLEFGLQPHKIPDTETLCYVMPSSSARCAQFPR AQDKVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLAQEDAKKMAVKEEKYDP GYEAAYGGAYGENPCSSEPCGFSSNGLIESVELRGESAFSGIPNGQWMTQSFTDQIPS FSNHCGTQEQEEESHA (SEQ ID NO: 3872)

[0603] В некоторых вариантах осуществления описанные здесь слитые белки могут включать один или несколько гетерологичных доменов белка (например приблизительно или более чем 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, или более доменов помимо других компонентов праймерного редактора). Слитый белок может включать любую дополнительную последовательность белка и, необязательно, последовательность линкера между любыми двумя доменами. Другие примеры элементов, которые могут

присутствовать, представляют собой последовательности локализации, например последовательности цитоплазматической локализации, последовательности экспорта, например последовательности ядерного экспорта, или другие последовательности локализации, а также метки последовательностей, которые могут использоваться для солюбилизации, очистки, или детектирования слитых белков.

[0604] Примеры доменов белка, которые могут быть слиты с праймерным редактором или его компонентом (например доменом napDNAbp, полимеразным доменом, или доменом NLS) включают, без ограничения, эпитопные метки и последовательности репортерных генов. Неограничивающие примеры эпитопных меток включают гистидиновые (His) метки, метки V5, метки FLAG, метки гемагглютинина гриппа (HA), метки Myc, метки VSV-G, и тиоредоксиновые (Trx) метки. Примеры репортерных генов включают, но без ограничения, глутатион-5-трансферазу (GST), пероксидазу хрена (HRP), хлорамфениколацетилтрансферазу (CAT), бета-галактозидазу, бета-глюкуронидазу, люциферазу, зеленый флуоресцентный белок (GFP), HcRed, DsRed, голубой флуоресцентный белок (CFP), желтый флуоресцентный белок (YFP), и аутофлуоресцентные белки, включая синий флуоресцентный белок (BFP). Праймерный редактор может быть слит с генной последовательностью, кодирующей белок или фрагмент белка, который связывается с молекулами ДНК или связывается с другими клеточными молекулами, включая, но без ограничения, белок связывания мальтозы (MBP), S-метку, продукты слияния ДНК-связывающего домена Lex A (DBD), продукты слияния ДНК-связывающего домена GAL4, и продукты слияния белка BP16 вируса простого герпеса (HSV). Дополнительные домены, которые могут образовывать часть праймерного редактора, описаны в публикации патента США 2011/0059502, опубликованной 10 марта 2011 и полностью приведенной здесь в качестве ссылки.

[0605] В одном аспекте изобретения репортерный ген, который включает, но без ограничения, глутатион-5-трансферазу (GST), пероксидазу хрена (HRP), хлорамфениколацетилтрансферазу (CAT), бета-галактозидазу, бета-глюкуронидазу, люциферазу, зеленый флуоресцентный белок (GFP), HcRed, DsRed, голубой флуоресцентный белок (CFP), желтый флуоресцентный белок (YFP), и аутофлуоресцентные белки, включая синий флуоресцентный белок (BFP), может быть внесен в клетку для кодирования генного продукта, который служит в качества маркера для измерения изменения или модифицирования экспрессии генного

продукта. В определенных вариантах осуществления изобретения генный продукт представляет собой люциферазу. В дополнительном варианте осуществления изобретения экспрессия генного продукта является пониженой.

[0606] Подходящие предусмотренные здесь белковые метки включают, но без ограничения, метки на основе белка-носителя карбоксила биотина (BCCP), myc-метки, метки на основе кальмодулина, FLAG-метки, гемагглютининовые (HA) метки, полигистидиновые метки, также называемые гистидиновыми метками или His-метками, метки на основе связывающего мальтозу белка (MBP), nus-метки, метки на основе глутатион-S-трансферазы (GST), метки на основе зеленого флуоресцентного белка (GFP), тиоредоксиновые метки, S-метки, Softag (например Softag 1, Softag 3), strep-метки, метки на основе биотинлигазы, метки FlAsH, метки V5 и SBP-метки. Дополнительные подходящие последовательности будут очевидны специалистам в данной области. В некоторых вариантах осуществления слитый белок включает одну или несколько His меток.

[0607] В некоторых вариантах осуществления настоящего изобретения активность системы праймерного редактирования может временно регулироваться за счет изменения времени нахождения, количества, и/или активности экспрессируемых компонентов системы PE. Например, как описано здесь, PE может быть слит с доменом белка, который способен модифицировать внутриклеточное время полужизни PE. В определенных вариантах осуществления с использованием двух или более векторов (например векторной системы, в рамках которой описанные здесь компоненты кодируются на двух или более отдельных векторах), активность системы PE может временно регулироваться путем контролирования времени, за которое доставляются векторы. Например, в некоторых вариантах осуществления вектор, кодирующий нуклеазную систему, может доставлять PE перед вектором, кодирующим матрицу. В других вариантах осуществления вектор, кодирующий PEgRNA, может доставлять направляющую перед вектором, кодирующим систему PE. В некоторых вариантах осуществления векторы, кодирующие систему PE и PEgRNA, доставляются одновременно. В определенных вариантах осуществления одновременно доставляемый векторы временно доставляют, например, PE, PEgRNA, и/или компоненты направляющей нРНК второй цепи. В дополнительных вариантах осуществления НРК (такая как, например, нуклеазный транскрипт), транскрибированная из кодирующей последовательности на векторах, может дополнительно включать, по меньшей мере,

один элемент, который способен модифицировать внутриклеточное время полужизни РНК и/или модулировать управление трансляцией. В некоторых вариантах осуществления время полужизни РНК может быть повышено. В некоторых вариантах осуществления время полужизни РНК может быть снижено. В некоторых вариантах осуществления элемент может быть способен повышать стабильность РНК. В некоторых вариантах осуществления элемент может быть способен понижать стабильность РНК. В некоторых вариантах осуществления элемент может располагаться в 3' UTR РНК. В некоторых вариантах осуществления элемент может включать сигнал полиаденилирования (PA). В некоторых вариантах осуществления элемент может включать кэп, например конец в виде расположенной выше иРНК или PEgRNA. В некоторых вариантах осуществления РНК может не включать PA, так что она подвергается более быстрому расщеплению в клетке после транскрипции. В некоторых вариантах осуществления элемент может включать, по меньшей мере, один элемент с высоким содержанием AU (ARE). ARE могут быть связаны связывающими ARE белками (ARE-BP) таким образом, который зависит от типа ткани, типа клетки, времени, клеточной локализации, и внешней среды. В некоторых вариантах осуществления дестабилизирующий элемент может промотировать распад РНК, влиять на стабильность РНК, или активировать трансляцию. В некоторых вариантах осуществления ARE может иметь длину в виде от 50 до 150 нуклеотидов. В некоторых вариантах осуществления ARE может включать, по меньшей мере, одну копию последовательности AUUUA. В некоторых вариантах осуществления, по меньшей мере, один ARE может быть добавлен к 3' UTR РНК. В некоторых вариантах осуществления элемент может представлять собой вирус гепатита Вудчака (WHP).

[0608] Посттрансляционный регуляторный элемент (WPRE), который создает третичную структуру для стимулирования экспрессии из транскрипта. В дополнительных вариантах осуществления элемент представляет собой модифицированную и/или усеченную последовательность WPRE, которая способна стимулировать экспрессию из транскрипта, как описано, например, в Zufferey et al., J Virol, 73(4): 2886-92 (1999) and Flajolet et al., J Virol, 72(7): 6175-80 (1998). В некоторых вариантах осуществления WPRE или его эквивалент может быть добавлен к 3' UTR РНК. В некоторых вариантах осуществления элемент может быть выбран из других мотивов последовательности РНК с высоким содержанием быстро или медленно разлагающихся транскриптов.

[0609] В некоторых вариантах осуществления вектор, кодирующий PE или PEgRNA, может подвергаться саморазрушению за счет расщепления последовательности-мишени, присутствующей на векторе, за счет системы PE. Расщепление может устранять непрерывную транскрипцию PE или PEgRNA из вектора. Несмотря на то, что транскрипция может осуществляться на линеаризованном векторе в течение некоторого количества времени, экспрессируемые транскрипты или белки, подвергаемые внутриклеточному расщеплению, будут иметь меньше времени для продуцирования нецелевых эффектов без непрерывной поддержки на основе экспрессии кодирующих векторов.

[6] PEgRNA

[0610] Описанная здесь система праймерного редактирования включает применение любых подходящих PEgRNA. Авторы обнаружили, что механизм мишень-праймированной обратной транскрипции (TPRT) может быть использован или адаптирован для осуществления точного и универсального геномного редактирования на основе CRISPR/Cas путем применения специальным образом сконфигурированной направляющей РНК, включающей последовательность матрицы обратной транскрипции (RT), которая кодирует требуемое нуклеотидное изменение. В рамках изобретения данная специальным образом сконфигурированная направляющая РНК обозначается как "удлиненная направляющая РНК" или "PEgRNA", поскольку последовательность RT матрицы может быть представлена в виде удлинения стандартной или традиционной молекулы направляющей РНК. Изобретение охватывает любую подходящую конфигурацию или строение удлиненной направляющей РНК.

Архитектура PEgRNA

[0611] На Фиг.3A показан один вариант осуществления удлиненной направляющей РНК, который может применяться в рамках описанной здесь системы праймерного редактирования, где традиционная направляющая РНК (зеленая часть) включает протоспейсерную последовательность длиной ~20 nt и участок в виде ядра нРНК, который связывается с napDNAbp. В данном варианте осуществления направляющая РНК включает удлиненный сегмент РНК на 5′-конце, то есть 5′ удлинение. В данном варианте осуществления 5′ удлинение включает последовательность матрицы обратной транскрипции, участок связывания праймера обратной транскрипции и необязательную последовательность линкера длиной 5-20 нуклеотидов. Как показано

на Фиг.1A-1B, участок связывания праймера RT гибридизируется со свободным 3′-концом, который образуется после осуществления разрыва в не являющейся мишенью цепи R-петли, тем самым обеспечивая праймирование обратной транскриптазы для полимеризации ДНК в направлении 5′-3′.

[0612] На Фиг.3B показан другой вариант осуществления удлиненной направляющей РНК, который может применяться в рамках описанной здесь системы праймерного редактирования, где традиционная направляющая РНК (зеленая часть) включает протоспейсерную последовательность длиной ~20 nt и ядро нРНК, которое связывается с napDNAbp. В данном варианте осуществления направляющая РНК включает удлиненный сегмент РНК на 3′-конце, то есть 3′ удлинение. В данном варианте осуществления 3′ удлинение включает последовательность матрицы обратной транскрипции, участок связывания праймера обратной транскрипции. Как показано на Фиг.1C-1D, участок связывания праймера RT гибридизируется со свободным 3′-концом, который образуется после осуществления разрыва в не являющейся мишенью цепи R-петли, тем самым обеспечивая праймирование обратной транскриптазы для полимеризации ДНК в направлении 5′-3′.

[0613] На Фиг.3C показан другой вариант осуществления удлиненной направляющей РНК, который может применяться в рамках описанной здесь системы праймерного редактирования, где традиционная направляющая РНК (зеленая часть) включает протоспейсерную последовательность длиной ~20 nt и ядро нРНК, которое связывается с napDNAbp. В данном варианте осуществления направляющая РНК включает удлиненный сегмент РНК в межмолекулярном положении в ядре нРНК, то есть внутримолекулярное удлинение. В данном варианте осуществления внутримолекулярное удлинение включает последовательность матрицы обратной транскрипции и участок связывания праймера обратной транскрипции. Участок связывания праймера RT гибридизируется со свободным 3′-концом, который образуется после осуществления разрыва в не являющейся мишенью цепи R-петли, тем самым обеспечивая праймирование обратной транскриптазы для полимеризации ДНК в направлении 5′-3′.

[0614] В одном варианте осуществления положение межмолекулярного удлинения РНК не соответствует протоспейсерной последовательности направляющей РНК. В другом варианте осуществления положение межмолекулярного удлинения РНК соответствует ядру нРНК. В другом варианте осуществления положение

межмолекулярного удлинения РНК соответствует любому положению в рамках молекулы направляющей РНК за исключением протоспейсерной последовательности или положения, которое нарушает протоспейсерную последовательность.

[0615] В одном варианте осуществления межмолекулярное удлинение РНК подвергается вставке ниже 3′-конца протоспейсерной последовательности. В другом варианте осуществления межмолекулярное удлинение РНК подвергается вставке, по меньшей мере, на 1 нуклеотид, по меньшей мере, на 2 нуклеотида, по меньшей мере, на 3 нуклеотида, по меньшей мере, на 4 нуклеотида, по меньшей мере, на 5 нуклеотидов, по меньшей мере, на 6 нуклеотидов, по меньшей мере, на 7 нуклеотидов, по меньшей мере, на 8 нуклеотидов, по меньшей мере, на 9 нуклеотидов, по меньшей мере, на 10 нуклеотидов, по меньшей мере, на 11 нуклеотидов, по меньшей мере, на 12 нуклеотидов, по меньшей мере, на 13 нуклеотидов, по меньшей мере, на 14 нуклеотидов, по меньшей мере, на 15 нуклеотидов, по меньшей мере, на 16 нуклеотидов, по меньшей мере, на 17 нуклеотидов, по меньшей мере, на 18 нуклеотидов, по меньшей мере, на 19 нуклеотидов, по меньшей мере, на 20 нуклеотидов, по меньшей мере, на 21 нуклеотид, по меньшей мере, на 22 нуклеотида, по меньшей мере, на 23 нуклеотида, по меньшей мере, на 24 нуклеотида, по меньшей мере, на 25 нуклеотидов ниже 3′-конца протоспейсерной последовательности.

[0616] В других вариантах осуществления межмолекулярное удлинение РНК подвергается вставке в нРНК, что соответствует части направляющей РНК, соответствующей или включающей tracrРНК, которая связывает и/или взаимодействует с белком Cas9 или его эквивалентом (то есть другим napDNAbp). Вставка межмолекулярного удлинения РНК предпочтительно не нарушает или минимально нарушает взаимодействие между частью в виде tracrРНК и napDNAbp.

[0617] Длина удлинения РНК может представлять собой любую подходящую длину. В различных вариантах осуществления удлинение РНК может иметь длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по

меньшей мере, 22 нуклеотида, по меньшей мере, 23 нуклеотида, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0618] Последовательность RT матрицы также может иметь любую подходящую длину. Например, последовательность RT матрицы может иметь длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0619] В других вариантах осуществления последовательность участка связывания праймера обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по

меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0620] В других вариантах осуществления необязательная линкерная или спейсерная последовательность имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 50 нуклеотидов, по меньшей мере, 60 нуклеотидов, по меньшей мере, 70 нуклеотидов, по меньшей мере, 80 нуклеотидов, по меньшей мере, 90 нуклеотидов, по меньшей мере, 100 нуклеотидов, по меньшей мере, 200 нуклеотидов, по меньшей мере, 300 нуклеотидов, по меньшей мере, 400 нуклеотидов, или, по меньшей мере, 500 нуклеотидов.

[0621] Последовательность RT матрицы в определенных вариантах осуществления кодирует молекулу одноцепочечной ДНК, которая является гомологичной не являющейся мишенью цепи (и, таким образом, комплементарной соответствующему участку цепи-мишени), но включает одно или несколько нуклеотидных изменений. Указанное, по меньшей мере, одно нуклеотидное изменение может включать одно или несколько нуклеотидных изменений одного основания, одну или несколько делеций, и одну или несколько вставок.

[0622] Как изображено на Фиг.1G, синтезированный продукт в виде одноцепочечной ДНК-последовательности RT матрицы является гомологичным не являющейся мишенью цепи и содержит одно или несколько нуклеотидных изменений. Продукт в виде одноцепочечной ДНК-последовательности RT матрицы гибридизируется равновесным образом с комплементарной последовательностью цепи-мишени, тем самым вытесняя гомологичную последовательность эндогенной цепи-мишени. Вытесненная эндогенная цепь может быть обозначена в некоторых вариантах осуществления как 5′ эндогенный ДНК-флэп. Данный 5′ эндогенный ДНК-флэп

может быть удален с помощью 5′ флэп-эндонуклеазы (например FEN1), при этом продукт в виде одноцепочечной ДНК, являющийся теперь гибридизированным с эндогенной цепью-мишенью, может быть лигирован, тем самым создавая несовпадение между эндогенной последовательностью и новой синтезированной цепью. Несовпадение может быть подвергнуто коррекции с помощью внутренних процессов репарации и/или репликации ДНК клетки.

[0623] В различных вариантах осуществления нуклеотидная последовательность последовательности RT матрицы соответствует нуклеотидной последовательность не являющейся мишенью цепи, которая становится вытесненной в виде 5′ флэпа и которая перекрывает подвергаемый редактированию участок.

[0624] В различных вариантах осуществления удлиненных направляющих РНК последовательность матрицы обратной транскрипции может кодировать одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение. Одноцепочечный ДНК-флэп может вытеснять эндогенную одноцепочечную ДНК в участке разрыва. Вытесненная эндогенная одноцепочечная ДНК в участке разрыва может иметь 5′-конец и образовывать эндогенный флэп, который может быть подвергнут эксцизии клеткой. В различных вариантах осуществления эксцизия 5′-концевого эндогенного флэпа может обеспечивать стимулирование образования продукта, поскольку удаление 5′-концевого эндогенного флэпа стимулирует гибридизирование одноцепочечного 3′ ДНК-флэпа с соответствующей комплементарной цепью ДНК и встраивание или ассимилирование требуемого нуклеотидного изменения, которое несет одноцепочечный 3′ ДНК-флэп, в ДНК-мишени.

[0625] В различных вариантах осуществления удлиненных направляющих РНК клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[0626] В других вариантах осуществления требуемое нуклеотидное изменение вносится в окне редактирования, которое соответствует приблизительно от -5 до +5 относительно участка разрыва, приблизительно от -10 до +10 относительно участка разрыва, приблизительно от -20 до +20 относительно участка разрыва, приблизительно от -30 до +30 относительно участка разрыва, приблизительно от -40 до +40 относительно участка разрыва, приблизительно от -50 до +50 относительно участка

разрыва, приблизительно от -60 до +60 относительно участка разрыва, приблизительно от -70 до +70 относительно участка разрыва, приблизительно от -80 до +80 относительно участка разрыва, приблизительно от -90 до +90 относительно участка разрыва, приблизительно от -100 до +100 относительно участка разрыва, приблизительно от -200 до +200 относительно участка разрыва.

В других вариантах осуществления требуемое нуклеотидное изменение вносится в окне редактирования, которое соответствует приблизительно от +1 до +2 от участка разрыва, или приблизительно от +1 до +3, от +1 до +4, от +1 до +5, от +1 до +6, от +1 до +7, от +1 до +8, от +1 до +9, от +1 до +10, от +1 до +11, от +1 до +12, от +1 до +13, от +1 до +14, от +1 до +15, от +1 до +16, от +1 до +17, от +1 до +18, +1 to +19, от +1 до +20, от +1 до +21, от +1 до +22, от +1 до +23, от +1 до +24, от +1 до +25, от +1 до +26, от +1 до +27, от +1 до +28, от +1 до +29, от +1 до +30, от +1 до +31, от +1 до +32, от +1 до +33, от +1 до +34, от +1 до +35, от +1 до +36, от +1 до +37, от +1 до +38, от +1 до +39, от +1 до +40, от +1 до +41, от +1 до +42, от +1 до +43, от +1 до +44, от +1 до +45, от +1 до +46, от +1 до +47, от +1 до +48, от +1 до +49, от +1 до +50, от +1 до +51, от +1 до +52, от +1 до +53, от +1 до +54, от +1 до +55, от +1 до +56, от +1 до +57, от +1 до +58, от +1 до +59, от +1 до +60, от +1 до +61, от +1 до +62, от +1 до +63, от +1 до +64, от +1 до +65, от +1 до +66, от +1 до +67, от +1 до +68, от +1 до +69, от +1 до +70, от +1 до +71, от +1 до +72, от +1 до +73, от +1 до +74, от +1 до +75, от +1 до +76, от +1 до +77, от +1 до +78, от +1 до +79, от +1 до +80, от +1 до +81, от +1 до +82, от +1 до +83, от +1 до +84, от +1 до +85, от +1 до +86, от +1 до +87, от +1 до +88, от +1 до +89, от +1 до +90, от +1 до +90, от +1 до +91, от +1 до +92, от +1 до +93, от +1 до +94, от +1 до +95, от +1 до +96, от +1 до +97, от +1 до +98, от +1 до +99, от +1 до +100, от +1 до +101, от +1 до +102, от +1 до +103, от +1 до +104, от +1 до +105, от +1 до +106, от +1 до +107, от +1 до +108, от +1 до +109, от +1 до +110, от +1 до +111, от +1 до +112, от +1 до +113, от +1 до +114, от +1 до +115, от +1 до +116, от +1 до +117, от +1 до +118, от +1 до +119, от +1 до +120, от +1 до +121, от +1 до +122, от +1 до +123, от +1 до +124, или от +1 до +125 от участка разрыва.

[0627] В других вариантах осуществления требуемое нуклеотидное изменение вносится в окне редактирования, которое соответствует приблизительно от +1 до +2 от участка разрыва, или приблизительно от +1 до +5, от +1 до +10, от +1 до +15, от +1 до +20, от +1 до +25, от +1 до +30, от +1 до +35, от +1 до +40, от +1 до +45, от +1 до +50, от +1 до +55, от +1 до +100, от +1 до +105, от +1 до +110, от +1 до +115, от +1 до +120,

от +1 до +125, +1 to +130, от +1 до +135, от +1 до +140, от +1 до +145, от +1 до +150, от +1 до +155, от +1 до +160, от +1 до +165, от +1 до +170, от +1 до +175, от +1 до +180, от +1 до +185, от +1 до +190, от +1 до +195, или от +1 до +200 от участка разрыва.

[0628] В различных аспектах удлиненные направляющие РНК представляют собой модифицированные версии направляющей РНК. Направляющие РНК могут являться встречающимися в природе, экспрессируемыми из кодирующей нуклеиновой кислоты, или химически синтезированными. В данной области хорошо известны способы получения или иным образом синтезирования направляющих РНК и определения подходящей последовательности направляющей РНК, включая протоспейсерную последовательность, которая взаимодействует и гибридизируется с цепью-мишенью представляющего интерес геномного участка-мишени.

[0629] В различных вариантах осуществления конкретные аспекты конструирования последовательности направляющей РНК будут зависеть от нуклеотидной последовательности представляющего интерес геномного участка-мишени (то есть требуемого подвергаемого редактированию участка) и типа napDNAbp (например белка Cas9), присутствующего в рамках описанных здесь систем праймерного редактирования, среди других факторов, таких как расположение последовательностей PAM, процентная величина содержания G/C в последовательности-мишени, величина участков микрогомологии, вторичные структуры и так далее.

[0630] В целом, направляющая последовательность представляет собой любую полинуклеотидную последовательность, имеющую достаточную комплементарность относительно полинуклеотидной последовательности-мишени для гибридизирования с последовательностью-мишенью и прямого специфичного в плане последовательности связывания napDNAbp (например Cas9, гомолога Cas9, или варианта Cas9) с последовательностью-мишенью. В некоторых вариантах осуществления степень комплементарности между направляющей последовательностью и соответствующей ей последовательностью-мишенью при оптимальном выравнивании с помощью подходящего алгоритма выравнивания составляет приблизительно или более чем приблизительно 50%, 60%, 75%, 80%, 85%, 90%, 95%, 99% или более. Оптимальное выравнивание может быть определено с помощью любого подходящего алгоритма для выравнивания последовательностей, неограничивающие примеры которых включают алгоритм Смита-Уотермана, алгоритм Нидлмана-Вунша, алгоритмы на основе

преобразования Барроуза-Уилера (например выравнивание Барроуза-Уилера), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (доступен на soap.genomics.org.cn), и Maq (доступен на maq.sourceforge.net) В некоторых вариантах осуществления направляющая последовательность имеет длину приблизительно или более чем приблизительно 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, или более нуклеотидов.

[0631] В некоторых вариантах осуществления направляющая последовательность имеет длину менее чем приблизительно 75 50, 45, 40, 35, 30, 25, 20, 15, 12 или менее нуклеотидов. Способность направляющей последовательности направлять специфичное в плане последовательности связывание праймерного редактора (PE) с последовательностью-мишенью может быть проанализирована с помощью любого подходящего анализа. Например, компоненты праймерного редактора (PE), включая подвергаемую тестированию направляющую последовательность, могут быть перенесены в клетку-хозяина, имеющую соответствующую последовательность-мишень, например с помощью трансфекции с использованием векторов, кодирующих компоненты описанного здесь праймерного редактора (PE), с последующим анализом предпочтительного расщепления в рамках последовательности-мишени, например с помощью анализа Surveyor, как описано здесь. Аналогичным образом, расщепление полинуклеотидной последовательности-мишени может быть проанализировано в аналитической пробирке путем добавления последовательности-мишени, компонентов праймерного редактора (PE), включая подвергаемую тестированию направляющую последовательность и исходную направляющую последовательность, отличную от подвергаемой тестированию направляющей последовательности, и сравнения связывания или скорости расщепления в последовательности-мишени между реакциями подвергаемой тестированию и исходной направляющей последовательности. Также возможны другие анализы, при этом они будут очевидны специалистам в данной области.

[0632] Направляющая последовательность может быть выбрана для нацеливания на любую последовательность-мишень. В некоторых вариантах осуществления последовательность-мишень представляет собой последовательность в рамках генома клетки. Примеры последовательностей-мишеней включают последовательности, которые являются уникальными в геноме-мишени. Например для Cas9 S. pyogenes

уникальная последовательность-мишень в геноме может включать участок-мишень Cas9 в виде MMMMMMMMNNNNNNNNNNNNXGG (SEQ ID NO: 204), где NNNNNNNNNNNNXGG (SEQ ID NO: 205) (N представляет собой A, G, T, или C; а X может быть любым). Уникальная последовательность-мишень в геноме может включать участок-мишень Cas9 S. pyogenes в виде MMMMMMMMMNNNNNNNNNNNXGG (SEQ ID NO: 206), где NNNNNNNNNNNXGG (SEQ ID NO: 207) (N представляет собой A, G, T, или C; а X может быть любым). Для CRISPR 1 Cas9 S. thermophilus уникальная последовательность-мишень в геноме может включать участок-мишень Cas9 в виде MMMMMMMMNNNNNNNNNNNNXXAGAAW (SEQ ID NO: 208), где NNNNNNNNNNNNXXAGAAW (SEQ ID NO: 209) (N представляет собой A, G, T, или C; X может быть любым; а W представляет собой A или T). Уникальная последовательность-мишень в геноме может включать участок-мишень CRISPR 1 Cas9 S. thermophilus в виде MMMMMMMMMNNNNNNNNNNNXXAGAAW (SEQ ID NO: 210), где NNNNNNNNNNNXXAGAAW (SEQ ID NO: 211) (N представляет собой A, G, T, или C; X может быть любым; п W представляет собой A или T). Для Cas9 S. pyogenes уникальная последовательность-мишень в геноме может включать участок-мишень Cas9 в виде MMMMMMMMNNNNNNNNNNNNXGGXG (SEQ ID NO: 212), где NNNNNNNNNNNNXGGXG (SEQ ID NO: 213) (N представляет собой A, G, T, или C; а X может быть любым). Уникальная последовательность-мишень в геноме может включать участок-мишень Cas9 S. pyogenes в виде MMMMMMMMMNNNNNNNNNNNXGGXG (SEQ ID NO: 214), где NNNNNNNNNNNXGGXG (SEQ ID NO: 215) (N представляет собой A, G, T, или C; а X может быть любым). В каждой из данных последовательностей "M" может представлять собой A, G, T, или C, при это он не должен рассматриваться при идентификации последовательности в качестве уникальной.

[0633] В некоторых вариантах осуществления направляющая последовательность выбрана для снижения величины вторичной структуры в направляющей последовательности. Вторичная структура может быть определена с помощью любого подходящего алгоритма полинуклеотидного фолдинга. Некоторые программы основаны на вычислении минимальной свободной энергии Гиббса. Пример одного подобного алгоритма представляет собой mFold, как описано Zuker и Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Другой пример алгоритма фолдинга представляет собой

интернет-вебсервер RNAfold, разработанный в Институте теоретической химии Венского университета, основанный на использовании алгоритма предсказания центроидной структуры (см., например, A. R. Gruber et al., 2008, Cell 106(1): 23-24; и PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62). Дополнительные алгоритмы описаны в заявке на патент США 61/836,080; широкая ссылка BI-2013/004A, которая приведена здесь в качестве ссылки.

[0634] В целом, последовательность tracr mate включает любую последовательность, которая имеет достаточную комплементарность относительно последовательности tracr для промотирования одного из следующих действий: (1) эксцизии направляющей последовательности, фланкируемой последовательностями tracr mate, в клетке, содержащей соответствующую последовательность tracr; и (2) образования комплекса в последовательности-мишени, при этом комплекс включает последовательность tracr mate, гибридизированную с последовательностью tracr. В целом, степень комплементарности соответствует оптимальному выравниванию последовательности tracr mate и последовательности tracr в рамках длины более короткой из данных двух последовательностей. Оптимальное выравнивание может быть определено с помощью любого подходящего алгоритма выравнивания и может дополнительно учитывать вторичные структуры, например самокомплементарность в последовательности tracr или последовательности tracr mate. В некоторых вариантах осуществления степень комплементарности между последовательностью tracr и последовательностью tracr mate в рамках длины более короткой из них при оптимальном выравнивании составляет приблизительно или более чем приблизительно 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99%, или более. В некоторых вариантах осуществления последовательность tracr имеет длину приблизительно или более чем приблизительно 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50 или более нуклеотидов. В некоторых вариантах осуществления последовательность tracr и последовательность tracr mate содержатся в рамках одного транскрипта, так что гибридизация между ними двумя продуцирует транскрипт, имеющий вторичную структуру, например шпильку. Предпочтительные образующие петли последовательности для применения в рамках структур в виде шпилек имеют длину четыре нуклеотида и, наиболее предпочтительно, имеют последовательность GAAA. Однако могут применяться более длинные или короткие последовательности петель, а также альтернативные последовательности. Последовательности предпочтительно включают нуклеотидный триплет (например

AAA) и дополнительный нуклеотид (например C или G). Примеры образующих петли последовательностей включают CAAA и AAAG. В одном варианте осуществления изобретения транскрипт или транскрибированная полинуклеотидная последовательность имеет, по меньшей мере, две или более шпильки. В предпочтительных вариантах осуществления транскрипт имеет две, три, четыре или пять шпилек. В дополнительном варианте осуществления транскрипт имеет не более пяти шпилек. В некоторых вариантах осуществления отдельный транскрипт дополнительно включает последовательности терминации транскрипции; она предпочтительно представляет собой последовательность polyT, например шесть нуклеотидов T. Дополнительные неограничивающие примеры отдельных полинуклеотидов, включающих направляющую последовательность, последовательность tracr mate, и последовательность tracr, являются следующими (перечислены по направлению от 5' к 3'), где "N" обозначает основание направляющей последовательности, первый блок строчных букв обозначает последовательность tracr mate, а второй блок строчных букв обозначает последовательность tracr, и конечная последовательность poly-T соответствует терминатору транскрипции:
(1)NNNNNNNNGTTTTTGTACTCTCAAGATTTAGAAATAAATCTTGCAGAAGCTACA AAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTC GTTATTTAATTTTTT (SEQ ID NO: 216);

[0635]

(2)NNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAAGCTACAAA GATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTCGTT ATTTAATTTTTT (SEQ ID NO: 217);

[0636] (3)NNNNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAAGC TACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGT TTTTT (SEQ ID NO: 218);

[0637] (4)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 219);

[0638] (5)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGTTTTTTT (SEQ ID NO: 220); AND

[0639] (6) NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCATTTTTTTT (SEQ ID NO: 221).

[0640] В некоторых вариантах осуществления последовательности от (1) до (3) используются в комбинации с Cas9 из CRISPR1 S. thermophilus. В некоторых вариантах осуществления последовательности от (4) до (6) используются в комбинации с Cas9 из S. pyogenes. В некоторых вариантах осуществления последовательность tracr представляет собой отдельный транскрипт из транскрипта, включающего последовательность tracr mate.

[0641] Специалистам в данной области будет очевидно, что для нацеливания на любые слитые белки, включающие домен Cas9 и связывающий одноцепочечную ДНК белок, как описано здесь, для нацеливания на участок, например участок, включающий повергаемую редактированию точечную мутацию, обычно необходимо совместно экспрессировать слитый белок вместе с направляющей РНК, например sgРНК. Как более подробно описано в других разделах, направляющая РНК обычно включает каркас tracrРНК, обеспечивающий связывание Cas9, и направляющую последовательность, которая придает специфичность в плане последовательности слитому белку Cas9:фермент/домен для редактирования нуклеиновой кислоты.

[0642] В некоторых вариантах осуществления направляющая РНК включает структуру 5′-[направляющая последовательнсть]-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUU-3′ (SEQ ID NO: 222), где направляющая последовательность включает последовательность, которая комплементарна последовательности-мишени. Направляющая последовательность обычно имеет длину 20 нуклеотидов. Последовательности подходящих направляющих РНК для нацеливания слитых белков Cas9:фермент/домен для редактирования нуклеиновой кислоты на определенные геномные участки-мишени будут очевидны специалиста в данной области на основе настоящего описания. Такие последовательности подходящих направляющих РНК обычно включают направляющие последовательности, которые комплементарны нуклеиновой последовательности на расстоянии до 50 нуклеотидов выше или ниже подвергаемого редактированию нуклеотида-мишени. Некоторые примеры последовательностей направляющих ДНК, подходящих для нацеливания любых указанных здесь слитых

белков на определенные последовательности-мишени, приведены здесь. Дополнительные направляющие последовательности хорошо известны в данной области и могут применяться с описанным здесь праймерным редактором (PE).

[0643] В других вариантах осуществления PEgRNA включают изображенные на Фиг.3D.

[0644] В других вариантах осуществления PEgRNA могут включать изображенные на Фиг.3E.

[0645] На Фиг.3D приведена структура варианта осуществления предусмотренной здесь PEgRNA, которая может быть сконструирована в соответствии с методологией, определенной в рамках Примера 2. PEgRNA включает три основных компонента в направлении от 5′ к 3′, а именно: спейсер, ядро нРНК, и плечо удлинения на 3′-конце. Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5′ к 3′, а именно: участок связывания праймера (A), матрицу для редактирования (B), и плечо гомологии (C). Кроме того, PEgRNA может включать дополнительный 3′-концевой модификаторный участок (e1) и необязательно 5′-концевой модификаторный участок (e2). Кроме того, PEgRNA может включать сигнал терминации транскрипции на 3′-конце PEgRNA (не показан). Данные структурные элементы дополнительно определены здесь. Изображение структуры PEgRNA не является ограничивающим и охватывает вариации в плане расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми из показанных участков без ограничения в виде расположения на 3′- или 5′-концах.

[0646] На Фиг.3E приведена структура другого варианта осуществления предусмотренной здесь PEgRNA, которая может быть сконструирована в соответствии с методологией, определенной в рамках Примера 2. PEgRNA включает три основных компонента в направлении от 5′ к 3′, а именно: спейсер, ядро нРНК, и плечо удлинения на 3′-конце. Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5′ к 3′, а именно: участок связывания праймера (A), матрицу для редактирования (B), и плечо гомологии (C). Кроме того, PEgRNA может включать необязательный 3′-концевой модификаторный участок (e1) и необязательный 5′-концевой модификаторный участок (e2). Кроме того, PEgRNA может включать сигнал терминации транскрипции на 3′-конце PEgRNA (не показан). Данные структурные элементы дополнительно определены здесь. Изображение

структуры PEgRNA не является ограничивающим и охватывает вариации в плане расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми из показанных участков без ограничения в виде расположения на 3′- или 5′-концах.

Усовершенствования PEgRNA

[0647] PEgRNA также могут включать дополнительные усовершенствования конструкции, связанные с модификацией свойств и/или характеристик PEgRNA, тем самым повышая эффективность праймерного редактирования. В различных вариантах осуществления данные усовершенствования могут относится к одной или нескольким из ряда различных категорий, включая, но без ограничения: (1) конструкции для обеспечения эффективной экспрессии функциональных PEgRNA из не относящихся к полимеразе III (pol III) промоторов, которые будут обеспечивать экспрессию более длинных PEgRNA без обременительных требований в плане последовательности; (2) усовершенствования ядра, Cas9-связывающего каркаса PEgRNA, которые могут повысить эффективность; (3) модификации PEgRNA для повышения процессирования RT, обеспечивающие вставку более длинных последовательностей в геномных локусах-мишенях; и (4) добавление мотивов РНК на 5' или 3'-концы PEgRNA, которые повышают стабильность PEgRNA, стимулируют процессирование RT, предотвращают неправильный фолдинг PEgRNA, или рекрутируют дополнительные факторы, являющиеся важными для геномного редактирования.

[0648] В одном варианте осуществления PEgRNA может быть сконструирована с pol III промоторами для повышения экспрессии более длинной PEgRNA с более крупными плечами удлинения. sgРНК обычно экспрессируются из промотора мяРНК U6. Данный промотор рекрутирует pol III для экспрессии ассоциированной РНК и может применяться для экспрессии коротких РНК, которые остаются в ядре. Однако pol III не является высокопроцессивной и не способна экспрессировать РНК с длиной более чем несколько сотен нуклеотидов на уровнях, требуемых для эффективного геномного редактирования. Кроме того, pol III может вызывать остановку или терминирование на длинных участках в виде U, что потенциально ограничивает разнообразие цепей, которые могут быть вставлены с помощью PEgRNA. Были проанализированы другие промоторы, которые рекрутируют полимеразу II (например pCMV) или полимеразу I (например промотор мяРНК U1) на их способность экспрессировать более длинные sgРНК. Однако данные промоторы обычно частично транскрибируются, что будет

приводить к дополнительной последовательности 5' спейсера в экспрессируемой PEgRNA, что, как было показано, приводит к заметному снижению активности Cas9:sgРНК в зависимости от участка. Кроме того, в то время как транскрибированные из pol III PEgRNA могут просто приводить к терминации на участке из 6-7 U, PEgRNA, транскрибированные из pol II или pol I, будут требовать другого сигнала терминации. Зачастую такие сигналы также приводят к полиаденилированию, которое приведет к нежелательному транспорту PEgRNA из ядра. Аналогичным образом, РНК, экспрессируемые из pol II промоторов, например pCMV, обычно являются 5'-кэпированными, что также приводит к их экспорту из ядра.

[0649] Ранее Rinn и коллеги провели скрининг ряда платформ экспрессии на предмет продуцирования sgРНК, меченных длинными некодирующими РНК (днкРНК)[183]. Данные платформы включают РНК, экспрессированные из pCMV, которые приводят к терминации в элементе ENE из нкРНК MALAT1 человека[184], элементе PAN ENE из KSHV[185], или 3' боксе из sgРНК U1[186]. Следует отметить, что нкРНК MALAT1 и PAN ENE образуют тройные спирали, защищающие polyA-хвост [184, 187]. Данные конструкты также могут повышать стабильность РНК. Предполагается, что данные системы экспрессии также обеспечат возможность экспрессии более длинных PEgRNA.

[0650] Кроме того, был разработан ряд способов для расщепления части pol II промотора, которая будет транскрибирована в виде части PEgRNA, обеспечивая добавление либо саморасщепляемого рибозима, например рибозимов типа "головка молотка" (hammerhead)[188], "пистолет" (pistol)[189], "топор" (hatchet)[189], "шпилька" (hairpin)[190], VS[191], "твистер"[192], или "твистер-систер" (twister sister)[192], или других саморасщепляемых элементов для процессирования транскрибированной направляющей, или шпильки, которая распознается Csy4[193] и также ведет к процессированию направляющей. Кроме того, существует гипотеза, что встраивание нескольких мотивов ENE может приводить к повышенной экспрессии PEgRNA и стабильности, как ранее было продемонстрировано для РНК KSHV PAN и элемента[185]. Также ожидается, что циркуляризация PEgRNA в форме циркулярной интронной РНК (циРНК) также может приводить к повышенной экспрессии РНК и стабильности, а также ядерной локализации[194].

[0651] В различных вариантах осуществления PEgRNA может включать различные указанные выше элементы, как показано на примере следующей последовательности.

[0652] Неограничивающий пример 1 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и MALAT1 ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTT
CTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGAT
GCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCT
TTGCTTTGACT (SEQ ID NO: 223)

[0653] Неограничивающий пример 2 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и PAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA
TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 224)

[0654] Неограничивающий пример 3 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3xPAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC

CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA
TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAAACACACTGTT
TTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGG
TTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCT
AGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGG
TCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAATCTCTCTGTTTTGGCTGGGTT
TTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTATCCCAG
TGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAAT
TAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGGTCAATGTTTAAT
CCATAAAAAAAAAAAAAAAAAA (SEQ ID NO: 225)

[0655] Неограничивающий пример 4 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3′ бокса

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTGGTTT
GTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 226)

[0656] Неограничивающий пример 5 - платформа экспрессии PEgRNA, состоящая из pU1, Csy4 шпильки, PEgRNA, и 3′ бокса

CTAAGGACCAGCTTCTTTGGGAGAGAACAGACGCAGGGGCGGGAGGGAAAAAG
GGAGAGGCAGACGTCACTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCGGTTGA
GTGGCAGAAAGGCAGACGGGGACTGGGCAAGGCACTGTCGGTGACATCACGGAC
AGGGCGACTTCTATGTAGATGAGGCAGCGCAGAGGCTGCTGCTTCGCCACTTGCT
GCTTCACCACGAAGGAGTTCCCGTGCCCTGGGAGCGGGTTCAGGACCGCTGATCG
GAAGTGAGAATCCCAGCTGTGTGTCAGGGCTGGAAAGGGCTCGGGAGTGCGCGG
GGCAAGTGACCGTGTGTGTAAAGAGTGAGGCGTATGAGGCTGTGTCGGGGCAGA

GGCCCAAGATCTCAGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG
TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAA
AGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTCAGCAAG
TTCAGAGAAATCTGAACTTGCTGGATTTTTGGAGCAGGGAGATGGAATAGGAGCT
TGCTCCGTCCACTCCACGCATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGC
ACCCACTTTCTGGAGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTT
GTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 227).

[0657] В различных других вариантах осуществления PEgRNA может быть усовершенствована путем введения усовершенствований в последовательности каркаса или ядра. Это может быть осуществлено путем внесения известных

[0658] Ядро, представляющее собой Cas9-связывающий каркас PEgRNA, может быть вероятно усовершенствовано для повышения активности PE. Некоторые подобные подходы уже были продемонстрированы. Например, первый парный элемент каркаса (P1) содержит парный элемент GTTTT-AAAAC. Было показано, что такие повторы T приводят к остановке pol III и преждевременной терминации РНК транскрипта. Было показано, что разумная мутация одной из пар T-A на пару G-C в этой части P1 повышает активность sgРНК, что свидетельствует о том, что данный подход также может применяться для PEgRNA[195]. Кроме того, также было показано, что повышение длины P1 стимулирует фолдинг sgРНК и ведет к повышенной активности[195], что свидетельствует о том, что оно является другой возможностью повысить активность PEgRNA. Примеры усовершенствований ядра могут включать:

[0659] PEgRNA, содержащую удлинение P1 длиной 6 nt

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGCTCATGAAAATGAGCTAGCAAG
TTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTC
TGCCATCAAAGCGTGCTCAGTCTGTTTTTTT (SEQ ID NO: 228)

[0660] PEgRNA, содержащую мутацию T-A на G-C в P1

GGCCCAGACTGAGCACGTGAGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC
GTGCTCAGTCTGTTTTTTT (SEQ ID NO: 229)

[0661] В различных других вариантах осуществления PEgRNA может быть усовершенствована путем внесения модификаций в участок матрицы для редактирования. С увеличением размера вставки, матрицей которого служит PEgRNA, она с большей вероятностью будет расщепляться эндонуклеазами, подвергаться спонтанному гидролизу или фолдингу во вторичные структуры, не способные подвергаться обратной транскрипции за счет RT или нарушающие фолдинг каркаса PEgRNA и последующее связывание Cas9-RT. Соответственно, является вероятным,

что для реализации крупных вставок может быть необходимым модифицирование матрицы PEgRNA, например для вставок целых генов. Некоторые стратегии для этого включают встраивание модифицированных нуклеотидов в синтетическую или полусинтетическую PEgRNA, которые делают РНК более устойчивой к расщеплению или гидролизу или менее подверженной включению ингибиторных вторичных структур[196]. Подобные модификации могут включать 8-аза-7-деазагуанозин, который будет снижать содержание вторичной структуры РНК в последовательностях с высоким содержанием G; "заблокированные" нуклеиновые кислоты (LNA), которые снижают расщепление и повышают определенные виды вторичной структуры РНК; модификации в виде 2'-O-метил, 2'-фтор, или 2'-O-метоксиэтокси, которые повышают стабильность РНК. Такие модификации также могут быть включены в другие части PEgRNA для повышения стабильности и активности. В качестве альтернативы или дополнительно, матрица PEgRNA может быть сконструирована таким образом, что она как кодирует требуемый продукт в виде белка, так и с большей вероятностью включает простые вторичные структуры, которые способны подвергаться обратному фолдингу за счет RT. Такие простые структуры будут действовать в качестве термодинамического стока, делая менее вероятным появление более сложных структур, предотвращающих обратную транскрипцию. Наконец, также является возможным разделить матрицу на две отдельные PEgRNA. При такой конструкции PE будет использоваться для инициирования транскрипции, а также рекрутирования отдельной матричной РНК к являющемуся мишенью участку с помощью РНК-связывающего белка, слитого с Cas9, или элемента распознавания РНК на самой PEgRNA, например аптамера MS2. RT может либо напрямую связываться с данной отдельной матричной РНК, или инициировать обратную транскрипцию на исходной PEgRNA перед заменой на вторую матрицу. Такой подход может обеспечивать возможность осуществления длинных вставок за счет как предотвращения неправильного фолдинга PEgRNA при добавлении длинной матрицы, так и за счет отсутствия необходимости в диссоциации Cas9 из генома для реализации длинных вставок, что вероятно может ингибировать длинные вставки на основе PE.

[0662] В других вариантах осуществления PEgRNA может быть усовершенствована путем внесения дополнительных мотивов РНК на 5′ и 3′-концах PEgRNA или даже в положениях между ними (например в области ядра нРНК или спейсере). Некоторые подобные мотивы, например PEN ENE из KSHV и ENE из MALAT1, были описаны

выше в качестве возможных способов терминации экспрессии длинных PEgRNA из не относящихся к pol III промоторов. Эти элементы образуют тройные спирали РНК, которые окружают polyA-хвост, что приводит к их удержанию в ядре[184,187]. Однако за счет образования сложных структур на 3′-конце PEgRNA, которые закрывают концевой нуклеотид, данные структуры также вероятно будут помогать предотвращать опосредованное экзонуклеазой расщепление PEgRNA.

[0663] Другие структурные элементы, вставленные на 3′-конце, также могут повышать стабильность РНК, хотя и без обеспечения терминации из не относящихся к pol III промоторов. Такие мотивы могут включать шпильки или квадруплексы РНК, которые будут закрывать 3′-конец[197], или саморасщепляющиеся рибозимы, например HDV, которые будут приводить к образованию 2’-3′-циклического фосфата на 3′-конце, а также потенциально делать PEgRNA менее подверженной расщеплению экзонуклеазами[198]. Индуцирование циклизации PEgRNA за счет неполного сплайсинга – для образования циРНК – также может повышать стабильность PEgRNA и приводить к удержанию PEgRNA в ядре[194].

[0664] Дополнительные мотивы РНК также могут повышать процессирование RT или повышать активность PEgRNA за счет повышения связывания RT с дуплексом ДНК-РНК. Добавление нативной последовательности, связанной RT в ее родственном ретровирусном геноме, может повышать активность RT[199]. Это может включать нативный участок связывания праймера (PBS), полипуриновый тракт (PPT), или целующиеся петли, принимающие участие в димеризациии ретровирусного генома и инициировании транскрипции[199].

[0665] Добавление мотивов димеризации, таких как целующиеся петли или пары тетрапетля GNRA/рецептор тетрапетли[200], на 5′ и 3′-концах PEgRNA также может приводить к эффективной циркуляризации PEgRNA, что повышает стабильность. Кроме того, предполагается, что добавление данных мотивов может обеспечивать физическое разделение спейсера и праймера PEgRNA, предотвращать закрытие спейсера, ухудшающее активность PE. Короткие 5′ удлинения или 3′ удлинения PEgRNA, которые образуют малую опорную шпильку в спейсерном участке или в рамках участка связывания праймера, также могут благоприятным образом препятствовать отжигу внутрикомплементарных участков в рамках длины PEgRNA, например взаимодействие между спейсером и участком связывания праймера, которое может возникать. Наконец, целующиеся петли также могут использоваться для

рекрутирования других матричных РНК в участок генома и обеспечения обмена RT активностью от одной РНК к другой. В качестве примера вариантов осуществления различных вторичных структур, в рамках PEgRNA, изображенной на Фиг.3D и Фиг.3E, перечислен ряд вторичных структур РНК, которые могут быть встроены в любой участок PEgRNA, включая концевые части плеча удлинения (например e1 и e2), как показано.

[0666] Примеры усовершенствований включают, но без ограничения:

[0667] Продукт слияния PEgRNA-HDV

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC
GTGCTCAGTCTGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAA
CATGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 230)

[0668] Целующуюся петля PEgRNA-MMLV

GGTGGGAGACGTCCCACCGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAAT
AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCG
GTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCT
TTTTTT (SEQ ID NO: 231)

[0669] Целующаяся петля PEgRNA-VS рибозим

GAGCAGCATGGCGTCGCTGCTCACGGCCCAGACTGAGCACGTGAGTTTTAGAGCT
AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACC
GAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTCCATCAGTTGACA
CCCTGAGGTTTTTTT (SEQ ID NO: 232)

[0670] PEgRNA-GNRA тетрапетля/рецептор тетрапетли

GCAGACCTAAGTGGUGACATATGGTCTGGGCCCAGACTGAGCACGTGAGTTTTAG
AGCTAUACGTAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTUACGAAGTGG
GACCGAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGCATGCGATT
AGAAATAATCGCATGTTTTTTT (SEQ ID NO: 233)

[0671] PEgRNA матрица с заменой вторичного продукта слияния в виде РНК-HDV

TCTGCCATCAAAGCTGCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCGGCCG
GCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCG
AATGGGACTTTTTTT (SEQ ID NO: 234)

[0672] Каркас PEgRNA может быть дополнительно усовершенствован с помощью направленной эволюции аналогично тому, как был усовершенствован SpCas9 и праймерный редактор (PE). Направленная эволюция может повышать степень распознавания PEgRNA Cas9 или подвергнутыми эволюции вариантами Cas9. Кроме того, является возможным, что различные последовательности каркаса PEgRNA будут оптимальными в рамках различных геномных локусов, обеспечивая повышение активности PE в рассматриваемом участке, снижая нецелевые виды активности, или осуществляя оба из данных действий. Наконец, эволюция каркасов PEgRNA, к

которым были добавлены другие мотивы РНК, будет практически наверняка повышать активность слитой PEgRNA по сравнению с не подвергавшейся эволюции слитой РНК. Например, эволюция аллостерических рибозимов, состоящих из аптамеров c-di-GMP-I и рибозимов типа "головка молотка" (hammerhead), приводила к значительному повышению активности[202], что свидетельствует о том, что эволюция также будет повышать активность PEgRNA типа "головка молотка" (hammerhead). Кроме того, в то время как Cas9 в настоящее время в целом не совместим с 5′ удлинением sgРНК, направленная эволюция вероятно обеспечит реализацию мутаций, которые устранят данную несовместимость, обеспечивая возможность использования дополнительных мотивов РНК.

[0673] Настоящее изобретение охватывает любые подобные подходы к дальнейшему усовершенствованию эффективности описанных здесь систем праймерного редактирования.

[0674] В различных вариантах осуществления может являться полезным ограничивать появление последовательности из повторяющихся T из плеча удлинения, поскольку повторяющиеся T могут ограничивать объем транскрибируемой PEgRNA. Например, последовательности из, по меньшей мере, трех последовательных T, по меньшей мере, четырех последовательных T, по меньшей мере, пяти последовательных T, по меньшей мере, шести последовательных T, по меньшей мере, семи последовательных T, по меньшей мере, восьми последовательных T, по меньшей мере, девяти последовательных T, по меньшей мере, десяти последовательных T, по меньшей мере, одиннадцати последовательных T, по меньшей мере, двенадцати последовательных T, по меньшей мере, тринадцати последовательных T, по меньшей мере, четырнадцати последовательных T, по меньшей мере, пятнадцати последовательных T должны избегаться при конструировании PEgRNA или должны быть, по меньшей мере, удалены из конечной сконструированной последовательности. В одном варианте осуществления является возможным избежать включения нежелательных последовательностей повторяющихся T в плечах удлинения PEgRNA за счет избегания целевых участков с высоким содержанием повторяющихся пар азотистых оснований в виде A:T.

Конструкции в виде разделенных PEgRNA для *trans* праймерного редактирования

[0675] Настоящее изобретение также охватывает *trans* праймерное редактирование, обозначающее модифицированную версию праймерного редактирования, основанную на разделении PEgRNA на две отдельные молекулы: направляющую РНК и молекулу tPERT. Молекула tPERT программируется для солокализации с комплексом праймерного редактора в участке-мишени ДНК, направляя участок связывания праймера и матрицу для синтеза ДНК к праймерному редактору *in trans*. Например, см. вариант осуществления *trans* праймерного редактора (tPE) на Фиг.3G в виде двухкомпонентной системы, включающей (1) комплекс рекрутирующий белок (RP)-PE:нРНК и (2) tPERT, которая включает участок связывания праймера и матрицу для синтеза ДНК, присоединенные к домену рекрутирования РНК-белка (например структуре типа "стебель-петля"), при этом компонент в виде рекрутирующего белка комплекса RP-PE:нРНК рекрутирует tPERT к подвергаемому редактированию участку-мишени, тем самым связывая PBS и матрицу для синтеза ДНК с праймерным редактором in *trans* Другими словами, tPERT конструируют для включения (полностью или частично) плеча удлинения PEgRNA, которое включает участок связывания праймера и матрицу для синтеза ДНК. Одно преимущество данного подхода заключается в отделении плеча удлинения PEgRNA от направляющей РНК, тем самым обеспечивая минимизацию взаимодействий в виде отжига, которые периодически возникают между PBS плеча удлинения и спейсерной последовательностью направляющей РНК.

[0676] Ключевой особенностью *trans* праймерного редактирования является способность *trans* праймерного редактора рекрутировать tPERT в участок редактирования ДНК, тем самым эффективным образом обеспечивая солокализацию всех функций PEgRNA в участке праймерного редактирования. Рекрутирование может быть достигнуто путем внесения домена рекрутирования РНК-белка, например аптамера MS2, в tPERT и слияния соответствующего рекрутирующего белка с праймерным редактором (например через линкер с napDNAbp или через линкер с полимеразой), который способен специфично связываться с доменом рекрутирования РНК-белка, тем самым обеспечивая рекрутирование молекулы tPERT к комплексу праймерного редактора. Как показано в рамках процесса, изображенного на Фиг.3H, комплекс RP-PE:нРНК связывает и никирует последовательность-мишень ДНК. Затем рекрутирующий белок (RP) рекрутирует tPERT для солокализации с комплексом праймерного редактора, связанным с участком-мишенью ДНК, тем самым обеспечивая

связывание участка связывания праймера, расположенного на tPERT, с праймерной последовательностью на никированной цепи и далее обеспечивая синтез полимеразой (например RT) одиночной цепи ДНК на матрице для синтеза ДНК, расположенной на tPERT, к 5′-концу tPERT.

[0677] В то время как tPERT показана на Фиг.3G и Фиг.3H в качестве включающей PBS и матрицу для синтеза ДНК на 5′-конце домена рекрутирования РНК-белка, tPERT в других конфигурациях может быть сконструирована с PBS и матрицей для синтеза ДНК, расположенными на 3′-конце домена рекрутирования РНК-белка. Однако tPERT с 5’ удлинением имеет преимущество, заключающееся в том, что синтез одиночной цепи ДНК будет естественным образом терминирован на 5'-конце tPERT и, таким образом, не будет связан с риском использования какой-либо части домена рекрутирования РНК-белка в качестве матрицы на этапе праймерного редактирования, связанного с синтезом ДНК.

Способ конструирования PEgRNA

[0678] Настоящее изобретение также относится к способам конструирования PEgRNA.

[0679] В одном аспекте конструирования подход к конструированию может включать учет конкретного варианта использования, для которого применяется праймерное редактирование. Например, как описано и приведено здесь в качестве примера, праймерное редактирование может применяться, без ограничения, для (a) внесения корректирующих мутации изменений в нуклеотидную последовательность, (b) внесения белковых и РНК меток, (c) внесения иммуноэпитопов в представляющие интерес белки, (d) внесения доменов индуцируемой димеризации в белки, (e) внесения или удаления последовательностей для изменения активности биомолекулы, (f) внесения участков-мишеней рекомбиназы для реализации определенных генетических изменения, и (g) мутагенеза последовательности-мишени за счет использования подверженной ошибкам RT. Помимо данных способов, которые, в целом, обеспечивают вставку, изменение, или делецию нуклеотидных последовательностей в представляющих интерес участках, праймерные редакторы также могут применяться для конструирования высокопрограммируемых библиотек, а также осуществления записи клеточных данных и исследований в виде отслеживания происхождения. В рамках данных различных вариантов использования могут присутствовать, как они были описаны здесь, различные аспекты конструирования, связанные с получением

PEgRNA, которая является особенно полезной в рамках любого из данных вариантов использования.

[0680] При конструировании PEgRNA для любого конкретного варианта использования или применения в рамках праймерного редактирования, следует учесть ряд аспектов, которые включают, но без ограничения:

(a) последовательность-мишень, то есть нуклеотидную последовательность, в которую требуется внести одну или несколько модификаций азотистых оснований с помощью праймерного редактора;

(b) расположение участка разреза в рамках последовательности-мишени, то есть конкретного положения азотистого основания, в котором праймерный редактор индуцирует одноцепочечный разрыв для создания 3′-концевой праймерной последовательности RT на одной стороне разрыва и 5′-концевого эндогенного флэпа на другой стороне разрыва (который в конечном итоге удаляется FEN1 или ее эквивалентом и замещается 3′ оцДНК-флэпом. Участок разрыва аналогичен "расположению изменения", поскольку он обеспечивает создание 3′-концевой праймерной последовательности RT, которая становится удлиненной за счет RT при РНК-зависимой полимеризации ДНК для создания 3′ оцДНК-фдэпа, содержащего требуемое изменение, который затем замещает 5′ эндогенный ДНК-флэп в последовательности-мишени.

(c) доступность последовательностей PAM (включая стандартные участки PAM SpCas9, а также нестандартные участки PAM, распознаваемые вариантами и эквивалентами Cas9 с расширенной или иной специфичностью в отношении PAM).

(d) промежуток между доступными последовательностями PAM и расположением участка разрыва в последовательности-мишени;

(e) конкретный Cas9, вариант Cas9, или эквивалент Cas9, используемый в рамках праймерного редактора;

(f) последовательность и длина участка связывания праймера;

(g) последовательность и длина матрицы для редактирования;

(h) последовательность и длина плеча гомологии;

(i) спейсерная последовательность и длина; и

(j) последовательность ядра.

[0681] Данные аспекты описаны выше.

[0682] В одном варианте осуществления здесь указан подход к конструированию подходящей PEgRNA и, необязательно, руководство по конструированию никирующей sgРНК для никирования второго участка. В рамках этого варианта осуществления предусмотрен пошаговый набор инструкций для конструирования PEgRNA и никирующих sgРНК для праймерного редактирования с учетом одного или нескольких из указанных выше аспектов. В рамках этапов приведены ссылки на примеры, показанные на Фиг.70A-70I.

1. **Определение последовательности-мишени и изменения.** Получить последовательность участка-мишени ДНК (~200bp) с центром вокруг месторасположения требуемого изменения (точечной мутации, вставки, делеции, или их комбинации). См. Фиг.70A.
2. **Поиск PAM-мишеней.** Идентифицировать PAM поблизости от требуемого месторасположения изменения. PAM могут быть идентифицированы на любой цепи ДНК поблизости от требуемого месторасположения изменения. В то время как являются предпочтительными PAM рядом с положением изменения (то есть где участок разрыва расположен менее чем в 30 nt от положения изменения или менее чем в 29 nt, 28 nt, 27 nt, 26 nt, 25 nt, 24 nt, 23 nt, 22 nt, 21 nt, 20 nt, 19 nt, 18 nt, 17 nt, 16 nt, 15 nt, 14 nt, 13 nt, 12 nt, 11 nt, 10 nt, 9 nt, 8 nt, 7 nt, 6 nt, 5 nt, 4 nt, 3 nt, или 2 nt от положения изменения), является возможным вносить изменения с помощью протоспейсеров и PAM, которые обеспечивают расположение разрыва на расстоянии ≥ 30 nt от положения изменения. См. Фиг.70B.
3. **Поиск участков разрыва.** Для каждого рассматриваемого PAM идентифицировать соответствующий участок разрыва и цепь. Для никазы Sp Cas9 H840A расщепление осуществляется в PAM-содержащей цепи между 3-м и 4-м основаниями 5′ относительно PAM NGG. Все подвергнутые редактированию нуклеотиды должны быть расположены 3′ относительно участка разрыва, так что подходящие PAM должны обеспечивать расположение разрыва 5′ относительно изменения-мишени на PAM-содержащей цепи. В показанном ниже примере присутствуют два возможных PAM. Для простоты остальные этапы продемонстрируют конструирование PEgRNA с использованием только PAM 1. См. Фиг.70C.

4. **Конструирование спейсерной последовательности.** Протоспейсер Sp Cas9 соответствует 20 нуклеотидам 5′ относительно PAM NGG на PAM-содержащей цепи. Эффективное инициирование транскрипции Pol III требует, чтобы G был первым транскрибированным нуклеотидом. Если первый нуклеотид протоспейсера представляет собой G, то спейсерная последовательность для PEgRNA будет просто представлять собой протоспейсерную последовательность. Если первый нуклеотид протоспейсера не представляет собой G, то спейсерная последовательность PEgRNA представляет собой G и далее протоспейсерную последовательность. См. Фиг.70D.
5. **Конструирование участка связывания праймера (PBS).** Идентифицировать праймер ДНК на PAM-содержащей цепи с помощью последовательности начального аллеля. 3′-конец праймера ДНК представляет собой нуклеотид, расположенный непосредственно выше участка разрыва (то есть 4-го основания 5′ относительно PAM NGG для Sp Cas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера (PBS) PEgRNA, содержащий от 12 до 13 нуклеотидов комплементарности относительно праймера ДНК, может использоваться для последовательностей, которые содержат ~40-60% содержимого в виде GC. Для последовательностей с низким содержанием GC должны быть протестированы более длинные PBS (от 14 до 15 nt). Для последовательностей с высоким содержанием GC должны быть протестированы более короткие PBS (от 8 до 11 nt). Оптимальные последовательности PBS должны быть определены эмпирически независимо от содержания GC. Для конструирования PBS последовательности длиной *p* следует взять обратный комплемент первых p нуклеотидов 5′ участка разрыва в PAM-содержащей цепи с использованием последовательности *начального* аллеля. См. Фиг.70E.
6. **Конструирование RT матрицы (или матрицы для синтеза ДНК).** RT матрица (или матрица для синтеза ДНК, где полимераза не является обратной транскриптазой) кодирует требуемое изменение и гомологию относительно последовательности, являющейся смежной относительно изменения. В одном варианте осуществления данные участки соответствуют матрице для синтеза ДНК согласно Фиг.3D и Фиг.3E, где матрица для синтеза ДНК включает "матрицу для редактирования" и "плечо гомологии". Оптимальные величины

длины RT матрицы варьируются в зависимости от участка-мишени. Для коротких изменений (положения от +1 до +6) рекомендуется тестировать короткую (от 9 до 12 nt), среднюю (от 13 до 16 nt), и длинную (от 17 до 20 nt) RT матрицу. Для длинных изменений (положения +7 и более) рекомендуется использовать RT матрицы, которые имеют длину, по меньше мере, на 5 nt превышающую (предпочтительно на 10 nt или более) положение изменения для обеспечения достаточной гомологии 3′ ДНК-флэпа. Для длинных изменений несколько RT матриц должны быть подвергнуты скринингу для идентификации функциональных конструкций. Для более крупных вставок и делеций (≥5 nt) рекомендуется встраивание большей 3′ гомологии (~20 nt или более) в RT матрицу. Эффективность редактирования обычно ухудшается, когда RT матрица кодирует синтез G в качестве последнего нуклеотида в обратно транскрибированном ДНК продукте (соответствует C в RT матрице PEgRNA). Поскольку многие RT матрицы поддерживают эффективное праймерное редактирование, при конструировании RT матриц рекомендуется не использовать G в качестве последнего синтезируемого нуклеотида Для конструирования последовательности RT матрицы длиной r следует использовать последовательность требуемого аллеля и взять обратный комплемент первых r нуклеотидов 3′ участка разрыва в цепи, которая первоначально содержала PAM. Следует обратить внимание, что по сравнению с изменениями SNP, изменения в виде вставки или делеции с использованием RT матриц той же длины не будут содержать идентичную гомологию. См. Фиг.70F.

7. **Сборка полной последовательности PEgRNA.** Следует конкатенировать компоненты PEgRNA в следующем порядке (от 5′ к 3′): спейсер, каркас, RT матрица и PBS. См. Фиг.70G.
8. **Конструирование никирующих sgРНК для PE3.** Идентифицировать PAM на не подвергаемой редактированию цепи выше и ниже изменения. Оптимальные положения никирования имеют высокую зависимость от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии от 40 до 90 нуклеотидов 5′ относительно положения PEgRNA-индуцированного разрыва, ведут к более высоким выходам редактирования и меньшему количеству инделей. Никирующая sgРНК имеет спейсерную

последовательность, которая совпадает с протоспейсером длиной 20 nt в *начальном* аллеле, с добавлением 5′ G, если протоспейсер не начинается на G. См. Фиг. 70H.

9. **Конструирование никирующих sgРНК PE3b** Если в комплементарной цепи присутствует PAM и его соответствующий протоспейсер накладывается на последовательность, являющуюся мишенью редактирования, то такое изменение может являться кандидатом для системы PE3b. В системе PE3b спейсерная последовательность никирующей sgРНК совпадает с последовательностью требуемого подвергаемого редактированию аллеля, но не начального аллеля. Система PE3b работает эффективно, когда подвергаемый редактированию нуклеотид (нуклеотиды) попадают в участок затравки (~10 nt от PAM) протоспейсера никирующей sgРНК. Это предотвращает никирование комплементарной цепи до готовности подвергнутой редактированию цепи, предотвращая конкурирование между PEgRNA и sgРНК за связывание с ДНК-мишенью. PE3b также обеспечивает избежание образования одновременных разрывов на обеих цепях, тем самым значительно снижая образование инделей, в то же время сохраняя высокую эффективность редактирования. sgРНК PE3b должны иметь спейсерную последовательность, которая совпадает с протоспейсером длиной 20 nt в *требуемом* аллеле, с добавлением 5′ G, если это необходимо. См. Фиг.70I.

[0683] Указанный выше пошаговый процесс конструирования подходящей PEgRNA и sgРНК для никирования второго участка не является ограничивающим каким-либо образом. Изобретение охватывает вариации описанного выше пошагового процесса, которые будут очевидны среднему специалисту в данной области.

[7] Варианты использования праймерного редактирования

[0684] Помимо развития описанной здесь системы праймерного редактирования в качестве новой технологии геномного редактирования типа "поиск-и-замена", которая обеспечивает возможность реализации нацеленных вставок, делеций, и всех 12 возможных конверсий оснований в локусах-мишенях клеток человека без необходимости двуцепочечных разрывов ДНК или донорских матриц ДНК, авторы также описывают применение праймерных редакторов в рамках широкого ряда конкретных вариантов использования. Например, как описано и приведено здесь в качестве примера, праймерное редактирование может применяться для (a) внесения

корректирующих мутации изменений в нуклеотидную последовательность, (b) внесения белковых и РНК меток, (c) внесения иммуноэпитопов в представляющие интерес белки, (d) внесения доменов индуцируемой димеризации в белки, (e) внесения или удаления последовательностей для изменения активности биомолекулы, (f) внесения участков-мишеней рекомбиназы для реализации определенных генетических изменения, и (g) мутагенеза последовательности-мишени за счет использования подверженной ошибкам RT. Помимо данных способов, которые, в целом, обеспечивают вставку, изменение, или делецию нуклеотидных последовательностей в представляющих интерес участках, праймерные редакторы также могут применяться для конструирования высокопрограммируемых библиотек, а также осуществления записи клеточных данных и исследований в виде отслеживания происхождения. Авторы также описывают дополнительные аспекты конструкции PEgRNA, которые направлены на повышение эффективности праймерного редактирования. Кроме того, авторы описывают способы успешной доставки праймерных редакторов с помощью систем доставки на основе векторов, которые включают разделение napDNAbp с помощью доменов интеина.

[0685] Данные конкретные примеры вариантов использования праймерного редактирования не являются ограничивающими каким-либо образом. Настоящее изобретение охватывает любой вариант использования праймерного редактирования, который включает, в целом. некоторую форму внесения, удаления, и/или модифицирования одного или нескольких азотистых оснований в участке-мишени в нуклеотидной последовательности, например геномной ДНК.

[0686] Для любого из данных примеров вариантов использования праймерного редактирования является возможным использовать любой описанный здесь праймерный редактор, включая PE1, PE2, PE3, и PE3b, или PE-short.

A. Механизм праймерного редактирования

[0687] В различных вариантах осуществления праймерное редактирование осуществляется путем контактирования молекулы-мишени ДНК (для которое требуется внести изменение в нуклеотидной последовательности) с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) в виде комплекса с удлиненной направляющей РНК. Со ссылкой на Фиг.1G, удлиненная направляющая РНК включает удлинение на 3′ или 5′-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК и кодирует требуемое

нуклеотидное изменение (например изменение одного нуклеотида, вставку, или делецию). На этапе (a) комплекс в виде napDNAbp/удлиненной нРНК контактирует с молекулой ДНК, при этом удлиненная нРНК направляет napDNAbp для связывания с локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК, то есть "не являющейся мишенью цепи". Однако разрыв должен быть осуществлен в любой одной из цепей. То есть разрыв должен быть осуществлен в "цепи-мишени" R-петли (то есть цепи, гибридизированной с последовательностью протоспейсера удлиненной нРНК) или "не являющейся мишенью цепи" (то есть цепи, образующей одноцепочечную часть R-петли, которая является комплементарной цепи-мишени). На этапе (c) 3′-конец цепи ДНК (образованный разрывом) взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции (то есть "праймированной мишенью RT"). В определенных вариантах осуществления 3′-конец цепи ДНК гибридизируется с определенной праймирующей последовательностью RT на удлиненной части направляющей РНК, то есть "праймирующей последовательностью обратной транскриптазы". На этапе (d) вносится обратная транскриптаза (в виде слитого белка с napDNAbp или *in trans*), которая синтезирует одиночную цепь ДНК от 3′-конца праймированного участка к 5′-концу удлиненной направляющей РНК. Это образует одноцепочечный ДНК-флэп, включающий требуемое нуклеотидное изменение (например изменение одного основания, вставку, или делецию, или их комбинацию) и являющийся иным образом гомологичным эндогенной ДНК в участке разрыва или являющейся смежной с ним. На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа, так что требуемое нуклеотидное изменение становится встроенным в локус-мишень. Данный процесс может быть смещен в сторону образования требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа (например с помощью FEN1 или аналогичного фермента, который предоставляется *in trans* в виде продукта слияния с праймерным редактором или предоставляется эндогенным образом), который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с

последовательностью эндогенной ДНК. Без связи с теорией, процессы репарации и репликации клеточной эндогенной ДНК обеспечивают коррекцию несовпадающей ДНК для встраивания нуклеотидного изменения (изменений) с образованием требуемого измененного продукта. Процесс также может быть смещен в сторону образования продукта с "никированием второй цепи", как показано в качестве примера на Фиг.1G, или "временны́м никированием второй цепи", как показано в качестве примера на Фиг.1I и описано здесь.

[0688] Процесс праймерного редактирования может обеспечивать внесение, по меньшей мере, одного или нескольких из следующих генетических изменений: трансверсий, транзиций, делеций, и вставок. Кроме того, праймерное редактирование может быть реализовано для определенных вариантов использования. Например, как описано и приведено здесь в качестве примера, праймерное редактирование может применяться для (a) внесения корректирующих мутации изменений в нуклеотидную последовательность, (b) внесения белковых и РНК меток, (c) внесения иммуноэпитопов в представляющие интерес белки, (d) внесения доменов индуцируемой димеризации в белки, (e) внесения или удаления последовательностей для изменения активности биомолекулы, (f) внесения участков-мишеней рекомбиназы для реализации определенных генетических изменения, и (g) мутагенеза последовательности-мишени за счет использования подверженной ошибкам RT. Помимо данных способов, которые, в целом, обеспечивают вставку, изменение, или делецию нуклеотидных последовательностей в представляющих интерес участках, праймерные редакторы также могут применяться для конструирования высокопрограммируемых библиотек, а также осуществления записи клеточных данных и исследований в виде отслеживания происхождения. Авторы также описывают дополнительные аспекты конструкции PEgRNA, которые направлены на повышение эффективности праймерного редактирования. Кроме того, авторы описывают способы успешной доставки праймерных редакторов с помощью систем доставки на основе векторов, которые включают разделение napDNAbp с помощью доменов интеина.

[0689] Термин "система праймерного редактирования" или "праймерный редактор (PE)" обозначает композиции, используемые в рамках способа геномного редактирования с применением описанной здесь мишень-инициированной обратной транскрипции (TPRT), включая, но без ограничения, napDNAbp, обратные транскриптазы, слитые белки (например включающие napDNAbp и обратные

транскриптазы), удлиненные направляющие РНК, и комплексы, включающие слитые белки и удлиненные направляющие РНК, а также вспомогательные элементы, например компоненты для никирования второй цепи и эндонуклеазы для удаления 5′ эндогенного ДНК-флэпа (например FEN1) для обеспечения сдвига процесса праймерного редактирования в сторону образования подвергнутого редактированию продукта.

[0690] В другом варианте осуществления схема на Фиг.3F показывает взаимодействие стандартной PEgRNA с участком-мишенью двуцепочечной ДНК и сопутствующее получение 3′ одноцепочечного ДНК-флэпа, содержащего представляющее интерес генетическое изменение. Двуцепочечная ДНК показана с верхней цепью в ориентации от 3′ к 5′ и нижней цепью в направлении от 5′ к 3′. Верхняя цепь включает "протоспейсер" и последовательность PAM и обозначена как "цепь-мишень". Комплементарная нижняя цепь обозначена как "не являющаяся мишенью цепь". Хотя это и не показано, изображенная PEgRNA будет образовывать комплекс с Cas9 или его эквивалентом. Как показано на схеме, спейсер PEgRNA отжигается на комплементарном участке на цепи-мишени, который обозначен как протоспейсер, который расположен непосредственно ниже последовательности PAM и имеет длину приблизительно 20 нуклеотидов. Это взаимодействие образует ДНК/РНК гибрид между спейсерной РНК и ДНК протоспейсера и индуцирует образование R-петли в участке напротив протоспейсера. Как описано в других абзацах, затем белок Cas9 (не показан) индуцирует разрыв в не являющейся мишенью цепи, как показано. Это затем ведет к образованию участка в виде 3′ оцДНК-флэпа, который, в соответствии с *z*, взаимодействует с 3′-концом PEgRNA в участке связывания праймера. 3′-конец оцДНК-флэпа (то есть последовательность праймера в виде обратной транскриптазы) отжигается на участке связывания праймера (A) на PEgRNA, тем самым обеспечивая праймирование обратной транскриптазы. Далее обратная транскриптаза (например поступающая *in trans* или поступающая *cis* в виде слитого белка, присоединенного к конструкту Cas9) затем обеспечивает полимеризацию одиночной цепи ДНК, которая кодируется матрицей для редактирования (B) и плечом гомологии (C). Полимеризация продолжается по направлению к 5′-концу плеча удлинения. Полимеризованная цепь оцДНК образует 3′-концевой оцДНК-флэп, который, как описано в других разделах (например на Фиг.1G) вносится в эндогенную ДНК, вытесняя соответствующую эндогенную цепь (которая удаляется в виде 5′-концевого ДНК-флэпа эндогенной ДНК)

и обеспечивая внесение требуемого нуклеотидного изменения (изменения одной пары азотистых оснований нуклеотида, делеций, вставок (включая целые гены) с помощью раундов естественной репарации/репликации ДНК).

[0691] Данный вариант использования праймерного редактирования дополнительно описан в Примере 1.

B. Мутагенез с помощью праймерного редактирования с подверженной ошибкам RT

[0692] В различных вариантах осуществления система праймерного редактирования может включать применение подверженной ошибкам обратной транскриптазы для реализации нацеленного мутагенеза, то есть мутирования только хорошо определенной последовательности ДНК в геноме или другого ДНК-элемента в клетке. На Фиг.22 показана схема примера процесса осуществления направленного мутагенеза с использованием подверженной ошибкам обратной транскриптазы в локусе-мишени с помощью программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) в виде комплекса с удлиненной направляющей РНК. Данный процесс может быть обозначен как вариант осуществления праймерного редактирования для направленного мутагенеза. Удлиненная направляющая РНК включает удлинение на 3′- или 5′-конце направляющей РНК или во внутримолекулярном расположении в направляющей РНК. На этапе (a) комплекс napDNAbp/нРНК контактирует с молекулой ДНК, при этом нРНК направляет napDNAbp для связывания с подвергаемым мутагенезу локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК. На этапе (c) 3′-концевая цепь ДНК взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции. В определенных вариантах осуществления 3′-концевая цепь ДНК гибридизируется с определенной последовательностью для праймирования RT на удлиненной части направляющей РНК. На этапе (d) вносится подверженная ошибкам обратная транскриптаза, которая синтезирует подвергнутую мутагенезу одиночную цепь ДНК от 3′-конца праймированного участка к 3′-концу направляющей РНК. Примеры мутаций указаны с помощью звездочки "*". Это обеспечивает образование

одноцепочечного ДНК-флэпа, включающего требуемый подвергнутый мутагенезу участку. На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа (включающего подвергнутый мутагенезу участок), так что требуемый подвергнутый мутагенезу участок становится встроенным в локус-мишень. Данный процесс может быть смещен в сторону образования требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа, который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с комплементарной последовательностью на другой цепи. Процесс также может быть смещен в сторону образования продукта с никированием второй цепи, как показано в качестве примера на Фиг.1F. После процессов репарации и/или репликации эндогенной ДНК мутагенизированный участок становится встроенным в обе цепи ДНК локуса ДНК.

[0693] Данный вариант использования праймерного редактирования дополнительно описан в Примере 2.

[0694] Ферменты в виде подверженной ошибкам или мутагенной RT известны в данной области. В соответствии с используемым здесь значением термин "подверженная ошибкам" обратная транскриптаза обозначает фермент в виде обратной транскриптазы, которая встречается в природе или которая была получена из другой обратной транскриптазы (например обратной транскриптазы M-MLV дикого типа), которая имеет частоту ошибок, которая меньше, чем частота ошибок обратной транскриптазы M-MLV дикого типа. Сообщается, что частота ошибок обратной транскриптазы M-MLV дикого типа соответствует диапазону в виде одной ошибки на от 15000 (верхняя граница) до 27000 (нижняя граница) встраиваний азотистых оснований. См. Boutabout et al. (2001) “DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1,” *Nucleic Acids Res* 29(11):2217–2222, который приведен здесь в качестве ссылки. Таким образом, для целей данного изобретения термин "подверженная ошибкам" означает RT, которые имеют частоту ошибок в виде более чем одной ошибки на 15000 встраиваний азотистых оснований (6,7 x 10^{-5} или более), например 1 ошибку в 14000 азотистых оснований (7,14 x 10^{-5} или более), 1 ошибку в 13000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 12000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 11000 азотистых оснований или менее (9,1 x 10^{-5} или более), 1 ошибку в 10000 азотистых оснований или менее (1 x 10^{-4} или 0,0001 или более), 1 ошибку в 9000 азотистых оснований или менее

(0,00011 или более), 1 ошибку в 8000 азотистых оснований или менее (0,00013 или более), 1 ошибку в 7000 азотистых оснований или менее (0,00014 или более), 1 ошибку в 6000 азотистых оснований или менее (0,00016 или более), 1 ошибку в 5000 азотистых оснований или менее (0,0002 или более), 1 ошибку в 4000 азотистых оснований или менее (0,00025 или более), 1 ошибку в 3000 азотистых оснований или менее (0,00033 или более), 1 ошибку в 2000 азотистых оснований или менее (0,00050 или более), или 1 ошибку в 1000 азотистых оснований или менее (0,001 или более), или 1 ошибку в 500 азотистых основаниях или менее (0,002 или более), или 1 ошибку в 250 азотистых основаниях или менее (0,004 или более).

[0695] Для получения последовательностей, подвергнутых мутагенезу с помощью праймерного редактирования, может быть определен ряд мутагенных RT. Двумя подобными примерами являются мутагенные обратные транскриптазы из фага *Bordetella* (см. Handa, S., et al. *Nucl Acids Res* 9711-25 (2018), который приведен здесь в качестве ссылки) и *Legionella pneumophila* (см. Arambula, D., et al. *Proc Natl Acad Sci USA* 8212-7 (2013), который приведен здесь в качестве ссылки). В случае RT из фага *Bordetella* (brt) может потребоваться также добавление к Cas9 вспомогательного белка (bavd) или его доставка *in trans*, а также дополнительных последовательностей РНК в PEgRNA для повышения связывания мутагенной RT с участком-мишенью (см. Handa, S., et al. *Nucl Acids Res* 9711-25 (2018)). При использовании мутагенных RT матричный участок PEgRNA может иметь высокое содержание аденозинов или кодонов AAY для повышения разнообразия.

[0696] Аминокислотная последовательность мутагенной RT из фага *Bordetella* приведена ниже. Подобно другим описанным здесь RT, белок Brt может быть слит с napDNAbp в виде слитого белка для образования функционального PE.

Название	Последовательность
Мутагенная RT brt	MGKRHRNLIDQITTWENLLDAYRKTSHGKRRTWGYLEFKEYDL ANLLALQAELKAGNYERGPYREFLVYEPKPRLISALEFKDRLVQH ALCNIVAPIFEAGLLPYTYACRPDKGTHAGVCHVQAELRRTRATH FLKSDFSKFFPSIDRAALYAMIDKKIHCAATRRLLRVVLPDEGVGI PIGSLTSQLFANVYGGAVDRLLHDELKQRHWARYMDDIVVLGD DPEELRAVFYRLRDFASERLGLKISHWQVAPVSRGINFLGYRIWP THKLLRKSSVKRAKRKVANFIKHGEDESLQRFLASWSGHAQWA DTHNLFTWMEEQYGIACH (SEQ Id no: 235)

[0697] В случае Brt из *Bodetella* продукт слияния PE также может включать дополнительный вспомогательный белок (Bavd). Вспомогательный белок может быть

слит со слитым белком PE или поступать *in trans*. Аминокислотная последовательность вспомогательного белка Bavd приведена ниже:

Название	Последовательность
Вспомогательный белок bavd для brt	MEPIEEATKCYDQMLIVERYERVISYLYPIAQSIPRKHGVAREMFL KCLLGQVELFIVAGKSNQVSKLYAADAGLAMLRFWLRFLAGIQK PHAMTPHQVETAQVLIAEVGRILGSWIARVNRKGQAGK (SEQ ID NO: 236)

[0698] В случае Brt из *Bodetella* PEgRNA может включать дополнительную нуклеотидную последовательность, добавленную к PEgRNA, например на 5′ или 3′-конце. Ниже приведен пример последовательности, которую обычно получают из генома фага *Bordetella*:

НАЗВАНИАЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 1	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGC CUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAA CACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAG GCGGGGCCGACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 237)

[0699] Данная последовательность дополнения к PEgRNA может быть укорочена различными способами для сокращения длины. Например, последовательность дополнения 1 к PEgRNA может быть укорочена до следующих примеров альтернативных последовательностей дополнения:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 2	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGC CUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAA CACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAG GCGGGGCCGACUUC (SEQ ID NO: 238)
ДОПОЛНЕНИЕ К PEGRNA 3	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGC CUUUCCUCAAUAGGUGGUCAAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 239)
ДОПОЛНЕНИЕ К PEGRNA 4	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGC CUUUCCUCAAUAGGUGGUCA (SEQ ID NO: 277)
ДОПОЛНЕНИЕ К PEGRNA 5	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCA AUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACA CGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCG ACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 240)
ДОПОЛНЕНИЕ К PEGRNA 6	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCA AUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACA CGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCG ACUUC (SEQ ID NO: 241)

ДОПОЛНЕНИЕ К PEGRNA 7	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 242)
ДОПОЛНЕНИЕ К PEGRNA 8	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCA (SEQ ID NO: 243)

[0700] В других вариантах осуществления последовательность дополнения PEgRNA также может быть подвержена мутированию. Например, последовательность дополнения к PEgRNA 1 может быть подвержена мутированию до следующих примеров альтернативных последовательностей дополнения:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 1 МУТИРОВАННОЕ	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGAUGAGCCGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCGACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 244)

[0701] В различных вариантах осуществления, связанных с применением PE для внесения мутаций, могут применяться специальные условия в плане PEgRNA. Например, без связи с теорией, могут потребоваться дополнительные описанные выше последовательности PEgRNA для обеспечения эффективного мутагенеза за счет мутагенных RT.

[0702] В рамках описанных здесь праймерных редакторов может использоваться любая мутагенная RT. Например, может использоваться подверженная ошибкам RT, описанная в нижеследующих ссылках, которые приведены в качестве ссылок:

[0703] Bebenek et al., “Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots.,” *J. Biol Chem*, 1993, 268: 10324-34; и

[0704] Menendez-Arias, “Mutation rates and instrinsic fidelity of retroviral reverse transcriptases,” 2009, *Viruses*, 1(3): 1137-1165.

[0705] Различные подверженные ошибкам RT могут включать, но без ограничения, следующие ферменты, описанные в Таблице 1 у Menendez-Arias et al. (полное содержание которого приведено здесь в качестве ссылки), следующим образом:

ПОДВЕРЖЕННАЯ ОШИБКА RT	ПРИВОДИМЫЙ ДИАПАЗОН ВЕЛИЧИНЫ ОШИБОК
RT ВИЧ-1 (ГРУППА M, ПОДТИП B)	ОТ 0,6 X 10^{-4} ДО 2,0 X 10^{-4}
RT ВИЧ-1 (ГРУППА O)	5,5 X 10^{-5}
RT SIV AGM	2,9 X 10^{-5}
RT SIV MNE	ОТ 1,6 X 10^{-5} ДО 1,2 X 10^{-4}

RT PFV	1,7 X 10^{-4}
RT FIV	6,2 X 10^{-5}
RT AMV	5,9 X 10^{-5}
RT MO-MLV	ОТ 2,7 X 10^{-5} ДО 3,3 X 10^{-5}

C. Применение праймерного редактирования для лечения заболеваний, связанных с экспансией тринуклеотидных повторов

[0706] Описанная здесь система праймерного редактирования (PE) может применяться для сокращения мутаций в виде тринуклеотидных повторов (или "заболеваний, связанных с экспансией тринуклеотидных повторов") для лечения состояний, таких как болезнь Хантингтона и другие заболевания, связанные с тринуклеотидными повторами). Заболевания, связанные с экспансией тринуклеотидных повторов, представляют собой комплексные, прогрессирующие заболевания, которые затрагивают связанные с развитием нейробиологические процессы и зачастую влияют на когнитивные способности, а также сенсорно-моторные функции. Данные заболевания демонстрируют наличие генетической связи (то есть повышение тяжести с каждым поколением). Экспансии или сокращения ДНК обычно происходят при мейозе (то есть во время гаметогенеза или при раннем эмбриональном развитии) и зачастую связаны с полом, что свидетельствует о том, что некоторые гены подвергаются экспансии только при наследовании от матери, а другие – только от отца. У человека связанные с тринуклеотидными повторами заболевания могут вызывать молчание генов либо на транскрипционном, или трансляционном уровне, что по существу блокирует функцию гена. В качестве альтернативы, связанные с тринуклеотидными повторами заболевания могут вызывать образование измененных белков, которые образуются за счет крупных аминокислотных последовательностей с повторами, которые нарушают или изменяют функцию белка, зачастую доминантно-отрицательным образом (например при полиглутаминовых заболеваниях).

[0707] Без связи с теорией, экспансия триплетов вызывается проскальзыванием при репликации ДНК или при репаративном синтезе ДНК. Поскольку тандемные повторы имеют идентичную последовательность относительно друг друга, спаривание оснований между двумя цепями ДНК может происходить в нескольких точках в рамках последовательности. Это может приводить к образованию "выступающих в виде петель" структур при репликации ДНК или репаративном синтезе ДНК. Это может приводить к повторяющемуся копированию повторенной последовательности, что увеличивает количество повторов. Были предложены дополнительные механизмы,

связанные с промежуточными продуктами в виде РНК:ДНК гибридов. Праймерное редактирование может применяться для снижения или устранения данных участков экспансии триплетов за счет делеции одного или нескольких триплетов в виде нежелательных повторяющихся кодонов. В рамках варианта осуществления, соответствующего данному варианту использования, на Фиг.23 приведена схема конструкции PEgRNA для сокращения или уменьшения последовательностей тринуклеотидных повторов с помощью праймерного редактирования.

[0708] Праймерное редактирование может быть реализовано для сокращения участков экспансии триплетов за счет никирования участка выше участка в виде триплетного повтора с помощью праймерного редактора, включающего PEgRNA, которая подходящим образом нацелена на участок разрыва. Затем праймерный редактор обеспечивает синтез новой цепи ДНК (оцДНК-флэпа) на основе PEgRNA в качестве матрицы (то есть матрицы для ее редактирования), которая кодирует здоровое количество триплетных повторов (которое зависит от конкретного гена и заболевания). Новая синтезированная цепь оцДНК, включающая здоровую последовательность триплетных повторов, также синтезируется таким образом, чтобы включать короткий участок гомологии (то есть плечо гомологии), который соответствует последовательности, являющейся смежной относительного другого конца повтора (красная цепь). Внесение новой синтезированной цепи и последующее замещение эндогенной ДНК новым синтезированным оцДНК-флэпом ведет к сокращению повторяющегося аллеля.

[0709] В зависимости от конкретного связанного с экспансией тринуклеотидов заболевания, дефектные экспансии триплетов могут располагаться в "связанных с экспансией тринуклеотидных повторов белках". Связанные с экспансией тринуклеотидных повторов белки представляют собой разнообразный набор белков, связанных с подверженностью к развитию заболевания, связанного с экспансией тринуклеотидных повторов, наличием заболевания, связанного с экспансией тринуклеотидных повторов, тяжестью заболевания, связанного с экспансией тринуклеотидных повторов, или любой их комбинацией. Заболевания, связанные с экспансией тринуклеотидных повторов, разделяются на две категории, определяемые на основе типа повтора. Наиболее частый повтор представляет собой повтор CAG, который при его наличии в кодирующем участке гена кодирует аминокислоту глутамин (Q). Таким образом, данные заболевания обозначаются как

полиглутаминовые (polyQ) заболевания и включают следующие заболевания: болезнь Хантингтона (HD); спинальную и бульбарную мышечную атрофию (SBMA); спинально-церебеллярную атаксию (SCA) (SCA типов 1, 2, 3, 6, 7 и 17); и дентато-рубро-паллидо-льюисову атрофию (Dentatorubro-Pallidoluysian Atrophy, DRPLA). Остальные заболевания, связанные с экспансией тринуклеотидных повторов, либо не включают триплет CAG, или триплет CAG не располагается в кодирующем участке гена и, таким образом, они обозначаются как неполиглутаминовые заболевания. Неполиглутаминовые заболевания включают синдром ломкой X-хромосомы (FRAXA), ломкую XE MR (FRAXE), атаксию Фрейдрейха (FRDA), миотоническую дистрофию (DM), и спиноцеребеллярные атаксии (SCA типов 8 и 12).

[0710] Белки, ассоциированные с заболеваниями, связанными с экспансией тринуклеотидных повторов, могут быть выбраны на основе экспериментального ассоциирования белка, ассоциированного с заболеванием, связанным с экспансией тринуклеотидных повторов, с заболеванием, связанным с экспансией тринуклеотидных повторов. Например, скорость продуцирования или концентрация в крови белка, ассоциированного с заболеванием, связанным с экспансией тринуклеотидных повторов, может быть повышена или понижена у группы, имеющей заболевание, связанное с экспансией тринуклеотидных повторов, по сравнению с группой без заболевания, связанного с экспансией тринуклеотидных повторов. Различия в плане уровней белка могут быть проанализированы с помощью протеомных способов, включая, но без ограничения, вестерн-блот, иммуногистохимическое окрашивание, иммуноферментный анализ (ELISA), и масс-спектометрию. В качестве альтернативы, белки, ассоциированные с заболеваниями, связанными с экспансией тринуклеотидных повторов, могут быть идентифицированы путем получения профилей экспрессии генов, кодирующих данные белки, с помощью геномных способов, включая, но без ограничения, анализы на основе ДНК-микрочипов, последовательные анализы экспрессии генов (SAGE), и количественную полимеразную цепную реакцию в реальном времени (кПЦР).

[0711] Неограничивающие примеры белков, ассоциированных с заболеваниями, связанными с экспансией тринуклеотидных повторов, которые могут быть скорректированы с помощью праймерного редактирования, включают AR (андрогенный рецептор), FMR1 (белок 1, ассоциированный с умственной отсталостью, сцепленной с ломкой X-хромосомой), HTT (гентингтин), DMPK (протеинкиназу,

ассоциированную с миотонической дистрофией), FXN (фратаксин), ATXN2 (атаксин 2), ATN1 (атрофии 1), FEN1 (структуроспецифичную флэп-эндонуклеазу 1), TNRC6A (белок, кодируемый геном 6A, содержащим тринуклеотидные повторы), PABPN1 (ядерный поли(A)-связывающий белок 1), JPH3 (юнктофилин 3), MED15 (субъединицу 15 медиаторного комплекса), ATXN1 (атаксин 1), ATXN3 (атаксин 3), TBP (TATA-бокс-связывающий белок), CACNA1A (альфа-1A-субъединицу потенциал-зависимого кальциевого канала P/Q-типа), ATXN80S (белок, синтезируемый с противоположной нити ATXN8 (не кодирующей белок)), PPP2R2B (бета-изоформу регуляторной субъединицы B протеинфосфатазы 2), ATXN7 (атаксин 7), TNRC6B (белок, кодируемый геном 6B, содержащим тринуклеотидные повторы), TNRC6C (белок, кодируемый геном 6C, содержащим тринуклеотидные повторы), CELF3 (CUGBP, член 3 семейства Elav-подобных белков), MAB21L1 (mab-21-подобный белок 1 (C.elegans)), MSH2 (гомолог 2 mutS, ассоциированный с неполипозным колоректальным раком типа 1 (E. coli)), TMEM185A (трансмембранный белок 185A), SIX5 (белок, кодируемый гомеобоксом 5 SIX), CNPY3 (гомолог Canopy 3 (данио-рерио)), FRAXE (белок, ассоциированный с "редким" ломким сайтом, проявляющимся при недостатке фолиевой кислоты, fra(X)(q28) E), GNB2 (бета полипептид 2 белка, связывающего гуаниновые нуклеотиды (G-белка)), RPL14 (рибосомный белок L14), ATXN8 (атаксин 8), INSR (инсулиновый рецептор), TTR (транстиретин), EP400 (E1A-связывающий белок p400), GIGYF2 (белок GYF 2, взаимодействующий с GRB10), OGG1 (8-оксогуанин-ДНК-гликозилазу), STC1 (станниокальцин 1), CNDP1 (карнозиндипептидазу 1 (металлопептидазу семейства M20)), C10orf2 (белок, кодируемый открытой рамкой считывания 2 хромосомы 10), MAML3 (mastermind-подобный белок 3 (Drosophila)), DKC1 (белок 1, ассоциированный с врожденным дискератозом, дискерин), PAXIP1 (белок 1, взаимодействующий с PAX (с доменом активации транскрипции)), CaSK (кальций/кальмодулин-зависимую сериновую протеинкиназу (семейства MAGUK)), MAPT (белок tau, ассоциированный с микротрубочками), SP1 (фактор транскрипции Sp1), POLG (полимеразу гамма (ДНК-направленную)), AFF2 (член 2 семейства AF4/FMR2), THBS1 (тромбоспондин 1), TP53 (опухолевый белок p53), ESR1 (эстрогеновый рецептор 1), CGGBP1 (белок 1, связывающий триплетный повтор CGG), ABT1 (активатор 1 базальной транскрипции), KLK3 (родственную калликреину пептидазу 3), PRNP (белок приона), JUN (онкоген jun), KCNN3 (кальций-активируемый калиевый канал средней/малой проводимости,

представитель 3 подсемейства N), BAX (BCL2-ассоциированный белок X), FRAXA (белок, ассоциированный с "редким" ломким сайтом, проявляющимся при недостатке фолиевой кислоты, fra(X)(q27.3) A (макроорхидизм, умственная отсталость)), KBTBD10 (белок 10, содержащий повтор Kelch и домен BTB (POZ)), MBNL1 (muscleblind-подобный белок (Drosophila)), RAD51 (гомолог RAD51 (гомолог RecA, E. coli) (S. cerevisiae)), NCOA3 (коактиватор 3 ядерных рецепторов), ERDA1 (белок с экспансией повторяющихся доменов, CAG/CTG 1), TSC1 (белок 1, ассоциированный с туберозным склерозом), COMP (олигомерный матриксный белок хряща), GCLC (каталитическую субъединицу глутаматцистеинлигазы), RRAD (Ras-родственный белок, ассоциированный с сахарным диабетом), MSH3 (гомолог 3 mutS (E. coli)), DRD2 (дофаминовый рецептор D2), CD44 (молекулу CD44 (система групп крови Indian)), CTCF (CCCTC-связывающий фактор (белок с "цинковыми пальцами")), CCND1 (циклин D1), CLSPN (гомолог класпина (Xenopus laevis)), MEF2A (энхансерный фактор 2A миоцитов), PTPRU (протеинтирозинфосфатазу рецепторного типа U), GAPDH (глицеральдегид-3-фосфатдегидрогеназу), TRIM22 (белок 22, содержащий тройной мотив), WT1 (белок 1 опухоли Вильмса), AHR (арил-углеводородный рецептор), GPX1 (глутатионпероксидазу 1), TPMT (тиопурин-S-метилтрансферазу), NDP (белок, ассоциированный с болезнью Норри (псевдоглиомой)), ARX (белок, кодируемый гомеобоксом гена, родственного aristaless), MUS81 (гомолог эндонуклеазы MUS81 (S. cerevisiae)), TYR (тирозиназу (глазокожный альбинизм IA)), EGR1 (белок 1 раннего ростового ответа), UNG (урацил-ДНК-гликозилазу), NUMBL (белок, подобный гомологу numb (Drosophila)), FABP2 (белок 2, связывающий жирные кислоты в кишечнике), EN2 (белок, кодируемый гомеобоксом engrailed 2), CRYGC (гамма-C-кристаллин), SRP14 (гомологичный РНК-связывающий белок Alu размером 14 кДа из частицы узнавания сигнала), CRYGB (гамма-B-кристаллин), PDCD1 (белок 1 запрограммированной гибели клеток), HOXA1 (белок, кодируемый гомеобоксом A1), ATXN2L (атаксин-2-подобный белок), PMS2 (PMS2, белок 2, противодействующий повышению уровня постмейотической сегрегации (S. cerevisiae)), GLA (альфа-галактозидазу), CBL (белок, кодируемый последовательностью, трансформирующей с экотропным ретровирусом Cas-Br-M (мышей)), FTH1 (полипептид 1 тяжелой субъединицы ферритина), IL12RB2 (бета-2-субъединицу рецептора интерлейкина 12), OTX2 (белок, кодируемый гомеобоксом orthodenticle 2), HOXA5 (белок, кодируемый гомеобоксом A5), POLG2 (вспомогательную гамма-2-субъединицу полимеразы (ДНК-

направленной)), DLX2 (белок, кодируемый гомеобоксом distal-less 2), SIRPA (сигнально-регуляторный белок альфа), OTX1 (белок, кодируемый гомеобоксом orthodenticle 1), AHRR (репрессор арил-углеводородного рецептора), MANF (мезэнцефальный нейротрофический фактор, происходящий из астроцитов), TMEM158 (трансмембранный белок 158 (ген/псевдоген)) и ENSG00000078687.

[0712] Описанные здесь праймерные редакторы могут применяться для сокращения участков экспансии триплетных повторов в рамках любых из указанных выше связанных с заболеваниями белков, включая следующие гены, ассоциированные с полиглутаминовыми заболеваниями, связанными с экспансией триплетов (которые демонстрируют конкретное расположение патогенных повторов, которые могут быть удалены полностью или частично с помощью праймерного редактирования).

ЗАБОЛЕВАНИЕ, СВЯЗАННОЕ С ЭКСПАНСИЕЙ ТРИПЛЕТОВ	**СВЯЗАННЫЙ ГЕН**	**ПОЛОЖЕНИЕ ПАТОГЕННЫХ ПОВТОРОВ**
ДЕНТАТО-РУБРО-ПАЛЛИДО-ЛЬЮИСОВА АТРОФИЯ	ATN1 АТРОФИН-1	49-88
БОЛЕЗНЬ ХАНТИНГТОНА	HTT ГЕН ХАНТИНГТИНА	36-250
СПИНАЛЬНАЯ И БУЛЬБАРНАЯ МЫШЕЧНАЯ АТРОФИЯ	AR АНДРОГЕНОВЫЙ РЕЦЕПТОР	38-62
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 1	ATXN1 АТАКСИН 1	49-88
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 2	ATXN2 АТАКСИН 2	33-77
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 3	ATXN3 АТАКСИН 3	55-86
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 6	CACNA1A	21-30
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 7	ATXN7 АТАКСИН 7	38-120
СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 17	TBP TATA-СВЯЗЫВАЮЩИЙ БЕЛОЙ	47-63

[0713] Описанные здесь праймерные редакторы также могут применяться для сокращения участков экспансии триплетных повторов, которые обычно располагаются в следующих генах, ассоциированных с неполиглутаминовыми заболеваниями, связанными с экспансией триплетов:

ЗАБОЛЕВАНИЕ, СВЯЗАННОЕ С ЭКСПАНСИЕЙ ТРИПЛЕТОВ	**СВЯЗАННЫЙ ГЕН**	**ПОЛОЖЕНИЕ ПАТОГЕННЫХ ПОВТОРОВ**

FRAXA (СИНДРОМ ЛОМКОЙ Х-ХРОМОСОМЫ)	FMR1 БЕЛОК УМСТВЕННОЙ ОТСТАЛОСТИ ПРИ ЛОМКОЙ Х-ХРОМОСОМЕ	230+
FXTAS (ТРЕМОРНЫЙ/АТАКСИЧЕСКИЙ СИНДРОМ, АССОЦИИРОВАННЫЙ С ЛОМКОЙ Х-ХРОМОСОМОЙ)	FMR1	55-200
FRAXE (УМСТВЕННАЯ ОТСТАЛОСТЬ, АССОЦИИРОВАННАЯ С ЛОМКОЙ XE)	AFF2	200+
FRDA (АТАКСИЯ ФРИДРЕЙХА)	FXN ФРАТАКСИН	100+
DM1 (МИОТОНИЧЕСКАЯ ДИСТРОФИЯ ТИП 1)	DMPK ПРОТЕИНКИНАЗА МИОТОНИЧЕСКОЙ ДИСТРОФИИ	50+
SCA8 (СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 8)	SCA8 АТАКСИН 8	110-250
SCA12 (СПИНАЛЬНО-ЦЕРЕБЕЛЛЯРНАЯ АТАКСИЯ ТИПА 12)	PPP2R2B СЕРИН/ТРЕОНИН ПРОТЕИНФОСФАТАЗА 2A	66-78

[0714] Праймерное редактирование может быть реализовано для сокращения участков экспансии триплетов с помощью PEgRNA с матрицей для редактирования, которая сконструирована для удаления, по меньшей мере, одного кодона участка экспансии триплетов. В других вариантах осуществления PEgRNA для применения в рамках праймерного редактирования используются для удаления, по меньшей мере, 1, или 2, или 3, или 4, или 5, или 6, или 7, или 8, или 9, или 10, или 11, или 12, или 13, или 14, или 15, или 16, или 17, или 18, или 19, или 20, или 21, или 22, или 23, или 24, или 25, или 26, или 27, или 28, или 29, или 30, или 31, или 32, или 33, или 34, или 35, или 36, или 37, или 38, или 39, или 40, или 41, или 42, или 43, или 44, или 45, или 46, или 47, или 48, или 49, или 50, или 51, или 52, или 53, или 54, или 55, или 56, или 57, или 58, или 59, или 60, или 61, или 62, или 63, или 64, или 65, или 66, или 67, или 68, или 69, или 70, или 71, или 72, или 73, или 74, или 75, или 76, или 77, или 78, или 79, или 80, или 81, или 82, или 83, или 84, или 85, или 86, или 87, или 88, или 89, или 90, или 91, или 92, или 93, или 94, или 95, или 96, или 97, или 98, или 99, или 100, или большего количества кодонов из участка экспансии триплетов для достижения здорового (то есть

не ассоциированного с продуцированием заболевания) количества триплетных повторов.

[0715] В других вариантах осуществления PEgRNA для применения в рамках праймерного редактирования используются для удаления, по меньшей мере, 1, или 2, или 3, или 4, или 5, или 6, или 7, или 8, или 9, или 10, или 11, или 12, или 13, или 14, или 15, или 16, или 17, или 18, или 19, или 20 или большего количества кодонов из участка экспансии триплетов для достижения здорового (то есть не ассоциированного с продуцированием заболевания) количества триплетных повторов.

[0716] В других вариантах осуществления PEgRNA для применения в рамках праймерного редактирования используются для удаления, по меньшей мере, 1, или 2, или 3, или 4, или 5, или 6, или 7, или 8, или 9, или 10, или 11, или 12, или 13, или 14, или 15 или большего количества кодонов из участка экспансии триплетов для достижения здорового (то есть не ассоциированного с продуцированием заболевания) количества триплетных повторов.

[0717] В других вариантах осуществления PEgRNA для применения в рамках праймерного редактирования используются для удаления, по меньшей мере, 1, или 2, или 3, или 4, или 5, или 6, или 7, или 8, или 9, или 10 или большего количества кодонов из участка экспансии триплетов для достижения здорового (то есть не ассоциированного с продуцированием заболевания) количества триплетных повторов.

[0718] Праймерное редактирование может быть реализовано для корректирования любого участка экспансии триплетов, например описанных в Budworth et al., "A Brief History of Triplet Repeat Diseases," *Methods Mol Biol*, 2013, 1010: 3-17, US 20011/00165540 A1 (Геномное редактирование генов, ассоциированных с заболеваниями, связанными с экспансией тринуклеотидных повторов, у животных), заявке на патент США 2016/0355796 A1 (Композиции и способы применения систем CRISPR-Cas при заболеваниях, связанных с нуклеотидными повторами.

[0719] В различных вариантах осуществления в рамках изобретения предусмотрен конструкт для праймерного редактирования, подходящий для применения в клетке, имеющей участок экспансии тринуклеотидных повторов в дефектном гене, включающий (a) продукт слияния в виде праймерного редактора, включающий napDNAbp и обратную транскриптазу, (b) PEgRNA, включающую спейсерную последовательность, которая нацелена на участок экспансии тринуклеотидных

повторов, и плечо удлинения, включающее матрицу для редактирования, которая кодирует удаление участка экспансии тринуклеотидных повторов.

[0720] В различных других вариантах осуществления в рамках изобретения предусмотрен способ удаления всего или части участка экспансии тринуклеотидных повторов в дефектном гене в клетке с помощью праймерного редактирования, включающий контактирование клетки с продуктом слияния праймерного редактора, включающего napDNAbp и обратную транскриптазу и PEgRNA, включающую спейсерную последовательность, которая нацелена на участок экспансии тринуклеотидных повторов, и плечо удлинения, включающее матрицу для редактирования, которая кодирует удаление участка экспансии тринуклеотидных повторов.

[0721] В различных вариантах осуществления тринуклеотидный повтор включает повторяющиеся тринуклеотиды CTG, CAG, CGG, CCG, GAA или TTC.

[0722] В различных других вариантах осуществления — тетрануклеотидные повторы, пентануклеотидные повторы, или гексануклеотидные повторы.

D. Применение праймерного редактирования для пептидного мечения

[0723] В другом аспекте изобретение предусматривает способ применение описанных здесь праймерных редакторов для генетического внесения одной или нескольких пептидных меток в белок с помощью праймерного редактирования. Более конкретно, изобретение предусматривает способ генетического внесения одной или нескольких пептидных меток в белок, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с праймерным редактором, сконструированным для внесения в нее второй нуклеотидной последовательности, кодирующей указанную одну или несколько пептидных меток для получения рекомбинантной нуклеотидной последовательности, которая кодирует слитый белок, включающий белок, слитый с белковой меткой.

[0724] В других вариантах осуществления изобретение предусматривает способ получения слитого белка, включающий представляющий интерес пептид и одну или несколько пептидных меток, при этом способ включает: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с праймерным редактором, сконструированным для внесения в нее второй нуклеотидной последовательности, кодирующей указанную одну или несколько пептидных меток

для получения рекомбинантной нуклеотидной последовательности, которая кодирует слитый белок, включающий белок, слитый с белковой меткой.

[0725] В различных вариантах осуществления нуклеотидная последовательность-мишень представляет собой определенный представляющий интерес ген в геномной ДНК. Представляющий интерес ген может кодировать представляющий интерес белок (например рецептор, фермент, терапевтический белок, мембранный белок, транспортный белок, белок передачи сигнала, или иммунологический белок, и так далее). Представляющий интерес ген может также кодировать молекулу РНК, включая, но без ограничения, информационную РНК (иРНК), транспортную РНК (тРНК), рибисомную РНК (рРНК), малую ядерную РНК (мяРНК), антисмысловую РНК, направляющую РНК, микроРНК (мРНК), малую интерферирующую РНК (миРНК) и бесклеточную РНК (бкРНК).

[0726] Пептидная метка может представлять собой любую пептидную метку или ее вариант, который придает одну или несколько функций белку для таких целей, как отделение, очистка, визуализация, солюбилизация или детектирование. Пептидные метки могут включать "аффинные метки" (для реализации очистки белка), "метки солюбилизации" (для обеспечения правильного фолдинга белков), "хроматографические метки" (для изменения хроматографических свойств белков), "эпитопные метки" (для связывания с антителами с высокой аффинностью), "флуоресцентные метки" (для обеспечения визуализации белков в клетке или in vitro). Примеры пептидных меток включают, но без ограничения, следующие метки:

НАЗВАНИЕ	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ**	**SEQ ID NO:**
AVITAG™	GLNDIFEAQKIEWHE	SEQ ID NO: 245
С-МЕТКА	EPEA	SEQ ID NO: 246
КАЛМОДУЛИНОВАЯ МЕТКА	KRRWKKNFIAVSAANRFKKISSSGAL	SEQ ID NO: 247
ПОЛИГЛУТАМАТНАЯ МЕТКА	EEEEEE	SEQ ID NO: 248
E-МЕТКА	GAPVPYPDPLEPR	SEQ ID NO: 249
FLAG-МЕТКА	DYKDDDDK	SEQ ID NO: 2
HA-МЕТКА	YPYDVPDYA	SEQ ID NO: 5
HIS-МЕТКА	H (HIS_1) HH (HIS_2) HHH (HIS_3) HHHH (HIS_4) HHHHH (HIS_5) HHHHHH (HIS_6)	SEQ ID NO: 252 SEQ ID NO: 253 SEQ ID NO: 254 SEQ ID NO: 255 SEQ ID NO: 256 SEQ ID NO: 257

	HHHHHHH (HIS_7) HHHHHHHH (HIS_8) HHHHHHHHH (HIS_9) HHHHHHHHHH (HIS_{10}) HHHHHHHHHH…H…(HIS_N, WHEREIN N = 1-25)	SEQ ID NO: 258 SEQ ID NO: 259 SEQ ID NO: 260 SEQ ID NO: 261 SEQ ID NO: 262
MYC-МЕТКА	EQKLISEEDL	SEQ ID NO: 6
NE-МЕТКА	TKENPRSNQEESYDDNES	SEQ ID NO: 264
RHO1D4-МЕТКА	TETSQVAPA	SEQ ID NO: 265
S-МЕТКА	KETAAAKFERQHMDS	SEQ ID NO: 266
SBP-МЕТКА	MDEKTTGWRGGHVVEGLAGELEQ LRARLEHHPQGQREP	SEQ ID NO: 267
SOFTAG-1	SLAELLNAGLGGS	SEQ ID NO: 268
SOFTAG-2	TQDPSRVG	SEQ ID NO: 269
SPOT-МЕТКА	PDRVRAVSHWSS	SEQ ID NO: 270
STREP-МЕТКА	WSHPQFEK	SEQ ID NO: 271
TC МЕТКА	CCPGCC	SEQ ID NO: 272
TY МЕТКА	EVHTNQDPLD	SEQ ID NO: 273
V5 МЕТКА	GKPIPNPLLGLDST	SEQ ID NO: 3
VSV-МЕТКА	YTDIEMNRLGK	SEQ ID NO: 275
МЕТКА XPRESS	DLYDDDDK	SEQ ID NO: 276

[0727] Пептидные метки также могут представлять собой следующие аффинные метки (для отделения и/или очистки белков) (как описано в Таблице 9.9.1 у Kimple et al., “Overview of Affinity Tags for Protein Purification,” *Curr Protoc Protein Sci*, 2013, 73:Unit-9.9, который приведен здесь в качестве ссылки).

НАЗВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	
ЭПИТОП AU1	DTYRYI	SEQ ID NO: 278
ЭПИТОП AU5	TDFYLK	SEQ ID NO: 279
ЭПИТОП БАКТЕРИОФАГА T7 (T7-МЕТКА)	MASMTGGQQMG	SEQ ID NO: 280
МЕТКА НА ОСНОВЕ ВИРУСА BLUETONGUE (B-МЕТКА)	QYPALT	SEQ ID NO: 281
ЭПИТОП E2	SSTSSDFRDR	SEQ ID NO: 282
АФФИННАЯ ГИСТИДИНОВАЯ МЕТКА (HAT)	KDHLIHNVHKEFHAHAHNK	SEQ ID NO: 283
ЭПИТОП HSV	QPELAPED	SEQ ID NO: 284
ПОЛИАРГИНИН (ARG-МЕТКА)	RRRRR	SEQ ID NO: 285
ПОЛИАСПАРТАТ (ASP-МЕТКА)	CCCC	SEQ ID NO: 286
ПОЛИФЕНИЛАЛАНИН (PHE-МЕТКА)	FFFFFFFFFF	SEQ ID NO: 287
S1-МЕТКА	NANNPDWDF	SEQ ID NO: 288

S-METKA	KETAAAKFERQHMDS	SEQ ID NO: 266
VSV-G	YTDIEMNRLGK	SEQ ID NO: 275

[0728] В определенных вариантах осуществления пептидные метки могут включать His^6-метку, FLAG-метку, V5-метку, GCN4-метку, HA-метку, Myc-метку, FIAsH/ReAsH-метку, сортазный субстрат, π-клэмп.

[0729] В различных вариантах осуществления пептидные метки могут применяться для вариантов использования, которые включают флуоресцентное мечение белка, иммуноосаждение, иммуноблоттинг, иммуногистохимию, рекрутирование белка, индуцируемые белковые дегроны, и скрининг в рамках всего генома.

[0730] В различных других вариантах осуществления пептидная метка может включать последовательность интеина для внесения функции в виде аутосплайсинга белка. В соответствии с используемым здесь значением термин "интеин" обозначает автопроцессирующие полипептидные домены, располагающиеся в организмах всех царств жизни. Интеин (intervening protein, промежуточный белок) обеспечивает реализацию уникального автопроцессирующего события, известного как сплайсинг белка, при котором он эксцизирует себя из более крупного прекурсорного полипептида за счет расщепления двух пептидных связей и, в рамках данного процесса, обеспечивает лигирование фланкирующих последовательностей экстеина (внешнего белка) за счет образования новой пептидной связи Данное перестроение осуществляется после трансляции (или возможно одновременно с трансляцией). поскольку гены интеина являются встроенными в каркас в рамках других кодирующих белок генов. Более того, опосредованный интеином сплайсинг белков является спонтанным; он не требует внешнего фактора или источника энергии, а лишь фолдинг домена интеина. Данный процесс также известен как цис-сплайсинг белка, в отличие от природного процесса транс-сплайсинга белка с использованием "разделенных интеинов". Интеины представляют собой эквивалент белка аутосплайсирующих РНК интронов (см. Perler et al., Nucleic Acids Res. 22:1125-1127 (1994)), которые катализируют свою собственную эксцизию из прекурсорного белка с одновременным слиянием фланкирующих последовательностей белка, известных как экстеины (описаны в Perler et al., Curr. Opin. Chem. Biol. 1:292-299 (1997); Perler, F. B. Cell 92(1):1-4 (1998); Xu et al., EMBO J. 15(19):5146-5153 (1996)).

[0731] Механизм процесса сплайсинга белка был подробно изучен (Chong, et al., J. Biol. Chem. 1996, 271, 22159-22168; Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15,

5146-5153), при этом консервативные аминокислоты были обнаружены в точках сплайсинга интеина и экстеина (Xu, et al., EMBO Journal, 1994, 13 5517-522).

[0732] Интеины также могут существовать в виде двух фрагментов, кодируемых двумя отдельно транскрибируемыми и транслируемыми генами. Данные так называемые разделенные интеины самоассоциируются и катализируют активность сплайсинга белка in trans. Разделенные интеины были идентифицированы у многих цианобактерий и архей (Caspi et al, Mol Microbiol. 50: 1569-1577 (2003); Choi J. et al, J Mol Biol. 556: 1093-1106 (2006.); Dassa B. et al, Biochemistry. 46:322-330 (2007.); Liu X. and Yang J., J Biol Chem. 275:26315-26318 (2003); Wu H. et al. Proc Natl Acad Sci USA. £5:9226-9231 (1998.); и Zettler J. et al, FEBS Letters. 553:909-914 (2009)), но до сих пор не были обнаружены у эукариотов. Недавно биоинформационный анализ природных метагеномных данных выявил 26 различных локусов с новым геномным строением. В каждом локусе консервативный кодирующий фермент участок прерывается разделенным интеином со свободным эндонуклеазным геном, вставленным между участками, кодирующими субдомены интеина. Среди них пять локусов были полностью собраны: ДНК хеликазы (gp41-l, gp41-8); инозин-5 '-монофосфатдегидрогеназа (IMPDH-1); и каталитические субъединицы рибонуклеотид-редуктазы (NrdA-2 и NrdJ-1). Похоже, что данная фрагментированная организация генов присутствует в основном у фагов (Dassa et al, Nucleic Acids Research. 57:2560-2573 (2009)).

[0733] В определенных вариантах осуществления описанные здесь праймерные редакторы могут применяться для вставки меток в виде разделенного интеина в два различных белка, вызывая их внутриклеточное лигирования при совместной экспрессии с образованием слитого белка. При транс-сплайсинге белка один белок-предшественник состоит из части в виде N-экстеина и далее N-интеина, другой белок-предшественник состоит из С-интеина и далее части в виде С-экстеина, при этом реакция транс-сплайсинга (катализируемая совместно N- и С-интеинами) обеспечивает эксцизию данных двух последовательностей интеина и соединение данных двух последовательностей экстеина с помощью пептидной связи. Транс-сплайсинг белка, являясь ферментативной реакцией, может работать при очень низких (например микромолярных) концентрациях белков и может осуществляться при физиологических условиях.

[0734] Разделенный интеин Npu DnaE был описан как имеющий наиболее высокую отмеченную скорость реакции транс-сплайсинга белка. Кроме того, реакция сплайсинга белка Npu DnaE считается устойчивой и связанной с высоким выходом в плане различных последовательностей экстеина, температурами от 6 до 37°C, и наличием до 6M мочевины (Zettler J. et al, FEBS Letters. 553:909-914 (2009); Iwai I. et al, FEBS Letters 550: 1853-1858 (2006)). Как ожидается, при внесении мутации Cysl Ala в N-домене данных интеинов первоначальный N-C-ацильный сдвиг и, таким образом, сплайсинг белка блокируется. К сожалению, реакция C-концевого расщепления также практически полностью ингибировалась. Зависимость циклизации аспарагина в C-концевом сплайсинговом соединении от ацильного сдвига в N-концевой неустойчивой пептидной связи вероятно является уникальным свойством, присущим аллелям естественным образом разделенного интеина DnaE (Zettler J. et al. FEBS Letters. 555:909-914 (2009).

[0735] Транс-сплайсинг белка, катализируемый разделенными интеинами, представляет собой полностью ферментативный способ лигирования белка. Разделенный интеин представляет собой по существу непрерывный интеин (например мини-интеин), разделенный на две части, называемые N-интеином и C-интеином, соответственно. N-интеин и C-интеин разделенного интеина могут нековалентно связываться для образования активного интеина и катализировать реакцию сплайсинга по существу так же, как и непрерывный интеин. Разделенные интеины были обнаружены в природе, а также сконструированы в лабораториях. В соответствии с используемым здесь значением термин "разделенный интеин" обозначает любой интеин, в котором присутствует один или несколько разрывов пептидных связей между N-концевыми и C-концевыми последовательностями аминокислот, так что N-концевые и C-концевые последовательности становятся отдельными молекулами, которые могут нековалентно связываться или восстанавливаться с образованием интеина, который является функциональным для реакций транс-сплайсинга белка. Для получения разделенного интеина для использования в рамках способов по изобретению может применяться любой каталитически активный интеин или его фрагмент. Например, в одном аспекте разделенный интеин может быть получен из эукариотического интеина. В другом аспекте разделенный интеин может быть получен из бактериального интеина. В другом аспекте разделенный интеин может быть получен из архейного интеина. Предпочтительно, полученный таким образом

разделенный интеин будет иметь только аминокислотные последовательности, необходимые для катализирования реакций транс-сплайсинга.

[0736] Разделенные интеины могут быть созданы из непрерывных интеинов путем конструирования одного или нескольких участков разделения в неструктурированной петле или внесения аминокислотной последовательности между 12 консервативными бета-цепями, находящимися в структуре мини-интеинов. Может присутствовать некоторая гибкость в плане расположения участка разрыва в участках между бета-цепями при условии, что создание разрыва не будет нарушать структуру интеина, в частности структурированных бета-цепей, до значительной степени, при которой теряется активность, связанная со сплайсингом белка.

[0737] Описанные здесь праймерные редакторы могут обеспечивать встраивание пептидных меток (включая интеины) в С-конец представляющего интерес белка. В других вариантах осуществления пептидные метки (включая интеины) могут быть встроены в N-конец представляющего интерес белка. Пептидные метки также могут быть встроены во внутреннюю часть представляющего интерес белка. Полученные слитые белки, созданные с помощью описанных здесь праймерных редакторов, могут иметь следующие структуры:

[0738] [представляющий интерес белок] – [пептидная метка];

[0739] [пептидная метка] – [представляющий интерес белок]; или

[0740] [представляющий интерес белок – N-концевой участок] – [пептидная метка] – [представляющий интерес белок– С-концевой участок].

[0741] Принципы конструирования направляющей РНК для применения в рамках пептидного мечения могут применяться при пептидном мечении. Например, в одном варианте осуществления PEgRNA для пептидного мечения может иметь следующую структуру: 5′-[спейсерная последовательность]-[ядро или каркас нРНК]-[плечо удлинения]-3′, где плечо удлинения включает в направлении от 5′ к 3′ плечо гомологии, матрицу для редактирования (включающую последовательность, которая кодирует пептидную метку) и участок связывания праймера. Данная конфигурация изображена на Фиг.3D и Фиг.24.

[0742] В другом варианте осуществления PEgRNA для пептидного мечения может иметь следующую структуру: 5′-[плечо удлинения]-[спейсерная последовательность]-[ядро или каркас нРНК]-3′, где плечо удлинения включает в направлении от 5′ к 3′ плечо гомологии, матрицу для редактирования (включающую последовательность,

которая кодирует пептидную метку) и участок связывания праймера. Данная конфигурация изображена на Фиг.3E.

[0743] Варианты осуществления пептидного мечения с помощью праймерного редактирования показаны на Фиг.25 и 26 и описаны в Примере 4.

E. Применение праймерного редактирования для предотвращения или лечения прионного заболевания

[0744] Праймерное редактирование также может применяться для предотвращения или остановки прогрессирования прионного заболевания за счет внесения одной или нескольких защитных мутаций в прионные белки (PRNP), которые подвергаются неправильному фолдингу при заболевании. Прионные болезни или трансмиссивные губчатые энцефалопатии (TSE) представляют собой семейство редких прогрессирующих нейродегенеративных заболеваний, которыми страдают как люди, так и животные. Они характеризуются длительными инкубационными периодами, характеризуются губчатыми изменениями, связанными с потерей нейронов, и неспособностью индуцировать воспалительный ответ.

[0745] У человека прионное заболевание включает болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, и куру. У животных прионное заболевание включает коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрепи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных. Праймерное редактирование может применяться для внесения защитных точечных мутаций в прионный белок для предотвращения или остановки прогрессирования любого из данных прионных заболеваний.

[0746] Классическая CJD представляет собой прионное заболевание человека. Она представляет собой нейродегенеративное заболевание с характерными клиническими и диагностическими признаками. Данное заболевание быстро прогрессирует и всегда является смертельным. Заражение данным заболеванием приводит к смерти обычно в течение 1 года после возникновения заболевания. CJD является быстро прогрессирующим неизменно смертельным нейродегенеративным заболеванием, которое, как считается, вызывается анормальной изформой клеточного гликопротеина, известного как прионный белок. CJD возникает по всему миру, при этом сообщается, что оцениваемая ежегодная частотность во многих странах, включая Соединенные

Штаты, составляет около одного случая на один миллион населения. Подавляющее большинство пациентов с CJD обычно умирают в течение 1 года после возникновения заболевания. CJD классифицируется как трансмиссивная губчатая энцефалопатия (TSE) наряду с другими прионными заболеваниями, которые возникают у человека и животных. У приблизительно 80% пациентов CJD возникает в виде случайного заболевания без определенного паттерна передачи. У меньшей части пациентов (от 5 до 15%) CJD развивается вследствие наследственных мутаций гена прионного белка. Данные наследственные формы включают синдром Герстмана-Штрауслера-Шейнкера и фатальную семейную бессонницу. Способы лечения CJD в настоящее время неизвестны.

[0747] Вариант болезни Крейтцфельдта-Якоба (vCJD) представляет собой прионное заболевание, которое было впервые описано в 1996 в Соединенном Королевстве. В настоящее время существует сильное научное свидетельство того, что агент, отвечающий за вспышку прионного заболевания у коров, коровьей губчатой энцефалопатии (BSE или "коровьего бешенства"), является тем же агентом, ответственным за вспышку vCJD у людей. Вариант CJD (vCJD) не является тем же заболеванием, что и классическая CJD (которую часто называют просто CJD). Она имеет клинические и патологические характеристики, отличающиеся от классической CJD. Каждое заболевание также обладает определенным генетическим профилем гена прионного белка. Оба заболевания являются необратимо смертельными заболеваниями мозга с необычно длительными инкубационными периодами, измеряемыми годами, и вызываются нестандартным переносимым агентом, называемым прионом. Способы лечения vCJD в настоящее время неизвестны.

[0748] BSE (коровья губчатая энцефалопатия или "коровье бешенство") представляет собой прогрессирующее неврологическое заболевание крупного рогатого скота, которое возникает вследствие инфицирования необычным переносимым агентом, называемым прионом. Природа данного переносимого агента не до конца понятна. В настоящее время наиболее распространенная теория заключается в том, что данный агент представляет собой модифицированную форму нормального белка, известную как прионный белок. Вследствие неизвестных в настоящее время причин нормальный прионный белок превращается в патогенную (вредоносную) форму, которая затем повреждает центральную нервную систему крупного рогатого скота. Существует значительное свидетельство того, что существуют различные виды BSE: типичный или

классический вид BSE, ответственный за вспышку в Соединенном Королевстве, и два атипичных вида (виды H и L). Способы лечения BSE в настоящее время неизвестны.

[0749] Хроническое истощающее заболевание (CWD) представляет собой прионное заболевание, которое поражает оленей, настоящих оленей, северных оленей, пятнистых оленей и лосей. Оно было обнаружено в некоторых областях Северной Америки, включая Канаду и Соединенные Штаты, Норвегию и Южную Корею. Может пройти более года прежде чем у инфицированного животного появятся симптомы, которые могут включать значительную потерю веса (истощение), спотыкание, вялость и другие неврологические симптомы. CWD может поражать животных всех возрастов, при этом некоторые инфицированные животные могут умирать даже без развития заболевания. CWD является смертельным для животных, при этом способы его лечения или вакцины отсутствуют.

[0750] Считается, что вызывающими TSE агентами являются прионы. Термин "прионы" обозначает анормальные патогенные агенты, которые являются трансмиссивными и способными индуцировать анормальный фолдинг определенных нормальных клеточных белков, называемых прионными белками и находящихся в основном в мозге. Функции данных нормальных прионных белков все еще не полностью известны. Анормальный фолдинг прионных белков приводит к повреждению мозга и характерным признакам и симптомам заболевания. Прионные заболевания обычно быстро прогрессируют и всегда являются смертельными.

[0751] В соответствии с используемым здесь значением, термин "прион" означает инфекционную частицу, в отношении которой известно, что она вызывает заболевания (губчатую энцефалопатию) у человека и животных. Термин "прион" является сокращением слов "белок" (протеин) и "инфекция", при этом данные частицы в основном (если не полностью) состоят из молекул $PRNP^{Sc}$, кодируемых геном PRNP, который экспрессирует $PRNP^{C}$, который изменяет конформацию с образованием $PRNP^{Sc}$. Прионы являются отличными от бактерий, вирусов и вироидов. Известные прионы включают те из них, которые инфицируют животных и вызывают скрепи, представляющее собой трансмиссивное дегенеративное заболевание нервной системы овец и коз, а также коровью губчатую энцефалопатию (BSE) или коровье бешенство и губчатую энцефалопатию кошек. Известно, что четырьмя описанными выше прионными заболеваниями, которые поражают человека, являются (1) куру, (2) болезнь Крейтцфельдта-Якоба (CJD), (3) болезнь Герстмана-Штрауслера-Шейнкера

(GSS), и (4) фатальная семейная бессонница (FFI). В соответствии с используемым здесь значением термин "прион" включает все формы прионов, вызывающие все или любые из данных заболеваний, или другие из них у любых животных, в частности у человека и домашних сельскохозяйственных животных.

[0752] В целом, и без связи с теорией, прионные заболевания вызываются неправильным фолдингом прионных белков. При таких заболеваниях, которые часто называются здесь связанными с отложением заболеваниями, неправильный фолдинг прионных белков может осуществляться следующим образом. Если A представляет собой нормально синтезированный генный продукт, который реализует требуемую физиологическую роль в мономерном или олигомерном состоянии, то A* представляет собой конформационно активированную форму A, которая способна подвергаться значительному конформационному изменению, B представляет собой конформационно измененное состояние, для которого являются предпочтительными мультимерные сборки (то есть подвергнутая неправильному фолдингу форма, которая образует отложения), а B_n представляет собой мультимерный материал, который является патогенным и относительно сложным в плане рециклизации. Для прионных заболеваний $PRNP^C$ и $PRNP^{Sc}$ соответствуют состояниям A и B_n, где A является по большей части спиральным и мономерным, а B_n имеет высокое содержание β и является мультимерным.

[0753] Известно, что определенные мутации в прионных белках могут быть ассоциированы с повышенным риском прионного заболевания. С другой стороны, определенные мутации в прионных белках могут являться защитными по своей природе. См. Bagynszky et al., “Characterization of mutations in PRNP (prion) gene and their possible roles in neurodegenerative diseases,” *Neuropsychiatr Dis Treat.*, 2018; 14: 2067-2085, содержание которого приведено здесь в качестве ссылки.

[0754] PRNP (NCBI RefSeq No. NP_000302.1 (SEQ ID NO: 291)) — прионный белок человека — кодируется геном длиной 16 kb, расположенным на хромосоме 20 (4686151-4701588). Он содержит два экзона, при этом экзон 2 несет открытую рамку считывания, которая кодирует белок PrP длиной 253 аминокислоты (AA). Экзон 1 представляет собой некодирующий экзон, который может служить в качестве участка инициации транскрипции. Посттрансляционные модификации приводят к удалению N-концевого фрагмента (NTF) в виде первых 22 AA и C-концевого фрагмента (CTF) в виде последних 23 AA. NTF отщепляется после транспорта PrP в эндоплазматическую

сеть (ER), в то время как CTF отщепляется GPI-якорем (сигнальный пептид [GPI-SP] гликозилфосфатидилинозитол [GPI]). GPI-якорь может принимать участие в транспорте белка PrP. Он также может принимать участие в присоединении прионного белка к внешней поверхности мембраны клетки. Нормальный PrP состоит из длинной N-концевой петли (которая содержит участок октапептидного повтора), двух коротких β-листов, трех α-спиралей, и C-концевого участка (который содержит GPI-якорь). Расщепление PrP приводит к получению гликопротеина длиной 208 AA, присоединенного к клеточной мембране.

[0755] Аминокислотная последовательность PRNP (NP_000302.1) является следующей:

MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGG
GWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKT
NMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQV
YYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQ
YERESQAYYQRGSSMVLFSSPPVILLISFLIFLIVG (SEQ ID NO: 291).

[0756] Аминокислотная последовательность PRNP (NP_000302.1) кодируется нуклеотидной последовательностью (NCBI Ref. Seq No. NM_000311.5, "прионный белок homo sapiens (PRNP), транскриптный вариант 1, иРНК)", которая является следующей:

GCGAACCTTGGCTGCTGGATGCTGGTTCTCTTTGTGGCCACATGGAGTGACCTGG
GCCTCTGCAAGAAGCGCCCGAAGCCTGGAGGATGGAACACTGGGGGCAGCCGAT
ACCCGGGGCAGGGCAGCCCTGGAGGCAACCGCTACCCACCTCAGGGCGGTGGTG
GCTGGGGGCAGCCTCATGGTGGTGGCTGGGGGCAGCCTCATGGTGGTGGCTGGG
GGCAGCCCCATGGTGGTGGCTGGGGACAGCCTCATGGTGGTGGCTGGGGTCAAG
GAGGTGGCACCCACAGTCAGTGGAACAAGCCGAGTAAGCCAAAAACCAACATGA
AGCACATGGCTGGTGCTGCAGCAGCTGGGGCAGTGGTGGGGGGCCTTGGCGGCT
ACATGCTGGGAAGTGCCATGAGCAGGCCCATCATACATTTCGGCAGTGACTATGAG
GACCGTTACTATCGTGAAAACATGCACCGTTACCCCAACCAAGTGTACTACAGGCC
CATGGATGAGTACAGCAACCAGAACAACTTTGTGCACGACTGCGTCAATATCACA
ATCAAGCAGCACACGGTCACCACAACCACCAAGGGGGAGAACTTCACCGAGACC
GACGTTAAGATGATGGAGCGCGTGGTTGAGCAGATGTGTATCACCCAGTACGAGA
GGGAATCTCAGGCCTATTACCAGAGAGGATCGAGCATGGTCCTCTTCTCCTCTCCA
CCTGTGATCCTCCTGATCTCTTTCCTCATCTTCCTGATAGTGGGATGAGGAAGGTCT
TCCTGTTTTCACCATCTTTCTAATCTTTTTCCAGCTTGAGGGAGGCGGTATCCACCT
GCAGCCCTTTTAGTGGTGGTGTCTCACTCTTTCTTCTCTCTTTGTCCCGGATAGGCT
AATCAATACCCTTGGCACTGATGGGCACTGGAAAACATAGAGTAGACCTGAGATG
CTGGTCAAGCCCCCTTTGATTGAGTTCATCATGAGCCGTTGCTAATGCCAGGCCAG
TAAAAGTATAACAGCAAATAACCATTGGTTAATCTGGACTTATTTTTGGACTTAGTG
CAACAGGTTGAGGCTAAAACAAATCTCAGAACAGTCTGAAATACCTTTGCCTGGA
TACCTCTGGCTCCTTCAGCAGCTAGAGCTCAGTATACTAATGCCCTATCTTAGTAGA
GATTTCATAGCTATTTAGAGATATTTTCCATTTTAAGAAAACCCGACAACATTTCTG
CCAGGTTTGTTAGGAGGCCACATGATACTTATTCAAAAAAATCCTAGAGATTCTTA
GCTCTTGGGATGCAGGCTCAGCCCGCTGGAGCATGAGCTCTGTGTGTACCGAGAA
CTGGGGTGATGTTTTACTTTTCACAGTATGGGCTACACAGCAGCTGTTCAACAAGA

GTAAATATTGTCACAACACTGAACCTCTGGCTAGAGGACATATTCACAGTGAACAT
AACTGTAACATATATGAAAGGCTTCTGGGACTTGAAATCAAATGTTTGGGAATGGT
GCCCTTGGAGGCAACCTCCCATTTTAGATGTTTAAAGGACCCTATATGTGGCATTCC
TTTCTTTAAACTATAGGTAATTAAGGCAGCTGAAAAGTAAATTGCCTTCTAGACAC
TGAAGGCAAATCTCCTTTGTCCATTTACCTGGAAACCAGAATGATTTTGACATACA
GGAGAGCTGCAGTTGTGAAAGCACCATCATCATAGAGGATGATGTAATTAAAAAAT
GGTCAGTGTGCAAAGAAAAGAACTGCTTGCATTTCTTTATTTCTGTCTCATAATTG
TCAAAAACCAGAATTAGGTCAAGTTCATAGTTTCTGTAATTGGCTTTTGAATCAAA
GAATAGGGAGACAATCTAAAAAATATCTTAGGTTGGAGATGACAGAAATATGATTG
ATTTGAAGTGGAAAAAGAAATTCTGTTAATGTTAATTAAAGTAAAATTATTCCCTG
AATTGTTTGATATTGTCACCTAGCAGATATGTATTACTTTTCTGCAATGTTATTATTG
GCTTGCACTTTGTGAGTATTCTATGTAAAAATATATATGTATATAAAATATATATTGCA
TAGGACAGACTTAGGAGTTTTGTTTAGAGCAGTTAACATCTGAAGTGTCTAATGCA
TTAACTTTTGTAAGGTACTGAATACTTAATATGTGGGAAACCCTTTTGCGTGGTCCT
TAGGCTTACAATGTGCACTGAATCGTTTCATGTAAGAATCCAAAGTGGACACCATT
AACAGGTCTTTGAAATATGCATGTACTTTATATTTTCTATATTTGTAACTTTGCATGT
TCTTGTTTTGTTATATAAAAAAATTGTAAATGTTTAATATCTGACTGAAATTAAACG
AGCGAAGATGAGCACCA (SEQ ID NO: 292)

[0757] Сообщается, что участки мутирования, которые относятся к PRNP (NP_000302.1), которые связаны с CJD и FFI, являются следующими. Данные мутации могут быть удалены или внесены с помощью описанных здесь праймерных редакторов.

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННОГО С ПРИОННЫМ ЗАБОЛЕВАНИЕМ CJD (СМ. ТАБЛИЦУ 1 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 PRNP NP_000302.1)
D178N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHNCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 293)
T188K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHKVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 294)
E196K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGKNFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 295)
E196A	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG

	GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGANFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 296)
E200K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTKTDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 297)
E200G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTGTDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 298)
V203I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDIKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPVI LLISFLIFLIVG (SEQ ID NO: 299)
R208H	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMEHVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 300)
V210I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVIEQMCITQYERESQAYYQRGSSMVLFSSPPVI LLISFLIFLIVG (SEQ ID NO: 301)
E211Q	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVQQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 302)
M232R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSRVLFSSPPVI LLISFLIFLIVG (SEQ ID NO: 303)

[0758] Сообщается, что участки мутирования, которые относятся к PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с GSS, являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННОГО С ПРИОННЫМ ЗАБОЛЕВАНИЕМ GSS (СМ. ТАБЛИЦУ 2 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 PRNP NP_000302.1)
P102L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKLSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 304)
P105L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKLKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 305)
A117V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAVAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 306)
G131V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLVSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 307)
V176G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFGHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 308)
H187R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQRTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 309)
	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 291)
F198S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF

	GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENSTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 311)
D202N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETNVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 312)
Q212P	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEPMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 313)
Q217R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITRYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 314)
M232T	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSTVLFSSPPVI LLISFLIFLIVG (SEQ ID NO: 315)

[0759] Участки мутирования, которые относятся к PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с возможной защитной природой против прионного заболевания, являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННАЯ С ЗАЩИТНОЙ ПРИРОДОЙ ПРОТИВ ПРИОННОГО ЗАБОЛЕВАНИЯ (СМ. ТАБЛИЦУ 4 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 OF PRNP NP_000302.1)
G127S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGSYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 316)
G127V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGVYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 317)
M129V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR

	YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYVLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 318)
D167G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMGEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 319)
D167N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMNEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 320)
N171S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSSQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 321)
E219K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYKRESQAYYQRGSSMVLFSSPPV ILLISFLIFLIVG (SEQ ID NO: 322)
P238S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSSPV ILLISFLIFLIVG (SEQ ID NO: 323)

[0760] Таким образом, в различных вариантах осуществления праймерное редактирование может применяться для удаления мутации в PRNP, которая связана с прионным заболеванием, или внесения мутации в PRNP, которая считается являющейся защитной против прионного заболевания. Например, праймерное редактирование может применяться для удаления или восстановления мутации D178N, V180I, T188K, E196K, E196A, E200K, E200G, V203I, R208H, V210I, E211Q, I215V, или M232R в белке PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). Например, праймерное редактирование может применяться для удаления или восстановления мутации P102L, P105L, A117V, G131V, V176G, H187R, F198S, D202N, Q212P, Q217R,

или M232T в белке PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). Путем удаления или корректирования наличия подобных мутаций в PRNP с помощью праймерного редактирования риск развития прионного заболевания может быть снижен или устранен.

[0761] В других вариантах осуществления праймерное редактирование может применяться для внесения защитной мутации в PRNP, которая связана с защитным эффектом против одного или нескольких прионных заболеваний. Например, праймерное редактирование может применяться для внесения защитной мутации G127S, G127V, M129V, D167G, D167N, N171S, E219K, или P238S в PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). В других вариантах осуществления защитная мутация может представлять собой любую альтернативную аминокислоту, внесенную в G127, G127, M129, D167, D167, N171, E219, или P238 в PRNP (относительно PRNP of NP_000302.1) (SEQ ID NO: 291).

[0762] В определенных вариантах осуществления праймерное редактирование может применяться для внесения защитной мутации G127V в PRNP, как показано на Фиг.27 и описано в Примере 5.

[0763] В другом варианте осуществления праймерное редактирование может применяться для внесения защитной мутации E219K в PRNP.

[0764] Белок *PRNP* и участок защитной мутации являются консервативными у млекопитающих, поэтому помимо лечения заболевания у человека он также может использоваться для получения коров и овец с иммунитетом против прионного заболевания или даже осуществления лечения диких популяций животных, которые страдают прионным заболеванием. Праймерное редактирование уже использовалось для достижения ~25% внесения встречающегося в природе защитного аллеля в клетки человека, при этом предшествующие эксперименты на мышах показывают, что данный уровень внесения является достаточным для возникновения иммунитета против большинства прионных заболеваний. Данный способ является первым и потенциально единственным доступным в настоящее время способом внесения данного аллеля с настолько высокой эффективностью в большинство типов клеток. Другая возможная стратегия лечения заключается в применении праймерного редактирования для снижения или устранения экспрессии *PRNP* путем внесения в ген раннего стоп-кодона.

[0765] С помощью описанных здесь способов конструирования PEgRNA могут быть сконструированы походящие PEgRNA для внесения требуемых защитных мутаций или

для удаления ассоциированных с прионным заболеванием мутаций из PRNP. Например, приведенный ниже перечень PEgRNA может применяться для внесения защитного аллеля G127V и защитного аллеля E219K в PRNP человека, а также защитного аллеля G127V в PRNP различных животных.

PEGRNA ЧЕЛОВЕКА ДЛЯ ВНЕСЕНИЯ G127V В PRNP ЧЕЛОВЕКА:	**GCAGTGGTGGGGGGCCTTGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACGCCAA GGCCCCCCAC(SEQ ID NO: 324)
PEGRNA ЧЕЛОВЕКА ДЛЯ ВНЕСЕНИЯ E219K В PRNP ЧЕЛОВЕКА	**TGTGTATCACCCAGTACGAG**GTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA AAAAGTGGCACCGAGTCGGTGC AGATTCTCTCTT GTACTGGGTGA(SEQ ID NO: 325)
PEGRNA КОРОВЫ (BOS TAURUS) ДЛЯ ВНЕСЕНИЯ G127V В PRNP КОРОВЫ	**GCAGTGGTAGGGGGCCTTGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACACCAA GGCCCCCTAC(SEQ ID NO: 326)
PEGRNA ХОМЯКА (MESOCRICETUS AURATUS) ДЛЯ ВНЕСЕНИЯ G127V В PRNP ХОМЯКА	**GCCGTGGTGGGGGGCCTTGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACACCAA GGCCCCCCAC(SEQ ID NO: 327)
PEGRNA МЫШИ (MUS MUSCULUS) ДЛЯ ВНЕСЕНИЯ G127V В PRNP МЫШИ	**GCAGTAGTGGGGGGCCTTGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACACCAA GGCCCCCCAC(SEQ ID NO: 328)
PEGRNA ОЛЕНЯ (ODOCOILEUS VIRGINIANUS) ДЛЯ ВНЕСЕНИЯ G127V В PNRP ОЛЕНЯ	**GCAGTGGTAGGGGGCCTTGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACACCAA GGCCCCCTAC(SEQ ID NO: 329)
PEGRNA ХОРЬКА (MUSTELA PUTORIUS FURO) ДЛЯ ВНЕСЕНИЯ G127V В PNRP ХОРЬКА	**GCGGTTGTGGGGGGCCTGGG**GTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCATGTAGACGCCCA GGCCCCCCAC(SEQ ID NO: 330)
ОБОЗНАЧЕНИЯ:	СПЕЙСЕР ОБОЗНАЧЕН ЖИРНЫМ ШРИФТОМ. КАРКАС SGPHK УКАЗАН В ВИДЕ ОБЫЧНОГО ТЕКСТА. RT МАТРИЦА ПОДЧЕРКНУТА ОДИН РАЗ. PBS ПОДЧЕРКНУТ ДВА РАЗА.

F. Применение праймерного редактирования для мечения РНК

[0766] Праймерное редактирование также может применяться для манипулирования, изменения, и иным образом модифицирования последовательностей ДНК, кодирующих функции РНК, за счет мечения РНК, при этом оно обеспечивает

возможность непрямой модификации структуры и функции РНК. Например, PE может применяться для вставки мотивов, которые являются функциональными на уровне РНК (далее — мотивов РНК) для мечения или иным образом манипулирования некодирующими РНК или иРНК. Данные мотивы могут служить для повышения генной экспрессии, снижения генной экспрессии, изменения сплайсинга, изменения посттрансляционной модификации, влияния на субклеточное расположение РНК, обеспечения выделения или определения внутри- или внешнеклеточного расположения РНК (с помощью, например, флуоресцентных РНК аптамеров, например Spinach, Spinach2, Baby Spinach, или Broccoli), рекрутирования эндогенного и экзогенного белка или связывающихся с РНК элементов, внесения sgРНК или индуцирования процессирования РНК за счет саморасщепления или РНКаз (см. дополнительные подробности на Фиг.28B и в Примере 6).

[0767] Следующие метки или мотивы РНК могут быть вставлены в представляющий интерес ген с помощью праймерного редактирования с использованием подходящей PEgRNA (сконструированной согласно приведенному здесь руководству) для воздействия на различные свойства РНК, включая транспорт РНК, уровень экспрессии, сплайсинг и детектирование.

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
ПОЛИОМАВИРУС ОБЕЗЬЯНЫ 40 (SV40) ТИП 1	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTA (SEQ ID NO: 331)	ТЕРМИНАЦИЯ ТРАНСКРИПЦИИ МЕЧЕННОГО ГЕНА; ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМО	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*A**CCTGAACCGTATATC*** **TAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
		ГО БЕЛКА	**TAACCATTATAAGC TGCAATAAACAAG TT**CTATG*GCCCTGA CTGGAA* (SEQ ID NO: 332)
ПОЛИОМАВИРУС ОБЕЗЬЯНЫ 40 (SV40) ТИП 2	CCATGGCCCAACTTGTTTA TTGCAGCTTATAATGGTTA CAAATAAAGCAATAGCAT CACAAATTTCACAAATAA AGCATTTTTTTCACTGCAT TCTAGTTGTGGTTTGTCCA AACTCATCAATGTATCTTA TCATGTCTGGATCTC (SEQ ID NO: 333)	ТЕРМИНАЦИЯ ТРАНСКРИПЦИИ МЕЧЕННОГО ГЕНА; ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GAGATCCAGACAT GATAAGATACATT GATGAGTTTGGAC AAACCACAACTAG AATGCAGTGAAAA AAATGCTTTATTT GTGAAATTTGTGA TGCTATTGCTTTAT TTGTAACCATTATA AGCTGCAATAAAC AAGTTGGGCCATG G**CTATG*GCCCTGAC TGGAA* (SEQ ID NO: 334)
ПОЛИОМАВИРУС ОБЕЗЬЯНЫ 40 (SV40) ТИП 3	TGATCATAATCAAGCCATA TCACATCTGTAGAGGTTTA CTTGCTTTAAAAAACCTC CACACCTCCCCCTGAACC TGAAACATAAAATGAATG CAATTGTTGTTGTTAACTT GTTTATTGCAGCTTATAAT GGTTACAAATAAAGCAAT AGCATCACAAATTTCACA AATAAAGCATTTTTTTCAC TGCATTCTAGTTGTGGTTT GTCCAAACTCATCAATGTA TCTTATCATGTCTGGATCT GC (SEQ ID NO: 335)	ТЕРМИНАЦИЯ ТРАНСКРИПЦИИ МЕЧЕННОГО ГЕНА; ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GCAGATCCAGACA TGATAAGATACATT GATGAGTTTGGAC AAACCACAACTAG AATGCAGTGAAAA AAATGCTTTATTT**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
		КОДИРУЕМОГО БЕЛКА	**GTGAAATTTGTGA TGCTATTGCTTTAT TTGTAACCATTATA AGCTGCAATAAAC AAGTTAACAACAA CAATTGCATTCAT TTTATGTTTCAGG TTCAGGGGGAGG TGTGGAGGTTTTT TAAAGCAAGTAAA CCTCTACAGATGT GATATGGCTTGAT TATGATCA**CTATG*G CCCTGACTGGAA* (SEQ ID NO: 336)
ГОРМОН РОСТА ЧЕЛОВЕКА (HGH)	ACGGGTGGCATCCCTGTG ACCCCTCCCCAGTGCCTC TCCTGGCCCTGGAAGTTG CCACTCCAGTGCCCACCA GCCTTGTCCTAATAAAATT AAGTTGCATCATTTTGTCT GACTAGGTGTCCTTCTATA ATATTATGGGGTGGAGGG GGGTGGTATGGAGCAAGG GGCAAGTTGGGAAGACA ACCTGTAGGGCCTGCGGG GTCTATTGGGAACCAAGC TGGAGTGCAGTGGCACAA TCTTGGCTCACTGCAATCT CCGCCTCCTGGGTTCAAG CGATTCTCCTGCCTCAGCC TCCCGAGTTGTTGGGATT CCAGGCATGCATGACCAG GCTCAGCTAATTTTTGTTT TTTTGGTAGAGACGGGGT TTCACCATATTGGCCAGGC TGGTCTCCAACTCCTAATC TCAGGTGATCTACCCACCT TGGCCTCCCAAATTGCTG GGATTACAGGCGTGAACC ACTGCTCCCTTCCCTGTCC	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAG GGCCAT<u>GTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC</u>*A CCTGAACCGTATATC* **AAGGACAGGGAA GGGAGCAGTGGT TCACGCCTGTAAT CCCAGCAATTTGG GAGGCCAAGGTG GGTAGATCACCTG AGATTAGGAGTTG GAGACCAGCCTG GCCAATATGGTGA AACCCCGTCTCTA CCAAAAAAACAAA AATTAGCTGAGCC TGGTCATGCATGC CTGGAATCCCAAC AACTCGGGAGGCT GAGGCAGGAGAAT CGCTTGAACCCAG GAGGCGGAGATTG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	TT (SEQ ID NO: 337)		**CAGTGAGCCAAGA TTGTGCCACTGCA CTCCAGCTTGGTT CCCAATAGACCCC GCAGGCCCTACAG GTTGTCTTCCCAA CTTGCCCCTTGCT CCATACCACCCCC CTCCACCCCATAA TATTATAGAAGGA CACCTAGTCAGAC AAAATGATGCAAC TTAATTTTATTAGG ACAAGGCTGGTG GGCACTGGAGTG GCAACTTCCAGGG CCAGGAGAGGCA CTGGGGAGGGGT CACAGGGATGCCA CCCGT**CTATG*GCCC TGACTGGAA* (SEQ ID NO: 338)
ГОРМОН РОСТА КОРОВЫ (BGH)	CGACTGTGCCTTCTAGTTG CCAGCCATCTGTTGTTTGC CCCTCCCCCGTGCCTTCCT TGACCCTGGAAGGTGCCA CTCCCACTGTCCTTTCCTA ATAAAATGAGGAAATTGC ATCGCATTGTCTGAGTAGG TGTCATTCTATTCTGGGGG GTGGGGTGGGGCAGGAC AGCAAGGGGGAGGATTG GGAAGACAATAGCAGGCA TGCTGGGGATGCGGTGGG CTCTATGG (SEQ ID NO: 339)	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAG GGCCAT<u>GTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC</u>*A CCTGAACCGTATATC* **CCATAGAGCCCAC CGCATCCCCAGCA TGCCTGCTATTGT CTTCCCAATCCTC CCCCTTGCTGTCC TGCCCCACCCCAC CCCCCAGAATAGA ATGACACCTACTC AGACAATGCGATG CAATTTCCTCATTT TATTAGGAAAGGA**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
			CAGTGGGAGTGG CACCTTCCAGGGT CAAGGAAGGCAC GGGGGAGGGGCA AACAACAGATGGC TGGCAACTAGAAG GCACAGTCGCTAT *GGCCCTGACTGGAA* (SEQ ID NO: 340)
БЕТА-ГЛОБИН КРОЛИКА (RBGLOB)	TTCACTCCTCAGGTGCAG GCTGCCTATCAGAAGGTG GTGGCTGGTGTGGCCAAT GCCCTGGCTCACAAATAC CACTGAGATCTTTTTCCCT CTGCCAAAAATTATGGGG ACATCATGAAGCCCCTTG AGCATCTGACTTCTGGCTA ATAAAGGAAATTTATTTTC ATTGCAATAGTGTGTTGGA ATTTTTTGTGTCTCTCACT CGGAAGGACATATGGGAG GGCAAATCATTTAAAACAT CAGAATGAGTATTTGGTTT AGAGTTTGGCAACATATG CCCATATGCTGGCTGCCAT GAACAAAGGTTGGCTATA AAGAGGTCATCAGTATATG AAACAGCCCCCTGCTGTC CATTCCTTATTCCATAGAA AAGCCTTGACTTGAGGTT AGATTTTTTTTATATTTTGT TTTGTGTTATTTTTTTCTTT AACATCCCTAAAATTTTCC TTACATGTTTTACTAGCCA GATTTTTCCTCCTCTCCTG ACTACTCCCAGTCATAGCT GTCCCTCTTCTCTTATGGA GATC (SEQ ID NO: 341)	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A* *CCTGAACCGTATATC* **GATCTCCATAAGA GAAGAGGGACAG CTATGACTGGGAG TAGTCAGGAGAGG AGGAAAAATCTGG CTAGTAAAACATG TAAGGAAAATTTT AGGGATGTTAAAG AAAAAAATAACAC AAAACAAAATATA AAAAAAATCTAAC CTCAAGTCAAGGC TTTTCTATGGAATA AGGAATGGACAGC AGGGGGCTGTTTC ATATACTGATGAC CTCTTTATAGCCA ACCTTTGTTCATG GCAGCCAGCATAT GGGCATATGTTGC CAAACTCTAAACC AAATACTCATTCT GATGTTTTAAATG ATTTGCCCTCCCA**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
			TATGTCCTTCCGA GTGAGAGACACAA AAAATTCCAACAC ACTATTGCAATGA AAATAAATTTCCTT TATTAGCCAGAAG TCAGATGCTCAAG GGGCTTCATGATG TCCCCATAATTTTT GGCAGAGGGAAA AAGATCTCAGTGG TATTTGTGAGCCA GGGCATTGGCCAC ACCAGCCACCACC TTCTGATAGGCAG CCTGCACCTGAGG AGTGAACTATG*GCC CTGACTGGAA* (SEQ ID NO: 342)
ТИМИДИНКИНАЗА (TK)	GGGGGAGGCTAACTGAAA CACGGAAGGAGACAATAC CGGAAGGAACCCGCGCTA TGACGGCAATAAAAAGAC AGAATAAAACGCACGGGT GTTGGGTCGTTTGTTCATA AACGCGGGGTTCGGTCCC AGGGCTGGCACTCTGTCG ATACCCCACCGAGACCCC ATTGGGGCCAATACGCCC GCGTTTCTTCCTTTTCCCC ACCCCACCCCCCAAGTTC GGGTGAAGGCCCAGGGCT CGCAGCCAACGTCGGGGC GGCAGGCCCTGCCATAG (SEQ ID NO: 343)	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ; ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAG GGCCAT<u>GTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC</u>*A CCTGAACCGTATATC* **CTATGGCAGGGCC TGCCGCCCCGACG TTGGCTGCGAGCC CTGGGCCTTCACC CGAACTTGGGGG GTGGGGTGGGGA AAAGGAAGAAAC GCGGGCGTATTGG CCCCAATGGGGTC TCGGTGGGGTATC GACAGAGTGCCA GCCCTGGGACCG AACCCCGCGTTTA TGAACAAACGACC**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
			CAACACCCGTGCG TTTTATTCTGTCTT TTTATTGCCGTCAT AGCGCGGGTTCCT TCCGGTATTGTCT CCTTCCGTGTTTC AGTTAGCCTCCCC CCTATG*GCCCTGAC TGGAA* (SEQ ID NO: 344)
MALAT1 ENE-MASCPHK	TAGGGTCATGAAGGTTTTT CTTTTCCTGAGAAAACAA CACGTATTGTTTTCTCAGG TTTTGCTTTTTGGCCTTTT TCTAGCTTAAAAAAAAAA AAAGCAAAAGATGCTGGT GGTTGGCACTCCTGGTTT CCAGGACGGGGTTCAAAT CCCTGCGGCGTCTTTGCTT TGACT (SEQ ID NO: 345)	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И ТРАНСКРИПЦИИ И СТАБИЛИЗАЦИИ	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **AGTCAAAGCAAAG ACGCCGCAGGGAT TTGAACCCCGTCC TGGAAACCAGGA GTGCCAACCACCA GCATCTTTTGCTT TTTTTTTTTTTAAG CTAGAAAAAGGCC AAAAAGCAAAACC TGAGAAAACAATA CGTGTTGTTTTCT CAGGAAAAGAAAA ACCTTCATGACCC TA**CTATG*GCCCTGA CTGGAA* (SEQ ID NO: 346)
KSHV PAN ENE	TGTTTTGGCTGGGTTTTTC CTTGTTCGCACCGGACAC CTCCAGTGACCAGACGGC AAGGTTTTTATCCCAGTGT ATATTGGAAAAACATGTTA TACTTTTGACAATTTAACG TGCCTAGAGCTCAAATTA	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И ТРАНСКРИПЦИИ И	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A*

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	AACTAATACCATAACGTAA TGCAACTTACAACATAAAT AAAGGTCAATGTTTAATCC ATAAAAAAAAAAAAAAA AAAA (SEQ ID NO: 347)	СТАБИЛИЗАЦИИ	*CCTGAACCGTATATC* **TTTTTTTTTTTTT TTTTTATGGATTAA ACATTGACCTTTA TTTATGTTGTAAG TTGCATTACGTTAT GGTATTAGTTTAAT TTGAGCTCTAGGC ACGTTAAATTGTC AAAAGTATAACAT GTTTTTCCAATATA CACTGGGATAAAA ACCTTGCCGTCTG GTCACTGGAGGTG TCCGGTGCGAACA AGGAAAAACCCAG CCAAAACA**CTATG *GCCCTGACTGGAA* (SEQ ID NO: 348)
ТРИ ПОСЛЕДОВАТЕЛЬНЫХ KSHV PAN ENE С КОРОТКИМИ НЕКОНСЕРВАТИВНЫМИ РНК ЛИНКЕРАМИ	TGTTTTGGCTGGGTTTTTC CTTGTTCGCACCGGACAC CTCCAGTGACCAGACGGC AAGGTTTTTATCCCAGTGT ATATTGGAAAAACATGTTA TACTTTTGACAATTTAACG TGCCTAGAGCTCAAATTA AACTAATACCATAACGTAA TGCAACTTACAACATAAAT AAAGGTCAATGTTTAATCC ATAAAAAAAAAAAAAAA AAAAACACACTGTTTTGG CTGGGTTTTTCCTTGTTCG CACCGGACACCTCCAGTG ACCAGACGGCAAGGTTTT TATCCCAGTGTATATTGGA AAAACATGTTATACTTTTG ACAATTTAACGTGCCTAG AGCTCAAATTAAACTAATA CCATAACGTAATGCAACTT ACAACATAAATAAAGGTC AATGTTTAATCCATAAAAA	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И ТРАНСКРИПЦИИ И СТАБИЛИЗАЦИИ, ПРЕДПОЛОЖИТЕЛЬНО ПРЕВЫШАЕТ ОДИНОЧНЫЙ PAN ENE	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A* *CCTGAACCGTATATC* **TTTTTTTTTTTTT TTTTTATGGATTAA ACATTGACCTTTA TTTATGTTGTAAG TTGCATTACGTTAT GGTATTAGTTTAAT TTGAGCTCTAGGC ACGTTAAATTGTC AAAAGTATAACAT GTTTTTCCAATATA CACTGGGATAAAA ACCTTGCCGTCTG GTCACTGGAGGTG TCCGGTGCGAACA**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	AAAAAAAAAAAAAATCTC TCTGTTTTGGCTGGGTTTT TCCTTGTTCGCACCGGAC ACCTCCAGTGACCAGACG GCAAGGTTTTTATCCCAGT GTATATTGGAAAAACATGT TATACTTTTGACAATTTAA CGTGCCTAGAGCTCAAAT TAAACTAATACCATAACGT AATGCAACTTACAACATA AATAAAGGTCAATGTTTAA TCCATAAAAAAAAAAAAA AAAAAA (SEQ ID NO: 349)		**AGGAAAAACCCAG CCAAAACAGAGAG ATTTTTTTTTTTTT TTTTTTATGGATTA AACATTGACCTTT ATTTATGTTGTAAG TTGCATTACGTTAT GGTATTAGTTTAAT TTGAGCTCTAGGC ACGTTAAATTGTC AAAAGTATAACAT GTTTTTCCAATATA CACTGGGATAAAA ACCTTGCCGTCTG GTCACTGGAGGTG TCCGGTGCGAACA AGGAAAAACCCAG CCAAAACAGTGTG TTTTTTTTTTTTTT TTTTTTATGGATTA AACATTGACCTTT ATTTATGTTGTAAG TTGCATTACGTTAT GGTATTAGTTTAAT TTGAGCTCTAGGC ACGTTAAATTGTC AAAAGTATAACAT GTTTTTCCAATATA CACTGGGATAAAA ACCTTGCCGTCTG GTCACTGGAGGTG TCCGGTGCGAACA AGGAAAAACCCAG CCAAAACA**CTATG *GCCCTGACTGGAA* (SEQ ID NO: 350)
SM-БОКС/БОКС МЯРНК U1	CAGCAAGTTCAGAGAAAT CTGAACTTGCTGGATTTTT GGAGCAGGGAGATGGAAT AGGAGCTTGCTCCGTCCA CTCCACGCATCGACCTGG	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ И ТЕРМИНАЦИ	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	TATTGCAGTACCTCCAGG AACGGTGCACCCACTTTC TGGAGTTTCAAAAGTAGA CTGTACGCTAAGGGTCATA TCTTTTTTTGTTTGGTTTG TGTCTTGGTTGGCGTCTTA AA (SEQ ID NO: 351)	И ТРАНСКРИПЦИИ	TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **TTTAAGACGCCAA CCAAGACACAAAC CAAACAAAAAAAG ATATGACCCTTAG CGTACAGTCTACT TTTGAAACTCCAG AAAGTGGGTGCAC CGTTCCTGGAGGT ACTGCAATACCAG GTCGATGCGTGGA GTGGACGGAGCA AGCTCCTATTCCA TCTCCCTGCTCCA AAAATCCAGCAAG TTCAGATTTCTCT GAACTTGCTG**CTA TG*GCCCTGACTGGA A* (SEQ ID NO: 352)
3′ БОКС МЯРНК U1	GTTTCAAAAGTAGACTGT ACGCTAAGGGTCATATCTT TTTTTGTTTGGTTTGTGTC TTGGTTGGCGTCTTAAA (SEQ ID NO: 353)	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ И ТЕРМИНАЦИИ И ТРАНСКРИПЦИИ	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **TTTAAGACGCCAA CCAAGACACAAAC CAAACAAAAAAAG ATATGACCCTTAG CGTACAGTCTACT TTTGAAAC**CTATG*G CCCTGACTGGAA* (SEQ ID NO: 354)
ТРНК-ЛИЗИН	GCCCGGCTAGCTCAGTCG GTAGAGCATGAGACTCTT AATCTCAGGGTCGTGGGT	ПРЕДПОЛОЖИТЕЛЬНО ОБЕСПЕЧИВ	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	TCGAGCCCCACGTTGGGCG (SEQ ID NO: 355)	АЕТ ТРАНСПОРТ РНК В МИТОХОНДРИИ	GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **CGCCCAACGTGG GGCTCGAACCCAC GACCCTGAGATTA AGAGTCTCATGCT CTACCGACTGAGC TAGCCGGGC**CTAT G*GCCCTGACTGGAA* (SEQ ID NO: 356)
АПТАМЕР BROCCOLI	GAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTC (SEQ ID NO: 357)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GAGCCCACACTCT ACTCGACAGATAC GAATATCTGGACC CGACCGTCTC**CTA TG*GCCCTGACTGGA A* (SEQ ID NO: 358)
АПТАМЕР SPINACH	GACGCAACTGAATGAAATGGTGAAGGACGGGTCCAGGTGTGGCTGCTTCGGCAGTGCAGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTAGTCGCGTC (SEQ ID NO: 359)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GACGCGACTAGTT ACGGAGCTCACAC TCTACTCAACAAG CTGCACTGCCGAA GCAGCCACACCTG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
			GACCCGTCCTTCA CCATTTCATTCAG TTGCGTCCTATG*GC CCTGACTGGAA* (SEQ ID NO: 360)
АПТАМЕР SPINACH2	GATGTAACTGAATGAAAT GGTGAAGGACGGGTCCA GTAGGCTGCTTCGGCAGC CTACTTGTTGAGTAGAGT GTGAGCTCCGTAACTAGTT ACATC (SEQ ID NO: 361)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GATGTAACTAGTT ACGGAGCTCACAC TCTACTCAACAAG TAGGCTGCCGAAG CAGCCTACTGGAC CCGTCCTTCACCA TTTCATTCAGTTA CATC**CTATG*GCCCT GACTGGAA* (SEQ ID NO: 362)
АПТАМЕР MANGO	GGCACGTACGAAGGGACG GTGCGGAGAGGAGAGTAC GTGC (SEQ ID NO: 363)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GCACGTACTCTCC TCTCCGCACCGTC CCTTCGTACGTGC C**CTATG*GCCCTGAC TGGAA* (SEQ ID NO: 364)
РИБОЗИМ HDV	GGCCGGCATGGTCCCAGC CTCCTCGCTGGCGCCGGC TGGGCAACATGCTTCGGC ATGGCGAATGGGAC (SEQ	ПРОЦЕССИРОВАНИЕ 3′-КОНЦЕВОЙ РНК	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНЦИЯ/ЭФФЕКТ	ПРИМЕР PEGRNA ДЛЯ ВСТАВКИ МОТИВА РНК В ПРИМЕР В ВИДЕ ГЕНА HEXA* С ПОМОЩЬЮ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ
	ID NO: 365)		AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **GTCCCATTCGCCA TGCCGAAGCATGT TGCCCAGCCGGC GCCAGCGAGGAG GCTGGGACCATGC CGGCC**CTATG*GCCC TGACTGGAA* (SEQ ID NO: 366)
N6-МЕТИЛАДЕНОЗИНОВЫЙ МАРКЕР (M6A)	GGACTCTAGGACTGGACT TTGGACT (SEQ ID NO: 367)	МИШЕНЬ ДЛЯ МЕТИЛИРОВАНИЯ (ПОДЧЕРКНУТЫЕ А ЯВЛЯЮТСЯ МЕТИЛИРОВАННЫМИ). МЕТИЛИРОВАНИЕ M6A МОЖЕТ ПРИВОДИТЬ К ПОВЫШЕНИЮ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИИ, НО ПОКА ПОЛНОСТЬЮ НЕ ИЗУЧЕНО	ATCCTTCCAGTCAG GGCCATGTTTGAGA GCTAGAAATAGCAA GTTTAAATAAGGCT AGTCCGTTATCAAC TTGAAAAAGTGGG ACCGAGTCGGTCC*A CCTGAACCGTATATC* **AGTCCAAAGTCCA GTCCTAGAGTCC**C TATG*GCCCTGACTG GAA* (SEQ ID NO: 368)

* каждая PEgRNA показана в направлении от 5′ к 3′ и имеет следующие структурные элементы, указанные на Фиг.3F, обозначенные с помощью шрифта следующим образом: 5′ – спейсерная последовательность (обычный шрифт) – ядро нРНК (подчеркнутая последовательность) – плечо гомологии (курсив) – RT матрица (жирный шрифт) – участок связывания праймера (курсив) – 3′.

[0768] PEgRNA в приведенной выше таблице сконструированы для специфической в плане участка вставки примеров приведенных выше мотивов в ген *HEXA* (являющийся дефектным при болезни Тея-Сакса) (например GenBank No. KR710351.1 (SEQ ID NO: 369), однако они приведены лишь для целей иллюстрирования. Применение праймерного редактирования при мечении РНК не ограничивается геном *HEXA* и действительно может являться любым. иРНК *HEXA* имеет следующую нуклеотидную последовательность:

GTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAA
AAAGTTGGCATGACAAGTTCCAGGCTTTGGTTTTCGCTGCTGCTGGCGGCAGCGT
TCGCAGGACGGGCGACGGCCCTCTGGCCCTGGCCTCAGAACTTCCAAACCTCCG
ACCAGCGCTACGTCCTTTACCCGAACAACTTTCAATTCCAGTACGATGTCAGCTCG
GCCGCGCAGCCCGGCTGCTCAGTCCTCGACGAGGCCTTCCAGCGCTATCGTGACC
TGCTTTTCGGTTCCGGGTCTTGGCCCCGTCCTTACCTCACAGGGAAACGGCATACA
CTGGAGAAGAATGTGTTGGTTGTCTCTGTAGTCACACCTGGATGTAACCAGCTTCC
TACTTTGGAGTCAGTGGAGAATTATACCCTGACCATAAATGATGACCAGTGTTTAC
TCCTCTCTGAGACTGTCTGGGGAGCTCTCCGAGGTCTGGAGACTTTTAGCCAGCT
TGTTTGGAAATCTGCTGAGGGCACATTCTTTATCAACAAGACTGAGATTGAGGACT
TTCCCCGCTTTCCTCACCGGGGCTTGCTGTTGGATACATCTCGCCATTACCTGCCAC
TCTCTAGCATCCTGGACACTCTGGATGTCATGGCGTACAATAAATTGAACGTGTTC
CACTGGCATCTGGTAGATGATCCTTCCTTCCCATATGAGAGCTTCACTTTTCCAGAG
CTCATGAGAAAGGGGTCCTACAACCCTGTCACCCACATCTACACAGCACAGGATG
TGAAGGAGGTCATTGAATACGCACGGCTCCGGGGTATCCGTGTGCTTGCAGAGTT
TGACACTCCTGGCCACACTTTGTCCTGGGGACCAGGTATCCCTGGATTACTGACTC
CTTGCTACCCTGGGTCTGAGCCCTCTGGCACCTTTGGACCAGTGAATCCCAGTCTC
AATAATACCTATGAGTTCATGAGCACATTCTTCTTAGAAGTCAGCTCTGTCTTCCCA
GATTTTTATCTTCATCTTGGAGGAGATGAGGTTGATTTCACCTGCTGGAAGTCCAA
CCCAGAGATCCAGGACTTTATGAGGAAGAAAGGCTTCGGTGAGGACTTCAAGCA
GCTGGAGTCCTTCTACATCCAGACGCTGCTGGACATCGTCTCTTCTTATGGCAAGG
GCTATGTGGTGTGGCAGGAGGTGTTTGATAATAAAGTAAAGATTCAGCCAGACAC
AATCATACAGGTGTGGCGAGAGGATATTCCAGTGAACTATATGAAGGAGCTGGAA
CTGGTCACCAAGGCCGGCTTCCGGGCCCTTCTCTCTGCCCCCTGGTACCTGAACC
GTATATCCTATGGCCCTGACTGGAAGGATTTCTACGTAGTGGAACCCCTGGCATTTG
AAGGTACCCCTGAGCAGAAGGCTCTGGTGATTGGTGGAGAGGCTTGTATGTGGGG

AGAATATGTGGACAACACAAACCTGGTCCCCAGGCTCTGGCCCAGAGCAGGGGC
TGTTGCCGAAAGGCTGTGGAGCAACAAGTTGACATCTGACCTGACATTTGCCTAT
GAACGTTTGTCACACTTCCGCTGTGAGTTGCTGAGGCGAGGTGTCCAGGCCCAAC
CCCTCAATGTAGGCTTCTGTGAGCAGGAG
TTTGAACAGACCTGCCCAACTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTG
CTTATCAATTTGTTGCAACGAAC (SEQ ID NO: 369).

[0769] Соответствующий белок HEXA имеет следующую аминокислотную последовательность:

MTSSRLWFSLLLAAAFAGRATALWPWPQNFQTSDQRYVLYPNNFQFQYDVSSAAQP
GCSVLDEAFQRYRDLLFGSGSWPRPYLTGKRHTLEKNVLVVSVVTPGCNQLPTLESV
ENYTLTINDDQCLLLSETVWGALRGLETFSQLVWKSAEGTFFINKTEIEDFPRFPHRGL
LLDTSRHYLPLSSILDTLDVMAYNKLNVFHWHLVDDPSFPYESFTFPELMRKGSYNP
VTHIYTAQDVKEVIEYARLRGIRVLAEFDTPGHTLSWGPGIPGLLTPCYPGSEPSGTFG
PVNPSLNNTYEFMSTFFLEVSSVFPDFYLHLGGDEVDFTCWKSNPEIQDFMRKKGFG
EDFKQLESFYIQTLLDIVSSYGKGYVVWQEVFDNKVKIQPDTIIQVWREDIPVNYMK
ELELVTKAGFRALLSAPWYLNRISYGPDWKDFYVVEPLAFEGTPEQKALVIGGEACM
WGEYVDNTNLVPRLWPRAGAVAERLWSNKLTSDLTFAYERLSHFRCELLRRGVQAQP
LNVGFCEQEFEQT (SEQ ID NO: 370).

[0770] Следует отметить, что получаемые мотивы РНК будут включены в транслируемый участок гена *HEXA*, нарушая функцию кодирующего белок гена. Вставляемые мотивы полиаденилирования будут приводить к преждевременной терминации транскрипции. Этот участок является лишь иллюстрацией потенциальных PEgRNA, которые могут приводить к вставке перечисленных мотивов РНК из приведенной выше таблицы в участок генома, который транскрибируется, и, таким образом, обеспечивать получение РНК продукта.

[0771] PEgRNA для применения с PE для мечения РНК могут быть экспрессированы из промотора U6 (в этом случае один гуанозин будет добавлен на 5′-конце PEgRNA для направляющих, которые включают протоспейсеры, которые не начинаются с G, и 6-7 тиминов будут добавлены на 3'-конец) или промотора pol II, такого как pCMV (в этом случая возможно будет необходимо удалить фактически транскрибируемую последовательность данного промотора на 5′-конце РНК с помощью саморасщепляемого элемента или мотива Csy4, и потребуется добавить мотив терминации на 3'-конце РНК, который не приводит к экспорту РНК из ядра, например

перечисленный выше мотив 3' бокса). Следует обратить внимание, что этот мотив не будет вставлен в геном в результате PE, поскольку он будет располагаться на 3' участка отжина). Ядро каркаса PEgRNA обозначено с помощью подчеркивания, участки гомологии и отжига обозначены курсивом, а вставляемая последовательность обозначена с помощью жирного шрифта. Следует обратить внимание, что вставленная последовательность представляет собой обратный комплемент приведенных выше примеров, как описано ниже, и, таким образом, потребуется нацелить данные PEgRNA на кодирующую цепь.

[0772] Кроме того, следует обратить внимание, что в некоторых вариантах осуществления необходимо, чтобы саморасщепляющиеся рибозимы, отличные от HDV, являлись подходящими для участка-мишени; то есть в то время как HDV расщепляет кодируемый транскрипт непосредственно 5′ относительно себя, участки разрыва для всех других саморасщепляемых рибозимов располагаются внутри самого рибозима. Таким образом, первые и последние приблизительно 5-10 нуклеотидов (и в некоторых случаях более чем 10) будут по существу представлять собой часть кодируемой последовательности. В качестве примера, для расщепления последовательности 5′NNNNNTCATCCTGATAAACTGCAAA3′ (SEQ ID NO: 371) после 5 N, где N представляет собой любой нуклеотид, с помощью саморасщепляемого рибозима типа "головка молотка" (hammerhead) будет вставлена следующая последовательность, где подчеркнутые последовательности образуют несовершенный элемент связывания РНК.

[0773] 5′NNNNNCAGTTTGTACGGATGACTGATGAGTCCCAAATAGGACGAAACGCGCTTCGGTGCGTCTCATCCTGATAAACTGCAAA-3′ (SEQ ID NO: 372).

[0774] Присутствует значительная гибкость в плане длины и природы данного элемента связывания, при этом она будет присутствовать для любого из перечисленных в исходной подаче не относящихся к HDV саморасщепляемых рибозимов. Для внесения рибозима типа "головка молотка" (hammerhead) для расщепления иРНК hexA с помощью PEgRNA с использованием того же протоспейсера, что и перечисленные выше конструкты, должна использоваться следующая последовательность PEgRNA (обозначения являются теми же самыми, что и выше):

[0775] 5′ACCTGAACCGTATATCGACGCACCGAAGCGCGTTTCGTCCTATTTGGGACTCATCAGGATATACGGTTCAGGTGATATACGGTTCAGGTGACGCACCGAAGCGCGT

TTCGTCCTATTTGGGACTCATCAGACCTGAACCGTATATCATCCTTCCAGTCAGGGC CATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGGACCGAGTCGGTCCACCTGAACCGTATATC**GACGCACCGAAGCGC GTTTCGTCCTATTTGGGACTCATCAGGATATACGGT**TCAGGTCTATGGCCCTGA CTGGAA-3′ (где ядро каркаса PEgRNA подчеркнуто, участки гомологии и отжига обозначены курсивом, а вставленная последовательность обозначена жирным шрифтом) (SEQ ID NO: 373).

[0776] Конструирование других PEgRNA для вставки мотивов РНК может соответствовать общему описанному здесь принципу. Однако следует отметить, что многие мотивы РНК являются потенциально высокоструктурированными, что может затруднять их обратную транскрипцию и вставку в геном. Однако для некоторых последовательностей РНК, например простых шпилек, как последовательность РНК сама по себе, так и ее комплемент являются структурированными. Однако это вероятно не будет так для приведенных выше последовательностей. Таким образом, при вставке данных мотивов вероятно будет лучше всего, если PEgRNA будет кодировать обратный комплемент данных последовательностей, приводя к вставке последовательности ДНК, фактически кодирующей данный мотив, в геном. Аналогичным образом, включение саморасщепляемого рибозима в матричный участок PEgRNA будет приводить к процессированию и неэффективной активности, в то время как включение его обратного комплемента не будет приводить к этому. Таким образом, данные PEgRNA вероятно должны будут быть нацелены на кодирующую цепь, в то время как PEgRNA, кодирующие другие типы вставок (например для терапевтического корректирования), будут теоретически способны нацеливаться на любую цепь.

[0777] Кроме того, следует отметить, что для многих вставляемых мотивов соответствующая PEgRNA может не иметь способности транскрибироваться из промотора U6, что делает необходимым использование других промоторов, например pCMV. Аналогичным образом, боле длинные PEgRNA также могут являться менее стабильными. Более короткие мотивы, например маркеры m^6A, не будут иметь подобной проблемы.

G. Применение праймерного редактирования для получения сложных генных библиотек

[0778] Праймерное редактирование также может применяться для получения сложных библиотек белок- или РНК-кодирующих генов с определенными или варьирующимися вставками, делециями, или определенными аминокислотными/нуклеотидными конверсиями, при этом здесь описано их применение в рамках высокопроизводительного скрининга и направленной эволюции. Данный вариант использования праймерного редактирования дополнительно описан в Примере 7.

[0779] Получение вариабельных генетических библиотек наиболее часто достигалось с помощью мутагенной ПЦР (см. Cadwell RC and Joyce GF. PCR Methods Appl. 1992). Данный способ основан либо на использовании реакционных условий, которые снижают точность ДНК-полимеразы, или использовании модифицированных ДНК-полимераз с более высокими показателями мутации. Таким образом, отклонения в плане данных полимераз отражаются на продукте в виде библиотеки (например предпочтение мутаций в виде транзиций против трансверсий). Врожденным ограничением данного подхода для конструирования библиотеки является относительная неспособность влиять на размер подвергаемого варьированию гена. Большинство ДНК-полимераз имеет крайне низкие показатели индельных мутаций (вставок и делеций), при этом большинство из них будут приводить к мутациям рамки считывания в кодирующих белок участках, делая элементы библиотеки неспособными пройти какой-либо последующий отбор (см. McInerney P, Adams P, and Hadi MZ. Mol Biol Int. 2014).

[0780] Кроме того, отклонения при ПЦР и клонировании могут усложнить получение отдельных библиотек, состоящих из генов различных размеров. Данные ограничения могут существенно ограничить эффективность направленной эволюции для стимулирования существующих или конструирования новых функций белков. При естественной эволюции крупные изменения функции или эффективности белка обычно связаны со вставкой и делецией мутаций, которые имеют малую вероятность возникновения при стандартном получении библиотеки для мутагенеза. Более того, данные мутации наиболее часто возникают в участках представляющего интерес белка, которые, как предполагается, будут образовывать петли, в отличие от гидрофобного ядра. Таким образом, большинство инделей, полученных с помощью традиционного подхода без отклонений, вероятно будут являться либо вредоносными, или неэффективными.

[0781] Библиотеки, которые могут обеспечивать образование данных мутаций в участках в белке, где они с наибольшей вероятностью будут являться полезными, например в участках в виде петель, будут иметь значительное преимущество по сравнению с традиционными библиотеками при условии, что все библиотеки имеют доступ лишь к небольшой части возможного пространства мутаций. Наконец, хотя является возможным получать генетические библиотеки со специфичными в плане участков мутациями в виде инделей с помощью многоэтапной ПЦР и клональной сборки с помощью праймеров NNK или с помощью тасования ДНК, данные библиотеки не могут подвергаться дополнительным раундам "образования инделей" при непрерывной эволюции. Непрерывная эволюция представляет собой тип направленной эволюции с минимальным вмешательством пользователя. Одним подобным примером является PACE (см. Esvelt KM, Carlson JC, and Liu DR. Nature. 2011). Поскольку непрерывная эволюция осуществляется при минимальном вмешательстве пользователя, любое повышение разнообразия библиотеки при эволюции должно осуществляться с помощью нативного механизма репликации. Таким образом, несмотря на то, что библиотеки генов со вставленными или удаленными кодонами в качестве специфических локусов могут быть получены и подвергнуты скринингу в рамках PACE, дополнительные раунды "образования инделей" являются невозможными.

[0782] Предполагается, что программируемость праймерного редактирования (PE) может быть использована для получения имеющих высокую сложность программируемых генетических библиотек для применения в рамках высокопроизводительного скрининга и направленной эволюции (см. Фиг.29A). PE может обеспечивать вставку, изменение или удаление определенных количеств нуклеотидов из определенных генетических локусов с помощью информации, кодируемой в направляющей РНК для праймерного редактирования (см. Фиг.29B). Это обеспечивает получение нацеленных библиотек с одной или несколькими аминокислотами, вставленными или удаленными из участков в виде петель, где мутации с наибольшей вероятностью вызовут изменения в плане функции без побочного внесения нефункциональных мутаций рамки считывания (см. Фиг.29C). PE может использоваться для внесения определенных наборов мутаций безотносительно отклонений, присущих ДНК-полимеразе или подвергаемой мутированию последовательности.

[0783] Например, в то время как конвертация кодона CCC в стоп-кодон будет иметь малую вероятность осуществления при стандартном получении библиотеки, поскольку она потребует трех последовательных мутаций, включая две трансверсии, PE может быть использован для конвертации любого необходимого кодона в стоп-кодон TGA за один этап. Они также могут применяться для внесения программируемого разнообразия в необходимых положениях, например путем внесения кодонов, кодирующих любую гидрофобную аминокислоту в необходимом участке, в то же время не кодирую какие-либо другие. Более того, вследствие программируемости PE несколько PEgRNA могут применяться для получения нескольких различных изменений в нескольких участках одновременно, обеспечивая получение высокопрограммируемых библиотек (см. Фиг.29D). Кроме того, является возможным использовать обратные транскриптазы с меньшей точностью для получения участков мутагенеза в рамках иным образом невариабельной библиотеки (например обратную транскриптазу ВИЧ-1 или обратную транскриптазу фага Bordetella) (см. Naorem SS, Hin J, Wang S, Lee WR, Heng X, Miller JF, Guo H. Proc Natl Acad Sci 2017 and Martinez MA, Vartanian JP, Wain-Hobson S. Proc Natl Acad Sci USA 1994).

[0784] Также предполагается возможность итеративных раундов PE в рамках одного участка, что обеспечивает, например, возможность повторной вставки кодонов в одном участке, например участке в виде петли. Кроме того, предполагается, что все из описанных выше подходов могут быть встроены в непрерывную эволюцию, обеспечивая получение новых подвергаемых эволюции *in situ* библиотек (см. Фиг.30). Они также могут применяться для конструирования этих библиотек в рамках других типов клеток, где иным образом будет сложно собрать крупные библиотеки, например в рамках клеток млекопитающих. Получение кодируемых PE бактериальных штаммов, которые были оптимизированы для направленной эволюции, будет являться полезным дополнительным инструментом для идентификации белков и РНК с улучшенной или новой функциональностью. Все из данных вариантов применения PE не являются очевидными вследствие новой природы PE. В заключение, получение библиотек с помощью PE будет являться крайне полезным инструментом в рамках синтетической биологии и направленной эволюции, а также для высокопроизводительного скрининга белка и комбинаторных мутантов РНК.

Альтернативные подходы

[0785] Основной способ, с помощью которого получают разнообразные библиотеки в настоящее время, представляет собой описанную здесь мутагенную ПЦР (см. Cadwell RC and Joyce GF. PCR Methods Appl. 1992). Вставки или делеции могут быть внесены с помощью расщепленных праймеров NNK в определенных участках во время ПЦР, однако внесение подобных мутаций в нескольких участках требует нескольких раундов итеративной ПЦР и клонирования перед конструированием более разнообразной библиотеки с помощью мутагенной ПЦР, что делает данный способ медленным. Альтернативный дополнительный способ представляет собой тасование ДНК, при котором фрагменты библиотеки генов, полученной с помощью обработки ДНКазой, вносят в ПЦР реакцию без праймеров, что приводит к отжигу различных фрагментов друг на друге и быстрому получению более разнообразных библиотек, чем при использовании мутагенной ПЦР самостоятельно (см. Meyer AJ, Ellefson JW, Ellington AD. Curr Protoc Mol Biol. 2014). Несмотря на то, что этот подход может теоретически обеспечивать получение индельных мутаций, он чаще приводит к мутациям рамки считывания, которые устраняют функцию гена. Более того, тасование ДНК требует высокой степени гомологии между фрагментами генов.

[0786] Оба данных способа должны быть реализованы *in vitro* с трансформацией полученной библиотеки в клетки, в то время как библиотеки, полученные с помощью PE, могут быть сконструированы *in situ*, обеспечивая возможность их применения в рамках непрерывной эволюции. В то время как библиотеки могут быть сконструированы *in situ* с помощью *in vivo* мутагенеза, данные библиотеки основаны на клеточном механизме хозяина и демонстрируют отклонения в плане инделей. Аналогичным образом, несмотря на то, что традиционные способы клонирования могут применяться для получения специфичных в плане участка мутационных профилей, они не могут применяться *in situ* и в целом собираются по одному за раз *in vitro* перед трансформацией в клетки. Эффективность и широкая функциональность PE как в прокариотических, так и эукариотических типах клеток дополнительно свидетельствует о том, что данные библиотеки могут быть сконструированы напрямую в представляющем интерес типе клеток, в отличие от клонирования в организм-модель, такой как *E. coli*, с последующим трансфицированием в представляющую интерес клетку или организм. Другой альтернативный подход для направленной диверсификации представляет собой автоматизированную мультиплексную геномную инженерию или MAGE, в рамках которой несколько одноцепочечных ДНК-

олигонуклеотидов могут быть встроены в репликационные вилки и приводят к программируемым мутациям[7]. Однако MAGE требует значительного модифицирования штамма хозяина и может приводить к 100-кратному повышению нецелевых или фоновых мутаций (см. Nyerges Á et al. Proc Natl Acad Sci USA. 2016), в то время как PE является более высокопрограммируемым и, как ожидается, будет приводить к меньшему количеству нецелевых эффектов. Кроме того, MAGE не была продемонстрирована в рамках широкого ряда типов клеток, включая клетки млекопитающих.

[0787] С другой стороны, праймерное редактирование представляет собой новый и неочевидный дополнительный способ получения библиотек.

Применение PE для конструирования генных библиотек

[0788] PE может применяться для конструирования генных библиотек программируемым образом.

[0789] В одном примере PE может применяться в рамках эксперимента на основе направленной эволюции для внесения вариантов белка в библиотеки генов в рамках эксперимента в виде непрерывной эволюции с помощью PACE, обеспечивая итеративное накопление как точечных мутаций, так и инделей таким образом, который не является возможным с использованием традиционных подходов.

[0790] Уже было показано, что PE может обеспечивать специфическую в плане участка и программируемую вставку нуклеотидов в генетическую последовательность *E. coli*. Направленная эволюция может применяться для идентификации монотел с повышенным связыванием в отношении определенного эпитопа с использованием модифицированного отбора в рамках PACE на основе двухгибридного связывания белок:белок. Специфичные и высоковариабельные петли в рамках данных монотел вносят значительный вклад в аффинность и специфичность. Повышенное связывание монотел может быть быстро достигнуто в рамках PACE путем варьирования длины и композиции данных петель определенным образом. Однако варьирование длины последовательности не является стандартной функциональностью PACE. В то время как библиотека с петлями различного размера может использоваться в качестве начальной точки для PACE, в рамках отбора на основе PACE не будут осуществляться последующие улучшения в плане длины, что ограничивает доступ к полезным синергетическим комбинациям точечных мутаций и индельных мутаций.

[0791] В рамках различных вариантов осуществления PE может применяться для улучшения отбора на основе PACE за счет обеспечения *in situ* получения и эволюции монотел с петлями различной длины. Для этого штамм *E. coli* для PACE может быть соединен с дополнительной плазмидой PE, которая кодирует фермент PE и одну или несколько PEgRNA. Экспрессия фермента PE и PEgRNA в *E. coli* будет контролироваться малой молекулой, доставляемой в "лагуну" PACE (lagoon, сосуд фиксированного объема) со скоростью, выбранной экспериментатором.

[0792] В различных вариантах осуществления компоненты PEgRNA будут содержать спейсер, направляющий PE к представляющему интерес участку на фаге для отбора, и будет сконструирована таким образом, что несколько из трех нуклеотидов могут быть вставлены в участок-мишень, так что будет осуществляться внесение нового участка связывания PEgRNA, обеспечивая итеративную вставку одного или нескольких кодонов в участке-мишени.

[0793] Параллельно с этим другой штамм-хозяин *E. coli* может включать PEgRNA, которые будут служить матрицей для удаления одного или нескольких кодонов, обеспечивая уменьшение размера петли при эволюции. В рамках эксперимента PACE может использоваться смесь обоих штаммов или обеспечиваться их изменение для реализации медленного и контролируемого добавления или удаления последовательностей петель.

[0794] Помимо применения PE и PACE для создания библиотек монотел, данный способ также может использоваться для эволюции антител с помощью PE и PACE. Принципы связывания, присущие антителам, являются очень похожими на принципы, присущие монотелам: длина петель, соответствующих участку определения комплементарности антитела, является критически важной для их функции связывания. Кроме того, было обнаружено, что более длинные петли являются критически важными в плане развития редких антител с широкой защитной активностью против ВИЧ-1 и других вирусных инфекций (см. Mascola JR, Haynes BF. Immunol Rev. 2013). Применение PE как описано выше в отношении антитела или полученной на основе антитела молекулы будет обеспечивать получение антитела с различной длиной петель и различной последовательностью петель. В комбинации с PACE подобный подход будет обеспечивать улучшенное связывание за счет геометрии петель, которая является недостижимой при стандартном PACE, и, таким образом, обеспечивать эволюцию высокофункциональных антител.

[0795] В качестве неограничивающего примера следующие PEgRNA могут использоваться для программирования модифицирования генома бактериофага, используемого в рамках эксперимента на основе непрерывной эволюции:

МОДИФИКАЦИЯ	ПОСЛЕДОВАТЕЛЬНОСТЬ PEGRNA
ВСТАВКА CCA	UACACCAUCACGGUCUAUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUAUCUUCGCCCCAUGCAUAGACCGUGAUGG (SEQ ID NO: 101)
1 NT ДЕЛЕЦИЯ	CGCGUCGCGCUCGUCAGAUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUCCGCCUACCUGCAUCUGACGAGCGCGA (SEQ ID NO: 102)
ТОЧЕЧНАЯ МУТАЦИЯ	AUCGGAGAAUACAUGAACAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCCGGAGAAUACAUGAACAUCGGACCCGCGCUAUCUUC (SEQ ID NO: 103)
ВСТАВКА NNN (N=A/T/G/C)	UACACCAUCACGGUCUAUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUAUCUUCGCCNNNUGCAUAGACCGUGAUGG (SEQ ID NO: 104)
ИТЕРАТИВНАЯ ВСТАВКА GGG	UACACCAUCACGGUCUAUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGGGGGGGGGGGGUGCAUAGACCGUGAUGG (SEQ ID NO: 105)

[0796] В различных вариантах осуществления применение PE для конструирования библиотек генов может быть основано на применении мутагенной активности подверженных ошибкам обратных транскриптаз. Применения подобной мутагенной обратной транскриптазы может обеспечивать получения подвергнутых мутагенезу программируемых библиотек вследствие низкой точности подверженных ошибкам RT. В соответствии с используемым здесь значением термин "подверженная ошибкам" обратная транскриптаза обозначает фермент в виде обратной транскриптазы, которая встречается в природе или которая была получена из другой обратной транскриптазы (например обратной транскриптазы M-MLV дикого типа), которая имеет частоту ошибок, которая меньше, чем частота ошибок обратной транскриптазы M-MLV дикого типа. Сообщается, что частота ошибок обратной транскриптазы M-MLV дикого типа соответствует диапазону в виде одной ошибки на от 15000 (верхняя граница) до 27000 (нижняя граница) встраиваний азотистых оснований. См. Boutabout et al. (2001) “DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1,” *Nucleic Acids Res* 29(11):2217–2222, который приведен здесь в качестве ссылки.

[0797] Таким образом, для целей данного изобретения термин "подверженная ошибкам" означает RT, которые имеют частоту ошибок в виде более чем одной ошибки на 15000 встраиваний азотистых оснований (6,7 x 10^{-5} или более), например 1 ошибку в 14000 азотистых оснований (7,14 x 10^{-5} или более), 1 ошибку в 13000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 12000 азотистых оснований или менее (7,7 x 10^{-5} или более), 1 ошибку в 11000 азотистых оснований или менее (9,1 x 10^{-5} или более), 1 ошибку в 10000 азотистых оснований или менее (1 x 10^{-4} или 0,0001 или более), 1 ошибку в 9000 азотистых оснований или менее (0,00011 или более), 1 ошибку в 8000 азотистых оснований или менее (0,00013 или более), 1 ошибку в 7000 азотистых оснований или менее (0,00014 или более), 1 ошибку в 6000 азотистых оснований или менее (0,00016 или более), 1 ошибку в 5000 азотистых оснований или менее (0,0002 или более), 1 ошибку в 4000 азотистых оснований или менее (0,00025 или более), 1 ошибку в 3000 азотистых оснований или менее (0,00033 или более), 1 ошибку в 2000 азотистых оснований или менее (0,00050 или более), или 1 ошибку в 1000 азотистых оснований или менее (0,001 или более), или 1 ошибку в 500 азотистых основаниях или менее (0,002 или более), или 1 ошибку в 250 азотистых основаниях или менее (0,004 или более).

[0798] Может быть предусмотрен ряд мутагенных RT для получения высокомутагенизированных программируемых библиотек. Двумя подобными примерами являются мутагенные обратные транскриптазы из фага *Bordetella* (см. Handa, S., et al. *Nucl Acids Res* 9711-25 (2018), который приведен здесь в качестве ссылки) и *Legionella pneumophila* (см. Arambula, D., et al. *Proc Natl Acad Sci USA* 8212-7 (2013), который приведен здесь в качестве ссылки). В случае RT из фага *Bordetella* (brt) может потребоваться также добавление к Cas9 вспомогательного белка (bavd) или его доставка *in trans*, а также дополнительных последовательностей РНК в PEgRNA для повышения связывания мутагенной RT с участком-мишенью (см. Handa, S., et al. *Nucl Acids Res* 9711-25 (2018)). При использовании мутагенных RT матричный участок PEgRNA может иметь высокое содержание аденозинов или кодонов AAY для повышения разнообразия.

[0799] Аминокислотная последовательность мутагенной RT из фага *Bordetella* приведена ниже. Подобно другим описанным здесь RT, белок Brt может быть слит с napDNAbp в виде слитого белка для образования функционального PE.

Название	Последовательность
Мутагенная RT brt	MGKRHRNLIDQITTWENLLDAYRKTSHGKRRTWGYLEFKE YDLANLLALQAELKAGNYERGPYREFLVYEPKPRLISALEF KDRLVQHALCNIVAPIFEAGLLPYTYACRPDKGTHAGVCHV QAELRRTRATHFLKSDFSKFFPSIDRAALYAMIDKKIHCAAT RRLLRVVLPDEGVGIPIGSLTSQLFANVYGGAVDRLLHDELK QRHWARYMDDIVVLGDDPEELRAVFYRLRDFASERLGLKIS HWQVAPVSRGINFLGYRIWPTHKLLRKSSVKRAKRKVANFI KHGEDESLQRFLASWSGHAQWADTHNLFTWMEEQYGIAC H (SEQ ID NO: 129)

[0800] В случае Brt из *Bodetella* продукт слияния PE также может включать дополнительный вспомогательный белок (Bavd). Вспомогательный белок может быть слит со слитым белком PE или поступать *in trans*. Аминокислотная последовательность вспомогательного белка Bavd приведена ниже:

Название	Последовательность
Вспомогательный белок bavd для brt	MEPIEEATKCYDQMLIVERYERVISYLYPIAQSIPRKHGVARE MFLKCLLGQVELFIVAGKSNQVSKLYAADAG LAMLRFWLRFLAGIQKPHAMTPHQVETAQVLIAEVGRILGS WIARVNRKGQAGK (SEQ ID NO: 236)

[0801] В случае Brt из *Bodetella* PEgRNA может включать дополнительную нуклеотидную последовательность, добавленную к PEgRNA, например на 5′ или 3′-конце. Ниже приведен пример последовательности, которую обычно получают из генома фага *Bordetella*:

НАЗВАНИАЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 1	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGC CUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAA CACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAG GCGGGGCCGACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 237)

[0802] Данная последовательность дополнения к PEgRNA может быть укорочена различными способами для сокращения длины. Например, последовательность

дополнения к PEgRNA 1 может быть укорочена до следующих примеров альтернативных последовательностей дополнения:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 2	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCGACUUC (SEQ ID NO: 238)
ДОПОЛНЕНИЕ К PEGRNA 3	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 239)
ДОПОЛНЕНИЕ К PEGRNA 4	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCA (SEQ ID NO: 277)
ДОПОЛНЕНИЕ К PEGRNA 5	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCGACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 240)
ДОПОЛНЕНИЕ К PEGRNA 6	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCGACUUC (SEQ ID NO: 241)
ДОПОЛНЕНИЕ К PEGRNA 7	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCAAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 242)
ДОПОЛНЕНИЕ К PEGRNA 8	CAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGGUGGUCA (SEQ ID NO: 243)

[0803] В других вариантах осуществления последовательность дополнения PEgRNA также может быть подвержена мутированию. Например, последовательность дополнения к PEgRNA 1 может быть подвержена мутированию до следующих примеров альтернативных последовательностей дополнения:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ДОПОЛНЕНИЕ К PEGRNA 1 МУТИРОВАННОЕ	ACCUUCUUGCAUGGCUCUGCCAACGCUACGGCUUGGCGGGCUGGCCUUUCCUCAAUAGAUGAGCCGCCGGUUCUGUCCUGCUUCGGCGAACACGUUACACGGUUCGGCAAAACGUCGAUUACUGAAAAUGGAAAGGCGGGGCCGACUUCAAGGGCAGGCUGGGAAAUAA (SEQ ID NO: 244)

[0804] В различных вариантах осуществления, связанных с применением PE для конструирования библиотек генов, могут применяться специальные условия в плане PEgRNA. Например, без связи с теорией, могут потребоваться дополнительные описанные выше последовательности PEgRNA для обеспечения эффективного мутагенеза за счет мутагенных RT. В другом варианте осуществления ставка

итеративного кодона с помощью PE может требовать определенных конструкций PEgRNA. Например, может требоваться, чтобы для итеративной вставки кодона GGG (глицин) полный участок гомологии PEgRNA состоял из G, как как показано выше. Это будет означать, что только определенные участки могут подвергаться итеративной вставке.Кроме того, PE не будет способен нарушать последовательность PAM.

H. Применение праймерного редактирования для вставки иммуноэпитопов

[0805] Праймерные редакторы также могут применяться в качестве средства вставки известных иммуногенетических эпитопов в эндогенную или чужеродную геномную ДНК, что приводит к модифицированию соответствующих белков для терапевтических или биотехнологических вариантов использования (см. Фиг.31 и 32). До изобретения праймерного редактирования такие вставки могли быть реализованы лишь неэффективным образом с высокими показателями образования инделей из DSB. Праймерное редактирование устраняет проблему в виде высокого образования инделей при изменениях в виде вставок, в то же время в целом обеспечивая более высокую эффективность, чем HDR. Данный низкий показатель образования инделей является значительным преимуществом праймерного редактирования над HDR в качестве способа нацеленных вставок ДНК, особенно в рамках описанного использования в виде вставки иммуногенных эпитопов. Длина эпитопов соответствует диапазону от нескольких оснований до сотен оснований. Праймерный редактор представляет собой эффективный подход для реализации таких нацеленных вставок в клетках млекопитающих.

[0806] Ключевая концепция изобретения заключается в применении праймерных редакторов для вставки нуклеотидной последовательности, содержащей ранее описанные иммуногенные эпитопы, в эндогенную или чужеродную геномную ДНК для снижения и/или устранения их продуктов в виде белков и/или экспрессирующих типов клеток. Нуклеотидные последовательности для вставки иммуногенных эпитопов будут нацелены на гены для получения слитых белков с белком, кодируемым геном, на который осуществляется нацеливание, и трансляции соответствующего белка вставленного иммуногенного эпитопа. Иммунная система пациента была ранее подвергнута тренировке для распознавания данных эпитопов в результате стандартной предшествующей иммунизации на основе обычной вакцинации против, например, столбняка, дифтерии, или кори. В результате иммуногенной природы слитых эпитопов ожидается, что иммунная система пациента будет распознавать и деактивировать

подвергнутый праймерному редактированию белок (а не только вставленный эпитоп) и потенциально клетки, из которых он был экспрессирован.

[0807] Технологии точного геномного нацеливания на основе использования системы CRISPR/Cas недавно были протестированы в рамках ряда вариантов использования, включая вставку сконструированных последовательностей ДНК в геномные локусы-мишени. Ранее для данного варианта использования применяли направляемую на основе гомологии репарацию (HDR), которая требовала донорской оцДНК матрицы и инициирования репарации с помощью двуцепочечного разрыва ДНК (DSB). Данная стратегия обеспечивает наиболее широкий диапазон возможных реализуемых изменений в клетках и является единственным способом, доступным для вставки крупных последовательностей ДНК в клетки млекопитающих. Однако HDR связана с затруднениями в виде нежелательных клеточных побочных эффектов, возникающих вследствие ее инициирующего DSB, таких как высокие уровни образования инделей, транслокация ДНК, крупные делеции и активация P53. Помимо данных недостатков, HDR ограничивается низкой эффективностью в рамках многих типов клеток (Т-клетки являются заметным исключением из данного наблюдения). Недавние попытки преодолеть данные недостатки включают слияния мутантов Rad51 человека с никазой D10A Cas9 (RDN), приводящие к получению HDR системы без DSB, которая включает улучшенные соотношения продукт HDR:индели и более низкое нецелевое редактирования, но все же ограничена зависимостями в плане типов клеток и лишь умеренной эффективностью HDR редактирования.

[0808] Недавно разработанные продукты слияния Cas9 с обратными транскриптазами ("праймерные редакторы"), соединенные с PEgRNA, представляют собой новую технологию геномного редактирования, которая обеспечивает ряд преимуществ по сравнению с существующими способами геномного редактирования, включая способность вносить любое нуклеотидное замещение и обеспечивать вставку или делецию любой короткой последовательности нуклеотидов (до, по меньшей мере, нескольких дюжин оснований) специфичным в плане участка образом. Следует отметить, что изменения на основе PE осуществляются со в целом низкими показателями нежелательных инделей. Таким образом, PE обеспечивает возможность нацеленного редактирования на основе вставок, которое ранее являлось невозможным или непрактичным.

[0809] Данный конкретный аспект относится к способу применения праймерного редактирования в качестве средства вставки известных иммуногенетических эпитопов в эндогенную или чужеродную геномную ДНК, что приводит к модифицированию соответствующих белков для терапевтических или биотехнологических вариантов использования (см. Фиг.31 и 32). До изобретения праймерного редактирования такие вставки могли быть реализованы лишь неэффективным образом с высокими показателями образования инделей из DSB. Праймерное редактирование устраняет проблему в виде высокого образования инделей при изменениях в виде вставок, в то же время в целом обеспечивая более высокую эффективность, чем HDR. Данный низкий показатель образования инделей является значительным преимуществом праймерного редактирования над HDR в качестве способа нацеленных вставок ДНК, особенно в рамках описанного использования в виде вставки иммуногенных эпитопов. Длина эпитопов соответствует диапазону от нескольких оснований до сотен оснований. Праймерный редактор представляет собой наиболее эффективную и точную технологию реализации таких нацеленных вставок в клетках млекопитающих.

[0810] Ключевая концепция данного аспекта заключается в применении праймерных редакторов для вставки нуклеотидной последовательности, содержащей ранее описанные иммуногенные эпитопы, в эндогенную или чужеродную геномную ДНК для снижения и/или устранения их продуктов в виде белков и/или экспрессирующих типов клеток. Нуклеотидные последовательности для вставки иммуногенных эпитопов будут нацелены на гены для получения слитых белков с белком, кодируемым геном, на который осуществляется нацеливание, и трансляции соответствующего белка вставленного иммуногенного эпитопа. Иммунная система пациента была ранее подвергнута тренировке для распознавания данных эпитопов в результате стандартной предшествующей иммунизации на основе обычной вакцинации против, например, столбняка, дифтерии, или кори. В результате иммуногенной природы слитых эпитопов ожидается, что иммунная система пациента будет распознавать и деактивировать подвергнутый праймерному редактированию белок (а не только вставленный эпитоп) и потенциально клетки, из которых он был экспрессирован.

[0811] Продукты слияния с генами-мишенями будут сконструированы соответствующим образом для обеспечения видимости трансляции вставленного эпитопа белка для распознавания иммунной системой. Это может включать нацеленные нуклеотидные вставки, приводящие к трансляции белка с образованием С-

концевых продуктов слияния иммуногенных эпитопов с генами-мишенями, N-концевых продуктов слияния иммуногенных эпитопов с генами-мишенями, или вставке нуклеотидов в гены, так что иммуногенные эпитопы кодируются внутри расположенных на поверхности участков структуры белка.

[0812] В качестве части данного изобретения может потребоваться сконструировать белковые линкеры, кодируемые в виде нуклеотидов, вставленных между последовательностью гена-мишени и нуклеотидной последовательностью вставленного иммуногенного эпитопа, для обеспечения распознавания иммунной системой, клеточного трафика, функции белка, или фолдинга белка, связанного с геном-мишенью. Данные вставленные кодируемые нуклеотидами белковые линкеры могут включать (но без ограничения) различную длину и последовательности линкера XTEN или различную длину и последовательности линкеров на основе глицина-серина. Данные сконструированные линкеры ранее применялись для успешной реализации слияния белков. Примеры линкеров могут включать любой из описанных здесь, включая аминокислотную последовательность (GGGGS)n (SEQ ID NO: 165), (G)n (SEQ ID NO: 166), (EAAAK)n (SEQ ID NO: 167), (GGS)n (SEQ ID NO: 168), (SGGS)n (SEQ ID NO: 169), (XP)n (SEQ ID NO: 170), или любую их комбинацию, где n независимо представляет собой целое число от 1 до 30, и где X представляет собой любую аминокислоту. В некоторых вариантах осуществления линкер включает аминокислотную последовательность (GGS)n (SEQ ID NO: 176), при этом n представляет собой 1, 3, или 7. В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). В некоторых вариантах осуществления линкер включает аминокислотную последовательность SGGS (SEQ ID NO: 174).

[0813] Отличительные признаки данного аспекта включают способность использовать ранее приобретенные иммунные ответы на определенные аминокислотные последовательности в качестве средства индуцирования иммунного ответа против иным образом неиммуногенных белков. Другой отличительный признак представляет собой способность обеспечивать вставку нуклеотидных последовательностей данных

иммуногенных эпитопов в мишени таким образом, при котором не осуществляется индуцирование высоких уровней нежелательных инделей в качестве сопутствующего продукта редактирования и который является эффективным при их вставке. Данный конкретный вариант использования PE имеет способность обеспечивать комбинирование специфичных в плане типа клеток способов доставки (например серотипов AAV) для вставки эпитопов в типы клеток, которые представляют интерес в плане запуска иммунного ответа на них.

[0814] Праймерное редактирование в качестве средства вставки иммуногенных эпитопов в патогенные гены может применяться для программирования иммунной системы пациента для борьбы с широким диапазоном заболеваний (не ограничиваясь раком, как в случае иммуно-онкологических стратегий). Очевидным вариантом применения данной технологии будет использование в качестве лекарства от рака, поскольку она может нарушать механизм опухоли, связанный с ускользанием от иммунитета, за счет индуцирования иммунного ответа на соответствующий онкоген, подобный HER2, или факторы роста, подобные EGFR. Подобный подход может выглядеть аналогичным модифицированию Т-клеток, однако одно новое преимущество данного подхода заключается в том, что он может применяться в рамках многих типов клеток и для заболеваний помимо рака без необходимости в получении и введении модифицированных Т-клеток пациентам.

[0815] PE используется для вставки иммуногенного эпитопа, против которого большинство пациентов уже являются вакцинированными (столбняк, коклюш, дифтерия, корь, эпидемический паротит, краснуха и так далее), в чужеродный или эндогенный ген, который вызывает заболевание, так что иммунная система пациента учится деактивировать данный белок.

[0816] Заболевания, которые могут быть связаны с потенциальной терапевтической пользой от вышеупомянутой стратегии, включают заболевания, вызываемые накоплением токсичных белков, например фатальную семейную бессонницу. Другие заболевания, которые могут быть связаны с пользой, включают заболевания, вызываемые патогенной сверхэкспрессией иным образом нетоксичного эндогенного белка, а также вызываемые чужеродными патогенами.

[0817] Первичные терапевтические показания включают те, которые были указаны выше, такие как способы лечения рака, прионных и иных нейродегенеративных заболеваний, инфекционных заболеваний, и профилактическую медицину. Вторичные

терапевтические показания могут включать профилактический уход для пациентов с поздно развившимися генетическими заболеваниями. Ожидается, что текущий стандарт в плане поддерживающих видов лечения может быть использован совместно с праймерным редактированием для некоторых заболеваний, таких как особо агрессивные виды рака, или в случаях, когда лечение помогает смягчать симптомы заболевания до полного излечения заболевания. Ниже приведены примеры иммуноэпитопов, которые могут быть вставлены в гены с помощью описанных здесь праймерных редакторов:

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
1	СТОЛБНЯЧНЫЙ ТОКСОИД	QYIKANSKFIGITEL (SEQ ID NO: 396)	CATGATATAAAAGCAAATTCTAAATTTATAGGTATAACTGAACTA (SEQ ID NO: 397)
2	МУТАНТ ДИФТЕРИЙНОГО ТОКСИНА CRM197	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVA	GGCGCCGACGACGTGGTGGACAGCAGCAAGAGCTTCGTGATGGAGAACTTCAGCAGCTACCACGGCACCAAGCCCGGCTACGTGGACAGCATCCAGAAGGGCATCCAGAAGCCCAAGAGCGGCACCCAGGGCAACTACGACGACGACTGGAAGGAGTTCTACAGCACCGACAACAAGTACGACGCCGCCGGCTACAGCGTGGACAACGAGAACCCCCTGAGCGGCAAGGCCGGCGGCGTGGTGAAGGTGACCTACCCCGGCCTGACCAAGGTGCTGGCCCTGAAGGTGGACAACGCCGAGACCATCAAGAAGGAGCTGGGCCTGAGCCTGACCGAGCCCCTGATGGAGCAGGTGGGCACCGAGGAGTTCATCAAGAGGTTCGGCGACGGCGCCAGCAGGGTGGTGCTGAGCCTGCCCTTCGCCGAGGGCAGCAGCAGCGTGGAGTACATCAACAACTGGGAGCAGGCCAAGGCCCTGAGCGTGGAGCTGGAGATCAACTTCGAGACCAGGGGCAAGAGGGGCCAGGACGCCATGTACGAGTACATGGCCCAGGCCTGCGCCGGCAACAGGGTGAGGAGGAGCGTGGGCAGCAGCCTGAGCTGCATCAACCTGGACTGGGACGTGATCAGGGACAAGACCAAGACCAAGATCGAGAGCCTGAAGGAGCACGGCCCCATCAAGAACAAGATGAGCGAGAGCCCCAA

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
		QAIPLVGELVDIGF AAYNFVESIINLFQ VVHNSYNRPAYSP GHKTQPFLHDGY AVSWNTVEDSIIR TGFQGESGHDIKI TAENTPLPIAGVL LPTIPGKLDVNKS KTHISVNGRKIRM RCRAIDGDVTFCR PKSPVYVGNGVH ANLHVAFHRSSSE KIHSNEISSDSIGV LGYQKTVDHTKV NSKLSLFFEIKS (SEQ ID NO: 398)	CAAGACCGTGAGCGAGGAGAAGGCC AAGCAGTACCTGGAGGAGTTCCACCA GACCGCCCTGGAGCACCCCGAGCTGA GCGAGCTGAAGACCGTGACCGGCACC AACCCCGTGTTCGCCGGCGCCAACTA CGCCGCCTGGGCCGTGAACGTGGCCC AGGTGATCGACAGCGAGACCGCCGAC AACCTGGAGAAGACCACCGCCGCCCT GAGCATCCTGCCCGGCATCGGCAGCGT GATGGGCATCGCCGACGGCGCCGTGC ACCACAACACCGAGGAGATCGTGGCC CAGAGCATCGCCCTGAGCAGCCTGAT GGTGGCCCAGGCCATCCCCCTGGTGG GCGAGCTGGTGGACATCGGCTTCGCC GCCTACAACTTCGTGGAGAGCATCATC AACCTGTTCCAGGTGGTGCACAACAG CTACAACAGGCCCGCCTACAGCCCCG GCCACAAGACCCAGCCCTTCCTGCAC GACGGCTACGCCGTGAGCTGGAACAC CGTGGAGGACAGCATCATCAGGACCG GCTTCCAGGGCGAGAGCGGCCACGAC ATCAAGATCACCGCCGAGAACACCCC CCTGCCCATCGCCGGCGTGCTGCTGCC CACCATCCCCGGCAAGCTGGACGTGA ACAAGAGCAAGACCCACATCAGCGTG AACGGCAGGAAGATCAGGATGAGGTG CAGGGCCATCGACGGCGACGTGACCT TCTGCAGGCCCAAGAGCCCCGTGTAC GTGGGCAACGGCGTGCACGCCAACCT GCACGTGGCCTTCCACAGGAGCAGCA GCGAGAAGATCCACAGCAACGAGATC AGCAGCGACAGCATCGGCGTGCTGGG CTACCAGAAGACCGTGGACCACACCA AGGTGAACAGCAAGCTGAGCCTGTTC TTCGAGATCAAGAGC (SEQ ID NO: 399)
3	ИММУНОЭПИТОП ПАРОТИТА 1	GTYRLIPNARANLTA (SEQ ID NO: 400)	GGCACCTACAGGCTGATCCCCAACGC CAGGGCCAACCTGACCGCC (SEQ ID NO: 401)
4	ИММУНОЭПИТОП	PSKFFTISDSATFA PGPVSNA (SEQ ID	CCGAGCAAATTTTTTACCATTAGCGAT AGCGCGACCTTTGCGCCGGGCCCGGT

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
	ПАРОТИТА 2 ИММУНОЭПИТОП ПАРОТИТА 1	NO: 402) PSKLFIMLDNATFAPGPVVNA (SEQ ID NO: 404)	GAGCAACGCG (SEQ ID NO: 403) CCGAGCAAACTGTTTATTATGCTGGATAACGCGACCTTTGCGCCGGGCCCGGTGGTGAACGCG (SEQ ID NO: 405) ИЗБРАННЫЕ ПРИМЕРЫ НА ОСНОВЕ РАЗНООБРАЗИЯ ГЕМАГГЛУТИНИН-НЕЙРАМИНИДАЗЫ (HN) СРЕДИ СВЯЗАННЫХ С ЗАБОЛЕВАНИЕМ ШТАММОВ (ТАБЛИЦА 1), РАСХОЖДЕНИЕ МЕЖДУ ШТАММОМ В РАМКАХ ВАКЦИНЫ JL5 И СВЯЗАННЫМИ С ЗАБОЛЕВАНИЕМ ШТАММАМИ (ТАБЛИЦА 2)
5	ВИРУС КРАСНУХИ (RV)	TPPPYQVSCGGESDRASARVIDPAAQS (SEQ ID NO: 406)	ACCCCCCCCCCCTACCAGGTGAGCTGCGGCGGCGAGAGCGACAGGGCCAGCGCCAGGGTGATCGACCCCGCCGCCCAGAGC (SEQ ID NO: 407)
6	ГЕМАГГЛУТИНИН	PEYAYKIVKNKKMEDGFLQGMVDGWYGHHSNEQGSGLMENERTLDKANPNNDLCSWSDHEASSNNTNQEDLLQRESRRKKRIGTSTLNQRGNCNTKCQTEEARLKREEVSLVKSDQCSNGSLQCRANNSTEQVD (SEQ ID NO: 408)	CCCGAGTACGCCTACAAGATCGTGAAGAACAAGAAGATGGAGGACGGCTTCCTGCAGGGCATGGTGGACGGCTGGTACGGCCACCACAGCAACGAGCAGGGCAGCGGCCTGATGGAGAACGAGAGGACCCTGGACAAGGCCAACCCCAACAACGACCTGTGCAGCTGGAGCGACCACGAGGCCAGCAGCAACAACACCAACCAGGAGGACCTGCTGCAGAGGGAGAGCAGGAGGAAGAAGAGGATCGGCACCAGCACCCTGAACCAGAGGGGCAACTGCAACACCAAGTGCCAGACCGAGGAGGCCAGGCTGAAGAGGGAGGAGGTGAGCCTGGTGAAGAGCGACCAGTGCAGCAACGGCAGCCTGCAGTGCAGGGCCAACAACAGCACCGAGCAGGTGGAC (SEQ ID NO: 409)
7	НЕЙРАМИНИДАЗА	TKSTNSRSGGISGPDNEAPVGEAPSPYGDNPRPNDGNN	ACCAAGAGCACCAACAGCAGGAGCGGCGGCATCAGCGGCCCCGACAACGAGGCCCCCGTGGGCGAGGCCCCCAGCCC

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
		IRIGSKGYNGIITD TIEESCSCYPDAK VVKSVELDSTIWT SGSSPNQKIITIGW DPNGWTGTPMSP NGAYGTDGPSNG QANQHQAESISA GNSSLCPIRDNWH GSNRSWSWPDGA E (SEQ ID NO: 410)	CTACGGCGACAACCCCAGGCCCAACG ACGGCAACAACATCAGGATCGGCAGC AAGGGCTACAACGGCATCATCACCGA CACCATCGAGGAGAGCTGCAGCTGCT ACCCCGACGCCAAGGTGGTGAAGAGC GTGGAGCTGGACAGCACCATCTGGAC CAGCGGCAGCAGCCCCAACCAGAAGA TCATCACCATCGGCTGGGACCCCAACG GCTGGACCGGCACCCCCATGAGCCCC AACGGCGCCTACGGCACCGACGGCCC CAGCAACGGCCAGGCCAACCAGCACC AGGCCGAGAGCATCAGCGCCGGCAAC AGCAGCCTGTGCCCCATCAGGGACAA CTGGCACGGCAGCAACAGGAGCTGGA GCTGGCCCGACGGCGCCGAG (SEQ ID NO: 411)
8	TAP1 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА ГЕМАГГЛУТИНИНЕ ВИРУСА H5N1	EKIVLLLAMMEKI VLLLAKCQTPMG AIKAVDGVTNKCP YLGSPSF (SEQ ID NO: 412)	GAGAAGATCGTGCTGCTGCTGGCCATG ATGGAGAAGATCGTGCTGCTGCTGGCC AAGTGCCAGACCCCCATGGGCGCCAT CAAGGCCGTGGACGGCGTGACCAACA AGTGCCCCTACCTGGGCAGCCCCAGC TTC (SEQ ID NO: 413)
9	TAP2 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА НЕЙРАМИНИДАЗЕ ВИРУСА H5N1	IRPCFWVELNPNQ KIITIRPCFWVELI CYPDAGEIT (SEQ ID NO: 414)	ATCAGGCCCTGCTTCTGGGTGGAGCTG AACCCCAACCAGAAGATCATCACCATC AGGCCCTGCTTCTGGGTGGAGCTGATC TGCTACCCCGACGCCGGCGAGATCAC C (SEQ ID NO: 415)
10	ЭПИТОПЫ ГЕМАГГЛ	MEKIVLLLAEKIV LLLAMCPYLGSPS	ATGGAGAAGATCGTGCTGCTGCTGGCC GAGAAGATCGTGCTGCTGCTGGCCATG

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
	УТИНИНАК HLA КЛАССА I	FKCQTPMGAIKAVDGVTNK (SEQ ID NO: 416)	TGCCCCTACCTGGGCAGCCCCAGCTTCAAGTGCCAGACCCCCATGGGCGCCATCAAGGCCGTGGACGGCGTGACCAACAAG (SEQ ID NO: 417)
11	ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА I	NPNQKIITICYPDAGEITIRPCFWVELRPCFWVELI (SEQ ID NO: 418)	AACCCCAACCAGAAGATCATCACCATCTGCTACCCCGACGCCGGCGAGATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGAGGCCCTGCTTCTGGGTGGAGCTGATC (SEQ ID NO: 419)
12	ЭПИТОПЫ ГЕМАГГЛУТИНИНАК HLA КЛАССА II	MVSLVKSDQIGTSTLNQR (SEQ ID NO: 420)	ATGGTGAGCCTGGTGAAGAGCGACCAGATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 421)
13	ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА II	YNGIITDTI (SEQ ID NO: 422)	TACAACGGCATCATCACCGACACCATC (SEQ ID NO: 423)
14	HLA КЛАССА I И КЛАССА II, СВЯЗАННЫЙ С ЭПИТОПОМ ГЕМАГГЛУТИНИНА H5N1	MEKIVLLLAEKIVLLLAMMVSLVKSDQCPYLGSPSFIGTSTLNQRKCQTPMGAIKAVDGVTNK (SEQ ID NO: 424)	ATGGAGAAGATCGTGCTGCTGCTGGCCGAGAAGATCGTGCTGCTGCTGGCCATGATGGTGAGCCTGGTGAAGAGCGACCAGTGCCCCTACCTGGGCAGCCCCAGCTTCATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 425)
15	HLA КЛАССА I И КЛАССА II, СВЯЗАННЫЙ С	NPNQKIITIYNGIITDTICYPDAGEITIRPCFWVELRPCFWVELI (SEQ ID NO: 426)	AACCCCAACCAGAAGATCATCACCATCTACAACGGCATCATCACCGACACCATCTGCTACCCCGACGCCGGCGAGATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGAGGCCCTGCTTCTGGGTGGAGCTGATC (SEQ ID NO: 427)

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ПРИМЕР ПОСЛЕДОВАТЕЛЬНОСТИ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
	ЭПИТОПОМ НЕЙРАМИНИДАЗЫ H5N1		

[0818] Также могут быть внесены дополнительные иммуноэпитопы, которые известны в данной области. С помощью описанных здесь праймерных редакторов могут быть внесены любые иммуноэпитопы, доступные в рамках Epitope Database and Analysis Resource (iedb.org/epitopedetails_v3.php) (содержание которого приведено здесь в качестве ссылки).

[0819] В некоторых вариантах осуществления иммуноэпитопы, которые могут быть внесены с помощью описанных здесь праймерных редакторов, могут включать любые из следующих эпитопов:

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1	СТОЛБНЯЧНЫЙ ТОКСОИД	QYIKANSKFIGITEL	396	НЕТ
2	МУТАНТ ДИФТЕРИЙНОГО ТОКСИНА CRM197	GADDVVDSSKSFVMENFSSYHGTKP GYVDSIQKGIQKPKSGTQGNYDDDW KEFYSTDNKYDAAGYSVDNENPLSG KAGGVVKVTYPGLTKVLALKVDNA ETIKKELGLSLTEPLMEQVGTEEFIKR FGDGASRVVLSLPFAEGSSSVEYINN WEQAKALSVELEINFETRGKRGQDA MYEYMAQACAGNRVRRSVGSSLSCI NLDWDVIRDKTKTKIESLKEHGPIKN KMSESPNKTVSEEKAKQYLEEFHQT ALEHPELSELKTVTGTNPVFAGANYA AWAVNVAQVIDSETADNLEKTTAALS ILPGIGSVMGIADGAVHHNTEEIVAQS IALSSLMVAQAIPLVGELVDIGFAAYN FVESIINLFQVVHNSYNRPAYSPGHKT QPFLHDGYAVSWNTVEDSIIRTGFQG ESGHDIKITAENTPLPIAGVLLPTIPGK LDVNKSKTHISVNGRKIRMRCRAIDG DVTFCRPKSPVYVGNGVHANLHVAF HRSSSEKIHSNEISSDSIGVLGYQKTV	428	НЕТ

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
		DHTKVNSKLSLFFEIKS		
3	ПАРОТИТ	GTYRLIPNARANLTA	SEQ ID NO: 400	НЕТ
4	ПАРОТИТ	PSKFFTISDSATFAPGPVSNA; PSKLFIMLDNATFAPGPVVNA	402; 404	НЕТ
5	ВИРУС РУБЕЛЛЫ (RV)	TPPPYQVSCGGESDRASARVIDPAAQS	406	НЕТ
6	ГЕМАГГЛУТИНИН	PEYAYKIVKNKKMEDGFLQGMVDGWYGHHSNEQGSGLMENERTLDKANPNNDLCSWSDHEASSNNTNQEDLLQRESRRKKRIGTSTLNQRGNCNTKCQTEEARLKREEVSLVKSDQCSNGSLQCRANNSTEQVD	408	НЕТ
7	НЕЙРАМИНИДАЗА	TKSTNSRSGGISGPDNEAPVGEAPSPYGDNPRPNDGNNIRIGSKGYNGIITDTIEESCSCYPDAKVVKSVELDSTIWTSGSSPNQKIITIGWDPNGWTGTPMSPNGAYGTDGPSNGQANQHQAESISAGNSSLCPIRDNWHGSNRSWSWPDGAE	410	НЕТ
8	TAP1 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА ГЕМАГГЛУТИНИНЕ ВИРУСА H5N1	EKIVLLLAMMEKIVLLLAKCQTPMGAIKAVDGVTNKCPYLGSPSF	412	НЕТ
9	TAP2 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА НЕЙРАМИНИДАЗЕ ВИРУСА H5N1	IRPCFWVELNPNQKIITIRPCFWVELICYPDAGEIT	414	НЕТ
10	ЭПИТОПЫ ГЕМАГГЛУТИНИНА К HLA КЛАССА I	MEKIVLLLAEKIVLLLAMCPYLGSPSFKCQTPMGAIKAVDGVTNK	416	НЕТ
11	ЭПИТОПЫ НЕЙРАМИНИДАЗ	NPNQKIITICYPDAGEITIRPCFWVELRPCFWVELI	418	НЕТ

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	Ы К HLA КЛАССА I			
12	ЭПИТОПЫ ГЕМАГГЛУТИНИНА К HLA КЛАССА II	MVSLVKSDQIGTSTLNQR	420	НЕТ
13	ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА II	YNGIITDTI	422	НЕТ
14	HLA КЛАССА I И КЛАССА II, СВЯЗАННЫЙ С ЭПИТОПОМ ГЕМАГГЛУТИНИНА H5N1	MEKIVLLLAEKIVLLLAMMVSLVKSDQCPYLGSPSFIGTSTLNQRKCQTPMGAIKAVDGVTNK	424	НЕТ
15	HLA КЛАССА I И КЛАССА II, СВЯЗАННЫЙ С ЭПИТОПОМ НЕЙРАМИНИДАЗЫ H5N1	NPNQKIITIYNGIITDTICYPDAGEITIRPCFWVELRPCFWVELI	426	НЕТ
16	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	AACAGNRVRRSVGSSLKC	899	SRC280292
17	ШТАММ ВИРУСА КОРИ EDMONSTON	AADHCPVVEVNGVTI	900	P69353.1
18	ШТАММ ВИРУСА КОРИ EDMONSTON	AAHLPTGTPLDID	901	P04851.1
19	BORDETELLA PERTUSSIS	AALAVWAGLAVQ	902	Q00879.1
20	ШТАММ ВИРУСА КОРИ EDMONSTON	AALGVATAAQITAGI	903	P69353.1
21	ШТАММ ВИРУСА КРАСНУХИ THERIEN	AALLNTPPPYQVSCGGESDRATAR	904	P07566.1
22	ВИРУС КРАСНУХИ	AAQSFTGVVYGTHTT	905	BAA28178.1
23	ВИРУС КРАСНУХИ	ACEVEPAFGHSDAAC	906	BAA28178.1
24	ВИРУС КРАСНУХИ	ACTFWAVNAYSSGGY	907	BAA28178.1
25	ВИРУС	ADDPLLR	908	BAA1989

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КРАСНУХИ			3.1
26	ВИРУС КРАСНУХИ	ADDPLLRT	909	CAJ88851.1
27	ШТАММ ВИРУСА КОРИ EDMONSTON	AEMICDIDTYIVEAG	910	P04851.1
28	ШТАММ ВИРУСА КОРИ EDMONSTON	AEMICDIDTYIVEAGLASFI	911	P04851.1
29	ШТАММ ВИРУСА КОРИ EDMONSTON	AEPLLSC	912	P04851.1
30	BORDETELLA PERTUSSIS	AFVSTSSSRRYTEVY	913	CAD44970.1
31	BORDETELLA PERTUSSIS	AGFIYRETFCITTIYKTGQPAADHYYSKVTA	914	P04979.1
32	ВИРУС КРАСНУХИ	AGLLACCAKCLYYLR	915	BAA28178.1
33	ВИРУС КРАСНУХИ	AHTTSDPWHPPG	916	BAA19893.1
34	ШТАММ ВИРУСА КОРИ EDMONSTON	AIAKLEDAKELLESS	917	P69353.1
35	ШТАММ ВИРУСА КОРИ EDMONSTON	AIDNLRASLETTNQA	918	P69353.1
36	BORDETELLA PERTUSSIS	AKGVEFR	919	ACI16088.1
37	ШТАММ ВИРУСА КОРИ EDMONSTON	AKWAVPTTRTDDKLR	920	P08362.1
38	ШТАММ ВИРУСА КОРИ EDMONSTON	ALAEVLKKPV	921	ABO69699.1
39	ШТАММ ВИРУСА КОРИ EDMONSTON	ALGVINTLEWIPRFK	922	P08362.1
40	ШТАММ ВИРУСА КОРИ EDMONSTON	ALHQSMLNSQAIDNL	923	P69353.1
41	ШТАММ ВИРУСА КОРИ EDMONSTON	ALIGILSLFV	924	ABI54110.1
42	ВИРУС КРАСНУХИ	ALLNTPPPYQVSCGGESDRA	925	CAJ88851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
43	ШТАММ ВИРУСА КРАСНУХИ M33	ALLNTPPPYQVSCGGESDRASARV	926	CAJ88851.1
44	ВИРУС КРАСНУХИ	ALVEGLAPGGGNCHL	927	BAA28178.1
45	BORDETELLA PERTUSSIS	AMAAWSERAGEA	928	P04977.1
46	ШТАММ ВИРУСА КОРИ EDMONSTON	ANCASILCKCYTTGT	929	P69353.1
47	BORDETELLA PERTUSSIS	ANPNPYTSRRSV	930	P04977.1
48	ВИРУС КРАСНУХИ	APGPGEVW	931	CAJ88851.1
49	ВИРУС КРАСНУХИ	APLPPHTTERIETRSARHPWRIR	932	ABD64214.1
50	ВАКЦИННЫЙ ШТАММ ВИРУСА КРАСНУХИ RA27/3	APPMPPQPPRAHGQHYGHHHHQLPFLG	933	CAA33016.1
51	ШТАММ ВИРУСА КРАСНУХИ THERIEN	APPTLPQPPCAHGQHYGHHHHQLPFLG	934	P07566.1
52	ВИРУС КРАСНУХИ	APPTLPQPPRAHGQHYGHHHHQLPFLG	935	ABD64214.1
53	BORDETELLA PERTUSSIS	APQPGPQPPQPPQPQPEAPAPQ	936	P14283.3
54	ВИРУС КРАСНУХИ	AQLASYFNPGGSYYK	937	BAA28178.1
55	ШТАММ ВИРУСА КОРИ EDMONSTON	ARAAHLPTGTPLD	938	P04851.1
56	ШТАММ ВИРУСА КОРИ EDMONSTON	ARLVSEIAMHTTEDK	939	P04851.1
57	ШТАММ ВИРУСА КОРИ EDMONSTON	ARLVSEIAMHTTEDKISRAV	940	P04851.1
58	ШТАММ ВИРУСА КОРИ EDMONSTON-B	ASDVETAEGGEIHELLR	941	P03422.1
59	МОРБИЛЛИВИРУС КОРИ	ASDVETAEGGEIHELLRLQ	942	ABO69699.1
60	ШТАММ ВИРУСА КОРИ EDMONSTON-B	ASDVETAEGGEIHELLRLQSR	943	P03422.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
61	МОРБИЛЛИВИРУС КОРИ	ASDVETAEGGEIHKLLRLQ	944	AAA63285.1
62	ШТАММ ВИРУСА КРАСНУХИ M33	ASDVLPGHWLQG	945	NP_740663.1
63	ШТАММ ВИРУСА КОРИ EDMONSTON	ASELGITAEDARLVS	946	P04851.1
64	ШТАММ ВИРУСА КОРИ EDMONSTON	ASELGITAEDARLVSEIAMH	947	P04851.1
65	ШТАММ ВИРУСА КОРИ EDMONSTON	ASESSQDPQDSRR	948	P04851.1
66	ШТАММ ВИРУСА КОРИ EDMONSTON	ASILCKCYTTGTIIN	949	P69353.1
67	ВИРУС КРАСНУХИ	ASPVCQRHSPDCSRL	950	BAA28178.1
68	BORDETELLA PERTUSSIS	ASQARWTGATRA	951	BAF35031.1
69	ШТАММ ВИРУСА КРАСНУХИ THERIEN	ASYFNPGGSYYKQYHPTACEVEPAFGHS	952	P07566.1
70	ШТАММ ВИРУСА КОРИ EDMONSTON	ASYKVMTRSSHQSLV	953	P69353.1
71	BORDETELLA PERTUSSIS	ASYVKKPKEDVD	954	ACI16088.1
72	ШТАММ ВИРУСА КОРИ EDMONSTON	ATAAQITAGIALHQS	955	P69353.1
73	ВИРУС КРАСНУХИ	ATPERPRL	956	CAJ88851.1
74	ШТАММ ВИРУСА КОРИ EDMONSTON	AVCLGGLIGIPALIC	957	P69353.1
75	ШТАММ ВИРУСА КОРИ EDMONSTON	AVGPRQAQVSF	958	P04851.1
76	ВИРУС КРАСНУХИ	AVNAYSSGGYAQLAS	959	BAA28178.1
77	ВИРУС КРАСНУХИ	AVSETRQTWAEWAAA	960	BAA28178.1
78	ШТАММ ВИРУСА КОРИ	AVTAPDTAADSELRR	961	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
79	ШТАММ ВИРУСА КОРИ EDMONSTON	AVTAPDTAADSELRRWIKYT	962	P04851.1
80	BORDETELLA PERTUSSIS	AYGGIIKDAPPGAGFIYRETFC	963	P04979.1
81	ВИРУС КРАСНУХИ	CALPLAGLLACCAKC	964	BAA28178.1
82	ВИРУС КРАСНУХИ	CARIWNGTQRACTFW	965	BAA28178.1
83	ШТАММ ВИРУСА КОРИ EDMONSTON	CARTLVSGSFGNRFI	966	P69353.1
84	BORDETELLA PERTUSSIS	CASPYEGRYRDMYDALRBRLLY	967	SRC280066
85	BORDETELLA PERTUSSIS	CAVFVRSGQPVIGA	968	AAA83981.1
86	ВИРУС КРАСНУХИ	CCAKCLYYLRGAIAPR	969	BAA28178.1
87	ШТАММ ВИРУСА КОРИ EDMONSTON	CCRGRCNKKGEQVGM	970	P69353.1
88	ВИРУС КРАСНУХИ	CEIPTDVSCEGLGAW	971	BAA28178.1
89	BORDETELLA PERTUSSIS	CFGKDLKRPGSSPMEV	972	P0A3R5.1
90	МОРБИЛЛИВИРУС КОРИ	CFQQACKGKIQALCE	973	P06830.1
91	МОРБИЛЛИВИРУС КОРИ	CFQQACKGKIQALCENPEWAPLKDNRIPS	974	AAR89413.1
92	ВИРУС КРАСНУХИ	CGGESDRASARVIDP	975	BAA28178.1
93	BORDETELLA PERTUSSIS	CITTIYKTGQPAADHYYSKVTA	976	P04979.1
94	МОРБИЛЛИВИРУС КОРИ	CKGKIQALCENPEWA	977	AAR89413.1
95	ШТАММ ВИРУСА КОРИ EDMONSTON	CKPWQESRKNKAQ	978	P04851.1
96	ШТАММ ВИРУСА КОРИ EDMONSTON	CNKKGEQVGMSRPGL	979	P69353.1
97	ВИРУС КРАСНУХИ	CNVTTEHPFCNTPHG	980	BAA28178.1
98	BORDETELLA	CQVGSSNSAF	981	P04977.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	PERTUSSIS			
99	МОРБИЛЛИВИРУС КОРИ	CSGPTTIRGQFS	982	P08362.1
100	BORDETELLA PERTUSSIS	CTSPYDGKYWSMYSRL	983	AAA83981.1
101	ШТАММ ВИРУСА КОРИ EDMONSTON	CVLADSESGGHITHS	984	P08362.1
102	МОРБИЛЛИВИРУС КОРИ	CYTTGTIINQDPDKILTYIAADHC	985	AAF02706.1
103	ВИРУС КРАСНУХИ	DADDPLLR	986	CAJ88851.1
104	ШТАММ ВИРУСА КОРИ EDMONSTON	DARAAHLPTGTPLDI	987	P04851.1
105	ШТАММ ВИРУСА КОРИ EDMONSTON	DARAAHLPTGTPLDIDTASE	988	P04851.1
106	ШТАММ ВИРУСА КОРИ EDMONSTON	DCHAPTYLPAEVDGD	989	P08362.1
107	ВИРУС КРАСНУХИ	DCSRLVGATPERPRL	990	BAA28178.1
108	ШТАММ ВИРУСА КОРИ EDMONSTON	DDKLRMETCFQQACK	991	P08362.1
109	ВИРУС КРАСНУХИ	DDPLLRTA	992	CAJ88851.1
110	ВИРУС КРАСНУХИ	DDPLLRTAPGPGEVW	993	BAA28178.1
111	BORDETELLA PERTUSSIS	DDPPATVYRYD	994	P04977.1
112	BORDETELLA PERTUSSIS	DDPPATVYRYDSRPPED	995	CAD44970.1
113	BORDETELLA PERTUSSIS	DDPPATVYRYDSRPPEDV	996	ACI04548.1
114	МОРБИЛЛИВИРУС КОРИ	DEVGLRTPQRFTDLV	997	P06830.1
115	ШТАММ ВИРУСА КОРИ EDMONSTON	DHCPVVEVNGVTIQV	998	P69353.1
116	ШТАММ ВИРУСА КОРИ EDMONSTON	DIDTASESSQDPQ	999	P04851.1
117	ШТАММ ВИРУСА	DINKVLEKLGYSGGD	1000	P69353.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
118	BORDETELLA PERTUSSIS	DLIAYKQ	1001	ACI16088.1
119	ШТАММ ВИРУСА КОРИ EDMONSTON	DLIGQKLGLKLLRYY	1002	P69353.1
120	ВИРУС КРАСНУХИ	DLQKALEAQSRALRAELAA	1003	P07566.1
121	МОРБИЛЛИВИРУС КОРИ	DLQYVLATYDTSRVE	1004	P06830.1
122	ШТАММ ВИРУСА КОРИ EDMONSTON	DLSLRRFMV	1005	P04851.1
123	ШТАММ ВИРУСА КОРИ EDMONSTON	DLSNCMVALGELKLA	1006	P08362.1
124	ВИРУС КРАСНУХИ	DLVEYIMNYTGNQQSRWGLGSPNC	1007	CAJ88851.1
125	МОРБИЛЛИВИРУС КОРИ	DLVKFISDKIKFLNP	1008	AAR89413.1
126	МОРБИЛЛИВИРУС КОРИ	DLVKFISTKIKFLNP	1009	SRC280117
127	ШТАММ ВИРУСА КОРИ EDMONSTON	DLYKSNHNNV	1010	P08362.1
128	BORDETELLA PERTUSSIS	DNVLDHLTGR	1011	ACI04548.1
129	BORDETELLA PERTUSSIS	DNVLDHLTGRSC	1012	P04977.1
130	BORDETELLA PERTUSSIS	DNVLDHLTGRSCQ	1013	P04977.1
131	МОРБИЛЛИВИРУС КОРИ	DPDKILTYIAA	1014	AAF02706.1
132	ШТАММ ВИРУСА КРАСНУХИ THERIEN	DPGDLVEYIMNYTGNQQSR	1015	P07566.1
133	ВИРУС КРАСНУХИ	DPLLRTAP	1016	CAJ88851.1
134	ВИРУС КРАСНУХИ	DPLLRTAPGPGEVWVTPVIGSQ	1017	CAJ88851.1
135	ШТАММ ВИРУСА КОРИ EDMONSTON	DPQDSRRSAEPLL	1018	P04851.1
136	ШТАММ ВИРУСА	DPVIDRLYLSSHRGV	1019	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
137	ШТАММ ВИРУСА КОРИ EDMONSTON	DQILRSMKGLSSTSI	1020	P69353.1
138	МОРБИЛЛИВИРУС КОРИ	DQYCADVAAEELMNA	1021	P06830.1
139	ШТАММ ВИРУСА КОРИ EDMONSTON	DSESGGHITH	1022	P08362.1
140	ШТАММ ВИРУСА КОРИ EDMONSTON	DTASESSQDPQDS	1023	P04851.1
141	ВИРУС КРАСНУХИ	DTVMSVFALASYVQH	1024	BAA28178.1
142	BORDETELLA PERTUSSIS	DVFQNGFTAWGNND	1025	P04977.1
143	ВИРУС КРАСНУХИ	DVGAVPPGKFVTAAL	1026	BAA28178.1
144	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	DVNKSKTHISVNGRKI	1027	CAE11230.1
145	ВИРУС КРАСНУХИ	DVSCEGLGAWVPAAP	1028	BAA28178.1
146	ШТАММ ВИРУСА КРАСНУХИ THERIEN	DWASPVCQRHSPDCSRLVGATPERPRL	1029	P07566.1
147	ШТАММ ВИРУСА КОРИ EDMONSTON	EARESYRETGPSR	1030	P04851.1
148	BORDETELLA PERTUSSIS	EAVEAERAGRGTG	1031	ACI04548.1
149	ШТАММ ВИРУСА КОРИ EDMONSTON	EDAKELLESSDQILR	1032	P69353.1
150	ШТАММ ВИРУСА КОРИ EDMONSTON	EDRRVKQSRGEAR	1033	P04851.1
151	ШТАММ ВИРУСА КОРИ EDMONSTON	EDSITIPYQGSGKGV	1034	P08362.1
152	ВИРУС КРАСНУХИ	EEAFTYLCTAPGCAT	1035	BAA28178.1
153	ШТАММ ВИРУСА КОРИ EDMONSTON	EGFNMILGTILAQIW	1036	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
154	ШТАММ ВИРУСА КОРИ EDMONSTON	EGFNMILGTILAQIWVLLAK	1037	P04851.1
155	ВИРУС КРАСНУХИ	EHPFCNTPHGQLEVQ	1038	BAA28178.1
156	ШТАММ ВИРУСА КОРИ EDMONSTON	EISDIEVQDPEGFNM	1039	P04851.1
157	ШТАММ ВИРУСА КОРИ EDMONSTON	EISDIEVQDPEGFNMILGTI	1040	P04851.1
158	ШТАММ ВИРУСА КОРИ EDMONSTON	EKPNLSSKRSE	1041	P08362.1
159	ШТАММ ВИРУСА КОРИ EDMONSTON	ELKLAALCHGEDSIT	1042	P08362.1
160	МОРБИЛЛИВИРУС КОРИ	ELMNALVNSTLLETR	1043	P06830.1
161	ШТАММ ВИРУСА КОРИ EDMONSTON	ELPRL	1044	P04851.1
162	МОРБИЛЛИВИРУС КОРИ	ENPEWAPLKDNRIPSYGVLSVDL	1045	AAR89413.1
163	ШТАММ ВИРУСА КОРИ EDMONSTON	EPIRDALNAMTQNIR	1046	P69353.1
164	ШТАММ ВИРУСА КОРИ EDMONSTON	EQVGMSRPGLKPDLT	1047	P69353.1
165	ВИРУС КРАСНУХИ	ERPRLRLV	1048	CAJ88851.1
166	ВИРУС КРАСНУХИ	ERPRLRLVDADDPLL	1049	BAA28178.1
167	ВИРУС КОРИ CAM/RB	ESPGQLIQRITDDPDVS	1050	P04851.1
168	ШТАММ ВИРУСА КОРИ EDMONSTON	ESRGIKARITHVDTE	1051	P69353.1
169	ШТАММ ВИРУСА КОРИ EDMONSTON	ESSCTFMPEGTVCSQ	1052	P69353.1
170	ШТАММ ВИРУСА КОРИ EDMONSTON	ESSQDPQDSRRSA	1053	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
171	ШТАММ ВИРУСА КОРИ EDMONSTON	ETRTTNQFLAVSKGN	1054	P08362.1
172	ШТАММ ВИРУСА КОРИ EDMONSTON	EVDGDVKLSSNLVIL	1055	P08362.1
173	ШТАММ ВИРУСА КОРИ EDMONSTON-B	EVNGVTIQV	1056	P26031.1
174	ВИРУС КРАСНУХИ	EVWVTPVI	1057	CAJ88851.1
175	ВИРУС КРАСНУХИ	EVWVTPVIGSQA	1058	BAA19893.1
176	ВИРУС КРАСНУХИ	EWAAAHWWQLTLGAT	1059	BAA28178.1
177	МОРБИЛЛИВИРУС КОРИ	EWIPRFKVSPYLFTV	1060	P06830.1
178	BORDETELLA PERTUSSIS	FEYVDTYGDNAG	1061	P04977.1
179	МОРБИЛЛИВИРУС КОРИ	FGPLITHGSGMDLYK	1062	P06830.1
180	ШТАММ ВИРУСА КОРИ EDMONSTON	FIFDALAEV	1063	ABK40531.1
181	ШТАММ ВИРУСА КОРИ EDMONSTON	FISDKIKFL	1064	P08362.1
182	ШТАММ ВИРУСА КОРИ EDMONSTON	FKRNKDKPPITSGSG	1065	P04851.1
183	ШТАММ ВИРУСА КОРИ EDMONSTON	FKRNKDKPPITSGSGGAIRG	1066	P04851.1
184	ВИРУС КРАСНУХИ	FKTVRPVALPRTLAP	1067	BAA28178.1
185	ШТАММ ВИРУСА КОРИ EDMONSTON	FLMDRHIIV	1068	ABK40531.1
186	ШТАММ ВИРУСА КОРИ EDMONSTON	FMAVLLTLQTPTGQI	1069	P69353.1
187	ШТАММ ВИРУСА КОРИ EDMONSTON	FMPEGTVCSQNALYP	1070	P69353.1
188	МОРБИЛЛИВИРУ	FMYMSLLGV	1071	AAN0980

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	С КОРИ			4.1
189	ШТАММ ВИРУСА КОРИ EDMONSTON	FNVPIKEAGEDCHAP	1072	P08362.1
190	МОРБИЛЛИВИРУС КОРИ	FRDLTWCINPPERIK	1073	AAC35876.2
191	ШТАММ ВИРУСА КОРИ EDMONSTON	FSHDDPISSDQSRFG	1074	P04851.1
192	ШТАММ ВИРУСА КОРИ EDMONSTON	FSHDDPISSDQSRFGWFENK	1075	P04851.1
193	МОРБИЛЛИВИРУС КОРИ	FTDLVKFISDKIKFL	1076	P06830.1
194	МОРБИЛЛИВИРУС КОРИ	FTWDQKLWCRHFCVL	1077	P06830.1
195	BORDETELLA PERTUSSIS	FVRDGQSVIGACASPYEGRYRDLYDALRRLLY	1078	SRC280066
196	BORDETELLA PERTUSSIS	FVRSGQPVIGACTSPYDGKYWSILYSRLRKMLY	1079	SRC280066
197	МОРБИЛЛИВИРУС КОРИ	FYKDNPHPKGSRIVI	1080	P06830.1
198	BORDETELLA PERTUSSIS	GAASSYFEYVDTYG	1081	ACI04548.1
199	BORDETELLA PERTUSSIS	GAFDLKTTFCIMTTRNTGQPA	1082	AAA83981.1
200	ШТАММ ВИРУСА КОРИ EDMONSTON	GALIGILSLFVESPG	1083	P04851.1
201	ШТАММ ВИРУСА КОРИ EDMONSTON	GALIGILSLFVESPGQLIQR	1084	P04851.1
202	ВИРУС КРАСНУХИ	GATPERPR	1085	CAJ88851.1
203	BORDETELLA PERTUSSIS	GAYGRCPNGTRALTVAELRGNAEL	1086	P04979.1
204	ВИРУС КРАСНУХИ	GCFAPWDLEATGACI	1087	BAA28178.1
205	МОРБИЛЛИВИРУС КОРИ	GDINKVLEKLGYS	1088	BAB60865.1
206	МОРБИЛЛИВИРУС КОРИ	GDINKVLEKLGYSGGDLLG	1089	AAL29688.1
207	ВИРУС КРАСНУХИ	GDLRAVHHRPVPA	1090	CAA28880.1
208	ВИРУС	GDLVEYIMNYTGNQQ	1091	BAA2817

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КРАСНУХИ			8.1
209	ШТАММ ВИРУСА КОРИ EDMONSTON	GDSSITTRSRLLDRL	1092	P04851.1
210	ШТАММ ВИРУСА КОРИ EDMONSTON	GDSSITTRSRLLDRLVRLIG	1093	P04851.1
211	МОРБИЛЛИВИРУС КОРИ	GEDCHAPTYLPAEVD	1094	P06830.1
212	ШТАММ ВИРУСА КОРИ EDMONSTON	GELSTLESLMNLYQQ	1095	P04851.1
213	ШТАММ ВИРУСА КОРИ EDMONSTON	GELSTLESLMNLYQQMGKPA	1096	P04851.1
214	РУБУЛАВИРУС ПАРОТИТА	GEQARYLALLEA	1097	P21186.1
215	ВИРУС КРАСНУХИ	GEVWVT	1098	BAA19893.1
216	ВИРУС КРАСНУХИ	GEVWVTPV	1099	CAJ88851.1
217	ВИРУС КРАСНУХИ	GEVWVTPVIGSQAR	1100	BAA19893.1
218	BORDETELLA PERTUSSIS	GEYGGVIKDGTPGGA	1101	AAA83981.1
219	ВИРУС КРАСНУХИ	GFLSGVGPMRLRHGADT	1102	SRC265968
220	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GFRASDVETAEGGEIHELLRLQ	1103	P03422.1
221	BORDETELLA PERTUSSIS	GGAVPGGAVPGGAVPGGFGPGGFGP	1104	P14283.3
222	BORDETELLA PERTUSSIS	GGAVPGGAVPGGFGPGGFGPGGFGP	1105	CAA09475.1
223	BORDETELLA PERTUSSIS	GGAVPGGAVPGGFGPGGFGPGGFGPGGFGP	1106	CAA09474.1
224	МОРБИЛЛИВИРУС КОРИ	GGHITHSGMVGMGVS	1107	P06830.1
225	МОРБИЛЛИВИРУС КОРИ	GILESRGIKARITHVDTESY	1108	P26032.1
226	BORDETELLA PERTUSSIS	GITGETTTTEYSNARYV	1109	CAD44970.1
227	ШТАММ ВИРУСА КОРИ EDMONSTON	GKEDRRVKQSRGE	1110	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
228	BORDETELLA PERTUSSIS	GKVTNGS	1111	ACI16088.1
229	ВИРУС КРАСНУХИ	GLGAWVPAAPCARIW	1112	BAA28178.1
230	ШТАММ ВИРУСА КОРИ EDMONSTON	GLIGIPALICCCRGR	1113	P69353.1
231	ШТАММ ВИРУСА КРАСНУХИ THERIEN	GLLACCAKCLYYLRGAIAPR	1114	P07566.1
232	МОРБИЛЛИВИРУС КОРИ	GLLAIAGIRLHRAAI	1115	P06830.1
233	ВИРУС КРАСНУХИ	GLQPRADMAAPPTLPQ	1116	NP_740663.1
234	МОРБИЛЛИВИРУС КОРИ	GMGVSCTVTREDGTNRR	1117	AAR89413.1
235	МОРБИЛЛИВИРУС КОРИ	GMYGGTYLVEKP	1118	AAR89413.1
236	BORDETELLA PERTUSSIS	GNAELQTYLRQITPGWSIYGLYDGTYLG	1119	P04979.1
237	ВИРУС КРАСНУХИ	GNCHLTVNGEDVGAV	1120	BAA28178.1
238	BORDETELLA PERTUSSIS	GNNDNVLDHLTGR	1121	P04977.1
239	BORDETELLA PERTUSSIS	GNNDNVLDHLTGRSC	1122	P04977.1
240	ШТАММ ВИРУСА КОРИ EDMONSTON	GNRFILSQGNLIANC	1123	P69353.1
241	ВИРУС КРАСНУХИ	GNRGRGQRRDWSRAPPPPEERQETRSQTPAPKPS	1124	P07566.1
242	ВИРУС КРАСНУХИ	GPGEVWVT	1125	CAJ88851.1
243	ШТАММ ВИРУСА КРАСНУХИ THERIEN	GPMRLRHGADTRCGRLI	1126	P07566.1
244	BORDETELLA PERTUSSIS	GPNHTKV	1127	ACI16083.1
245	ШТАММ ВИРУСА КОРИ HALLE	GPRQAQVSF	1128	P10050.1
246	ШТАММ ВИРУСА КОРИ EDMONSTON	GPRQAQVSFLQGDQS	1129	P04851.1
247	ШТАММ ВИРУСА КОРИ	GPRQAQVSFLQGDQSENELP	1130	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
248	МОРБИЛЛИВИРУС КОРИ	GRGYNVSSIVTMTSQ	1131	P06830.1
249	BORDETELLA PERTUSSIS	GRTPFII	1132	ACI16083.1
250	ВИРУС КРАСНУХИ	GSPNCHGPDWASPVC	1133	BAA28178.1
251	ВИРУС КРАСНУХИ	GSQARKCGLHIRAGP	1134	BAA28178.1
252	BORDETELLA PERTUSSIS	GSSNSAFVSTSSSRR	1135	P04977.1
253	ШТАММ ВИРУСА КОРИ EDMONSTON	GSTKSCARTLVSGSF	1136	P69353.1
254	ВИРУС КРАСНУХИ	GSYYKQYHPTACEVE	1137	BAA28178.1
255	ВИРУС КРАСНУХИ	GTHTTAVSETRQTWA	1138	BAA28178.1
256	ШТАММ ВИРУСА КОРИ EDMONSTON	GTIINQDPDKILTYI	1139	P69353.1
257	BORDETELLA PERTUSSIS	GTLVRMAPVIG	1140	ADA85124.1
258	ШТАММ ВИРУСА КОРИ EDMONSTON	GTPLDIDTASESS	1141	P04851.1
259	BORDETELLA PERTUSSIS	GTYLGQAYGGIIKDAPPGAGFIYRETFC	1142	P04979.1
260	BORDETELLA PERTUSSIS	GVATKGLGVHAKSSDWG	1143	P15318.2
261	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GVLLPTIPGKLDVNKSKTHI	1144	AAV70486.1
262	МОРБИЛЛИВИРУС КОРИ	GVLSVDLSLTVELKI	1145	P06830.1
263	МОРБИЛЛИВИРУС КОРИ	GVPIELQVECFTWDQ	1146	P06830.1
264	ШТАММ ВИРУСА КОРИ EDMONSTON	GVSCTVTREDGTNRR	1147	P08362.1
265	ШТАММ ВИРУСА КОРИ EDMONSTON	GVSYNIGSQEWYTTV	1148	P69353.1
266	ШТАММ ВИРУСА КОРИ EDMONSTON	GYNVSSIVTMTSQGM	1149	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
267	МОРБИЛЛИВИРУС КОРИ	HFCVLADSESGGHIT	1150	P06830.1
268	МОРБИЛЛИВИРУС КОРИ	HGEDSITIPYQGSGK	1151	P06830.1
269	ВИРУС КРАСНУХИ	HGPDWASP	1152	BAA19893.1
270	ВИРУС КРАСНУХИ	HGPDWASPVCQRHSP	1153	BAA28178.1
271	ВИРУС КРАСНУХИ	HGPDWASPVCQRHSPDCSRLVG	1154	CAJ88851.1
272	ШТАММ ВИРУСА КРАСНУХИ M33	HGPDWASPVCQRHSPDCSRLVGATPERPRLRLV	1155	CAJ88851.1
273	ШТАММ ВИРУСА КОРИ EDMONSTON	HITHSGMEGMGVSCT	1156	P08362.1
274	МОРБИЛЛИВИРУС КОРИ	HKSLSTNLDVTNSIE	1157	P06830.1
275	ШТАММ ВИРУСА КОРИ EDMONSTON	HLMIDRPYV	1158	P08362.1
276	ШТАММ ВИРУСА КОРИ EDMONSTON	HLPTGTPLDIDTA	1159	P04851.1
277	ШТАММ ВИРУСА КОРИ EDMONSTON-B	HLPTGTPLDIDTATESSQDPQDSR	1160	Q77M43.1
278	МОРБИЛЛИВИРУС КОРИ	HMTNYLEQPVSNDLS	1161	P06830.1
279	ШТАММ ВИРУСА КОРИ EDMONSTON	HQSLVIKLMPNITLL	1162	P69353.1
280	МОРБИЛЛИВИРУС КОРИ	HRAAIYTAEIHKSLS	1163	P06830.1
281	BORDETELLA PERTUSSIS	HRMQEAVEAERAGRGTGH	1164	P04977.1
282	ШТАММ ВИРУСА КОРИ EDMONSTON	HVDTESYFIVLSIAY	1165	P69353.1
283	ШТАММ ВИРУСА КОРИ EDMONSTON	HWGNLSKIGVVGIGS	1166	P69353.1
284	ВИРУС КРАСНУХИ	HWWQLTLGATCALPL	1167	BAA28178.1
285	ВИРУС КРАСНУХИ	HYRNASDVLPGHWLQGGWGCYNL	1168	NP_740663.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
286	ШТАММ ВИРУСА КОРИ EDMONSTON	IDLGPPISLERLDVG	1169	P69353.1
287	ШТАММ ВИРУСА КОРИ EDMONSTON	IEAIRQAGQEMILAV	1170	P69353.1
288	ШТАММ ВИРУСА КРАСНУХИ M33	IETRSARHP	1171	CAA28880.1
289	ШТАММ ВИРУСА КОРИ EDMONSTON	IGSQEWYTTVPKYVA	1172	P69353.1
290	ШТАММ ВИРУСА КОРИ EDMONSTON	IKGVIVHRLEGVSYN	1173	P69353.1
291	ШТАММ ВИРУСА КОРИ EDMONSTON	IKHIIIVPIPGDSSI	1174	P04851.1
292	ШТАММ ВИРУСА КОРИ EDMONSTON	IKHIIIVPIPGDSSITTRSR	1175	P04851.1
293	BORDETELLA PERTUSSIS	IKLKDCP	1176	ACI16083.1
294	ШТАММ ВИРУСА КОРИ EDMONSTON	IKLMPNITLLNNCTR	1177	P69353.1
295	МОРБИЛЛИВИРУС КОРИ	ILLERLDVGT	1178	AAF85664.1
296	МОРБИЛЛИВИРУС КОРИ	ILPGQDLQYV	1179	P08362.1
297	ШТАММ ВИРУСА КОРИ EDMONSTON	ILTYIAADHCPVVEV	1180	P69353.1
298	МОРБИЛЛИВИРУС КОРИ	INQDPDKILTY	1181	AAL29688.1
299	ШТАММ ВИРУСА КОРИ EDMONSTON	IPRFKVSPYLFNVPI	1182	P08362.1
300	ШТАММ ВИРУСА КОРИ EDMONSTON	IQALSYALGGDINKV	1183	P69353.1
301	ВИРУС КРАСНУХИ	IRAGPYGHATVEMPE	1184	BAA28178.1
302	BORDETELLA PERTUSSIS	IRMGTDK	1185	ACI16088.1
303	ШТАММ ВИРУСА	ISNFDESSCTFMPEG	1186	P69353.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
304	ШТАММ ВИРУСА КОРИ EDMONSTON	ITAGIALHQSMLNSQ	1187	P69353.1
305	ШТАММ ВИРУСА КОРИ EDMONSTON	ITDDPDVSIRLLEVV	1188	P04851.1
306	ШТАММ ВИРУСА КОРИ EDMONSTON	ITDDPDVSIRLLEVVQSDQS	1189	P04851.1
307	BORDETELLA PERTUSSIS	ITTYV	1190	ACI16083.1
308	ШТАММ ВИРУСА КОРИ EDMONSTON	IVEAGLASFILTIKF	1191	P04851.1
309	ШТАММ ВИРУСА КОРИ EDMONSTON	IVEAGLASFILTIKFGIETM	1192	P04851.1
310	ШТАММ ВИРУСА КОРИ EDMONSTON	IVEAGLASFILTIKFGIETMYPALG	1193	P04851.1
311	ВИРУС КРАСНУХИ	KALEAQSRALRAELAA	1194	P07566.1
312	ШТАММ ВИРУСА КОРИ EDMONSTON	KARITHVDTESYFIV	1195	P69353.1
313	ВИРУС КРАСНУХИ	KCGLHIRAGPYGHAT	1196	BAA28178.1
314	ШТАММ ВИРУСА КОРИ EDMONSTON	KCYTTGTIINQDPDK	1197	P69353.1
315	ШТАММ ВИРУСА КОРИ EDMONSTON	KDNPHPKGSR	1198	P08362.1
316	ШТАММ ВИРУСА КОРИ EDMONSTON	KDNRIPSYGVLSVDL	1199	P08362.1
317	МОРБИЛЛИВИРУС КОРИ	KFLNPDREYDFRDLT	1200	AAC35876.2
318	ВИРУС КРАСНУХИ	KFVTAALLN	1201	BAA28178.1
319	ШТАММ ВИРУСА КОРИ EDMONSTON	KGNCSGPTTIR	1202	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
320	МОРБИЛЛИВИРУС КОРИ	KIKFLNPDREYDFRD	1203	P06830.1
321	ВИРУС КРАСНУХИ	KIVDGGCFAPWDLEA	1204	BAA28178.1
322	ШТАММ ВИРУСА КОРИ EDMONSTON	KLGLKLLRYYTEILS	1205	P69353.1
323	ШТАММ ВИРУСА КОРИ EDMONSTON	KLGVWKSPTDMQSWV	1206	P08362.1
324	BORDETELLA PERTUSSIS	KLKECPQ	1207	ACI16088.1
325	МОРБИЛЛИВИРУС КОРИ	KLLRYYTEI	1208	P26031.1
326	ШТАММ ВИРУСА КОРИ EDMONSTON	KLMPFSGDFV	1209	ABK40531.1
327	МОРБИЛЛИВИРУС КОРИ	KLMPNITLL	1210	P26031.1
328	МОРБИЛЛИВИРУС КОРИ	KLRMETCFQQACKGKIQALCENPEWA	1211	AAR89413.1
329	МОРБИЛЛИВИРУС КОРИ	KLWCRHFCV	1212	P08362.1
330	ШТАММ ВИРУСА КОРИ EDMONSTON	KLWCRHFCVL	1213	P08362.1
331	МОРБИЛЛИВИРУС КОРИ	KLWESPQEI	1214	BAB60863.1
332	ШТАММ ВИРУСА КОРИ EDMONSTON	KMSSAVGFV	1215	ABO69699.1
333	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KMSSAVGFVPDTGPASR	1216	P03422.1
334	BORDETELLA PERTUSSIS	KMVYATN	1217	ACI16083.1
335	ШТАММ ВИРУСА КОРИ EDMONSTON	KPDLTGTSKSYVRSL	1218	P69353.1
336	МОРБИЛЛИВИРУС КОРИ	KPNLSSKRSELSQLS	1219	P08362.1
337	ШТАММ ВИРУСА КОРИ EDMONSTON	KPNLSSKRSELSQLSMYRVF	1220	P08362.1
338	ШТАММ ВИРУСА	KQSRGEARESYRETG	1221	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
339	ШТАММ ВИРУСА КОРИ EDMONSTON	KQSRGEARESYRETGPSRAS	1222	P04851.1
340	ШТАММ ВИРУСА КОРИ EDMONSTON	KRFAGVVLAGAALGV	1223	P69353.1
341	ШТАММ ВИРУСА КОРИ EDMONSTON	KRTPGNKPRIAEMIC	1224	P04851.1
342	ШТАММ ВИРУСА КОРИ EDMONSTON	KRTPGNKPRIAEMICDIDTY	1225	P04851.1
343	МОРБИЛЛИВИРУС КОРИ	KSNHNNVYWLTIPPMKNLALGVINTL	1226	AAR89413.1
344	МОРБИЛЛИВИРУС КОРИ	KVSPYLFNV	1227	P08362.1
345	BORDETELLA PERTUSSIS	KVVQLPKISKNALKANG	1228	ACI16083.1
346	BORDETELLA PERTUSSIS	KVVQLPKISKNALRNDG	1229	ACI16087.1
347	BORDETELLA PERTUSSIS	LAHRRIPPENIR	1230	P04977.1
348	BORDETELLA PERTUSSIS	LALALWAGFALS	1231	P11092.1
349	ВИРУС КРАСНУХИ	LAPGGGNCHLTVNGE	1232	BAA28178.1
350	ШТАММ ВИРУСА КОРИ EDMONSTON	LAQIWVLLAKAVTAP	1233	P04851.1
351	ШТАММ ВИРУСА КОРИ EDMONSTON	LAQIWVLLAKAVTAPDTAAD	1234	P04851.1
352	ВИРУС КРАСНУХИ	LASYFNPGGSYYKQYHPTACEVEPAFGHS	1235	BAA19893.1
353	МОРБИЛЛИВИРУС КОРИ	LAVSKGNCSGPTTIR	1236	P06830.1
354	ШТАММ ВИРУСА КОРИ EDMONSTON	LCENPEWAPLKDNRI	1237	P08362.1
355	МОРБИЛЛИВИРУС КОРИ	LDRLVRLIG	1238	ABI54110.1
356	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	LEEEGVTPL	1239	P33120.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
357	BORDETELLA PERTUSSIS	LEHRMQEAVEAERAGRGTGHFI	1240	CAD44970.1
358	ШТАММ ВИРУСА КОРИ EDMONSTON	LEKLGYSGGDLLGIL	1241	P69353.1
359	ШТАММ ВИРУСА КОРИ EDMONSTON	LEQPVSNDLS	1242	P08362.1
360	ШТАММ ВИРУСА КОРИ EDMONSTON	LERKWLDVVRNIIAE	1243	P04851.1
361	ШТАММ ВИРУСА КОРИ EDMONSTON	LERKWLDVVRNIIAEDLSLR	1244	P04851.1
362	ШТАММ ВИРУСА КОРИ EDMONSTON	LFGPSLRDPISAEIS	1245	P69353.1
363	МОРБИЛЛИВИРУС КОРИ	LGELKLAALCHGEDS	1246	P06830.1
364	ШТАММ ВИРУСА КОРИ EDMONSTON	LGGKEDRRVKQSR	1247	P04851.1
365	ВИРУС КРАСНУХИ	LGHDGHHGGTLRVGQHHRNASDVL	1248	ABD64214.1
366	ШТАММ ВИРУСА КРАСНУХИ THERIEN	LGSPNCHGPDWASPVCQRHS	1249	P07566.1
367	ШТАММ ВИРУСА КРАСНУХИ THERIEN	LGSPNCHGPDWASPVCQRHSPDCSRLV	1250	P07566.1
368	ВИРУС КРАСНУХИ	LHDPDTEAPTEACVTSWL	1251	ABD64214.1
369	ШТАММ ВИРУСА КОРИ EDMONSTON	LIANCASILCKCYTT	1252	P69353.1
370	МОРБИЛЛИВИРУС КОРИ	LIGLLAIAGIRLHRAAIYTAEIHK	1253	AAR89413.1
371	ШТАММ ВИРУСА КОРИ EDMONSTON	LIPSMNQLSCDLIGQ	1254	P69353.1
372	ШТАММ ВИРУСА КОРИ EDMONSTON	LKIKIASGFGPLITH	1255	P08362.1
373	МОРБИЛЛИВИРУС КОРИ	LKIKIASGFGPLITHGSGMDLYK	1256	AAR89413.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
374	BORDETELLA PERTUSSIS	LKLYFEP	1257	ACI16088.1
375	ШТАММ ВИРУСА КОРИ EDMONSTON	LLAVLFVMFL	1258	P08362.1
376	ШТАММ ВИРУСА КОРИ EDMONSTON	LLDRLVRLIGNPDVS	1259	P04851.1
377	ШТАММ ВИРУСА КОРИ EDMONSTON	LLDRLVRLIGNPDVSGPKLT	1260	P04851.1
378	ШТАММ ВИРУСА КОРИ EDMONSTON	LLESSDQILRSMKGL	1261	P69353.1
379	МОРБИЛЛИВИРУС КОРИ	LLETRTTNQFLAVSK	1262	P06830.1
380	ШТАММ ВИРУСА КОРИ EDMONSTON	LLEVVQSDQSQSGLT	1263	P04851.1
381	ШТАММ ВИРУСА КОРИ EDMONSTON	LLEVVQSDQSQSGLTFASR	1264	P04851.1
382	ШТАММ ВИРУСА КОРИ EDMONSTON	LLEVVQSDQSQSGLTFASRG	1265	P04851.1
383	МОРБИЛЛИВИРУС КОРИ	LLGILESRGIKARIT	1266	AAL29688.1
384	ВИРУС КРАСНУХИ	LLRTAPGP	1267	CAJ88851.1
385	ШТАММ ВИРУСА КОРИ EDMONSTON	LLRYYTEILSLFGPS	1268	P69353.1
386	ШТАММ ВИРУСА КРАСНУХИ THERIEN	LLVPWVLIFMVCRRACRRRG	1269	P07566.1
387	ШТАММ ВИРУСА КОРИ EDMONSTON	LLWRSRCKIV	1270	ABK40528.1
388	МОРБИЛЛИВИРУС КОРИ	LLWSYAMGV	1271	P04851.1
389	ШТАММ ВИРУСА КОРИ EDMONSTON	LLWSYAMGVGVELEN	1272	P04851.1
390	ШТАММ ВИРУСА КОРИ	LLWSYAMGVGVELENSMGGL	1273	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
391	МОРБИЛЛИВИРУС КОРИ	LMIDRPYVL	1274	P08362.1
392	ШТАММ ВИРУСА КОРИ EDMONSTON	LNAMTQNIRPVQSVA	1275	P69353.1
393	ВИРУС КРАСНУХИ	LNTPPPYQVSCGGES	1276	BAA28178.1
394	ШТАММ ВИРУСА КОРИ EDMONSTON	LRDPISAEISIQALS	1277	P69353.1
395	BORDETELLA PERTUSSIS	LRGSGDLQEYLRHVTR	1278	AAA83981.1
396	ВИРУС КРАСНУХИ	LRLVDADD	1279	CAJ88851.1
397	ВИРУС КРАСНУХИ	LRLVDADDPLLR	1280	BAA19893.1
398	ВИРУС КРАСНУХИ	LRLVDADDPLLRTAPGPGEVWVTPVIGSQAR	1281	BAA19893.1
399	BORDETELLA PERTUSSIS	LRRLLYMIYMSGLAVRVHVSKEEQYYDY	1282	P04979.1
400	ВИРУС КРАСНУХИ	LRTAPGPG	1283	CAJ88851.1
401	ВИРУС КРАСНУХИ	LRVGQHYRNASDVLPGHWLQ	1284	NP_740663.1
402	BORDETELLA PERTUSSIS	LRYLA	1285	ACI16088.1
403	ШТАММ ВИРУСА КОРИ EDMONSTON	LSCKPWQESRKNK	1286	P04851.1
404	МОРБИЛЛИВИРУС КОРИ	LSEIKGVIVHRLEGV	1287	AAL29688.1
405	ШТАММ ВИРУСА КОРИ EDMONSTON	LSIAYPTLSEIKGVI	1288	P69353.1
406	ШТАММ ВИРУСА КОРИ EDMONSTON	LSLLDLYLGRGYNVS	1289	P08362.1
407	ШТАММ ВИРУСА КОРИ EDMONSTON	LSQGNLIANCASILC	1290	P69353.1
408	МОРБИЛЛИВИРУС КОРИ	LSSHRGVIADNQAKW	1291	P06830.1
409	ШТАММ ВИРУСА КОРИ	LSVDLSLTVELKIKI	1292	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
410	BORDETELLA PERTUSSIS	LTGISICNPGSSLC	1293	AAA83981.1
411	ШТАММ ВИРУСА КОРИ EDMONSTON	LTIKFGIETMYPALG	1294	P04851.1
412	ШТАММ ВИРУСА КОРИ EDMONSTON	LTIKFGIETMYPALGLHEFA	1295	P04851.1
413	ШТАММ ВИРУСА КОРИ EDMONSTON	LTLQTPTGQIHWGNL	1296	P69353.1
414	ВИРУС КРАСНУХИ	LVDADDPL	1297	CAJ88851.1
415	ВИРУС КРАСНУХИ	LVDADDPLLR	1298	BAA19893.1
416	ШТАММ ВИРУСА КОРИ EDMONSTON	LVEKPNLSSKRSELS	1299	P08362.1
417	ВИРУС КРАСНУХИ	LVGATPE	1300	BAA19893.1
418	ВИРУС КРАСНУХИ	LVGATPER	1301	CAJ88851.1
419	МОРБИЛЛИВИРУС КОРИ	LVKLGVWKSPTGMQS	1302	P06830.1
420	ШТАММ ВИРУСА КОРИ EDMONSTON-B	LVSGSFGNRFILSQGNLI	1303	P26031.1
421	ШТАММ ВИРУСА КОРИ EDMONSTON	LYKSNHNNVYWLTIP	1304	P08362.1
422	ШТАММ ВИРУСА КОРИ EDMONSTON	LYPMSPLLQECLRGSTKSCARTLVS	1305	P69353.1
423	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEA	1306	P07566.1
424	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSR	1307	ABD64200.1
425	ШТАММ ВИРУСА КРАСНУХИ THERIEN	MASTTPITMEDLQKALEAQSRALRAELAA	1308	P07566.1
426	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSRALRAGLAA	1309	ABD64200.1
427	ВАКЦИННЫЙ ШТАММ ВИРУСА	MASTTPITMEDLQKALETQSRVLRAGLTA	1310	CAA33016.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КРАСНУХИ RA27/3			
428	ШТАММ ВИРУСА КОРИ EDMONSTON	MATLLRSLALFKRNK	1311	P04851.1
429	ШТАММ ВИРУСА КОРИ EDMONSTON	MATLLRSLALFKRNKDKPPI	1312	P04851.1
430	МОРБИЛЛИВИРУС КОРИ	MDLYKSNHNNVYWLT	1313	P06830.1
431	ВИРУС КРАСНУХИ	MEDLQKALEAQSRA	1314	P07566.1
432	ВИРУС КРАСНУХИ	MEDLQKALEAQSRALRAELAA	1315	P07566.1
433	ШТАММ ВИРУСА КОРИ EDMONSTON	MGLKVNVSAIFMAVL	1316	P69353.1
434	МОРБИЛЛИВИРУС КОРИ	MIDRPYVLLAVLFVM	1317	P06830.1
435	ШТАММ ВИРУСА КОРИ EDMONSTON	MILAVQGVQDYINNE	1318	P69353.1
436	ШТАММ ВИРУСА КОРИ EDMONSTON	MLNSQAIDNLRASLE	1319	P69353.1
437	ШТАММ ВИРУСА КОРИ EDMONSTON	MNALVNSTLLETRTT	1320	P08362.1
438	ВИРУС КРАСНУХИ	MNYTGNQQSRWGLGSPNCHGPDWASPVCQRHS	1321	BAA19893.1
439	ШТАММ ВИРУСА КОРИ EDMONSTON	MQSWVPLSTDDPVID	1322	P08362.1
440	МОРБИЛЛИВИРУС КОРИ	MSLSLLDLYLGRGYN	1323	P06830.1
441	ШТАММ ВИРУСА КОРИ EDMONSTON	MSPLLQECLRGSTKS	1324	P69353.1
442	МОРБИЛЛИВИРУС КОРИ	MSPQRDRINAFYKDN	1325	P06830.1
443	МОРБИЛЛИВИРУС КОРИ	MYRVFEVSVIRNPGL	1326	P06830.1
444	ШТАММ ВИРУСА КОРИ EDMONSTON	NALYPMSPLLQECLR	1327	P69353.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
445	ШТАММ ВИРУСА КРАСНУХИ THERIEN	NCHGPDWASPVCQRHSPDCSRLVGAT	1328	P07566.1
446	ШТАММ ВИРУСА КОРИ EDMONSTON	NFGRSYFDPAYFRLG	1329	P04851.1
447	ШТАММ ВИРУСА КОРИ EDMONSTON	NFGRSYFDPAYFRLGQEMVR	1330	P04851.1
448	ВИРУС КРАСНУХИ	NGTQRACTFWAVNAY	1331	BAA28178.1
449	ШТАММ ВИРУСА КОРИ EDMONSTON	NGVTIQVGSRRYPDA	1332	P69353.1
450	ШТАММ ВИРУСА КОРИ EDMONSTON	NIIAEDLSLRRFMVA	1333	P04851.1
451	ШТАММ ВИРУСА КОРИ EDMONSTON	NIIAEDLSLRRFMVALILDI	1334	P04851.1
452	ШТАММ ВИРУСА КОРИ EDMONSTON	NITLLNNCTRVEIAE	1335	P69353.1
453	МОРБИЛЛИВИРУС КОРИ	NLALGVINTLEWIPR	1336	P06830.1
454	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	NLFQVVHWSYNRPAYSPG	1337	SRC280292
455	ШТАММ ВИРУСА КОРИ EDMONSTON	NLVILPGQDLQYVLA	1338	P08362.1
456	ШТАММ ВИРУСА КОРИ EDMONSTON	NLYQQMGKPAPYMVN	1339	P04851.1
457	ШТАММ ВИРУСА КОРИ EDMONSTON	NLYQQMGKPAPYMVNLENSI	1340	P04851.1
458	ШТАММ ВИРУСА КОРИ EDMONSTON	NNCTRVEIAEYRRLL	1341	P69353.1
459	ШТАММ ВИРУСА КОРИ EDMONSTON	NPDREYDFRD	1342	P08362.1
460	ШТАММ ВИРУСА КОРИ HALLE	NPDVSGPKL	1343	P10050.1
461	ШТАММ ВИРУСА	NPDVSGPKLTGALIG	1344	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
462	ШТАММ ВИРУСА КОРИ EDMONSTON	NPDVSGPKLTGALIGILSLF	1345	P04851.1
463	МОРБИЛЛИВИРУС КОРИ	NPPERIKLDYDQYCA	1346	P06830.1
464	МОРБИЛЛИВИРУС КОРИ	NQAKWAVPTTRTDDK	1347	P06830.1
465	МОРБИЛЛИВИРУС КОРИ	NQDPDKILTYIAADH	1348	AAF02706.1
466	ШТАММ ВИРУСА КОРИ EDMONSTON	NQLSCDLIGQKLGLK	1349	P69353.1
467	ВИРУС КРАСНУХИ	NQQSRWGLGSPNCHGPDWASPVCQRHS	1350	ABD64214.1
468	РУБУЛАВИРУС ПАРОТИТА	NSTLGVKSAREF	1351	ABP48111.1
469	ВИРУС КРАСНУХИ	NTPHGQLEVQVPPDP	1352	BAA28178.1
470	ШТАММ ВИРУСА КОРИ EDMONSTON	NVSAIFMAVLLTLQT	1353	P69353.1
471	МОРБИЛЛИВИРУС КОРИ	PAEVDGDVKLSSNLV	1354	P06830.1
472	ВИРУС КРАСНУХИ	PAFGHSDAACWGFPT	1355	BAA28178.1
473	ШТАММ ВИРУСА КОРИ EDMONSTON	PALICCCRGRCNKKG	1356	P69353.1
474	МОРБИЛЛИВИРУС КОРИ	PDKILTYIAADHC	1357	AAF02706.1
475	ШТАММ ВИРУСА КОРИ EDMONSTON	PERIKLDYDQYCADV	1358	P08362.1
476	ВИРУС КРАСНУХИ	PERPRLRL	1359	CAJ88851.1
477	ВИРУС КРАСНУХИ	PGCATQAPVPVRLAG	1360	BAA28178.1
478	ВАКЦИННЫЙ ШТАММ ВИРУСА КРАСНУХИ RA27/3	PGCATQAPVPVRLAGVRFESKIVDGGCFA	1361	CAJ88851.1
479	ВИРУС КРАСНУХИ	PGEVWVTP	1362	CAJ88851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
480	ВИРУС КРАСНУХИ	PGEVWVTPVIGSQAR	1363	BAA28178.1
481	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	PGKLDVNKSKTHISVN	1364	CAE11230.1
482	ШТАММ ВИРУСА КОРИ EDMONSTON	PGLGAPVFHMTNYLE	1365	P08362.1
483	ВИРУС КРАСНУХИ	PGPGEVWV	1366	CAJ88851.1
484	ВИРУС КРАСНУХИ	PHKTVRVKFHTETRT	1367	BAA28178.1
485	МОРБИЛЛИВИРУС КОРИ	PIELQVECFTWDQKL	1368	AAR89413.1
486	ШТАММ ВИРУСА КОРИ EDMONSTON-B	PISLERLDVG	1369	P26031.1
487	ШТАММ ВИРУСА КОРИ EDMONSTON	PISLERLDVGTNLGN	1370	P69353.1
488	BORDETELLA PERTUSSIS	PKALFTQQGGAYGRC	1371	P04979.1
489	ШТАММ ВИРУСА КОРИ EDMONSTON	PKYVATQGYLISNFD	1372	P69353.1
490	ШТАММ ВИРУСА КОРИ EDMONSTON	PLDIDTASESSQD	1373	P04851.1
491	ВИРУС КРАСНУХИ	PLGLKFKTVRPVALP	1374	BAA28178.1
492	ШТАММ ВИРУСА КОРИ EDMONSTON	PLITHGSGMDLYKSN	1375	P08362.1
493	МОРБИЛЛИВИРУС КОРИ	PLKDNRIPSYGVLSV	1376	P06830.1
494	ВИРУС КРАСНУХИ	PLLRTAPG	1377	CAJ88851.1
495	ШТАММ ВИРУСА КОРИ EDMONSTON	PLLSCKPWQESRK	1378	P04851.1
496	BORDETELLA PERTUSSIS	PPATVYRYDSRPPE	1379	P04977.1
497	МОРБИЛЛИВИРУС КОРИ	PPISLERLDVGT	1380	AAL29688.1
498	ВИРУС КРАСНУХИ	PPPPEERQETRSQTPAPKPS	1381	P07566.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
499	BORDETELLA PERTUSSIS	PQEQITQHGSPYGRC	1382	AAA83981.1
500	BORDETELLA PERTUSSIS	PQEQITQHGSPYGRCANK	1383	AAA83981.1
501	BORDETELLA PERTUSSIS	PQPGPQPPQPPQPQPEAPAPQPPAG	1384	P14283.3
502	ШТАММ ВИРУСА КОРИ EDMONSTON	PRLGGKEDRRVKQ	1385	P04851.1
503	ВИРУС КРАСНУХИ	PRLRLVDA	1386	CAJ88851.1
504	ВИРУС КРАСНУХИ	PRNVRVTGCYQCGTP	1387	BAA28178.1
505	ШТАММ ВИРУСА КОРИ EDMONSTON	PSRASDARAAHLP	1388	P04851.1
506	ШТАММ ВИРУСА КОРИ EDMONSTON	PTGQIHWGNLSKIGV	1389	P69353.1
507	ШТАММ ВИРУСА КОРИ EDMONSTON	PTGTPLDIDTASE	1390	P04851.1
508	ШТАММ ВИРУСА КОРИ EDMONSTON	PTLSEIKGVIVHRLE	1391	P69353.1
509	МОРБИЛЛИВИРУС КОРИ	PTTIRGQFSNMSLSL	1392	P06830.1
510	МОРБИЛЛИВИРУС КОРИ	PTTRTDDKLR	1393	AAR89413.1
511	МОРБИЛЛИВИРУС КОРИ	PTTRTDDKLRMETCFQQACKG	1394	AAR89413.1
512	ВИРУС КРАСНУХИ	PVALPRTLAPPRNVR	1395	BAA28178.1
513	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	PVFAGANYAAWAVNVAQVI	1396	AAV70486.1
514	ВИРУС КРАСНУХИ	PVIGSQAR	1397	CAJ88851.1
515	МОРБИЛЛИВИРУС КОРИ	PVVEVNGVTIQVGSR	1398	AAL29688.1
516	ШТАММ ВИРУСА КРАСНУХИ THERIEN	PWELVVLTARPEDGWTCRGV	1399	P07566.1
517	ШТАММ ВИРУСА КОРИ EDMONSTON	PWQESRKNKAQTR	1400	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
518	ШТАММ ВИРУСА КОРИ EDMONSTON	PYMVNLENSIQNKFS	1401	P04851.1
519	ШТАММ ВИРУСА КОРИ EDMONSTON	PYMVNLENSIQNKFSAGSYP	1402	P04851.1
520	ШТАММ ВИРУСА КОРИ EDMONSTON	PYQGSGKGVS	1403	P08362.1
521	ВИРУС КРАСНУХИ	PYQVSCGGESDRASA	1404	BAA28178.1
522	МОРБИЛЛИВИРУС КОРИ	QACKGKIQALCEN	1405	P08362.1
523	ШТАММ ВИРУСА КОРИ EDMONSTON	QAGQEMILAVQGVQD	1406	P69353.1
524	МОРБИЛЛИВИРУС КОРИ	QALCENPECVPLKDN	1407	P06830.1
525	BORDETELLA PERTUSSIS	QALGALK	1408	ACI16088.1
526	ВИРУС КРАСНУХИ	QAPVPVRLAGVRFES	1409	BAA28178.1
527	ВИРУС КРАСНУХИ	QCGTPALVEGLAPGG	1410	BAA28178.1
528	ШТАММ ВИРУСА КОРИ EDMONSTON	QDPDKILTYIAADHC	1411	P69353.1
529	ШТАММ ВИРУСА КОРИ EDMONSTON	QECLRGSTKSCARTL	1412	P69353.1
530	BORDETELLA PERTUSSIS	QEQITQHGSPYGRC	1413	AAA83981.1
531	ШТАММ ВИРУСА КОРИ EDMONSTON	QESRKNKAQTRTP	1414	P04851.1
532	ШТАММ ВИРУСА КОРИ EDMONSTON	QGDQSENELPRLGGK	1415	P04851.1
533	ШТАММ ВИРУСА КОРИ EDMONSTON	QGDQSENELPRLGGKEDRRV	1416	P04851.1
534	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	QGESGHDIKITAENTPLPIA	1417	AAV70486.1
535	МОРБИЛЛИВИРУС КОРИ	QGSGKGVSFQLVKLG	1418	P06830.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
536	ШТАММ ВИРУСА КОРИ EDMONSTON	QGVQDYINNELIPSM	1419	P69353.1
537	ВИРУС КРАСНУХИ	QLEVQVPPDPGDLVE	1420	BAA28178.1
538	МОРБИЛЛИВИРУС КОРИ	QLPEATFMV	1421	ABK40528.1
539	ВИРУС КРАСНУХИ	QLPFLGHDGHHGGTLRVGQHYRNAS	1422	NP_740663.1
540	ШТАММ ВИРУСА КОРИ EDMONSTON	QLSMYRVFEV	1423	P08362.1
541	BORDETELLA PERTUSSIS	QLSNIT	1424	ACI16083.1
542	ШТАММ ВИРУСА КОРИ EDMONSTON	QNIRPVQSVASSRRH	1425	P69353.1
543	ШТАММ ВИРУСА КОРИ EDMONSTON	QNKFSAGSYPLLWSY	1426	P04851.1
544	ШТАММ ВИРУСА КОРИ EDMONSTON	QNKFSAGSYPLLWSYAMGVG	1427	P04851.1
545	ШТАММ ВИРУСА КОРИ EDMONSTON	QNKFSAGSYPLLWSYAMGVGVELEN	1428	P04851.1
546	ШТАММ ВИРУСА КОРИ EDMONSTON	QQACKGKIQALCENP	1429	P08362.1
547	ШТАММ ВИРУСА КОРИ EDMONSTON	QQRRVVGEFRLERKW	1430	P04851.1
548	ШТАММ ВИРУСА КОРИ EDMONSTON	QQRRVVGEFRLERKWLDVVR	1431	P04851.1
549	BORDETELLA PERTUSSIS	QQTRANPNPYTSRRSVAS	1432	P04977.1
550	ВИРУС КРАСНУХИ	QRHSPDCSRLVGATP	1433	BAA28178.1
551	ШТАММ ВИРУСА КОРИ EDMONSTON	QSGLTFASRGTNMED	1434	P04851.1
552	ШТАММ ВИРУСА КОРИ EDMONSTON	QSGLTFASRGTNMEDEADQY	1435	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
553	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	QSIALSSLMVAQAIPLVGEL	1436	AAV70486.1
554	ШТАММ ВИРУСА КОРИ EDMONSTON	QSRFGWFENKEISDI	1437	P04851.1
555	ШТАММ ВИРУСА КОРИ EDMONSTON	QSRFGWFENKEISDIEVQDP	1438	P04851.1
556	ШТАММ ВИРУСА КОРИ EDMONSTON	QSRGEAR	1439	P04851.1
557	ШТАММ ВИРУСА КОРИ EDMONSTON	QSRGEARESYRETGPSRA	1440	P04851.1
558	ШТАММ ВИРУСА КРАСНУХИ THERIEN	QTGRGGSAPRPELGPPTN	1441	P07566.1
559	ШТАММ ВИРУСА КРАСНУХИ THERIEN	QTPAPKPSRAPPQQPQPPRMQTGRG	1442	P07566.1
560	ШТАММ ВИРУСА КОРИ EDMONSTON	QVGSRRYPDAVYLHR	1443	P69353.1
561	ШТАММ ВИРУСА КОРИ EDMONSTON	QVSFLQGDQSENE	1444	P04851.1
562	ВИРУС КРАСНУХИ	QYHPTACEVEPAFGH	1445	BAA28178.1
563	ШТАММ ВИРУСА КОРИ EDMONSTON	QYVLATYDTSRVEHA	1446	P08362.1
564	BORDETELLA PERTUSSIS	RANPNPYTSRRSV	1447	ACI04548.1
565	ШТАММ ВИРУСА КОРИ EDMONSTON	RASDARAAHLPTG	1448	P04851.1
566	ШТАММ ВИРУСА КОРИ EDMONSTON	RASLETTNQAIEAIR	1449	P69353.1
567	ШТАММ ВИРУСА КРАСНУХИ THERIEN	RCGRLICGLSTTAQYPPTRF	1450	P07566.1
568	BORDETELLA PERTUSSIS	RDGQSVIGACASPYEGRYR	1451	P04979.1
569	ШТАММ ВИРУСА	RESYRETGPSRAS	1452	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
570	ШТАММ ВИРУСА КОРИ EDMONSTON	RETGPSRASDARA	1453	P04851.1
571	РУБУЛАВИРУС ПАРОТИТА	RFAKYQQQGRLEAR	1454	P21186.1
572	ШТАММ ВИРУСА КРАСНУХИ THERIEN	RFGAPQAFLAGLLLATVAVGTARA	1455	P07566.1
573	ШТАММ ВИРУСА КОРИ EDMONSTON	RFMVALILDIKRTPG	1456	P04851.1
574	ШТАММ ВИРУСА КОРИ EDMONSTON	RFMVALILDIKRTPGNKPRI	1457	P04851.1
575	ШТАММ ВИРУСА КОРИ EDMONSTON	RGEARESYRETGP	1458	P04851.1
576	ВИРУС КРАСНУХИ	RGTTPPAYG	1459	CAA28880.1
577	ШТАММ ВИРУСА КРАСНУХИ M33	RIETRSARH	1460	ABD64214.1
578	BORDETELLA PERTUSSIS	RILAGALATYQ	1461	P04977.1
579	BORDETELLA PERTUSSIS	RIPPENIRRVT	1462	ACI04548.1
580	BORDETELLA PERTUSSIS	RISNLND	1463	ACI16083.1
581	BORDETELLA PERTUSSIS	RLANLNG	1464	ACI16088.1
582	ШТАММ ВИРУСА КОРИ EDMONSTON	RLDVGTNLGNAIAKL	1465	P69353.1
583	ШТАММ ВИРУСА КОРИ EDMONSTON	RLERKWLDV	1466	P04851.1
584	ШТАММ ВИРУСА КОРИ EDMONSTON	RLGGKEDRRVKQSRG	1467	P04851.1
585	ШТАММ ВИРУСА КОРИ EDMONSTON	RLGGKEDRRVKQSRGEARES	1468	P04851.1
586	ШТАММ ВИРУСА КОРИ	RLLDRLVRL	1469	ABI54110.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
587	ВИРУС КРАСНУХИ	RLRLVDAD	1470	CAJ88851.1
588	ВИРУС КРАСНУХИ	RLRLVDADDPLLR	1471	BAA19893.1
589	ВИРУС КРАСНУХИ	RLRLVDADDPLLRTAPGPGEVWVTPVIGSQA	1472	BAA19893.1
590	ВИРУС КРАСНУХИ	RLRLVQDADDPLLRIAPGPGEVWVTPVIGSQA	1473	SRC265968
591	ШТАММ ВИРУСА КОРИ EDMONSTON	RLSDNGYYTV	1474	ABK40528.1
592	ВИРУС КРАСНУХИ	RLVDADDP	1475	CAJ88851.1
593	ВИРУС КРАСНУХИ	RLVDADDPLLRTAPG	1476	BAA28178.1
594	ВИРУС КРАСНУХИ	RLVGATPE	1477	CAJ88851.1
595	ВИРУС КРАСНУХИ	RMQTGRGGSAPRPELGPPTNPFQAAVA	1478	ABD64216.1
596	МОРБИЛЛИВИРУС КОРИ	RMSKGVFKV	1479	ABY21184.1
597	МОРБИЛЛИВИРУС КОРИ	RNPGLGAPVFHMTNY	1480	P06830.1
598	ВИРУС КРАСНУХИ	RPRLRLVD	1481	CAJ88851.1
599	BORDETELLA PERTUSSIS	RQAESSEAMAAWSERAGEA	1482	P04977.1
600	ВИРУС КРАСНУХИ	RQTWAEWAAAHWWQL	1483	BAA28178.1
601	ШТАММ ВИРУСА КОРИ EDMONSTON	RRVKQSRGEARES	1484	P04851.1
602	МОРБИЛЛИВИРУС КОРИ	RRYPDAVYL	1485	ACA09725.1
603	ШТАММ ВИРУСА КОРИ EDMONSTON	RRYPDAVYLHRIDLG	1486	P69353.1
604	ШТАММ ВИРУСА КОРИ EDMONSTON	RSAGKVSSTLASELG	1487	P04851.1
605	ШТАММ ВИРУСА КОРИ EDMONSTON	RSAGKVSSTLASELGITAED	1488	P04851.1
606	ШТАММ ВИРУСА	RSELSQLS	1489	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
607	ШТАММ ВИРУСА КОРИ EDMONSTON	RSELSQLSMYRVFEV	1490	P08362.1
608	ВИРУС КРАСНУХИ	RSQTPAPKPSRAPPQQPQPPRMQT	1491	ABD64214.1
609	ВИРУС КРАСНУХИ	RTAPGPGE	1492	CAJ88851.1
610	ВИРУС КРАСНУХИ	RTAPGPGEVWVTPVI	1493	BAA28178.1
611	МОРБИЛЛИВИРУС КОРИ	RTDDKLRMETCFQQA	1494	P06830.1
612	ВИРУС КРАСНУХИ	RTLAPPRNVRVTGCY	1495	BAA28178.1
613	ШТАММ ВИРУСА КОРИ EDMONSTON	RTVLEPIRDALNAMT	1496	P69353.1
614	ШТАММ ВИРУСА КОРИ EDMONSTON	RVEHAVVYYVYSPSR	1497	P08362.1
615	ШТАММ ВИРУСА КОРИ EDMONSTON	RVFEVGVIRNPGLGA	1498	P08362.1
616	BORDETELLA PERTUSSIS	RVHVSKEEQYYDYEDATFE	1499	P04978.2
617	ВИРУС КРАСНУХИ	RVIDPAAQ	1500	BAA28178.1
618	ВИРУС КРАСНУХИ	RVKFHTETRTVWQLS	1501	BAA28178.1
619	BORDETELLA PERTUSSIS	RVYHNGITGET	1502	ACI04548.1
620	BORDETELLA PERTUSSIS	RYDSRPPEDVF	1503	ACI04548.1
621	ШТАММ ВИРУСА КОРИ EDMONSTON	RYPDAVYLHRIDLGP	1504	P69353.1
622	ШТАММ ВИРУСА КОРИ EDMONSTON	SAEISIQALSYALGG	1505	P69353.1
623	ШТАММ ВИРУСА КОРИ EDMONSTON	SAEPLLSCKPWQESR	1506	P04851.1
624	ШТАММ ВИРУСА КОРИ	SAEPLLSCKPWQESRKNKAQ	1507	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
625	МОРБИЛЛИВИРУС КОРИ	SAGKVSSTLASELG	1508	P04851.1
626	ШТАММ ВИРУСА КОРИ EDMONSTON	SAGKVSSTLASELGITAEDARLVS	1509	P04851.1
627	ШТАММ ВИРУСА КОРИ EDMONSTON	SCTVTREDGT	1510	P08362.1
628	ВИРУС КРАСНУХИ	SDAACWGFPTDTVMS	1511	BAA28178.1
629	ШТАММ ВИРУСА КОРИ EDMONSTON	SDARAAHLPTGTP	1512	P04851.1
630	ШТАММ ВИРУСА КРАСНУХИ THERIEN	SDWHQGTHVCHTKHMDFWCVEHD	1513	P07566.1
631	ШТАММ ВИРУСА КОРИ EDMONSTON	SELRRWIKYTQQRRV	1514	P04851.1
632	ШТАММ ВИРУСА КОРИ EDMONSTON	SELRRWIKYTQQRRVVGEFR	1515	P04851.1
633	ШТАММ ВИРУСА КОРИ EDMONSTON	SELSQL	1516	P08362.1
634	ШТАММ ВИРУСА КОРИ EDMONSTON	SELSQLS	1517	P08362.1
635	BORDETELLA PERTUSSIS	SEYLAHRRIPPENIRRVTRV	1518	CAD44970.1
636	ШТАММ ВИРУСА КОРИ EDMONSTON	SFLQGDQSENELP	1519	P04851.1
637	ШТАММ ВИРУСА КОРИ EDMONSTON	SGKGVSFQLVKLGVW	1520	P08362.1
638	ШТАММ ВИРУСА КОРИ EDMONSTON	SHRGVIADNQAKWAV	1521	P08362.1
639	ШТАММ ВИРУСА КОРИ EDMONSTON	SIEHQVKDVLTPLFK	1522	P08362.1
640	ШТАММ ВИРУСА КОРИ	SKIGVVGIGSASYKV	1523	P69353.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
641	МОРБИЛЛИВИРУС КОРИ	SKRSELSQLSMYRVF	1524	P06830.1
642	МОРБИЛЛИВИРУС КОРИ	SLFVESPGQLIQRITDDPDVS	1525	ABI54110.1
643	ШТАММ ВИРУСА КОРИ EDMONSTON	SLSTNLDVTNSIEHQ	1526	P08362.1
644	ШТАММ ВИРУСА КОРИ EDMONSTON	SLSTNLDVTNSIEHQVKDVLTPLFK	1527	P08362.1
645	ШТАММ ВИРУСА КОРИ EDMONSTON	SLWGSGLLML	1528	BAE98296.1
646	ШТАММ ВИРУСА КОРИ EDMONSTON	SMKGLSSTSIVYILI	1529	P69353.1
647	МОРБИЛЛИВИРУС КОРИ	SMYRVFEVGV	1530	P08362.1
648	МОРБИЛЛИВИРУС КОРИ	SNDLSNCMVALGELK	1531	P06830.1
649	ШТАММ ВИРУСА КОРИ HALLE	SPGQLIQR	1532	P10050.1
650	ШТАММ ВИРУСА КОРИ EDMONSTON	SQDPQDSRRSAEP	1533	P04851.1
651	МОРБИЛЛИВИРУС КОРИ	SRIVINREHLMIDRP	1534	P06830.1
652	ШТАММ ВИРУСА КОРИ EDMONSTON	SRKNKAQTRTPLQ	1535	P04851.1
653	ВИРУС КРАСНУХИ	SRLVGATP	1536	CAJ88851.1
654	ВИРУС КРАСНУХИ	SRLVGATPERPRLRLVDADDPLLR	1537	CAJ88851.1
655	ШТАММ ВИРУСА КОРИ EDMONSTON	SRPGLKPDLTGTSKS	1538	P69353.1
656	BORDETELLA PERTUSSIS	SRRSVASIVGTLVRM	1539	CAD44970.1
657	ВИРУС КРАСНУХИ	SRWGLGSPNCHGPDW	1540	BAA28178.1
658	BORDETELLA PERTUSSIS	SSATK	1541	ACI16088.1
659	ВИРУС	SSGGYAQLASYFNPG	1542	BAA2817

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КРАСНУХИ			8.1
660	BORDETELLA PERTUSSIS	SSLGNGV	1543	ACI16083.1
661	МОРБИЛЛИВИРУС КОРИ	SSNLVILPGQDLQYV	1544	P06830.1
662	ШТАММ ВИРУСА КОРИ EDMONSTON	SSQDPQDSRRSAEPL	1545	P04851.1
663	ШТАММ ВИРУСА КОРИ EDMONSTON	SSQDPQDSRRSAEPLLSCKP	1546	P04851.1
664	ШТАММ ВИРУСА КОРИ EDMONSTON	SSRRHKRFAGVVLAG	1547	P69353.1
665	ШТАММ ВИРУСА КОРИ EDMONSTON	SSTSIVYILIAVCLG	1548	P69353.1
666	BORDETELLA PERTUSSIS	STPGIVI	1549	AAA83981.1
667	BORDETELLA PERTUSSIS	STPGIVIPPQEQITQHGSPYGRC	1550	AAA83981.1
668	BORDETELLA PERTUSSIS	STSSSRRYTEVY	1551	P04977.1
669	ШТАММ ВИРУСА КОРИ EDMONSTON	SYFIVLSIAYPTLSE	1552	P69353.1
670	ШТАММ ВИРУСА КОРИ EDMONSTON	SYRETGPSRASDA	1553	P04851.1
671	BORDETELLA PERTUSSIS	SYVK	1554	ACI16083.1
672	ВИРУС КРАСНУХИ	SYVQHPHKTVRVKFH	1555	BAA28178.1
673	ВИРУС КРАСНУХИ	TAPGPGEV	1556	CAJ88851.1
674	BORDETELLA PERTUSSIS	TATRLLSSTNSRLC	1557	AAA83981.1
675	МОРБИЛЛИВИРУС КОРИ	TDDPVIDRLYLSSHR	1558	P06830.1
676	ШТАММ ВИРУСА КОРИ EDMONSTON	TEILSLFGPSLRDPI	1559	P69353.1
677	ВИРУС КРАСНУХИ	TETRTVWQLSVAGVS	1560	BAA28178.1
678	BORDETELLA	TEVYLEHRMQEAVE	1561	P04977.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	PERTUSSIS			
679	ШТАММ ВИРУСА КОРИ EDMONSTON	TFMPEGTVCSQNALY	1562	P69353.1
680	ВИРУС КРАСНУХИ	TGACICEIPTDVSCE	1563	BAA28178.1
681	BORDETELLA PERTUSSIS	TGDLRAY	1564	ACI16083.1
682	МОРБИЛЛИВИРУС КОРИ	TGMQSWVPLSTDDPV	1565	P06830.1
683	ВИРУС КРАСНУХИ	TGNQQSRWGLGSPNC	1566	BAA28178.1
684	ШТАММ ВИРУСА КОРИ EDMONSTON	TGPSRASDARAAH	1567	P04851.1
685	МОРБИЛЛИВИРУС КОРИ	TGTIINQDPDKILTY	1568	AAF02706.1
686	ВИРУС КРАСНУХИ	TGVVYGTHTTAVSET	1569	BAA28178.1
687	ШТАММ ВИРУСА КОРИ EDMONSTON	TIRGQFSNMSLSLLD	1570	P08362.1
688	ВИРУС КРАСНУХИ	TLGATCALPLAGLLA	1571	BAA28178.1
689	МОРБИЛЛИВИРУС КОРИ	TLLNNCTRV	1572	P26031.1
690	ШТАММ ВИРУСА КОРИ EDMONSTON-B	TLNVPPPPDPGR	1573	P03422.1
691	ШТАММ ВИРУСА КОРИ EDMONSTON-B	TLNVPPPPDPGRASTSGTPIKK	1574	P03422.1
692	МОРБИЛЛИВИРУС КОРИ	TMTSQGMYGGTYPVE	1575	P06830.1
693	ШТАММ ВИРУСА КОРИ EDMONSTON	TNLGNAIAKLEDAKE	1576	P69353.1
694	ШТАММ ВИРУСА КОРИ EDMONSTON	TNMEDEADQYFSHDD	1577	P04851.1
695	ШТАММ ВИРУСА КОРИ EDMONSTON	TNMEDEADQYFSHDDPISSD	1578	P04851.1
696	ШТАММ ВИРУСА КРАСНУХИ M33	TNPFQAAVARGLRPP	1579	CAA28880.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
697	МОРБИЛЛИВИРУС КОРИ	TNSIEHQVKDVLTPL	1580	P06830.1
698	ШТАММ ВИРУСА КОРИ EDMONSTON	TNYLEQPVSNDLSNC	1581	P08362.1
699	МОРБИЛЛИВИРУС КОРИ	TNYLEQPVSNDLSNCMVALGELKLAAL	1582	AAR89413.1
700	ВИРУС КРАСНУХИ	TPERPRLR	1583	CAJ88851.1
701	ШТАММ ВИРУСА КРАСНУХИ THERIEN	TPERPRLRLVDADDPLLRTA	1584	P07566.1
702	ШТАММ ВИРУСА КОРИ HALLE	TPGNKPRIA	1585	P10050.1
703	ШТАММ ВИРУСА КОРИ EDMONSTON	TPLDIDTASESSQDP	1586	P04851.1
704	ШТАММ ВИРУСА КОРИ EDMONSTON	TPLDIDTASESSQDPQDSRR	1587	P04851.1
705	ШТАММ ВИРУСА КОРИ EDMONSTON	TPLFKIIGDEVGLRT	1588	P08362.1
706	ШТАММ ВИРУСА КОРИ EDMONSTON	TPLQCTM	1589	P04851.1
707	ВИРУС КРАСНУХИ	TPVIGSQA	1590	CAJ88851.1
708	ВИРУС КРАСНУХИ	TPVIGSQARK	1591	BAA19893.1
709	ШТАММ ВИРУСА КОРИ EDMONSTON	TQGYLISNFDESSCT	1592	P69353.1
710	BORDETELLA PERTUSSIS	TRANPNPYTSRRSVASIVGTLVRM	1593	P04977.1
711	ВИРУС КРАСНУХИ	TRFGCAMRWGLPP	1594	NP_740663.1
712	BORDETELLA PERTUSSIS	TRNTGQPATDHYYSNV	1595	AAA83981.1
713	ШТАММ ВИРУСА КОРИ EDMONSTON	TRTPLQCTMTEIF	1596	P04851.1
714	ВИРУС КРАСНУХИ	TRWHRLLRMPVR	1597	ABD64216.1
715	ШТАММ ВИРУСА	TSGSGGAIRGIKHII	1598	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
716	ШТАММ ВИРУСА КОРИ EDMONSTON	TSGSGGAIRGIKHIIIVPIP	1599	P04851.1
717	ШТАММ ВИРУСА КОРИ EDMONSTON	TSQGMYGGTYLVEKP	1600	P08362.1
718	МОРБИЛЛИВИРУС КОРИ	TSRVEHAVVYYVYSP	1601	P06830.1
719	BORDETELLA PERTUSSIS	TSSSRRYTEVYL	1602	ACI04548.1
720	BORDETELLA PERTUSSIS	TSYVG	1603	ACI16088.1
721	ШТАММ ВИРУСА КОРИ EDMONSTON	TTEDKISRAVGPRQA	1604	P04851.1
722	ШТАММ ВИРУСА КОРИ EDMONSTON	TTEDKISRAVGPRQAQVSFL	1605	P04851.1
723	ШТАММ ВИРУСА КРАСНУХИ M33	TTERIETRSARHP	1606	ABD64214.1
724	ШТАММ ВИРУСА КОРИ EDMONSTON	TTNQAIEAIRQAGQE	1607	P69353.1
725	ВИРУС КРАСНУХИ	TTSDPWHPPGPLGLK	1608	BAA28178.1
726	ШТАММ ВИРУСА КОРИ EDMONSTON	TVCSQNALYPMSPLL	1609	P69353.1
727	ВИРУС КРАСНУХИ	TVNGEDVGAVPPGKF	1610	BAA28178.1
728	МОРБИЛЛИВИРУС КОРИ	TYPVEKPNLSSKRSE	1611	P06830.1
729	ВИРУС КРАСНУХИ	VAGVSCNVTTEHPFC	1612	BAA28178.1
730	BORDETELLA PERTUSSIS	VAPGIVIPPKALFTQQGGAYGRC	1613	P04979.1
731	ВИРУС КРАСНУХИ	VCHTKHMDFWCVEHDRPPPATPTPL	1614	NP_740663.1
732	ВИРУС КРАСНУХИ	VCQRHSPDCSRLVGATPER	1615	BAA19893.1
733	ВИРУС КРАСНУХИ	VDADDPLL	1616	CAJ88851.1
734	ВИРУС	VDADDPLLRTAPGPGEVWVT	1617	BAA1989

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КРАСНУХИ			3.1
735	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VDIGFAAYNFVESIINLFQV	1618	AAV70486.1
736	ШТАММ ВИРУСА КОРИ EDMONSTON	VEIAEYRRLLRTVLE	1619	P69353.1
737	ШТАММ ВИРУСА КОРИ EDMONSTON	VELENSMGGLNFGRS	1620	P04851.1
738	ШТАММ ВИРУСА КОРИ EDMONSTON	VELENSMGGLNFGRSYFDPA	1621	P04851.1
739	МОРБИЛЛИВИРУС КОРИ	VELKIKIASGFGPLI	1622	P06830.1
740	ВИРУС КРАСНУХИ	VEMDEWIHAHTTSD	1623	SRC265968
741	ВИРУС КРАСНУХИ	VEMPEWIHAHTTSDP	1624	BAA28178.1
742	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VERRLVKVL	1625	P33120.2
743	МОРБИЛЛИВИРУС КОРИ	VESPGQLI	1626	ABI54110.1
744	ШТАММ ВИРУСА КОРИ EDMONSTON	VESPGQLIQRITDDP	1627	P04851.1
745	ШТАММ ВИРУСА КОРИ EDMONSTON	VESPGQLIQRITDDPDVSIR	1628	P04851.1
746	ВИРУС КРАСНУХИ	VFALASYVQHPHKTV	1629	BAA28178.1
747	ВИРУС КРАСНУХИ	VGATPERP	1630	CAJ88851.1
748	ВИРУС КРАСНУХИ	VGATPERPRL	1631	BAA19893.1
749	ВИРУС КРАСНУХИ	VGATPERPRLRLVDA	1632	BAA28178.1
750	ШТАММ ВИРУСА КОРИ EDMONSTON	VGIGSASYKVMTRSS	1633	P69353.1
751	ШТАММ ВИРУСА КОРИ EDMONSTON	VGLRTPQRFTDLVKF	1634	P08362.1
752	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VHHNTEEIVAQSIALSSLMV	1635	AAV70486.1
753	ШТАММ ВИРУСА	VHRLEGVSYNIGSQE	1636	P69353.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
754	ВИРУС КРАСНУХИ	VIGSQARK	1637	CAJ88851.1
755	BORDETELLA PERTUSSIS	VITGSI	1638	ACI16088.1
756	BORDETELLA PERTUSSIS	VITGTI	1639	ACI16083.1
757	ШТАММ ВИРУСА КОРИ EDMONSTON	VKQSRGEA	1640	P04851.1
758	ШТАММ ВИРУСА КОРИ EDMONSTON	VLFVMFLSLI	1641	P08362.1
759	МОРБИЛЛИВИРУС КОРИ	VLFVMFLSLIGLLAI	1642	P06830.1
760	МОРБИЛЛИВИРУС КОРИ	VLTPLFKIIGDEVGL	1643	P06830.1
761	ШТАММ ВИРУСА КОРИ EDMONSTON	VMTRSSHQSLVIKLM	1644	P69353.1
762	ВИРУС КРАСНУХИ	VPAAPCARIWNGTQR	1645	BAA28178.1
763	ВИРУС КРАСНУХИ	VPPDPGDLVEYIMNY	1646	BAA28178.1
764	ШТАММ ВИРУСА КОРИ EDMONSTON	VQSVASSRRHKRFAG	1647	P69353.1
765	BORDETELLA PERTUSSIS	VQTGGTSRTVTMRYLAS	1648	ACI16083.1
766	BORDETELLA PERTUSSIS	VQVRI	1649	ACI16083.1
767	ВИРУС КРАСНУХИ	VRAYNQPAGDV	1650	NP_740662.1
768	ВИРУС КРАСНУХИ	VRFESKIVDGGCFAP	1651	BAA28178.1
769	ВИРУС КРАСНУХИ	VRLAGVRFESKIVDG	1652	BAA28178.1
770	ШТАММ ВИРУСА КОРИ EDMONSTON	VSGSFGNRFILSQGN	1653	P69353.1
771	BORDETELLA PERTUSSIS	VSKEEQYYDYEDAT	1654	AAA83981.1
772	ШТАММ ВИРУСА КОРИ	VSKGNCSGPTTIRGQ	1655	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
773	ВИРУС КРАСНУХИ	VTAALLNTPPPYQVS	1656	BAA28178.1
774	ВИРУС КРАСНУХИ	VTGCYQCGTPALVEG	1657	BAA28178.1
775	ВИРУС КРАСНУХИ	VTPVIGSQ	1658	CAJ88851.1
776	ШТАММ ВИРУСА КРАСНУХИ THERIEN	VTTEHPFCNTPHGQLEVQVPPD	1659	P07566.1
777	ШТАММ ВИРУСА КОРИ EDMONSTON	VVLAGAALGVATAAQ	1660	P69353.1
778	ВИРУС КРАСНУХИ	VWQLSVAGVSCNVTT	1661	BAA28178.1
779	ВИРУС КРАСНУХИ	VWVTPVIG	1662	CAJ88851.1
780	ВИРУС КРАСНУХИ	VWVTPVIGSQAR	1663	BAA19893.1
781	ШТАММ ВИРУСА КОРИ EDMONSTON	VYILIAVCLGGLIGI	1664	P69353.1
782	ШТАММ ВИРУСА КОРИ EDMONSTON	VYLHRIDLGPPISLE	1665	P69353.1
783	BORDETELLA PERTUSSIS	VYRYDSRP	1666	P04977.1
784	BORDETELLA PERTUSSIS	VYRYDSRPPEDV	1667	P04977.1
785	МОРБИЛЛИВИРУС КОРИ	VYWLTIPPMKNLALG	1668	P06830.1
786	ВИРУС КРАСНУХИ	WDLEATGACICEIPT	1669	BAA28178.1
787	МОРБИЛЛИВИРУС КОРИ	WDQKLWCRHFCVLAD	1670	AAR89413.1
788	ВИРУС КРАСНУХИ	WGFPTDTVMSVFALA	1671	BAA28178.1
789	ВИРУС КРАСНУХИ	WHPPGPLGLKFKTVR	1672	BAA28178.1
790	ВИРУС КРАСНУХИ	WIHAHTTSDPWHPPG	1673	BAA28178.1
791	ШТАММ ВИРУСА КОРИ EDMONSTON	WLTIPPMKNLALGVI	1674	P08362.1
792	ШТАММ ВИРУСА	WQESRKNKAQTRTPLQCTMT	1675	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
793	ВИРУС КРАСНУХИ	WVCIFMVCRRACR	1676	SRC265968
794	ВИРУС КРАСНУХИ	WVTPVIGS	1677	CAJ88851.1
795	ШТАММ ВИРУСА КОРИ EDMONSTON	WYTTVPKYVATQGYL	1678	P69353.1
796	ШТАММ ВИРУСА КОРИ EDMONSTON	YALGGDINKVLEKLG	1679	P69353.1
797	ШТАММ ВИРУСА КОРИ EDMONSTON	YAMGVGVELE	1680	P04851.1
798	МОРБИЛЛИВИРУС КОРИ	YAMGVGVELEN	1681	ABI54110.1
799	МОРБИЛЛИВИРУС КОРИ	YCADVAAEELMNALV	1682	AAR89413.1
800	МОРБИЛЛИВИРУС КОРИ	YDFRDLTWCINPPER	1683	P06830.1
801	BORDETELLA PERTUSSIS	YFEPGPT	1684	ACI16083.1
802	ВИРУС КРАСНУХИ	YFNPGGSYYKQYHPT	1685	BAA28178.1
803	ШТАММ ВИРУСА КОРИ EDMONSTON	YFRLGQEMVRRSAGK	1686	P04851.1
804	ШТАММ ВИРУСА КОРИ EDMONSTON	YFRLGQEMVRRSAGKVSSTL	1687	P04851.1
805	МОРБИЛЛИВИРУС КОРИ	YFYPFRLPIKGVPIE	1688	P06830.1
806	BORDETELLA PERTUSSIS	YGDNAGRILAGALAT	1689	P04977.1
807	ВИРУС КРАСНУХИ	YGHATVEMPEWIHAH	1690	BAA28178.1
808	ВИРУС КРАСНУХИ	YIMNYTGNQQSRWGL	1691	BAA28178.1
809	ШТАММ ВИРУСА КОРИ EDMONSTON	YINNELIPSMNQLSC	1692	P69353.1
810	ВИРУС КРАСНУХИ	YLCTAPGCATQAPVP	1693	BAA28178.1
811	МОРБИЛЛИВИРУ	YLFTVPIKEAGEDCH	1694	P06830.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	С КОРИ			
812	ШТАММ ВИРУСА КОРИ EDMONSTON	YLHDPEFNL	1695	ABK40531.1
813	ШТАММ ВИРУСА КОРИ EDMONSTON	YLNMSRLFV	1696	ABK40531.1
814	ШТАММ ВИРУСА КОРИ HALLE	YPALGLHEF	1697	P10050.1
815	МОРБИЛЛИВИРУС КОРИ	YPALGLHEFA	1698	ABI54110.1
816	ШТАММ ВИРУСА КОРИ EDMONSTON	YPALGLHEFAGELST	1699	P04851.1
817	ШТАММ ВИРУСА КОРИ EDMONSTON	YPALGLHEFAGELSTLESLM	1700	P04851.1
818	ШТАММ ВИРУСА КОРИ EDMONSTON	YPFRLPIKGVPIELQ	1701	P08362.1
819	ШТАММ ВИРУСА КОРИ HALLE	YPLLWSYAM	1702	P10050.1
820	РУБУЛАВИРУС ПАРОТИТА	YQQQGRL	1703	P21186.1
821	BORDETELLA PERTUSSIS	YQSEYLAHRR	1704	P04977.1
822	ШТАММ ВИРУСА КОРИ EDMONSTON	YRETGPSRASDARAA	1705	P04851.1
823	ШТАММ ВИРУСА КОРИ EDMONSTON	YRETGPSRASDARAAHLPTG	1706	P04851.1
824	ВИРУС КРАСНУХИ	YRNASDVLPGHWLQGGWGCYNLSDW	1707	NP_740663.1
825	ШТАММ ВИРУСА КОРИ EDMONSTON	YRRLLRTVLEPIRDA	1708	P69353.1
826	BORDETELLA PERTUSSIS	YRYDSRPP	1709	P04977.1
827	ШТАММ ВИРУСА КОРИ EDMONSTON	YSGGDLLGILESRGI	1710	P69353.1
828	BORDETELLA PERTUSSIS	YSKVTATBLLASTNSRLCAVFVRDG	1711	SRC280066
829	ШТАММ ВИРУСА	YSPSRSFSYFYPFRL	1712	P08362.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	КОРИ EDMONSTON			
830	ШТАММ ВИРУСА КРАСНУХИ THERIEN	YTGNQQSRWGLGSPNCHGPDWASPV	1713	P07566.1
831	МОРБИЛЛИВИРУС КОРИ	YVLLAVLFV	1714	P08362.1
832	МОРБИЛЛИВИРУС КОРИ	YVYSPGRSFSYFYPF	1715	P06830.1
833	BORDETELLA PERTUSSIS	YYDYEDATFQTYALTGISLCNPAASIC	1716	P04979.1
834	МОРБИЛЛИВИРУС КОРИ	GDLLGILESRGIKAR	1717	AAF02706.1
835	МОРБИЛЛИВИРУС КОРИ	TVPKYVATQGYLISN	1718	AAL29688.1
836	МОРБИЛЛИВИРУС КОРИ	KPWDSPQEI	1719	P26035.1
837	МОРБИЛЛИВИРУС КОРИ	KPWESPQEI	1720	CAA34579.1
838	BORDETELLA PERTUSSIS	ATYQSEYLAHRRIPP	1721	ACI04548.1
839	BORDETELLA PERTUSSIS	CMARQAESSEAMAAWSERAGEAMVLVYYESIAYSF	1722	ACI04548.1
840	BORDETELLA PERTUSSIS	CQVGSSNSAFVSTSSSRRYTEVYL	1723	ACI04548.1
841	BORDETELLA PERTUSSIS	DDPPATVYRYDSRPP	1724	ACI04548.1
842	BORDETELLA PERTUSSIS	GALATYQSEYLAHRRIPP	1725	ACI04548.1
843	BORDETELLA PERTUSSIS	MAAWSERAGEAMVLVYYESIAYSF	1726	ACI04548.1
844	BORDETELLA PERTUSSIS	MVLVYYESIAYSF	1727	ACI04548.1
845	BORDETELLA PERTUSSIS	PATVYRYDSRPPEDV	1728	ACI04548.1
846	BORDETELLA PERTUSSIS	YDSRPPEDV	1729	ACI04548.1
847	BORDETELLA PERTUSSIS	EPGITTNYDT	1730	ACI16087.1
848	BORDETELLA PERTUSSIS	GDLRAYKMVYATNPQTQLSN	1731	ACI16083.1
849	BORDETELLA PERTUSSIS	KNGDVEASAITTYVGFSVVYP	1732	ACI16083.1
850	BORDETELLA PERTUSSIS	KVTNGSKSYTLRYLASYVK	1733	ACI16088.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
851	BORDETELLA PERTUSSIS	QALGALKLYFEPGITTNYDTGDLIAY KQTYNASGN	1734	ACI16088 .1
852	BORDETELLA PERTUSSIS	YATNPQTQLS	1735	ACI16083 .1
853	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	DNENPLSGKAGGVVKVTYPGLTKV	1736	AAV7048 6.1
854	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	ENFSSYHGTKPGYVDSI	1737	AAV7048 6.1
855	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	KVDNAETIKKELGLSLTEP	1738	AAV7048 6.1
856	ШТАММ ВИРУСА КРАСНУХИ M33	MEDLQKALEAQSRALRAGLAA	1739	CAA2888 0.1
857	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	QKGIQKPKSGTQGNYDDDWKGFY	1740	AAV7048 6.1
858	ШТАММ ВИРУСА КРАСНУХИ M33	RTGAWQRKDWSRAPPPPEERQESRS QTPAPKPSR	1741	CAA2888 0.1
859	ВИРУС КРАСНУХИ	AAGASQSRRPRPPRHARAQHLPEMTP AVT	1742	SRC2659 68
860	ВИРУС КРАСНУХИ	CVTSWLWSEGEGAVFYRVDLHFINL GTP	1743	CAA2888 0.1
861	ВИРУС КРАСНУХИ	FRVGGTRWHRLLRMPVRGLDGDTAP LP	1744	CAA2888 0.1
862	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GRKIRMRCRAIDGDVTFCRPKSPVYV GN	1745	1007216A
863	ВИРУС КРАСНУХИ	GTPPLDEDGRWDPALMYNPCGPEPPA HV	1746	CAA2888 0.1
864	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GVHANLHVAFHRSSSEKIHSNEISSDS IGVLGYQKTVDHTKVNSKLSLFFEIK S	1747	AAV7048 6.1
865	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSRALRA GLAAG	1748	ABD6420 0.1
866	ВИРУС КРАСНУХИ	PELGPPTNPFQAAVARGLRPPLHDPD TEAPTEAC	1749	CAA2888 0.1
867	ВИРУС КРАСНУХИ	PLPPHTTERIETRSARHPWRIRFGAP	1750	CAA2888 0.1
868	ВИРУС КРАСНУХИ	SRAPPPPEERQESRSQTPAPKPSRAPP	1751	CAA2888 0.1
869	ВИРУС КРАСНУХИ	SRAPPQQPQPPRMQTGRGGSAPRPEL GP	1752	CAA2888 0.1
870	ВИРУС КРАСНУХИ	TPAVTPEGPAPPRTGAWQRKDWSRAP P	1753	CAA2888 0.1
871	ВИРУС КРАСНУХИ	VRAYNQPAGDVRGVWGKGERTYAE QDFRV	1754	CAA2888 0.1
872	ВИРУС КРАСНУХИ	AFGHSDAACWGFPTDTVMSV	1755	CAA2888 0.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
873	ВИРУС КРАСНУХИ	CARIWNGTQRACTFWAVNAYS	1756	CAA28880.1
874	ВИРУС КРАСНУХИ	EEAFTYLCTAPGCATQTPVPVR	1757	CAA28880.1
875	ВИРУС КРАСНУХИ	FAPWDLEATGACICEIPTDV	1758	CAA28880.1
876	ВИРУС КРАСНУХИ	GEDVGAFPPGKFVTAAL	1759	CAA28880.1
877	ВИРУС КРАСНУХИ	GEVWVTPVIGSQARKCGLHI	1760	CAA28880.1
878	ВИРУС КРАСНУХИ	GQLEVQVPPDPGDLVEYIMN	1761	CAA28880.1
879	ВИРУС КРАСНУХИ	GSYYKQYHPTACEVEPAFGH	1762	CAA28880.1
880	ВИРУС КРАСНУХИ	IHAHTTSDPWHPPGPLGLKF	1763	CAA28880.1
881	ВИРУС КРАСНУХИ	IMNYTGNQQSRWGLGSPNCH	1764	CAA28880.1
882	ВИРУС КРАСНУХИ	LHIRAGPYGHATVEMPEWIH	1765	CAA28880.1
883	ВИРУС КРАСНУХИ	LKFKTVRPVALPRALAPPRN	1766	CAA28880.1
884	ВИРУС КРАСНУХИ	LNTPPPYQVSCGGESDRASAGH	1767	CAA28880.1
885	ВИРУС КРАСНУХИ	NCHGPDWASPVCQRHSPDCS	1768	CAA28880.1
886	ВИРУС КРАСНУХИ	PDCSRLVGATPERPRLRLVD	1769	CAA28880.1
887	ВИРУС КРАСНУХИ	PRNVRVTGCYQCGTPALVEG	1770	CAA28880.1
888	ВИРУС КРАСНУХИ	PTDVSCEGLGAWVPTAPCARI	1771	CAA28880.1
889	ВИРУС КРАСНУХИ	RLVDADDPLLRTAPGPGEVW	1772	CAA28880.1
890	ВИРУС КРАСНУХИ	SVFALASYVQHPHKTVRVKF	1773	CAA28880.1
891	ВИРУС КРАСНУХИ	VEGLAPGGGNCHLTVNGEDV	1774	CAA28880.1
892	ВИРУС КРАСНУХИ	VKFHTETRTVWQLSVAGVSC	1775	CAA28880.1
893	ВИРУС КРАСНУХИ	VPVRLAGVGFESKIVDGGCF	1776	CAA28880.1
894	ВИРУС КРАСНУХИ	VSCNVTTEHPFCNTPHGQLE	1777	CAA28880.1
895	BORDETELLA	AAASSPDAHVPF	1778	AAA2298

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	PERTUSSIS			3.1
896	BORDETELLA PERTUSSIS	AASSPDA	1779	AAA22983.1
897	BORDETELLA PERTUSSIS	AKLGAAASSPDA	1780	AAA22983.1
898	BORDETELLA PERTUSSIS	AMKPYEVTPTRM	1781	AAA22983.1
899	BORDETELLA PERTUSSIS	AMTHLSPALADVPYVLVKTNMVVTS	1782	AAA22983.1
900	BORDETELLA PERTUSSIS	ASSPDAHVPFCF	1783	AAA22983.1
901	BORDETELLA PERTUSSIS	ASSPDAHVPFCFGKDLKRPGSSPME	1784	AAA22983.1
902	BORDETELLA PERTUSSIS	CFGKDLKRPGSS	1785	AAA22983.1
903	BORDETELLA PERTUSSIS	CFGKDLKRPGSSPMEVMLRAVFMQQ	1786	AAA22983.1
904	BORDETELLA PERTUSSIS	CGIAAKLGAAAS	1787	AAA22983.1
905	BORDETELLA PERTUSSIS	CGIAAKLGAAASSPDAHVPFCFGKD	1788	AAA22983.1
906	BORDETELLA PERTUSSIS	DAHVPFCFGKDL	1789	AAA22983.1
907	BORDETELLA PERTUSSIS	DLKRPGSSPMEV	1790	AAA22983.1
908	BORDETELLA PERTUSSIS	DVPYVLVKTNMV	1791	AAA22983.1
909	BORDETELLA PERTUSSIS	DVPYVLVKTNMVVTSVAMKPYEVTPT	1792	AAA22983.1
910	BORDETELLA PERTUSSIS	EVMLRAVFMQQR	1793	AAA22983.1
911	BORDETELLA PERTUSSIS	FEGKPALELIRM	1794	AAA22983.1
912	BORDETELLA PERTUSSIS	FLGPKQLTFEGK	1795	AAA22983.1
913	BORDETELLA PERTUSSIS	FLGPKQLTFEGKPALELIRMVECSG	1796	AAA22983.1
914	BORDETELLA PERTUSSIS	FMQQRPLRMFLGPKQLT	1797	AAA22983.1
915	BORDETELLA PERTUSSIS	GKDLKRPGSSPM	1798	AAA22983.1
916	BORDETELLA PERTUSSIS	GKDLKRPGSSPME	1799	AAA22983.1
917	BORDETELLA PERTUSSIS	GKPALELIRMVE	1800	AAA22983.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
918	BORDETELLA PERTUSSIS	GPKQLTFEGKPA	1801	AAA22983.1
919	BORDETELLA PERTUSSIS	HVPFCFGKDLKR	1802	AAA22983.1
920	BORDETELLA PERTUSSIS	IAAKLGAAASSP	1803	AAA22983.1
921	BORDETELLA PERTUSSIS	KPYEVTPTRMLV	1804	AAA22983.1
922	BORDETELLA PERTUSSIS	KQLTFEGKPALE	1805	AAA22983.1
923	BORDETELLA PERTUSSIS	KRPGSSPMEVML	1806	AAA22983.1
924	BORDETELLA PERTUSSIS	LELIRMVECSGK	1807	AAA22983.1
925	BORDETELLA PERTUSSIS	LGAAASSPDAHV	1808	AAA22983.1
926	BORDETELLA PERTUSSIS	LIRMVECSGKQD	1809	AAA22983.1
927	BORDETELLA PERTUSSIS	LVCGIAAKLGAA	1810	AAA22983.1
928	BORDETELLA PERTUSSIS	MKPYEVTPTRM	1811	AAA22983.1
929	BORDETELLA PERTUSSIS	MLRAVFMQQRPL	1812	AAA22983.1
930	BORDETELLA PERTUSSIS	MQQRPLRM	1813	AAA22983.1
931	BORDETELLA PERTUSSIS	MQQRPLRMFLGP	1814	AAA22983.1
932	BORDETELLA PERTUSSIS	MVVTSVAMKPYE	1815	AAA22983.1
933	BORDETELLA PERTUSSIS	MVVTSVAMKPYEVTPTRMLVCGIAA	1816	AAA22983.1
934	BORDETELLA PERTUSSIS	PALELIRMVECS	1817	AAA22983.1
935	BORDETELLA PERTUSSIS	PALELIRMVECSGK	1818	AAA22983.1
936	BORDETELLA PERTUSSIS	PFCFGKDLKRPG	1819	AAA22983.1
937	BORDETELLA PERTUSSIS	PGSSPMEVMLRA	1820	AAA22983.1
938	BORDETELLA PERTUSSIS	PGSSPMEVMLRAVF	1821	AAA22983.1
939	BORDETELLA PERTUSSIS	PKQLTFEGK	1822	AAA22983.1
940	BORDETELLA	PLRMFLGPKQLT	1823	AAA2298

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	PERTUSSIS			3.1
941	BORDETELLA PERTUSSIS	PTRMLVCGIAAK	1824	AAA22983.1
942	BORDETELLA PERTUSSIS	PYVLVKTNMVVT	1825	AAA22983.1
943	BORDETELLA PERTUSSIS	QLTFEGKPALELIRMVECSGKQDCP	1826	AAA22983.1
944	BORDETELLA PERTUSSIS	QRPLRMFLGPKQ	1827	AAA22983.1
945	BORDETELLA PERTUSSIS	RAVFMQQRPLRM	1828	AAA22983.1
946	BORDETELLA PERTUSSIS	RMFLGPKQLTFE	1829	AAA22983.1
947	BORDETELLA PERTUSSIS	RMLVCGIAAKLG	1830	AAA22983.1
948	BORDETELLA PERTUSSIS	RMVECSGKQDCP	1831	AAA22983.1
949	BORDETELLA PERTUSSIS	SPDAHVPFCFGK	1832	AAA22983.1
950	BORDETELLA PERTUSSIS	SSPMEVMLRAVF	1833	AAA22983.1
951	BORDETELLA PERTUSSIS	SSPMEVMLRAVFMQQRPLRMFLGPK	1834	AAA22983.1
952	BORDETELLA PERTUSSIS	SVAMKPYEVTPT	1835	AAA22983.1
953	BORDETELLA PERTUSSIS	VECSGKQDCP	1836	AAA22983.1
954	BORDETELLA PERTUSSIS	VFMQQRPLRMFL	1837	AAA22983.1
955	BORDETELLA PERTUSSIS	VFMQQRPLRMFLGPKQLTFEGKPAL	1838	AAA22983.1
956	BORDETELLA PERTUSSIS	VKTNMVVTSVAM	1839	AAA22983.1
957	BORDETELLA PERTUSSIS	VLVKTNMVVTSV	1840	AAA22983.1
958	BORDETELLA PERTUSSIS	VTPTRMLVCGIA	1841	AAA22983.1
959	BORDETELLA PERTUSSIS	VTSVAMKPYEVT	1842	AAA22983.1
960	BORDETELLA PERTUSSIS	YEVTPTRMLVCG	1843	AAA22983.1
961	BORDETELLA PERTUSSIS	YEVTPTRMLVCGIAAKLGAAASSPD	1844	AAA22983.1
962	BORDETELLA PERTUSSIS	CASPYEGRYRDMYDALR	1845	P04979.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
963	BORDETELLA PERTUSSIS	CAVFVRDGQSV	1846	P04979.1
964	BORDETELLA PERTUSSIS	CITTIYKTG	1847	P04979.1
965	BORDETELLA PERTUSSIS	CPNGTRALTV	1848	P04979.1
966	BORDETELLA PERTUSSIS	DALRRLLYMIYMSG	1849	P04979.1
967	ШТАММ ВИРУСА КРАСНУХИ THERIEN	GNRGRGQRRDWSRAPPPPEERQETRS	1850	P07566.1
968	BORDETELLA PERTUSSIS	GQPAADHYYSKVT	1851	P04979.1
969	ВИРУС КРАСНУХИ	GSPNCHGPDWASPVCQRHS	1852	ABD64214.1
970	ШТАММ ВИРУСА КОРИ EDMONSTON	HKSLSTNLDVTNSIEHQ	1853	P08362.1
971	BORDETELLA PERTUSSIS	LFTQQGGAYGRC	1854	P04979.1
972	ШТАММ ВИРУСА КОРИ EDMONSTON	LIGLLAIAGIRLHRAAIYTAEI	1855	P08362.1
973	МОРБИЛЛИВИРУС КОРИ	PDTAADSELRRWIKY	1856	ABI54110.1
974	ШТАММ ВИРУСА КРАСНУХИ THERIEN	PNCHGPDWASPVCQRHS	1857	P07566.1
975	ВИРУС КРАСНУХИ	QTPAPKPSRAPPQQPQPPRMQTGR	1858	ABD64216.1
976	ВАКЦИННЫЙ ШТАММ ВИРУСА КРАСНУХИ RA27/3	RAGLTAGASQSRRPRPPR	1859	CAA33016.1
977	ВАКЦИННЫЙ ШТАММ ВИРУСА КРАСНУХИ RA27/3	RFGAPQAFLAGLLLAAVAVGTARA	1860	ABD64214.1
978	BORDETELLA PERTUSSIS	RGNAELQTYLRQITPG	1861	P04979.1
979	BORDETELLA PERTUSSIS	RVHVSKEEQYYDYED	1862	P04979.1
980	BORDETELLA PERTUSSIS	SIYGLYDGTYL	1863	P04979.1
981	BORDETELLA	SKVTATRLLASTNS	1864	P04979.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	PERTUSSIS			
982	BORDETELLA PERTUSSIS	TQQGGAYGRCPNGTRA	1865	P04979.1
983	BORDETELLA PERTUSSIS	VAPGIVIPPKAL	1866	P04979.1
984	BORDETELLA PERTUSSIS	DSRPPEDVFQNGFTAWG	1867	ACI04548.1
985	BORDETELLA PERTUSSIS	EHRMQEAVEAERAGR	1868	ACI04548.1
986	ШТАММ ВИРУСА КОРИ EDMONSTON	ETCFQQACKGKIQALCENPEWA	1869	P08362.1
987	BORDETELLA PERTUSSIS	EYVDTYGDNAGRILAGALATYQ	1870	ACI04548.1
988	BORDETELLA PERTUSSIS	HRRIPPENIRRVTR	1871	ACI04548.1
989	BORDETELLA PERTUSSIS	MARQAESSE	1872	ACI04548.1
990	BORDETELLA PERTUSSIS	MQEAVEAERAGR	1873	ACI04548.1
991	BORDETELLA PERTUSSIS	SQQTRANPNPYTSRR	1874	ACI04548.1
992	BORDETELLA PERTUSSIS	TRANPNPYTSRRSVASIVGTLVHG	1875	SRC280066
993	BORDETELLA PERTUSSIS	TVYRYDSRPPED	1876	ACI04548.1
994	МОРБИЛЛИВИРУС КОРИ	NDRNLLD	1877	P10050.1
995	МОРБИЛЛИВИРУС КОРИ	NMEDEADQYFSHDDPISSDQSRFGWFENK	1878	P04851.1
996	МОРБИЛЛИВИРУС КОРИ	SRASDARAAHLPTGTPLDID	1879	P04851.1
997	BORDETELLA PERTUSSIS	EDVFQNGFTAW	1880	ACI04548.1
998	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	AEGSSSVEYINNWEQAK	1881	AAV70486.1
999	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GPIKNKMSESPNKT	1882	AAV70486.1
1000	ШТАММ ВИРУСА КОРИ HALLE	GPKLTGALIGILSLFVESPGQLIQRITDDPDV	1883	P10050.1
1001	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GYQKTVDHTKVNSK	1884	AAV70486.1
1002	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	KTVDH	1885	AAV70486.1
100	КОРИНЕБАКТЕРИ	SESPNK	1886	AAV7048

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
3	Я ДИФТЕРИИ			6.1
1004	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	AEGSSSVEYINNWEQAKALS	1887	AAV70486.1
1005	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	AQAIPLVGELVDIGFAAYNF	1888	AAV70486.1
1006	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	ASRVVLSLPFAEGSSSVEYI	1889	AAV70486.1
1007	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	CINLDWDVIRDKTKTKIESL	1890	AAV70486.1
1008	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	CRAIDGDVTFCRPKSPVYVG	1891	AAV70486.1
1009	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	CRPKSPVYVGNGVHANLHVA	1892	AAV70486.1
1010	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	DAAGYSVDNENPLSGKAGGV	1893	AAV70486.1
1011	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	DKTKTKIESLKEHGPIKNKM	1894	AAV70486.1
1012	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	EEFIKRFGDGASRVVLSLPF	1895	AAV70486.1
1013	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	EKAKQYLEEFHQTALEHPEL	1896	AAV70486.1
1014	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	FHRSSSEKIHSNEISSDSIG	1897	AAV70486.1
1015	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GADDVVDSSKSFVMENFSSY	1898	CAE11230.1
1016	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GKRGQDAMYEYMAQACAGNR	1899	AAV70486.1
1017	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GSVMGIADGAVHHNTEEIVA	1900	AAV70486.1
1018	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	GTQGNYDDDWKGFYSTDNKY	1901	AAV70486.1
1019	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	HDGYAVSWNTVEDSIIRTGF	1902	AAV70486.1
1020	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	HGTKPGYVDSIQKGIQKPKS	1903	AAV70486.1
1021	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	HQTALEHPELSELKTVTGTN	1904	AAV70486.1
1022	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	IQKGIQKPKSGTQGNYDDDW	1905	AAV70486.1
1023	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	KEHGPIKNKMSESPNKTVSE	1906	AAV70486.1
1024	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	KGFYSTDNKYDAAGYSVDNE	1907	AAV70486.1
1025	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	LDVNKSKTHISVNGRKIRMR	1908	AAV70486.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1026	ВИРУС КРАСНУХИ	MASTIPITMEDLQKALEA	1909	SRC265968
1027	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	NGVHANLHVAFHRSSSEKIH	1910	AAV70486.1
1028	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	NNWEQAKALSVELEINFETR	1911	AAV70486.1
1029	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	NPLSGKAGGVVKVTYPGLTK	1912	AAV70486.1
1030	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	PVFAGANYAAWAVNVAQVID	1913	AAV70486.1
1031	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SELKTVTGTNPVFAGANYAA	1914	AAV70486.1
1032	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SESPNKTVSEEKAKQYLEEF	1915	AAV70486.1
1033	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SETADNLEKTTAALSILPGI	1916	AAV70486.1
1034	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SFVMENFSSYHGTKPGYVDS	1917	AAV70486.1
1035	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SNEISSDSIGVLGYQKTVDH	1918	AAV70486.1
1036	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SPGHKTQPFLHDGYAVSWNT	1919	AAV70486.1
1037	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SVNGRKIRMRCRAIDGDVTF	1920	AAV70486.1
1038	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	TAALSILPGIGSVMGIADGA	1921	AAV70486.1
1039	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	TAENTPLPIAGVLLPTIPGK	1922	AAV70486.1
1040	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	TEPLMEQVGTEEFIKRFGDG	1923	AAV70486.1
1041	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	TIKKELGLSLTEPLMEQVGT	1924	AAV70486.1
1042	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	TKVNSKLSLFFEIKS	1925	AAV70486.1
1043	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VEDSIIRTGFQGESGHDIKI	1926	AAV70486.1
1044	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VELEINFETRGKRGQDAMYE	1927	AAV70486.1
1045	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VESIINLFQVVHNSYNRPAY	1928	AAV70486.1
1046	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VHNSYNRPAYSPGHKTQPFL	1929	AAV70486.1
1047	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VKVTYPGLTKVLALKVDNAE	1930	AAV70486.1
104	КОРИНЕБАКТЕРИ	VLALKVDNAETIKKELGLSL	1931	AAV7048

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
8	Я ДИФТЕРИИ			6.1
1049	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VLGYQKTVDHTKVNSKLSLF	1932	AAV70486.1
1050	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	VRRSVGSSLSCINLDWDVIR	1933	AAV70486.1
1051	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	WAVNVAQVIDSETADNLEKT	1934	AAV70486.1
1052	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	YMAQACAGNRVRRSVGSSLS	1935	AAV70486.1
1053	BORDETELLA PERTUSSIS	AKAPPAPKPAPQPGP	1936	ABO77783.1
1054	BORDETELLA PERTUSSIS	APKPAPQPGP	1937	ABO77783.1
1055	BORDETELLA PERTUSSIS	APKPAPQPGPQPPQP	1938	ABO77783.1
1056	ШТАММ ВИРУСА КОРИ EDMONSTON	AQTRTPLQCTMTEIF	1939	P04851.1
1057	ШТАММ ВИРУСА КОРИ EDMONSTON	ASRGTNMEDEADQYFSHDD	1940	P04851.1
1058	BORDETELLA PERTUSSIS	ATIRR	1941	ABO77783.1
1059	BORDETELLA PERTUSSIS	DNRAG	1942	ABO77783.1
1060	BORDETELLA PERTUSSIS	EAPAPQPPAGRELSA	1943	ABO77783.1
1061	ШТАММ ВИРУСА КОРИ EDMONSTON	EMVRRSAGKVSSTLASELGI	1944	P04851.1
1062	BORDETELLA PERTUSSIS	GASEL	1945	ABO77783.1
1063	BORDETELLA PERTUSSIS	GDALAGGAVP	1946	AAA22980.1
1064	BORDETELLA PERTUSSIS	GDAPAGGAVP	1947	ABO77783.1
1065	BORDETELLA PERTUSSIS	GDTWDDD	1948	ABO77783.1
1066	BORDETELLA PERTUSSIS	GERQH	1949	ABO77783.1
1067	BORDETELLA PERTUSSIS	GGAVP	1950	ABO77783.1
1068	BORDETELLA PERTUSSIS	GGFGP	1951	P14283.3
106	BORDETELLA	GGFGPGGFGP	1952	BAF3503

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
9	PERTUSSIS			1.1
1070	BORDETELLA PERTUSSIS	GGFGPVLDGW	1953	ABO77783.1
1071	ШТАММ ВИРУСА КОРИ EDMONSTON	GGKEDRRVKQSRGEARESYR	1954	P04851.1
1072	BORDETELLA PERTUSSIS	GILLEN	1955	ABO77783.1
1073	BORDETELLA PERTUSSIS	GIRRFL	1956	ABO77783.1
1074	ШТАММ ВИРУСА КОРИ EDMONSTON	HDDPISSDQSRFGWFENKEI	1957	P04851.1
1075	ШТАММ ВИРУСА КОРИ EDMONSTON	HEFAGELSTLESLMNLY	1958	P04851.1
1076	BORDETELLA PERTUSSIS	HLGGLAGY	1959	ABO77783.1
1077	ШТАММ ВИРУСА КОРИ EDMONSTON	HTTEDKISRAVGPRQAQVSFL	1960	P04851.1
1078	ШТАММ ВИРУСА КОРИ EDMONSTON	ICDIDTYIVEAGLASFILTI	1961	P04851.1
1079	ШТАММ ВИРУСА КОРИ EDMONSTON	IIIVPIPGDSSITTRSRLLD	1962	P04851.1
1080	ШТАММ ВИРУСА КОРИ EDMONSTON	ILDIKRTPGNKPRIAEMICD	1963	P04851.1
1081	BORDETELLA PERTUSSIS	KALLYR	1964	ABO77783.1
1082	ШТАММ ВИРУСА КОРИ EDMONSTON	KPPITSGSGGAIRGIKHIII	1965	P04851.1
1083	BORDETELLA PERTUSSIS	LAGSGL	1966	ABO77783.1
1084	ШТАММ ВИРУСА КОРИ EDMONSTON	LGITAEDARLVSEIAMHTTE	1967	P04851.1
1085	ШТАММ ВИРУСА КОРИ EDMONSTON	LGTILAQIWVLLAKAVTA	1968	P04851.1
1086	ШТАММ ВИРУСА КОРИ	LPTGTPLDIDTASESSQD	1969	P04851.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
	EDMONSTON			
1087	ШТАММ ВИРУСА КОРИ EDMONSTON	MATLLRSLALFKRNKDK	1970	P04851.1
1088	BORDETELLA PERTUSSIS	PAPQPP	1971	ABO77783.1
1089	ШТАММ ВИРУСА КОРИ EDMONSTON	PDTAADSELRRWIKYTQQRR	1972	P04851.1
1090	ШТАММ ВИРУСА КОРИ EDMONSTON	PKLTGALIGILSLFVESPGQ	1973	P04851.1
1091	BORDETELLA PERTUSSIS	PQP	1974	ABO77783.1
1092	BORDETELLA PERTUSSIS	PQPGP	1975	ABO77783.1
1093	BORDETELLA PERTUSSIS	PQPGPQPPQPPQPQP	1976	ABO77783.1
1094	ШТАММ ВИРУСА КОРИ EDMONSTON	QDPQDSRRSAEPLLSCKPWQ	1977	P04851.1
1095	ШТАММ ВИРУСА КОРИ EDMONSTON	QRRVVGEFRLERKWLDVVR	1978	P04851.1
1096	BORDETELLA PERTUSSIS	RELSA	1979	ABO77783.1
1097	BORDETELLA PERTUSSIS	RFAPQ	1980	ABO77783.1
1098	ШТАММ ВИРУСА КОРИ EDMONSTON	SIQNKFSAGSYPLLWSYAMG	1981	P04851.1
1099	BORDETELLA PERTUSSIS	SITLQAGAH	1982	ABO77783.1
1100	BORDETELLA PERTUSSIS	SLQPED	1983	ABO77783.1
1101	BORDETELLA PERTUSSIS	SNALSKRL	1984	ABO77783.1
1102	ШТАММ ВИРУСА КОРИ EDMONSTON	SPGQLIQRITDDPDVSIRLL	1985	P04851.1
1103	BORDETELLA PERTUSSIS	TELPSIPG	1986	ABO77783.1
1104	BORDETELLA PERTUSSIS	TFTLANK	1987	ABO77783.1
110	BORDETELLA	TWDDD	1988	ABO7778

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
5	PERTUSSIS			3.1
1106	ШТАММ ВИРУСА КОРИ EDMONSTON	VSFLQGDQSENELPRLGGKE	1989	P04851.1
1107	ШТАММ ВИРУСА КОРИ EDMONSTON	WQESRKNKAQTRTPLQC	1990	P04851.1
1108	ШТАММ ВИРУСА КОРИ EDMONSTON	YAMGVGVELENSMGGLNFGR	1991	P04851.1
1109	ШТАММ ВИРУСА КОРИ EDMONSTON	YQQMGKPAPYMVNLENSI	1992	P04851.1
1110	ШТАММ ВИРУСА КОРИ HALLE	MTRSSHQSLVIKLMP	1993	P69355.1
1111	ШТАММ ВИРУСА КОРИ HALLE	PIRDALNAMTQNIRP	1994	P69355.1
1112	ШТАММ ВИРУСА КРАСНУХИ M33	ALLNTPPPYQVSCGGESDRASAGH	1995	CAA28880.1
1113	ШТАММ ВИРУСА КРАСНУХИ THERIEN	GLGSPNCHGPDWASPVCQRHS	1996	P07566.1
1114	ШТАММ ВИРУСА КРАСНУХИ THERIEN	GLGSPNCHGPDWASPVCQRHSPDCSRLV	1997	P07566.1
1115	ШТАММ ВИРУСА КРАСНУХИ THERIEN	NYTGNQQSRWGLGSPNCHGPDWASPV	1998	P07566.1
1116	ШТАММ ВИРУСА КРАСНУХИ THERIEN	TLPQPPCAHGQHYGHHHHQL	1999	P07566.1
1117	ШТАММ ВИРУСА КРАСНУХИ THERIEN	TVRVKFHTETRTVWQLSVAGVSCNVT	2000	P07566.1
1118	ВИРУС КРАСНУХИ	ASYFNPGGSYYKQYH	2001	BAA28178.1
1119	ВИРУС КРАСНУХИ	FALASYVQHPHKTVR	2002	BAA28178.1
1120	ВИРУС КРАСНУХИ	GGESDRASARVIDPAAQSFTG	2003	BAA28178.1
1121	ВИРУС КРАСНУХИ	GPGEVWVTPVIGSQARKC	2004	BAA28178.1
1122	ВИРУС КРАСНУХИ	GSQARKCGLHIRAG	2005	BAA28178.1
112	ВИРУС	LVGATPERPRLRLVDADDPLLRTAP	2006	BAA2817

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
3	КРАСНУХИ			8.1
1124	ВИРУС КРАСНУХИ	TAPGPGEVWVTPVI	2007	BAA28178.1
1125	МОРБИЛЛИВИРУС КОРИ	FIVLSIAYPTLSEIK	2008	AAL29688.1
1126	ШТАММ ВИРУСА КОРИ EDMONSTON-ZAGREB	ETCFQQACKGKIQALCENPEWAPLKDNRIPSY	2009	P08362.1
1127	МОРБИЛЛИВИРУС КОРИ	LPRLGGKEDR	2010	P04851.1
1128	МОРБИЛЛИВИРУС КОРИ	MSKTEWNASQ	2011	SRC280080
1129	МОРБИЛЛИВИРУС КОРИ	SRFGWFENKE	2012	P04851.1
1130	ВИРУС КРАСНУХИ	EACVTSWLWSEGEGAVFYRVDLHFINLGT	2013	CAA28880.1
1131	ВИРУС КРАСНУХИ	MDFWCVEHDRPPPATPTSLTT	2014	CAA33016.1
1132	ВИРУС КРАСНУХИ	PFLGHDGHHGGTLRVGQHHRNASDV	2015	CAA33016.1
1133	ВИРУС КРАСНУХИ	RVKFHTETRTVWQLSVAGVSC	2016	BAA19893.1
1134	МОРБИЛЛИВИРУС КОРИ	AEEQARHVKNGLE	2017	ABO69699.1
1135	МОРБИЛЛИВИРУС КОРИ	ESPQEISKHQALG	2018	SRC280080
1136	МОРБИЛЛИВИРУС КОРИ	GVGVELENSMGGLNF	2019	ABI54110.1
1137	МОРБИЛЛИВИРУС КОРИ	IKGANDLAKFHQMLMKIIMK	2020	ABO69699.1
1138	МОРБИЛЛИВИРУС КОРИ	MSKTLHAQLGFKKT	2021	ABK40528.1
1139	МОРБИЛЛИВИРУС КОРИ	NASGLSRPSPSAH	2022	BAE98296.1
1140	МОРБИЛЛИВИРУС КОРИ	VRVIDPSLGDRKDE	2023	SRC280148
1141	BORDETELLA PERTUSSIS	DLSDGDLLV	2024	AAC31207.1
1142	BORDETELLA PERTUSSIS	EAERAGRGTG	2025	ACI04548.1
1143	BORDETELLA PERTUSSIS	YRYDSRPPEDV	2026	ACI04548.1
1144	BORDETELLA PERTUSSIS	YVDTYGDNAG	2027	ACI04548.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1145	МОРБИЛЛИВИРУС КОРИ	SFSYFYPFR	2028	CAB43772.1
1146	РУБУЛАВИРУС ПАРОТИТА	DIFIVSPR	2029	ADF49557.1
1147	МОРБИЛЛИВИРУС КОРИ	QDSRRSADALLRLQAMAGISEEQGSDTDTPIVYNDRN	2030	CAA59302.1
1148	МОРБИЛЛИВИРУС КОРИ	SAEALLRLQA	2031	BAH22350.1
1149	МОРБИЛЛИВИРУС КОРИ	RIVINREHL	2032	BAB39835.1
1150	ГЕНОТИП ВИРУСА КОРИ А	IPRFK	2033	BAB39848.1
1151	BORDETELLA PERTUSSIS	AAALSPMEI	2034	P15318.2
1152	BORDETELLA PERTUSSIS	AAASVVGAPV	2035	P15318.2
1153	BORDETELLA PERTUSSIS	AALGRQDSI	2036	P15318.2
1154	BORDETELLA PERTUSSIS	AAQRLVHAIA	2037	P15318.2
1155	BORDETELLA PERTUSSIS	AAVEAAEL	2038	P15318.2
1156	BORDETELLA PERTUSSIS	AGANVLNGL	2039	P15318.2
1157	BORDETELLA PERTUSSIS	AGYANAAD	2040	P15318.2
1158	BORDETELLA PERTUSSIS	AGYEQFEFRV	2041	P15318.2
1159	BORDETELLA PERTUSSIS	AITGNADNL	2042	P15318.2
1160	BORDETELLA PERTUSSIS	AKEKNATLM	2043	P15318.2
1161	BORDETELLA PERTUSSIS	AKGVFLSL	2044	P15318.2
1162	BORDETELLA PERTUSSIS	APHEYGFGI	2045	P15318.2
1163	BORDETELLA PERTUSSIS	ARQGNDLEI	2046	P15318.2
1164	BORDETELLA PERTUSSIS	ASVVGAPV	2047	P15318.2
1165	BORDETELLA PERTUSSIS	ATLMFRLV	2048	P15318.2
1166	BORDETELLA PERTUSSIS	AVAAAQRL	2049	P15318.2
116	BORDETELLA	DAGANVLNGL	2050	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			
1168	BORDETELLA PERTUSSIS	DALLAQLYR	2051	P15318.2
1169	BORDETELLA PERTUSSIS	DANGVLKHSI	2052	P15318.2
1170	BORDETELLA PERTUSSIS	DGDMNIGVI	2053	P15318.2
1171	BORDETELLA PERTUSSIS	DHVKNIENL	2054	P15318.2
1172	BORDETELLA PERTUSSIS	DIDMFAIM	2055	P15318.2
1173	BORDETELLA PERTUSSIS	DMFAIMPHL	2056	P15318.2
1174	BORDETELLA PERTUSSIS	DNVRNVENV	2057	P15318.2
1175	BORDETELLA PERTUSSIS	DNVRNVENVI	2058	P15318.2
1176	BORDETELLA PERTUSSIS	DTVDYSAM	2059	P15318.2
1177	BORDETELLA PERTUSSIS	DTVDYSAMI	2060	P15318.2
1178	BORDETELLA PERTUSSIS	DYLRQAGL	2061	P15318.2
1179	BORDETELLA PERTUSSIS	DYYDNVRNV	2062	P15318.2
1180	BORDETELLA PERTUSSIS	EFTTFVEI	2063	P15318.2
1181	BORDETELLA PERTUSSIS	EFTTFVEIV	2064	P15318.2
1182	BORDETELLA PERTUSSIS	EGYVFYEN	2065	P15318.2
1183	BORDETELLA PERTUSSIS	ENVQYRHV	2066	P15318.2
1184	BORDETELLA PERTUSSIS	EQLANSDGL	2067	P15318.2
1185	BORDETELLA PERTUSSIS	FGVGYGHDTI	2068	P15318.2
1186	BORDETELLA PERTUSSIS	FSPDVLETVP	2069	P15318.2
1187	BORDETELLA PERTUSSIS	FSVDHVKNI	2070	P15318.2
1188	BORDETELLA PERTUSSIS	GDDTYLFGV	2071	P15318.2
1189	BORDETELLA PERTUSSIS	GDDVFLQDL	2072	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1190	BORDETELLA PERTUSSIS	GEDGNDIFL	2073	P15318.2
1191	BORDETELLA PERTUSSIS	GERFNVRKQL	2074	P15318.2
1192	BORDETELLA PERTUSSIS	GGAGNDTLV	2075	P15318.2
1193	BORDETELLA PERTUSSIS	GGDDFEAV	2076	P15318.2
1194	BORDETELLA PERTUSSIS	GKSEFTTFV	2077	P15318.2
1195	BORDETELLA PERTUSSIS	GKSLFDDGL	2078	P15318.2
1196	BORDETELLA PERTUSSIS	GNADNLKSV	2079	P15318.2
1197	BORDETELLA PERTUSSIS	GQLVEVDTL	2080	P15318.2
1198	BORDETELLA PERTUSSIS	GRSKFSPDV	2081	P15318.2
1199	BORDETELLA PERTUSSIS	GSSAYDTV	2082	P15318.2
1200	BORDETELLA PERTUSSIS	GTVEKWPAL	2083	P15318.2
1201	BORDETELLA PERTUSSIS	GVDYYDNV	2084	P15318.2
1202	BORDETELLA PERTUSSIS	GYEQFEFRV	2085	P15318.2
1203	BORDETELLA PERTUSSIS	HAVGAQDVV	2086	P15318.2
1204	BORDETELLA PERTUSSIS	IAAGRIGLGI	2087	P15318.2
1205	BORDETELLA PERTUSSIS	IGDAQANTL	2088	P15318.2
1206	BORDETELLA PERTUSSIS	IGLGILADL	2089	P15318.2
1207	BORDETELLA PERTUSSIS	IGNAAGIPL	2090	P15318.2
1208	BORDETELLA PERTUSSIS	IGTSMKDVL	2091	P15318.2
1209	BORDETELLA PERTUSSIS	IGVITDFEL	2092	P15318.2
1210	BORDETELLA PERTUSSIS	IPLTADIDM	2093	P15318.2
1211	BORDETELLA PERTUSSIS	ISKSALEL	2094	P15318.2
121	BORDETELLA	ITGNADNL	2095	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			
1213	BORDETELLA PERTUSSIS	KIFVVSAT	2096	P15318.2
1214	BORDETELLA PERTUSSIS	KQLNNANVYR	2097	P15318.2
1215	BORDETELLA PERTUSSIS	KVIGNAAGI	2098	P15318.2
1216	BORDETELLA PERTUSSIS	LAKVVSQL	2099	P15318.2
1217	BORDETELLA PERTUSSIS	LANDYARKI	2100	P15318.2
1218	BORDETELLA PERTUSSIS	LDYLRQAGL	2101	P15318.2
1219	BORDETELLA PERTUSSIS	LGKGFASL	2102	P15318.2
1220	BORDETELLA PERTUSSIS	LGKGFASLM	2103	P15318.2
1221	BORDETELLA PERTUSSIS	LGVDYYDN	2104	P15318.2
1222	BORDETELLA PERTUSSIS	LGVDYYDNV	2105	P15318.2
1223	BORDETELLA PERTUSSIS	LKHSIKLDVI	2106	P15318.2
1224	BORDETELLA PERTUSSIS	LQAGYIPV	2107	P15318.2
1225	BORDETELLA PERTUSSIS	LQLTGGTVE	2108	P15318.2
1226	BORDETELLA PERTUSSIS	LSAAVFGL	2109	P15318.2
1227	BORDETELLA PERTUSSIS	LSLGKGFASL	2110	P15318.2
1228	BORDETELLA PERTUSSIS	LSPMEIYGL	2111	P15318.2
1229	BORDETELLA PERTUSSIS	NAHDNFLAGG	2112	P15318.2
1230	BORDETELLA PERTUSSIS	NANVYREGV	2113	P15318.2
1231	BORDETELLA PERTUSSIS	NDTLYGGL	2114	P15318.2
1232	BORDETELLA PERTUSSIS	NGLAGNDVL	2115	P15318.2
1233	BORDETELLA PERTUSSIS	NNANVYREGV	2116	P15318.2
1234	BORDETELLA PERTUSSIS	NTVSYAAL	2117	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1235	BORDETELLA PERTUSSIS	NVLRNIENAV	2118	P15318.2
1236	BORDETELLA PERTUSSIS	PALTFITPL	2119	P15318.2
1237	BORDETELLA PERTUSSIS	PETSNVLRNI	2120	P15318.2
1238	BORDETELLA PERTUSSIS	PMEIYGLV	2121	P15318.2
1239	BORDETELLA PERTUSSIS	PQAYFEKNL	2122	P15318.2
1240	BORDETELLA PERTUSSIS	PVNPNLSKL	2123	P15318.2
1241	BORDETELLA PERTUSSIS	QAGWNASSV	2124	P15318.2
1242	BORDETELLA PERTUSSIS	QAGWNASSVI	2125	P15318.2
1243	BORDETELLA PERTUSSIS	QDAANAGNL	2126	P15318.2
1244	BORDETELLA PERTUSSIS	QDAANAGNLL	2127	P15318.2
1245	BORDETELLA PERTUSSIS	QDSGYDSL	2128	P15318.2
1246	BORDETELLA PERTUSSIS	QQSHYADQL	2129	P15318.2
1247	BORDETELLA PERTUSSIS	RALQGAQAV	2130	P15318.2
1248	BORDETELLA PERTUSSIS	RGGLGLDTL	2131	P15318.2
1249	BORDETELLA PERTUSSIS	RKQLNNANV	2132	P15318.2
1250	BORDETELLA PERTUSSIS	RQDSGYDSL	2133	P15318.2
1251	BORDETELLA PERTUSSIS	RQFRYDGDM	2134	P15318.2
1252	BORDETELLA PERTUSSIS	RSKFSPDVL	2135	P15318.2
1253	BORDETELLA PERTUSSIS	SAGAAAGAL	2136	P15318.2
1254	BORDETELLA PERTUSSIS	SAHWGQRAL	2137	P15318.2
1255	BORDETELLA PERTUSSIS	SAMIHPGRI	2138	P15318.2
1256	BORDETELLA PERTUSSIS	SAMIHPGRIV	2139	P15318.2
125	BORDETELLA	SAYDTVSGI	2140	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			
1258	BORDETELLA PERTUSSIS	SAYGYEGD	2141	P15318.2
1259	BORDETELLA PERTUSSIS	SGGAGDDVL	2142	P15318.2
1260	BORDETELLA PERTUSSIS	SGLQVAGA	2143	P15318.2
1261	BORDETELLA PERTUSSIS	SGYDSLDGV	2144	P15318.2
1262	BORDETELLA PERTUSSIS	SLLTGALNGI	2145	P15318.2
1263	BORDETELLA PERTUSSIS	SPMEIYGL	2146	P15318.2
1264	BORDETELLA PERTUSSIS	SQMLTRGQL	2147	P15318.2
1265	BORDETELLA PERTUSSIS	SSAYDTVSGI	2148	P15318.2
1266	BORDETELLA PERTUSSIS	SSLAHGHTA	2149	P15318.2
1267	BORDETELLA PERTUSSIS	SSVTSGDSV	2150	P15318.2
1268	BORDETELLA PERTUSSIS	SVIGVQTTEI	2151	P15318.2
1269	BORDETELLA PERTUSSIS	TNTVSYAAL	2152	P15318.2
1270	BORDETELLA PERTUSSIS	TSLIAEGV	2153	P15318.2
1271	BORDETELLA PERTUSSIS	TSLLTGAL	2154	P15318.2
1272	BORDETELLA PERTUSSIS	TVPASPGL	2155	P15318.2
1273	BORDETELLA PERTUSSIS	VAKEKNATL	2156	P15318.2
1274	BORDETELLA PERTUSSIS	VAPHEYGFGI	2157	P15318.2
1275	BORDETELLA PERTUSSIS	VAVVTSLL	2158	P15318.2
1276	BORDETELLA PERTUSSIS	VFYENRAYG	2159	P15318.2
1277	BORDETELLA PERTUSSIS	VFYENRAYGV	2160	P15318.2
1278	BORDETELLA PERTUSSIS	VITDFELEV	2161	P15318.2
1279	BORDETELLA PERTUSSIS	VNPHSTSL	2162	P15318.2

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1280	BORDETELLA PERTUSSIS	VNPHSTSLI	2163	P15318.2
1281	BORDETELLA PERTUSSIS	VNPNLSKL	2164	P15318.2
1282	BORDETELLA PERTUSSIS	VNPNLSKLF	2165	P15318.2
1283	BORDETELLA PERTUSSIS	VQQPIIEKL	2166	P15318.2
1284	BORDETELLA PERTUSSIS	VQYRHVEL	2167	P15318.2
1285	BORDETELLA PERTUSSIS	VSIAAAASV	2168	P15318.2
1286	BORDETELLA PERTUSSIS	VSIAAAASVV	2169	P15318.2
1287	BORDETELLA PERTUSSIS	VTSLLTGAL	2170	P15318.2
1288	BORDETELLA PERTUSSIS	VVLANASRI	2171	P15318.2
1289	BORDETELLA PERTUSSIS	WPALNLFSV	2172	P15318.2
1290	BORDETELLA PERTUSSIS	WVRKASAL	2173	P15318.2
1291	BORDETELLA PERTUSSIS	YAVQYRRKGG	2174	P15318.2
1292	BORDETELLA PERTUSSIS	YGGLGDDTL	2175	P15318.2
1293	BORDETELLA PERTUSSIS	YGLVQQSHYA	2176	P15318.2
1294	BORDETELLA PERTUSSIS	YGYEGDALL	2177	P15318.2
1295	BORDETELLA PERTUSSIS	YIPVNPNL	2178	P15318.2
1296	BORDETELLA PERTUSSIS	YSAMIHPGRI	2179	P15318.2
1297	BORDETELLA PERTUSSIS	YSQTGAHAGI	2180	P15318.2
1298	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	AYNFVESIINLFQVVHNSYN	2181	CAE11230.1
1299	BORDETELLA PERTUSSIS	SGTTIK	2182	BAF35031.1
1300	BORDETELLA PERTUSSIS	RGHTLESAEGRKIFG	2183	AAA22974.1
1301	BORDETELLA PERTUSSIS	AGAMTVRDVAAAADLALQAGDA	2184	AAA22974.1
130	BORDETELLA	AGAMTVRDVAAAADLALQAGDAL	2185	AAA2297

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			4.1
1303	BORDETELLA PERTUSSIS	ALAAVLVNPHIFTRIGAAQTSLADGAAGPA	2186	AAA22974.1
1304	BORDETELLA PERTUSSIS	ALSIDSMTALGA	2187	AAA22974.1
1305	BORDETELLA PERTUSSIS	DLSAARGADISGEGR	2188	AAA22974.1
1306	BORDETELLA PERTUSSIS	DQNRYEYIWGLY	2189	AAA22974.1
1307	BORDETELLA PERTUSSIS	DYTVSADAIALA	2190	AAA22974.1
1308	BORDETELLA PERTUSSIS	GPIVVEAGELVSHAGG	2191	AAA22974.1
1309	BORDETELLA PERTUSSIS	GRPEGLKIGAHSATSVSGSFDAL	2192	AAA22974.1
1310	BORDETELLA PERTUSSIS	ITVTSRGGFDNEGKMESNK	2193	AAA22974.1
1311	BORDETELLA PERTUSSIS	LDQNRYEYIWGLYP	2194	AAA22974.1
1312	BORDETELLA PERTUSSIS	LSAARGADISG	2195	AAA22974.1
1313	BORDETELLA PERTUSSIS	NKIRLMGPLQ	2196	AAA22974.1
1314	BORDETELLA PERTUSSIS	NKLGRIRAGEDM	2197	AAA22974.1
1315	BORDETELLA PERTUSSIS	NKLGRIRAGEDMHLDAPRIE	2198	AAA22974.1
1316	BORDETELLA PERTUSSIS	PHLRNTGQVVAG	2199	AAA22974.1
1317	BORDETELLA PERTUSSIS	QVDLHDLSAARGADISG	2200	AAA22974.1
1318	BORDETELLA PERTUSSIS	RDVAAAADLALQ	2201	AAA22974.1
1319	BORDETELLA PERTUSSIS	SAARGADISGEG	2202	AAA22974.1
1320	BORDETELLA PERTUSSIS	TKGEMQIAGKGGGSP	2203	AAA22974.1
1321	BORDETELLA PERTUSSIS	TVSADAIALAAQ	2204	AAA22974.1
1322	BORDETELLA PERTUSSIS	VVAGHDIHI	2205	AAA22974.1
1323	BORDETELLA PERTUSSIS	PSGPNHTKVVQLPKISKNALKANG	2206	CAD12823.1
1324	ВИРУС КРАСНУХИ	LVGATPERPRLRLVDADDPLLRTAPGPGEVWVTPVIGSQAR	2207	BAA19902.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1325	ВИРУС КРАСНУХИ	QQSRWGLGSPNCHGPDWASPVCQRHSP	2208	BAA19902.1
1326	BORDETELLA PERTUSSIS	AGEAMVLVYYESIAYSF	2209	ACI04548.1
1327	BORDETELLA PERTUSSIS	GGVGLASTLWYAESNALSKRLGEL	2210	AAZ74322.1
1328	BORDETELLA PERTUSSIS	GTLVRIAPVIGACMARQA	2211	ACI04548.1
1329	BORDETELLA PERTUSSIS	IRRVTRVYHNGITGETTT	2212	ACI04548.1
1330	BORDETELLA PERTUSSIS	IVKTGERQHGIHIQGSDP	2213	AAZ74322.1
1331	BORDETELLA PERTUSSIS	IVKTGERQHGIHIQGSDPGGVRTA	2214	AAZ74338.1
1332	BORDETELLA PERTUSSIS	LRDTNVTAVPASGAPAAVSVLGAS	2215	AAZ74338.1
1333	BORDETELLA PERTUSSIS	PEAPAPQPPAGRELSAAANAAVNT	2216	AAZ74322.1
1334	BORDETELLA PERTUSSIS	AAADFAHAE	2217	WP_019247158.1
1335	BORDETELLA PERTUSSIS	AAAEVAGAL	2218	WP_019249248.1
1336	BORDETELLA PERTUSSIS	AAESTFESY	2219	WP_019247158.1
1337	BORDETELLA PERTUSSIS	AAGFDPEVQ	2220	WP_019248145.1
1338	BORDETELLA PERTUSSIS	AALGRGHSL	2221	AGS56996.1
1339	BORDETELLA PERTUSSIS	AAMQGAVVH	2222	AGT50936.1
1340	BORDETELLA PERTUSSIS	AAPAAHADW	2223	AGS56996.1
1341	BORDETELLA PERTUSSIS	AAQATVVQR	2224	WP_019247158.1
1342	BORDETELLA PERTUSSIS	AARVAGDNY	2225	WP_019249248.1
1343	BORDETELLA PERTUSSIS	AAVALLNKL	2226	WP_019249248.1
1344	BORDETELLA PERTUSSIS	ADDPPATVY	2227	AAW72734.1
1345	BORDETELLA PERTUSSIS	AEAGRFKVL	2228	AGS56996.1
1346	BORDETELLA PERTUSSIS	AEATQLVTA	2229	WP_019247158.1
134	BORDETELLA	AEGGATLGA	2230	WP_0192

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			49248.1
1348	BORDETELLA PERTUSSIS	AEHGEVSIQ	2231	WP_019249248.1
1349	BORDETELLA PERTUSSIS	AEIAFYPKE	2232	WP_019249248.1
1350	BORDETELLA PERTUSSIS	AEKVTTPAV	2233	WP_019247158.1
1351	BORDETELLA PERTUSSIS	AELQTYLRQ	2234	1BCP_C
1352	BORDETELLA PERTUSSIS	AEQSLIEVG	2235	WP_019249248.1
1353	BORDETELLA PERTUSSIS	AESNALSKR	2236	AGS56996.1
1354	BORDETELLA PERTUSSIS	AESSEAMAA	2237	AFK26302.1
1355	BORDETELLA PERTUSSIS	AEVKVGYRA	2238	WP_019247158.1
1356	BORDETELLA PERTUSSIS	AEVTDTSPS	2239	WP_019249248.1
1357	BORDETELLA PERTUSSIS	AGKSLKKKN	2240	WP_019247158.1
1358	BORDETELLA PERTUSSIS	AGLAGPSAV	2241	WP_019249248.1
1359	BORDETELLA PERTUSSIS	AIRVGRGAR	2242	AGT50936.1
1360	BORDETELLA PERTUSSIS	ALAAIASAA	2243	WP_019248145.1
1361	BORDETELLA PERTUSSIS	ALADVPYVL	2244	AAA22983.1
1362	BORDETELLA PERTUSSIS	ALANDGTIV	2245	WP_019248658.1
1363	BORDETELLA PERTUSSIS	ALGRGHSLY	2246	AGS56996.1
1364	BORDETELLA PERTUSSIS	ALILAASPV	2247	WP_019248658.1
1365	BORDETELLA PERTUSSIS	ALMLACTGL	2248	AAA22974.1
1366	BORDETELLA PERTUSSIS	AMQGAVVHL	2249	AGT50936.1
1367	BORDETELLA PERTUSSIS	AMTHLSPAL	2250	AAA22983.1
1368	BORDETELLA PERTUSSIS	AMYGKHITL	2251	WP_019249248.1
1369	BORDETELLA PERTUSSIS	ANEANALLW	2252	WP_019249248.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1370	BORDETELLA PERTUSSIS	APLSITLQA	2253	AGT50936.1
1371	BORDETELLA PERTUSSIS	APNALAWAL	2254	AAA22974.1
1372	BORDETELLA PERTUSSIS	APPAPKPAP	2255	AGS56996.1
1373	BORDETELLA PERTUSSIS	APPGAGFIY	2256	1BCP_C
1374	BORDETELLA PERTUSSIS	APQAAPLSI	2257	AGT50936.1
1375	BORDETELLA PERTUSSIS	APRIENTAK	2258	WP_019249248.1
1376	BORDETELLA PERTUSSIS	AQGKALLYR	2259	AGT50936.1
1377	BORDETELLA PERTUSSIS	AQITSYVGF	2260	WP_019248658.1
1378	BORDETELLA PERTUSSIS	AQLEVRGQR	2261	WP_019249248.1
1379	BORDETELLA PERTUSSIS	AQQLKQADR	2262	WP_019247699.1
1380	BORDETELLA PERTUSSIS	AQVTVAGRY	2263	WP_019249248.1
1381	BORDETELLA PERTUSSIS	ARRSRVRAL	2264	NP_882284.1
1382	BORDETELLA PERTUSSIS	ASPRRARRA	2265	WP_019249248.1
1383	BORDETELLA PERTUSSIS	ASSPDAHVP	2266	AAA22983.1
1384	BORDETELLA PERTUSSIS	ASVSNPGTF	2267	WP_019249248.1
1385	BORDETELLA PERTUSSIS	ATWNFQSTY	2268	WP_019249248.1
1386	BORDETELLA PERTUSSIS	ATYIADSGF	2269	AGS56996.1
1387	BORDETELLA PERTUSSIS	AVAAPAVGA	2270	WP_019249248.1
1388	BORDETELLA PERTUSSIS	AVFMQQRPL	2271	AAA22983.1
1389	BORDETELLA PERTUSSIS	AVLVNPHIF	2272	WP_019249248.1
1390	BORDETELLA PERTUSSIS	CFGKDLKRP	2273	AAA22983.1
1391	BORDETELLA PERTUSSIS	CPSSLGNGV	2274	WP_019248145.1
139	BORDETELLA	DAGHEHDTW	2275	AAA2298

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			4.1
1393	BORDETELLA PERTUSSIS	DAKHDLTVT	2276	WP_019249248.1
1394	BORDETELLA PERTUSSIS	DASGQHRLW	2277	AGS56996.1
1395	BORDETELLA PERTUSSIS	DATFETYAL	2278	YP_006628018.1
1396	BORDETELLA PERTUSSIS	DATFQTYAL	2279	1BCP_C
1397	BORDETELLA PERTUSSIS	DATLVGAKF	2280	WP_019247158.1
1398	BORDETELLA PERTUSSIS	DDEVDVSGR	2281	AAA22974.1
1399	BORDETELLA PERTUSSIS	DENGKPQTY	2282	WP_019247158.1
1400	BORDETELLA PERTUSSIS	DGPPSRPTT	2283	WP_019247158.1
1401	BORDETELLA PERTUSSIS	DHLTGRSCQ	2284	AAW72734.1
1402	BORDETELLA PERTUSSIS	DNEGKMESN	2285	WP_019249248.1
1403	BORDETELLA PERTUSSIS	DPPATVYRY	2286	AAW72734.1
1404	BORDETELLA PERTUSSIS	EATEGDATL	2287	WP_019247158.1
1405	BORDETELLA PERTUSSIS	EATQQAAGF	2288	WP_019248145.1
1406	BORDETELLA PERTUSSIS	ECSGKQDCP	2289	AAA22983.1
1407	BORDETELLA PERTUSSIS	EGGKLRGKD	2290	AAA22974.1
1408	BORDETELLA PERTUSSIS	EGKMESNKD	2291	WP_019249248.1
1409	BORDETELLA PERTUSSIS	EHRMQEAVE	2292	AAW72734.1
1410	BORDETELLA PERTUSSIS	EKRLDIDDA	2293	WP_019249248.1
1411	BORDETELLA PERTUSSIS	EPASANTLL	2294	AGS56996.1
1412	BORDETELLA PERTUSSIS	EPQAELAVF	2295	AGS56996.1
1413	BORDETELLA PERTUSSIS	EPVKLTLTG	2296	AGT50936.1
1414	BORDETELLA PERTUSSIS	ESAEGRKIF	2297	WP_019249248.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1415	BORDETELLA PERTUSSIS	ESYSESHNF	2298	WP_019247158.1
1416	BORDETELLA PERTUSSIS	ETFCITTIY	2299	1BCP_C
1417	BORDETELLA PERTUSSIS	EVAGALELS	2300	WP_019249248.1
1418	BORDETELLA PERTUSSIS	EVAKVEVVP	2301	WP_019247158.1
1419	BORDETELLA PERTUSSIS	EVDGIIQEF	2302	WP_019249248.1
1420	BORDETELLA PERTUSSIS	EVRADNNFY	2303	WP_019248344.1
1421	BORDETELLA PERTUSSIS	FAILSSTTE	2304	WP_019247158.1
1422	BORDETELLA PERTUSSIS	FAISAYALK	2305	AAA22984.1
1423	BORDETELLA PERTUSSIS	FALYDGTYL	2306	AFK26303.1
1424	BORDETELLA PERTUSSIS	FDTMLGFAI	2307	AAA22984.1
1425	BORDETELLA PERTUSSIS	FEGKPALEL	2308	AAA22983.1
1426	BORDETELLA PERTUSSIS	FELGADHAV	2309	AGS56996.1
1427	BORDETELLA PERTUSSIS	FEPGITTNY	2310	WP_019248658.1
1428	BORDETELLA PERTUSSIS	FETYALTGI	2311	YP_006628018.1
1429	BORDETELLA PERTUSSIS	FIYRETFCI	2312	1BCP_C
1430	BORDETELLA PERTUSSIS	FPTRTTAPG	2313	NP_882284.1
1431	BORDETELLA PERTUSSIS	FQTYALTGI	2314	1BCP_C
1432	BORDETELLA PERTUSSIS	FTHADGWFL	2315	AGS56996.1
1433	BORDETELLA PERTUSSIS	FVRDGQSVI	2316	1BCP_C
1434	BORDETELLA PERTUSSIS	FVRSGQPVI	2317	YP_006628018.1
1435	BORDETELLA PERTUSSIS	FVWYVDTVI	2318	WP_019248866.1
1436	BORDETELLA PERTUSSIS	GAASSRQAL	2319	WP_019249248.1
143	BORDETELLA	GAASSYFEY	2320	AAW7273

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			4.1
1438	BORDETELLA PERTUSSIS	GAFDLKTTF	2321	AFK26303.1
1439	BORDETELLA PERTUSSIS	GAPAAVSVL	2322	AGS56996.1
1440	BORDETELLA PERTUSSIS	GATRAVDSL	2323	AGS56996.1
1441	BORDETELLA PERTUSSIS	GAVPGGAVP	2324	AGT50936.1
1442	BORDETELLA PERTUSSIS	GEAMVLVYY	2325	AFK26302.1
1443	BORDETELLA PERTUSSIS	GEIALGDAT	2326	WP_019249248.1
1444	BORDETELLA PERTUSSIS	GELMAAQVA	2327	WP_019247158.1
1445	BORDETELLA PERTUSSIS	GGVPGGAVP	2328	AAZ74338.1
1446	BORDETELLA PERTUSSIS	GHEHDTWFD	2329	AAA22984.1
1447	BORDETELLA PERTUSSIS	GIGALKAGA	2330	WP_019249248.1
1448	BORDETELLA PERTUSSIS	GIVIPPKAL	2331	1BCP_C
1449	BORDETELLA PERTUSSIS	GKDLKRPGS	2332	AAA22983.1
1450	BORDETELLA PERTUSSIS	GKLPKPVTV	2333	WP_019247158.1
1451	BORDETELLA PERTUSSIS	GKSLKKKNQ	2334	WP_019247158.1
1452	BORDETELLA PERTUSSIS	GLDVQQGTV	2335	WP_019249248.1
1453	BORDETELLA PERTUSSIS	GLTDGVSRI	2336	WP_019249248.1
1454	BORDETELLA PERTUSSIS	GLYPTYTEW	2337	WP_019249248.1
1455	BORDETELLA PERTUSSIS	GLYQTYTEW	2338	YP_006626470.1
1456	BORDETELLA PERTUSSIS	GPPSRPTTP	2339	WP_019247158.1
1457	BORDETELLA PERTUSSIS	GPSAVAAPA	2340	WP_019249248.1
1458	BORDETELLA PERTUSSIS	GVAPTAQQL	2341	WP_019248866.1
1459	BORDETELLA PERTUSSIS	GVGLASTLW	2342	AGS56996.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1460	BORDETELLA PERTUSSIS	GYEAGFSLG	2343	WP_019247158.1
1461	BORDETELLA PERTUSSIS	HADDGTIVI	2344	WP_019248145.1
1462	BORDETELLA PERTUSSIS	HADWNNQSI	2345	AGS56996.1
1463	BORDETELLA PERTUSSIS	HAEHEKDVR	2346	WP_019247158.1
1464	BORDETELLA PERTUSSIS	HANHYGTRI	2347	WP_019247158.1
1465	BORDETELLA PERTUSSIS	HAQGKALLY	2348	AGT50936.1
1466	BORDETELLA PERTUSSIS	HFIGYIYEV	2349	AAW72734.1
1467	BORDETELLA PERTUSSIS	HLSPALADV	2350	AAA22983.1
1468	BORDETELLA PERTUSSIS	HSLYASYEY	2351	AGS56996.1
1469	BORDETELLA PERTUSSIS	HVRGMLVPV	2352	AAA22974.1
1470	BORDETELLA PERTUSSIS	HVSKEEQYY	2353	YP_006628018.1
1471	BORDETELLA PERTUSSIS	HVTRGWSIF	2354	AFK26303.1
1472	BORDETELLA PERTUSSIS	IADSGFYLD	2355	AGS56996.1
1473	BORDETELLA PERTUSSIS	IAHRTELRG	2356	AGS56996.1
1474	BORDETELLA PERTUSSIS	IENTAKLSG	2357	WP_019249248.1
1475	BORDETELLA PERTUSSIS	IESKISQSV	2358	WP_019249248.1
1476	BORDETELLA PERTUSSIS	IETGGARRF	2359	AGT50936.1
1477	BORDETELLA PERTUSSIS	IIKDAPPGA	2360	1BCP_C
1478	BORDETELLA PERTUSSIS	IIQEFAADL	2361	WP_019249248.1
1479	BORDETELLA PERTUSSIS	ILAGALATY	2362	AAW72734.1
1480	BORDETELLA PERTUSSIS	ILLENPAAE	2363	AGS56996.1
1481	BORDETELLA PERTUSSIS	ILPILVLAL	2364	NP_882286.1
148	BORDETELLA	IPFQRALRL	2365	WP_0192

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			48145.1
1483	BORDETELLA PERTUSSIS	ISVRVHVSK	2366	YP_006628018.1
1484	BORDETELLA PERTUSSIS	ITNETGKTY	2367	WP_019247158.1
1485	BORDETELLA PERTUSSIS	ITNKRAALI	2368	WP_019249248.1
1486	BORDETELLA PERTUSSIS	ITSYVGFSV	2369	WP_019248658.1
1487	BORDETELLA PERTUSSIS	ITVTSRGGF	2370	WP_019249248.1
1488	BORDETELLA PERTUSSIS	IVIPPKALF	2371	1BCP_C
1489	BORDETELLA PERTUSSIS	IVVEAGELV	2372	WP_019249248.1
1490	BORDETELLA PERTUSSIS	KAAKSVNLM	2373	WP_019247158.1
1491	BORDETELLA PERTUSSIS	KAAPLRRTT	2374	AGS56996.1
1492	BORDETELLA PERTUSSIS	KAGKLSATG	2375	WP_019249248.1
1493	BORDETELLA PERTUSSIS	KAGTIAAPW	2376	WP_019249248.1
1494	BORDETELLA PERTUSSIS	KAKSLTTEI	2377	WP_019249248.1
1495	BORDETELLA PERTUSSIS	KATVTTVQV	2378	WP_019247158.1
1496	BORDETELLA PERTUSSIS	KDYRDKDGG	2379	WP_019247158.1
1497	BORDETELLA PERTUSSIS	KEAATIVAA	2380	WP_019249248.1
1498	BORDETELLA PERTUSSIS	KEDVDAAQI	2381	WP_019248658.1
1499	BORDETELLA PERTUSSIS	KEVDGIIQE	2382	WP_019249248.1
1500	BORDETELLA PERTUSSIS	KGPKLAMPW	2383	AGS56996.1
1501	BORDETELLA PERTUSSIS	KLASGGGAV	2384	WP_019249248.1
1502	BORDETELLA PERTUSSIS	KLKGKNQEF	2385	AAA22984.1
1503	BORDETELLA PERTUSSIS	KLLHHILPI	2386	NP_882286.1
1504	BORDETELLA PERTUSSIS	KPAPQPGPQ	2387	AGS56996.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1505	BORDETELLA PERTUSSIS	KPAPTAPPM	2388	WP_019249248.1
1506	BORDETELLA PERTUSSIS	KPAVSVKVA	2389	WP_019249248.1
1507	BORDETELLA PERTUSSIS	KPDRAARVA	2390	WP_019249248.1
1508	BORDETELLA PERTUSSIS	KPLADIAVI	2391	YP_006626470.1
1509	BORDETELLA PERTUSSIS	KPLADIAVV	2392	WP_019249248.1
1510	BORDETELLA PERTUSSIS	KPLPKPLPV	2393	WP_019247158.1
1511	BORDETELLA PERTUSSIS	KQADRDFVW	2394	WP_019247699.1
1512	BORDETELLA PERTUSSIS	KSLPGGKLP	2395	WP_019247158.1
1513	BORDETELLA PERTUSSIS	KSYTLRYLA	2396	WP_019248658.1
1514	BORDETELLA PERTUSSIS	KTNMVVTSV	2397	AAA22983.1
1515	BORDETELLA PERTUSSIS	KVLAPRLYL	2398	AAA22974.1
1516	BORDETELLA PERTUSSIS	KVLSTKTTL	2399	WP_019247158.1
1517	BORDETELLA PERTUSSIS	LAAGAGLTL	2400	WP_019249248.1
1518	BORDETELLA PERTUSSIS	LAAIASAAH	2401	WP_019248145.1
1519	BORDETELLA PERTUSSIS	LAANGNGQW	2402	AGS56996.1
1520	BORDETELLA PERTUSSIS	LAAQVTQRG	2403	WP_019249248.1
1521	BORDETELLA PERTUSSIS	LAARGDGAL	2404	AAA22974.1
1522	BORDETELLA PERTUSSIS	LAAVLVNPH	2405	WP_019249248.1
1523	BORDETELLA PERTUSSIS	LAGSGLFRM	2406	AGS56996.1
1524	BORDETELLA PERTUSSIS	LAKALSAAL	2407	WP_019249248.1
1525	BORDETELLA PERTUSSIS	LALQAGDAL	2408	WP_019249248.1
1526	BORDETELLA PERTUSSIS	LAMPWTFHA	2409	AGS56996.1
152	BORDETELLA	LANDGTIVI	2410	WP_0192

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			48658.1
1528	BORDETELLA PERTUSSIS	LAPTVGVAF	2411	WP_019247158.1
1529	BORDETELLA PERTUSSIS	LASDGSVDF	2412	AGS56996.1
1530	BORDETELLA PERTUSSIS	LATYQSEYL	2413	AAW72734.1
1531	BORDETELLA PERTUSSIS	LAWALMLAC	2414	AAA22974.1
1532	BORDETELLA PERTUSSIS	LEAGRRFTH	2415	AGS56996.1
1533	BORDETELLA PERTUSSIS	LFTQQGGAY	2416	1BCP_C
1534	BORDETELLA PERTUSSIS	LKRPGSSPM	2417	AAA22983.1
1535	BORDETELLA PERTUSSIS	LLGSHVARA	2418	AFK26303.1
1536	BORDETELLA PERTUSSIS	LLHHILPIL	2419	NP_882286.1
1537	BORDETELLA PERTUSSIS	LLNAGGTLL	2420	WP_019249248.1
1538	BORDETELLA PERTUSSIS	LMGPLQVNA	2421	WP_019249248.1
1539	BORDETELLA PERTUSSIS	LNDSKITMG	2422	WP_019248145.1
1540	BORDETELLA PERTUSSIS	LPEPVKLTL	2423	AGT50936.1
1541	BORDETELLA PERTUSSIS	LPILVLALL	2424	NP_882286.1
1542	BORDETELLA PERTUSSIS	LPKISKNAL	2425	WP_019248145.1
1543	BORDETELLA PERTUSSIS	LPKPVTVKL	2426	WP_019247158.1
1544	BORDETELLA PERTUSSIS	LPLKANPMH	2427	NP_882285.1
1545	BORDETELLA PERTUSSIS	LPPRPVVAE	2428	WP_019247158.1
1546	BORDETELLA PERTUSSIS	LPSIPGTSI	2429	AGS56996.1
1547	BORDETELLA PERTUSSIS	LPTHLYKNF	2430	AAA22984.1
1548	BORDETELLA PERTUSSIS	LPVRGVALV	2431	WP_019249248.1
1549	BORDETELLA PERTUSSIS	LPVSLTALD	2432	WP_019249248.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1550	BORDETELLA PERTUSSIS	LQGGNKVPV	2433	WP_019249248.1
1551	BORDETELLA PERTUSSIS	LSAALGADW	2434	WP_019249248.1
1552	BORDETELLA PERTUSSIS	LSDAGHEHD	2435	AAA22984.1
1553	BORDETELLA PERTUSSIS	LSGEVQRKG	2436	WP_019249248.1
1554	BORDETELLA PERTUSSIS	LSSPSAITV	2437	WP_019249248.1
1555	BORDETELLA PERTUSSIS	LTWLAILAV	2438	AAW72734.1
1556	BORDETELLA PERTUSSIS	LVFSHVRGM	2439	AAA22974.1
1557	BORDETELLA PERTUSSIS	LVSDAGADL	2440	WP_019249248.1
1558	BORDETELLA PERTUSSIS	LVYYESIAY	2441	AFK26302.1
1559	BORDETELLA PERTUSSIS	MAAESTFES	2442	WP_019247158.1
1560	BORDETELLA PERTUSSIS	MAAGHDATL	2443	WP_019249248.1
1561	BORDETELLA PERTUSSIS	MAAWSERAG	2444	AFK26302.1
1562	BORDETELLA PERTUSSIS	MALGALGAA	2445	AGS56996.1
1563	BORDETELLA PERTUSSIS	MAPVMGACM	2446	ADA85123.1
1564	BORDETELLA PERTUSSIS	MATKGEMQI	2447	WP_019249248.1
1565	BORDETELLA PERTUSSIS	MDAKGGTLL	2448	WP_019249248.1
1566	BORDETELLA PERTUSSIS	MESNKDIVI	2449	WP_019249248.1
1567	BORDETELLA PERTUSSIS	MEVMLRAVF	2450	AAA22983.1
1568	BORDETELLA PERTUSSIS	MEYFKTPLP	2451	WP_019249248.1
1569	BORDETELLA PERTUSSIS	MHTIASILL	2452	AAA22984.1
1570	BORDETELLA PERTUSSIS	MIYMSGLAV	2453	1BCP_C
1571	BORDETELLA PERTUSSIS	MLACTGLPL	2454	AAA22974.1
157	BORDETELLA	MPIDRKTLC	2455	AFK2630

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			3.1
1573	BORDETELLA PERTUSSIS	MPKAPELDL	2456	WP_019249248.1
1574	BORDETELLA PERTUSSIS	MQRQAGLPL	2457	NP_882285.1
1575	BORDETELLA PERTUSSIS	NALAWALML	2458	AAA22974.1
1576	BORDETELLA PERTUSSIS	NITNKRAAL	2459	WP_019249248.1
1577	BORDETELLA PERTUSSIS	NLMAAESTF	2460	WP_019247158.1
1578	BORDETELLA PERTUSSIS	NNETMSGRQ	2461	WP_019249248.1
1579	BORDETELLA PERTUSSIS	NPGSLIAEV	2462	WP_019249248.1
1580	BORDETELLA PERTUSSIS	NPMHTIASI	2463	NP_882285.1
1581	BORDETELLA PERTUSSIS	NPQTQLSNI	2464	WP_019248145.1
1582	BORDETELLA PERTUSSIS	NPYTSRRSV	2465	AFK26302.1
1583	BORDETELLA PERTUSSIS	PAPTAPPMP	2466	WP_019249248.1
1584	BORDETELLA PERTUSSIS	PASANTLLL	2467	AGS56996.1
1585	BORDETELLA PERTUSSIS	PAVALPRPL	2468	WP_019249248.1
1586	BORDETELLA PERTUSSIS	PDAHVPFCF	2469	AAA22983.1
1587	BORDETELLA PERTUSSIS	PELGAAIRV	2470	AGT50936.1
1588	BORDETELLA PERTUSSIS	PFIIKLKDC	2471	WP_019248145.1
1589	BORDETELLA PERTUSSIS	PGPQPPQPP	2472	AGS56996.1
1590	BORDETELLA PERTUSSIS	PGPQPPQPQ	2473	AAZ74338.1
1591	BORDETELLA PERTUSSIS	PGPTTDYST	2474	WP_019248145.1
1592	BORDETELLA PERTUSSIS	PGTFTAGKD	2475	WP_019249248.1
1593	BORDETELLA PERTUSSIS	PGTPGDLLE	2476	AAA22984.1
1594	BORDETELLA PERTUSSIS	PKPKPKAER	2477	WP_019247158.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1595	BORDETELLA PERTUSSIS	PKPKPKPKA	2478	WP_019247158.1
1596	BORDETELLA PERTUSSIS	PKPKPKPKP	2479	WP_019247158.1
1597	BORDETELLA PERTUSSIS	PLPPRPVVA	2480	WP_019247158.1
1598	BORDETELLA PERTUSSIS	PPAPKPAPQ	2481	AGS56996.1
1599	BORDETELLA PERTUSSIS	PPKPAPVAK	2482	WP_019247158.1
1600	BORDETELLA PERTUSSIS	PPRPVAAQV	2483	WP_019247158.1
1601	BORDETELLA PERTUSSIS	PPRPVVAEK	2484	WP_019247158.1
1602	BORDETELLA PERTUSSIS	PPSRPTTPP	2485	WP_019247158.1
1603	BORDETELLA PERTUSSIS	PRRARRALR	2486	WP_019249248.1
1604	BORDETELLA PERTUSSIS	QAAPLSITL	2487	AGT50936.1
1605	BORDETELLA PERTUSSIS	QADRDFVWY	2488	WP_019247699.1
1606	BORDETELLA PERTUSSIS	QAIVVGKDL	2489	WP_019249248.1
1607	BORDETELLA PERTUSSIS	QALGALKLY	2490	ACI16088.1
1608	BORDETELLA PERTUSSIS	QELALKLKG	2491	AAA22984.1
1609	BORDETELLA PERTUSSIS	QITQHGGPY	2492	NP_882283.1
1610	BORDETELLA PERTUSSIS	QITQHGSPY	2493	AFK26303.1
1611	BORDETELLA PERTUSSIS	QPLPPRPVA	2494	WP_019247158.1
1612	BORDETELLA PERTUSSIS	QPPAGRELS	2495	AGS56996.1
1613	BORDETELLA PERTUSSIS	QQLKQADRD	2496	WP_019247699.1
1614	BORDETELLA PERTUSSIS	QQVQVLQRQ	2497	WP_019247158.1
1615	BORDETELLA PERTUSSIS	QSIVEAPEL	2498	AGT50936.1
1616	BORDETELLA PERTUSSIS	QVGSSNSAF	2499	AAW72734.1
161	BORDETELLA	RAGLSPATW	2500	WP_0192

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			49248.1
1618	BORDETELLA PERTUSSIS	RARRALRQD	2501	WP_019249248.1
1619	BORDETELLA PERTUSSIS	RASASRARI	2502	WP_019249248.1
1620	BORDETELLA PERTUSSIS	RELSAAANA	2503	AGS56996.1
1621	BORDETELLA PERTUSSIS	RETFCITTI	2504	1BCP_C
1622	BORDETELLA PERTUSSIS	RGFAQRQQL	2505	AGS56996.1
1623	BORDETELLA PERTUSSIS	RGSAATFTL	2506	WP_019248295.1
1624	BORDETELLA PERTUSSIS	RGWSIFALY	2507	AFK263033.1
1625	BORDETELLA PERTUSSIS	RKMLYLIYV	2508	YP_006628018.1
1626	BORDETELLA PERTUSSIS	RLRKMLYLI	2509	YP_006628018.1
1627	BORDETELLA PERTUSSIS	RPQITDAVT	2510	WP_019249248.1
1628	BORDETELLA PERTUSSIS	RPSVNGGRI	2511	WP_019249248.1
1629	BORDETELLA PERTUSSIS	RRFTHADGW	2512	AGS56996.1
1630	BORDETELLA PERTUSSIS	RSGARATSL	2513	AAA229744.1
1631	BORDETELLA PERTUSSIS	RSRVRALAW	2514	NP_882284.1
1632	BORDETELLA PERTUSSIS	RSRVRALSW	2515	YP_006628019.1
1633	BORDETELLA PERTUSSIS	RTHGVGASL	2516	AGS56996.1
1634	BORDETELLA PERTUSSIS	RTRGQARSG	2517	AAA229744.1
1635	BORDETELLA PERTUSSIS	RVAPPAVAL	2518	WP_019249248.1
1636	BORDETELLA PERTUSSIS	RVLPEPVKL	2519	AGT50936.1
1637	BORDETELLA PERTUSSIS	RVRALAWLL	2520	NP_882284.1
1638	BORDETELLA PERTUSSIS	RVRALSWLL	2521	YP_006628019.1
1639	BORDETELLA PERTUSSIS	RVTVSGGSL	2522	AGT50936.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1640	BORDETELLA PERTUSSIS	SEAMAAWSE	2523	AFK26302.1
1641	BORDETELLA PERTUSSIS	SESHNFHAS	2524	WP_019247158.1
1642	BORDETELLA PERTUSSIS	SGEGRVNIG	2525	WP_019249248.1
1643	BORDETELLA PERTUSSIS	SGLAVRVHV	2526	1BCP_C
1644	BORDETELLA PERTUSSIS	SLADISLGA	2527	WP_019249248.1
1645	BORDETELLA PERTUSSIS	SLFAILSST	2528	WP_019247158.1
1646	BORDETELLA PERTUSSIS	SLFAPHGNV	2529	AAZ74338.1
1647	BORDETELLA PERTUSSIS	SLSIDNATW	2530	AGS56996.1
1648	BORDETELLA PERTUSSIS	SPMEVMLRA	2531	AAA22983.1
1649	BORDETELLA PERTUSSIS	SPQPIRATV	2532	WP_019247158.1
1650	BORDETELLA PERTUSSIS	SPRRARRAL	2533	WP_019249248.1
1651	BORDETELLA PERTUSSIS	SPSRLAGTL	2534	WP_019249248.1
1652	BORDETELLA PERTUSSIS	SSTPLGSLF	2535	WP_019247158.1
1653	BORDETELLA PERTUSSIS	STYELLDYL	2536	WP_019249248.1
1654	BORDETELLA PERTUSSIS	SVAMKPYEV	2537	AAA22983.1
1655	BORDETELLA PERTUSSIS	SVAPNALAW	2538	AAA22974.1
1656	BORDETELLA PERTUSSIS	SVKVAKKLF	2539	WP_019249248.1
1657	BORDETELLA PERTUSSIS	TAFMSGRSL	2540	AAA22984.1
1658	BORDETELLA PERTUSSIS	TAGATPFDI	2541	WP_019248658.1
1659	BORDETELLA PERTUSSIS	TAPVTSPAW	2542	AAW72734.1
1660	BORDETELLA PERTUSSIS	TARTGWLTW	2543	AAW72734.1
1661	BORDETELLA PERTUSSIS	TEAQGVQVR	2544	WP_019248145.1
166	BORDETELLA	TEVYLEHRM	2545	AAW7273

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
2	PERTUSSIS			4.1
1663	BORDETELLA PERTUSSIS	TFEGKPALE	2546	AAA22983.1
1664	BORDETELLA PERTUSSIS	TFTGKVTNG	2547	WP_019248658.1
1665	BORDETELLA PERTUSSIS	TGDGGGHTD	2548	AGS56996.1
1666	BORDETELLA PERTUSSIS	TLAKALSAA	2549	WP_019249248.1
1667	BORDETELLA PERTUSSIS	TLANVGDTW	2550	AGS56996.1
1668	BORDETELLA PERTUSSIS	TLNASNLTL	2551	WP_019249248.1
1669	BORDETELLA PERTUSSIS	TLSSAHGNV	2552	WP_019249248.1
1670	BORDETELLA PERTUSSIS	TPFDIKLKE	2553	WP_019248658.1
1671	BORDETELLA PERTUSSIS	TPFIIKLKD	2554	WP_019248145.1
1672	BORDETELLA PERTUSSIS	TPGWSIYGL	2555	1BCP_C
1673	BORDETELLA PERTUSSIS	TPLGSAATF	2556	AGS56996.1
1674	BORDETELLA PERTUSSIS	TPLGSLFAI	2557	WP_019247158.1
1675	BORDETELLA PERTUSSIS	TPLPVSLTA	2558	WP_019249248.1
1676	BORDETELLA PERTUSSIS	TRQGIMDQY	2559	YP_006626873.1
1677	BORDETELLA PERTUSSIS	TSKQDERNY	2560	WP_019247158.1
1678	BORDETELLA PERTUSSIS	TSPYDGKYW	2561	YP_006628018.1
1679	BORDETELLA PERTUSSIS	TSRRSVASI	2562	AFK26302.1
1680	BORDETELLA PERTUSSIS	TSRTVTMRY	2563	NP_879898.1
1681	BORDETELLA PERTUSSIS	TTEYPNARY	2564	ADA85123.1
1682	BORDETELLA PERTUSSIS	TTEYSNARY	2565	AFK26302.1
1683	BORDETELLA PERTUSSIS	TTLGLEQTF	2566	WP_019247158.1
1684	BORDETELLA PERTUSSIS	TVLAAGAGL	2567	WP_019249248.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1685	BORDETELLA PERTUSSIS	TVQELALKL	2568	AAA22984.1
1686	BORDETELLA PERTUSSIS	TVVQRNKHW	2569	WP_019247158.1
1687	BORDETELLA PERTUSSIS	VAAAADLAL	2570	WP_019249248.1
1688	BORDETELLA PERTUSSIS	VALASQARW	2571	AGS56996.1
1689	BORDETELLA PERTUSSIS	VAMKPYEVT	2572	AAA22983.1
1690	BORDETELLA PERTUSSIS	VARLVKLQG	2573	WP_019249248.1
1691	BORDETELLA PERTUSSIS	VAVAGGRWH	2574	AGS56996.1
1692	BORDETELLA PERTUSSIS	VEASAITTY	2575	WP_019248145.1
1693	BORDETELLA PERTUSSIS	VEDIGGKNY	2576	WP_019247158.1
1694	BORDETELLA PERTUSSIS	VEVSSPPPV	2577	WP_019247158.1
1695	BORDETELLA PERTUSSIS	VGAGGYEAG	2578	WP_019247158.1
1696	BORDETELLA PERTUSSIS	VGGGEHGRW	2579	WP_019249248.1
1697	BORDETELLA PERTUSSIS	VHVSKEEQY	2580	YP_006628018.1
1698	BORDETELLA PERTUSSIS	VIDGQKVLA	2581	AAA22974.1
1699	BORDETELLA PERTUSSIS	VIGACTSPY	2582	YP_006628018.1
1700	BORDETELLA PERTUSSIS	VKLGGVYEA	2583	AAA22974.1
1701	BORDETELLA PERTUSSIS	VLAPRLYLT	2584	AAA22974.1
1702	BORDETELLA PERTUSSIS	VLVKTNMVV	2585	AAA22983.1
1703	BORDETELLA PERTUSSIS	VPASGAPAA	2586	AGS56996.1
1704	BORDETELLA PERTUSSIS	VPFCFGKDL	2587	AAA22983.1
1705	BORDETELLA PERTUSSIS	VPVSEHCTV	2588	AAA22974.1
1706	BORDETELLA PERTUSSIS	VPVTPPKVE	2589	WP_019247158.1
170	BORDETELLA	VRTVSAMEY	2590	WP_0192

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
7	PERTUSSIS			49248.1
1708	BORDETELLA PERTUSSIS	VSGGSLSAP	2591	AGT50936.1
1709	BORDETELLA PERTUSSIS	VSSATKAKG	2592	WP_019248658.1
1710	BORDETELLA PERTUSSIS	VSSPPPVSV	2593	WP_019247158.1
1711	BORDETELLA PERTUSSIS	VSVKVAKKL	2594	WP_019249248.1
1712	BORDETELLA PERTUSSIS	VTMRYLASY	2595	WP_019248145.1
1713	BORDETELLA PERTUSSIS	VTSVAMKPY	2596	AAA22983.1
1714	BORDETELLA PERTUSSIS	VVAEKVTTP	2597	WP_019247158.1
1715	BORDETELLA PERTUSSIS	VVDGPPSRP	2598	WP_019247158.1
1716	BORDETELLA PERTUSSIS	VVETAQPLP	2599	WP_019247158.1
1717	BORDETELLA PERTUSSIS	WLTWLAILA	2600	AAW72734.1
1718	BORDETELLA PERTUSSIS	WTFHAGYRY	2601	AGS56996.1
1719	BORDETELLA PERTUSSIS	WVMTDNSNV	2602	AGS56996.1
1720	BORDETELLA PERTUSSIS	YAEHGEVSI	2603	WP_019249248.1
1721	BORDETELLA PERTUSSIS	YAIDGTAAG	2604	WP_019249248.1
1722	BORDETELLA PERTUSSIS	YALKSRIAL	2605	AAA22984.1
1723	BORDETELLA PERTUSSIS	YATNPQTQL	2606	WP_019248145.1
1724	BORDETELLA PERTUSSIS	YDTGDLIAY	2607	WP_019248658.1
1725	BORDETELLA PERTUSSIS	YEAGFSLGS	2608	WP_019247158.1
1726	BORDETELLA PERTUSSIS	YEDATFETY	2609	YP_006628018.1
1727	BORDETELLA PERTUSSIS	YENKSSTPL	2610	WP_019247158.1
1728	BORDETELLA PERTUSSIS	YEVTPTRML	2611	AAA22983.1
1729	BORDETELLA PERTUSSIS	YEYIWGLYP	2612	WP_019249248.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1730	BORDETELLA PERTUSSIS	YEYIWGLYQ	2613	YP_006626470.1
1731	BORDETELLA PERTUSSIS	YEYSKGPKL	2614	AGS56996.1
1732	BORDETELLA PERTUSSIS	YFEPGPTTD	2615	WP_019248145.1
1733	BORDETELLA PERTUSSIS	YIWGLYPTY	2616	WP_019249248.1
1734	BORDETELLA PERTUSSIS	YIWGLYQTY	2617	YP_006626470.1
1735	BORDETELLA PERTUSSIS	YLRQITPGW	2618	1BCP_C
1736	BORDETELLA PERTUSSIS	YMIYMSGLA	2619	1BCP_C
1737	BORDETELLA PERTUSSIS	YPALRAALI	2620	WP_019248658.1
1738	BORDETELLA PERTUSSIS	YPGTPGDLL	2621	AAA22984.1
1739	BORDETELLA PERTUSSIS	YPTYTEWSV	2622	WP_019249248.1
1740	BORDETELLA PERTUSSIS	YQTYTEWSV	2623	YP_006626470.1
1741	BORDETELLA PERTUSSIS	YSTGDLRAY	2624	WP_019248145.1
1742	BORDETELLA PERTUSSIS	YTLRYLASY	2625	WP_019248658.1
1743	BORDETELLA PERTUSSIS	YVLVKTNMV	2626	AAA22983.1
1744	BORDETELLA PERTUSSIS	YYDYEDATF	2627	YP_006628018.1
1745	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	AAFIALYPNSQLAPT	2628	Q7VU05
1746	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	GGAEYNLALGQRRA	2629	Q7VU04
1747	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	GGAEYNLALGQRRADA	2630	Q7VU04
1748	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	IALYPNSQLAPT	2631	Q7VU05
1749	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	KPDQGEVVAVGPGKKTED	2632	P0A339.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1750	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	KPDQGEVVAVGPGKKTEDG	2633	P0A339.1
1751	БОРДЕТЕЛЛА КОКЛЮШНАЯ 509	LAEVLDYHNFVLTQ	2634	Q7VWM1.1
1752	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	QSIALSSLMVAQAIP	2635	AAV70486.1
1753	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	SIGVLGYQKTVDHTKVNSKLSLF	2636	AAV70486.1
1754	BORDETELLA PERTUSSIS	AAHADWNNQSIVKT	2637	ABO77783.1
1755	BORDETELLA PERTUSSIS	AALGRG	2638	ABO77783.1
1756	BORDETELLA PERTUSSIS	AALGRGHSLYASYE	2639	ABO77783.1
1757	BORDETELLA PERTUSSIS	AAPLRRTTLAMALG	2640	ABO77783.1
1758	BORDETELLA PERTUSSIS	AAPLSITLQAGAHA	2641	ABO77783.1
1759	BORDETELLA PERTUSSIS	ADAQGDIVATELPS	2642	ABO77783.1
1760	BORDETELLA PERTUSSIS	ADSGFYLDATLRAS	2643	ABO77783.1
1761	BORDETELLA PERTUSSIS	AELA	2644	ABO77783.1
1762	BORDETELLA PERTUSSIS	AELAVFRAGGGAYR	2645	ABO77783.1
1763	BORDETELLA PERTUSSIS	AELQFRNGSVTSSG	2646	ABO77783.1
1764	BORDETELLA PERTUSSIS	AGGRWHLGGLAGYT	2647	ABO77783.1
1765	BORDETELLA PERTUSSIS	AGVAAMQGAVVHLQ	2648	ABO77783.1
1766	BORDETELLA PERTUSSIS	AGYTRGDRGFTGDG	2649	ABO77783.1
1767	BORDETELLA PERTUSSIS	ALASQARWTGATRA	2650	ABO77783.1
1768	BORDETELLA PERTUSSIS	AMPWTFHAGYRYSW	2651	ABO77783.1
1769	BORDETELLA PERTUSSIS	AMQGAVVHLQRATIRRGDAP	2652	ABO77783.1
1770	BORDETELLA PERTUSSIS	ANGLRVRDE	2653	ABO77783.1
177	BORDETELLA	ANGLRVRDEGGSSV	2654	ABO7778

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1	PERTUSSIS			3.1
1772	BORDETELLA PERTUSSIS	ANKDGKVDIGTYRY	2655	ABO77783.1
1773	BORDETELLA PERTUSSIS	APAAVSVLGASELT	2656	ABO77783.1
1774	BORDETELLA PERTUSSIS	APPAPKPAPQPGPQ	2657	ABO77783.1
1775	BORDETELLA PERTUSSIS	AQGILLENPAAELQ	2658	ABO77783.1
1776	BORDETELLA PERTUSSIS	ARWTGATRAVDSLS	2659	ABO77783.1
1777	BORDETELLA PERTUSSIS	ASLEAGRRFTHADG	2660	ABO77783.1
1778	BORDETELLA PERTUSSIS	ASYEYSKGPKLAMP	2661	ABO77783.1
1779	BORDETELLA PERTUSSIS	ATFTLANKD	2662	ABO77783.1
1780	BORDETELLA PERTUSSIS	ATFTLANKDGKVDI	2663	ABO77783.1
1781	BORDETELLA PERTUSSIS	ATRAVDSLSIDNAT	2664	ABO77783.1
1782	BORDETELLA PERTUSSIS	DDDGIALYVAGEQAQ	2665	ABO77783.1
1783	BORDETELLA PERTUSSIS	DDGIALYVAGEQAQ	2666	ABO77783.1
1784	BORDETELLA PERTUSSIS	DGGHITGGRAAGVA	2667	ABO77783.1
1785	BORDETELLA PERTUSSIS	DGIRRFLGTVTVKAGK	2668	ABO77783.1
1786	BORDETELLA PERTUSSIS	DGSVDFQQPAEAGR	2669	ABO77783.1
1787	BORDETELLA PERTUSSIS	DGYAVKGKYRTHGV	2670	ABO77783.1
1788	BORDETELLA PERTUSSIS	DIVATELPSIPGTS	2671	ABO77783.1
1789	BORDETELLA PERTUSSIS	DKLVVMQDASGQHR	2672	ABO77783.1
1790	BORDETELLA PERTUSSIS	DLGLSDKLVVMQDA	2673	ABO77783.1
1791	BORDETELLA PERTUSSIS	DNATWVMTDNSNVGA	2674	ABO77783.1
1792	BORDETELLA PERTUSSIS	DNATWVMTDNSNVGALRLA	2675	ABO77783.1
1793	BORDETELLA PERTUSSIS	DNRAGRRFDQKVAG	2676	ABO77783.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1794	BORDETELLA PERTUSSIS	EAGRFKVLTVNTLA	2677	ABO77783.1
1795	BORDETELLA PERTUSSIS	ELAQSIVEAPELGA	2678	ABO77783.1
1796	BORDETELLA PERTUSSIS	ELGAAIRVGRGARV	2679	ABO77783.1
1797	BORDETELLA PERTUSSIS	ELGADHAVAVAGGR	2680	ABO77783.1
1798	BORDETELLA PERTUSSIS	ELPSIPGTSIGPLD	2681	ABO77783.1
1799	BORDETELLA PERTUSSIS	EPVKLTLTGGADAQ	2682	ABO77783.1
1800	BORDETELLA PERTUSSIS	EQAQASIADSTLQG	2683	ABO77783.1
1801	BORDETELLA PERTUSSIS	ERGANVTVQRSAIV	2684	ABO77783.1
1802	BORDETELLA PERTUSSIS	ERQHGIHIQGSDPG	2685	ABO77783.1
1803	BORDETELLA PERTUSSIS	EVGKRIELAGGRQV	2686	ABO77783.1
1804	BORDETELLA PERTUSSIS	FDGAGTVHTNGIAH	2687	ABO77783.1
1805	BORDETELLA PERTUSSIS	FQQPAEAGRFKVLT	2688	ABO77783.1
1806	BORDETELLA PERTUSSIS	FRAGGGAYRAANGL	2689	ABO77783.1
1807	BORDETELLA PERTUSSIS	GAHAQGKALLYRVL	2690	ABO77783.1
1808	BORDETELLA PERTUSSIS	GARVTVSGGSLSAP	2691	ABO77783.1
1809	BORDETELLA PERTUSSIS	GAYRAANGLRVRDE	2692	ABO77783.1
1810	BORDETELLA PERTUSSIS	GDAPAGGAVPGGAV	2693	ABO77783.1
1811	BORDETELLA PERTUSSIS	GGAVPGGAVPGGFG	2694	ABO77783.1
1812	BORDETELLA PERTUSSIS	GGAVPGGFGPVLDG	2695	ABO77783.1
1813	BORDETELLA PERTUSSIS	GGFGPVLDGWYGVD	2696	ABO77783.1
1814	BORDETELLA PERTUSSIS	GGLHIGALQSLQPE	2697	ABO77783.1
1815	BORDETELLA PERTUSSIS	GGVQIERGANVTVQ	2698	ABO77783.1
181	BORDETELLA	GHSLYASYEYSKGP	2699	ABO7778

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
6	PERTUSSIS			3.1
1817	BORDETELLA PERTUSSIS	GHTDSVHVGGYATY	2700	ABO77783.1
1818	BORDETELLA PERTUSSIS	GIAHRTELRGTRAE	2701	ABO77783.1
1819	BORDETELLA PERTUSSIS	GKALLYRVLPEPVK	2702	ABO77783.1
1820	BORDETELLA PERTUSSIS	GLGMAAALGRGHSL	2703	ABO77783.1
1821	BORDETELLA PERTUSSIS	GNVIETGGARRFAP	2704	ABO77783.1
1822	BORDETELLA PERTUSSIS	GPLDVALASQARWT	2705	ABO77783.1
1823	BORDETELLA PERTUSSIS	GQHRLWVRN	2706	ABO77783.1
1824	BORDETELLA PERTUSSIS	GRGFAQRQQLDNRA	2707	ABO77783.1
1825	BORDETELLA PERTUSSIS	GRLGLEVGKRIELA	2708	ABO77783.1
1826	BORDETELLA PERTUSSIS	GRQVQPYIKASVLQ	2709	ABO77783.1
1827	BORDETELLA PERTUSSIS	GRRFTHADGWFLEPQAELA	2710	ABO77783.1
1828	BORDETELLA PERTUSSIS	GSEPASANTLLLVQ	2711	ABO77783.1
1829	BORDETELLA PERTUSSIS	GSSVLGRLGLEVGK	2712	ABO77783.1
1830	BORDETELLA PERTUSSIS	GTTIKVSGRQAQGI	2713	ABO77783.1
1831	BORDETELLA PERTUSSIS	GTVTVKAGKLVADH	2714	ABO77783.1
1832	BORDETELLA PERTUSSIS	HAVAVAGGRWHLGG	2715	ABO77783.1
1833	BORDETELLA PERTUSSIS	IELAGGRQVQPYIK	2716	ABO77783.1
1834	BORDETELLA PERTUSSIS	IHIQGSDPGGVRTA	2717	ABO77783.1
1835	BORDETELLA PERTUSSIS	IRRFLGTVTVKAGK	2718	ABO77783.1
1836	BORDETELLA PERTUSSIS	IRVGRGARVTVSGG	2719	ABO77783.1
1837	BORDETELLA PERTUSSIS	ITLQAGAHA	2720	ABO77783.1
1838	BORDETELLA PERTUSSIS	ITLQAGAHAQGKAL	2721	ABO77783.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1839	BORDETELLA PERTUSSIS	IVEAPELGAAIRVG	2722	ABO77783.1
1840	BORDETELLA PERTUSSIS	IVKTGERQHGIHIQ	2723	ABO77783.1
1841	BORDETELLA PERTUSSIS	KAGKLVADHATLAN	2724	ABO77783.1
1842	BORDETELLA PERTUSSIS	KGKYRTHGVGASLE	2725	ABO77783.1
1843	BORDETELLA PERTUSSIS	KPAPQPGPQPPQPP	2726	ABO77783.1
1844	BORDETELLA PERTUSSIS	KVAGFELGADHAVA	2727	ABO77783.1
1845	BORDETELLA PERTUSSIS	KVAGSDGYAVKGKY	2728	ABO77783.1
1846	BORDETELLA PERTUSSIS	KVDIGTYRYRLAAN	2729	ABO77783.1
1847	BORDETELLA PERTUSSIS	KVLTVNTLAGSGLF	2730	ABO77783.1
1848	BORDETELLA PERTUSSIS	LAANGNGQWSLVGA	2731	ABO77783.1
1849	BORDETELLA PERTUSSIS	LAMPWTFHAGYRYS	2732	ABO77783.1
1850	BORDETELLA PERTUSSIS	LASTLWYAESNALS	2733	ABO77783.1
1851	BORDETELLA PERTUSSIS	LENDFKVAGSDGYA	2734	ABO77783.1
1852	BORDETELLA PERTUSSIS	LENPAAELQFRNGS	2735	ABO77783.1
1853	BORDETELLA PERTUSSIS	LGAAPAAHADWNNQ	2736	ABO77783.1
1854	BORDETELLA PERTUSSIS	LGGLAGYTRGDRGFTGDG	2737	ABO77783.1
1855	BORDETELLA PERTUSSIS	LLENP	2738	ABO77783.1
1856	BORDETELLA PERTUSSIS	LLVQTPLGSAATFT	2739	ABO77783.1
1857	BORDETELLA PERTUSSIS	LPPSRVVLRDTNVT	2740	ABO77783.1
1858	BORDETELLA PERTUSSIS	LQPEDLPPS	2741	ABO77783.1
1859	BORDETELLA PERTUSSIS	LQPEDLPPSRVVLR	2742	ABO77783.1
1860	BORDETELLA PERTUSSIS	LRASRLENDFKVAG	2743	ABO77783.1
186	BORDETELLA	LRLASDGSVDFQQP	2744	ABO7778

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1	PERTUSSIS			3.1
1862	BORDETELLA PERTUSSIS	LSAAANAAVNTGGV	2745	ABO77783.1
1863	BORDETELLA PERTUSSIS	LSAPHGNVIETGGA	2746	ABO77783.1
1864	BORDETELLA PERTUSSIS	LSDDGIRRFLGTVT	2747	ABO77783.1
1865	BORDETELLA PERTUSSIS	LVGAKAPPAPKPAP	2748	ABO77783.1
1866	BORDETELLA PERTUSSIS	LYVAGEQAQASIAD	2749	ABO77783.1
1867	BORDETELLA PERTUSSIS	MALGALGAAPAAHA	2750	ABO77783.1
1868	BORDETELLA PERTUSSIS	MNMSLSRIVKAAPL	2751	ABO77783.1
1869	BORDETELLA PERTUSSIS	MQDASGQHR	2752	ABO77783.1
1870	BORDETELLA PERTUSSIS	MQGAVVHLQRATIR	2753	ABO77783.1
1871	BORDETELLA PERTUSSIS	NAAVNTGGVGLAST	2754	ABO77783.1
1872	BORDETELLA PERTUSSIS	NALSKRLGELRLNP	2755	ABO77783.1
1873	BORDETELLA PERTUSSIS	NGQWSLVGAKAPPA	2756	ABO77783.1
1874	BORDETELLA PERTUSSIS	NTLAGSGLFRMNVF	2757	ABO77783.1
1875	BORDETELLA PERTUSSIS	PAGRELSAAANAAV	2758	ABO77783.1
1876	BORDETELLA PERTUSSIS	PAPQPPAGRELSAA	2759	ABO77783.1
1877	BORDETELLA PERTUSSIS	PGPQPPQPPQPQPE	2760	ABO77783.1
1878	BORDETELLA PERTUSSIS	PGTSIGPLDVALAS	2761	ABO77783.1
1879	BORDETELLA PERTUSSIS	PLGSAATFTLANKD	2762	ABO77783.1
1880	BORDETELLA PERTUSSIS	PQPEAPAPQPPAGR	2763	ABO77783.1
1881	BORDETELLA PERTUSSIS	PQPPQPQPEAPAPQ	2764	ABO77783.1
1882	BORDETELLA PERTUSSIS	PYIKASVLQEFDGA	2765	ABO77783.1
1883	BORDETELLA PERTUSSIS	RFAPQAAPLSITLQ	2766	ABO77783.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1884	BORDETELLA PERTUSSIS	RLGELRLNPDAGGA	2767	ABO77783.1
1885	BORDETELLA PERTUSSIS	RLNPDAGGAWGRGF	2768	ABO77783.1
1886	BORDETELLA PERTUSSIS	RNGSVTSSGQLSDD	2769	ABO77783.1
1887	BORDETELLA PERTUSSIS	RRFDQKVAGFELGA	2770	ABO77783.1
1888	BORDETELLA PERTUSSIS	RTTLAMALGALGAA	2771	ABO77783.1
1889	BORDETELLA PERTUSSIS	SAIVDGGLHIGALQ	2772	ABO77783.1
1890	BORDETELLA PERTUSSIS	SANTLLLVQTPLGS	2773	ABO77783.1
1891	BORDETELLA PERTUSSIS	SDPGGVRTASGTTI	2774	ABO77783.1
1892	BORDETELLA PERTUSSIS	SELTLDGGHITGGR	2775	ABO77783.1
1893	BORDETELLA PERTUSSIS	SGLFRMNVF	2776	ABO77783.1
1894	BORDETELLA PERTUSSIS	SGLFRMNVFADLGL	2777	ABO77783.1
1895	BORDETELLA PERTUSSIS	SGSSVELAQSIVEA	2778	ABO77783.1
1896	BORDETELLA PERTUSSIS	SIADSTLQGAGGVQ	2779	ABO77783.1
1897	BORDETELLA PERTUSSIS	SKGPKLAMPWTFHA	2780	ABO77783.1
1898	BORDETELLA PERTUSSIS	SNVGALRLASDGSV	2781	ABO77783.1
1899	BORDETELLA PERTUSSIS	SRIVKAAPLRRTTL	2782	ABO77783.1
1900	BORDETELLA PERTUSSIS	SVLGASELTLDGGH	2783	ABO77783.1
1901	BORDETELLA PERTUSSIS	SVLQEFDGA	2784	ABO77783.1
1902	BORDETELLA PERTUSSIS	SVLQEFDGAGTVHT	2785	ABO77783.1
1903	BORDETELLA PERTUSSIS	TELR	2786	ABO77783.1
1904	BORDETELLA PERTUSSIS	TELRGTRAELGLGM	2787	ABO77783.1
1905	BORDETELLA PERTUSSIS	TGDGGGHTDSVHVG	2788	ABO77783.1
190	BORDETELLA	TGGARRFAPQAAPL	2789	ABO7778

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
6	PERTUSSIS			3.1
1907	BORDETELLA PERTUSSIS	TGGRAAGVAAMQGA	2790	ABO77783.1
1908	BORDETELLA PERTUSSIS	TGGVGLASTLWYAE	2791	ABO77783.1
1909	BORDETELLA PERTUSSIS	THGVGASLEAGRRF	2792	ABO77783.1
1910	BORDETELLA PERTUSSIS	TIRRGDAPA	2793	ABO77783.1
1911	BORDETELLA PERTUSSIS	TLANVGDTWDDDGI	2794	ABO77783.1
1912	BORDETELLA PERTUSSIS	TLQGAGGVQIERGA	2795	ABO77783.1
1913	BORDETELLA PERTUSSIS	TLTGGADAQGDIVA	2796	ABO77783.1
1914	BORDETELLA PERTUSSIS	TNVTAVPASGAPAA	2797	ABO77783.1
1915	BORDETELLA PERTUSSIS	TRAELGLGMAAALG	2798	ABO77783.1
1916	BORDETELLA PERTUSSIS	TSSGQLSDDGIRRF	2799	ABO77783.1
1917	BORDETELLA PERTUSSIS	TVHTNGIAHRTELR	2800	ABO77783.1
1918	BORDETELLA PERTUSSIS	TYRYRLAANGNGQW	2801	ABO77783.1
1919	BORDETELLA PERTUSSIS	VADHATLANVGDTW	2802	ABO77783.1
1920	BORDETELLA PERTUSSIS	VHVGGYATYIADSG	2803	ABO77783.1
1921	BORDETELLA PERTUSSIS	VLDGWYGVD	2804	ABO77783.1
1922	BORDETELLA PERTUSSIS	VPASGAPAAVSVLG	2805	ABO77783.1
1923	BORDETELLA PERTUSSIS	VRDEGGSSVLGRLG	2806	ABO77783.1
1924	BORDETELLA PERTUSSIS	VRTASGTTIKVSGR	2807	ABO77783.1
1925	BORDETELLA PERTUSSIS	VSGGSLSAPHGNVI	2808	ABO77783.1
1926	BORDETELLA PERTUSSIS	VSGRQAQGILLENP	2809	ABO77783.1
1927	BORDETELLA PERTUSSIS	VTVQRSAIVDGGLH	2810	ABO77783.1
1928	BORDETELLA PERTUSSIS	VVLRDTNVTAVPAS	2811	ABO77783.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
1929	BORDETELLA PERTUSSIS	WNNQSIVKTGERQH	2812	ABO77783.1
1930	BORDETELLA PERTUSSIS	WVRNSGSEPASANT	2813	ABO77783.1
1931	BORDETELLA PERTUSSIS	WYAESNALSKRLGE	2814	ABO77783.1
1932	BORDETELLA PERTUSSIS	YATYIADSGFYLDA	2815	ABO77783.1
1933	BORDETELLA PERTUSSIS	YGVDVSGSS	2816	ABO77783.1
1934	BORDETELLA PERTUSSIS	YGVDVSGSSVELAQ	2817	ABO77783.1
1935	BORDETELLA PERTUSSIS	YLDATLRASRLEND	2818	ABO77783.1
1936	BORDETELLA PERTUSSIS	YRVLPEPVKLTLTG	2819	ABO77783.1
1937	BORDETELLA PERTUSSIS	VKAQNITNKRAALIEA	2820	AAA22974.1
1938	BORDETELLA PERTUSSIS	YYSNVTATRLLSSTNS	2821	AAA83981.1
1939	BORDETELLA PERTUSSIS	SPNLTDERAAQAGVT	2822	CPP72976.1
1940	МОРБИЛЛИВИРУС КОРИ	SSRASDERAAHLPTS	2823	BAA33867.1
1941	КОРИНЕБАКТЕРИЯ ДИФТЕРИИ	QVVHNSYNRPAYSPG	2824	1007216A
1942	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AEGGEIHEL	2825	AAF85692.1
1943	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AENLISNGIGKY	2826	AAF85698.1
1944	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AEVDGDVKL	2827	CAB43772.1
1945	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AIYTAEIHK	2828	AAF85697.1
1946	ШТАММ ВИРУСА КОРИ EDMONSTON-B	APVFHMTNY	2829	CAB43772.1
1947	ШТАММ ВИРУСА КОРИ EDMONSTON-B	APVFHMTNYLEQPVSN	2830	AAR89413.1
194	ШТАММ ВИРУСА	AQRLNEIY	2831	AAF8569

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
8	КОРИ EDMONSTON-B			8.1
1949	ШТАММ ВИРУСА КОРИ EDMONSTON-B	ARVPHAYSL	2832	AAF85698.1
1950	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AVRDLERAM	2833	P03424.1
1951	ШТАММ ВИРУСА КОРИ EDMONSTON-B	AVRDLERAMTTLK	2834	P03424.1
1952	ШТАММ ВИРУСА КОРИ EDMONSTON-B	DALLRLQAM	2835	Q89933.1
1953	ШТАММ ВИРУСА КОРИ EDMONSTON-B	DIKEKVINL	2836	AAF85698.1
1954	ШТАММ ВИРУСА КОРИ EDMONSTON-B	DQGLFKVL	2837	AAF85695.1
1955	ШТАММ ВИРУСА КОРИ EDMONSTON-B	DTGVDTRIW	2838	Q9EMA9.1
1956	ШТАММ ВИРУСА КОРИ EDMONSTON-B	EPIGSLAIEEAM	2839	AAF85692.1
1957	ШТАММ ВИРУСА КОРИ EDMONSTON-B	EPIRDALNAM	2840	P69354.1
1958	ШТАММ ВИРУСА КОРИ EDMONSTON-B	FPKLGKTL	2841	AAF85692.1
1959	ШТАММ ВИРУСА КОРИ EDMONSTON-B	FRSVNAVAF	2842	AAF85695.1
1960	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GKIIDNTEQL	2843	AAF85695.1
1961	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GLNEKLVFY	2844	AAF85695.1
1962	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GMYGGTYLVEK	2845	AAC35876.2
196	ШТАММ ВИРУСА	GPPISLERLDVGTN	2846	P69354.1

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
3	КОРИ EDMONSTON-B			
1964	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GPRQAQVSFL	2847	Q89933.1
1965	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GRLVPQVRVID	2848	AAF85695.1
1966	ШТАММ ВИРУСА КОРИ EDMONSTON-B	GSAPISMGFR	2849	AAF85692.1
1967	ШТАММ ВИРУСА КОРИ EDMONSTON-B	HILAKSTAL	2850	AAF85698.1
1968	ШТАММ ВИРУСА КОРИ EDMONSTON-B	HYREVNLVY	2851	AAF85698.1
1969	ШТАММ ВИРУСА КОРИ EDMONSTON-B	IPPMKNLAL	2852	AAC35876.2
1970	ШТАММ ВИРУСА КОРИ EDMONSTON-B	IPYQGSGKGVSF	2853	CAB43772.1
1971	ШТАММ ВИРУСА КОРИ EDMONSTON-B	ISKESQHVY	2854	AAF85698.1
1972	ШТАММ ВИРУСА КОРИ EDMONSTON-B	IVSSHFFVY	2855	AAF85698.1
1973	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KEIKETGRLF	2856	AAF85698.1
1974	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KESQHVYYL	2857	AAF85698.1
1975	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KIIDNTEQL	2858	AAF85695.1
1976	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KKQINRQN	2859	AAA63285.1
1977	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KKVDTNFIYQ	2860	AAF85698.1
197	ШТАММ ВИРУСА	KLIDGFFPA	2861	AAF8569

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
8	КОРИ EDMONSTON-B			8.1
1979	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KPNLSSKRSEL	2862	BAB39848.1
1980	ШТАММ ВИРУСА КОРИ EDMONSTON-B	KVSPYLFTV	2863	AAR89413.1
1981	ШТАММ ВИРУСА КОРИ EDMONSTON-B	LETRTTNQFL	2864	CAB43772.1
1982	ШТАММ ВИРУСА КОРИ EDMONSTON-B	LLKEATEL	2865	AAF85695.1
1983	ШТАММ ВИРУСА КОРИ EDMONSTON-B	LLKKGNSLY	2866	AAF85698.1
1984	ШТАММ ВИРУСА КОРИ EDMONSTON-B	LPAPIGGMNY	2867	AAF85698.1
1985	ШТАММ ВИРУСА КОРИ EDMONSTON-B	MPEETLHQVM	2868	AAF85698.1
1986	ШТАММ ВИРУСА КОРИ EDMONSTON-B	PTTIRGQFS	2869	CAB43772.1
1987	ШТАММ ВИРУСА КОРИ EDMONSTON-B	QEISRHQALGY	2870	P03424.1
1988	ШТАММ ВИРУСА КОРИ EDMONSTON-B	RITHVDTESY	2871	P69354.1
1989	ШТАММ ВИРУСА КОРИ EDMONSTON-B	RPGLKPDL	2872	P69354.1
1990	ШТАММ ВИРУСА КОРИ EDMONSTON-B	RPIYGLEV	2873	AAF85698.1
1991	ШТАММ ВИРУСА КОРИ EDMONSTON-B	RQAGQEMILAV	2874	P69354.1
1992	ШТАММ ВИРУСА КОРИ EDMONSTON-B	SAVRIATVY	2875	AAF85698.1
199	ШТАММ ВИРУСА	SLMPEETLHQV	2876	AAF8569

#	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	ВСПОМ. #
3	КОРИ EDMONSTON-B			8.1
1994	ШТАММ ВИРУСА КОРИ EDMONSTON-B	SMIDLVTKF	2877	AAF85698.1
1995	ШТАММ ВИРУСА КОРИ EDMONSTON-B	SMLNSQAIDNLRA	2878	P69354.1
1996	ШТАММ ВИРУСА КОРИ EDMONSTON-B	SMYRVFEV	2879	CAB43772.1
1997	ШТАММ ВИРУСА КОРИ EDMONSTON-B	SQQGMFHAY	2880	AAF85698.1
1998	ШТАММ ВИРУСА КОРИ EDMONSTON-B	TDTPIVYNDRNLLD	2881	Q89933.1
1999	ШТАММ ВИРУСА КОРИ EDMONSTON-B	VIINDDQGLFKV	2882	AAF85695.1
2000	ШТАММ ВИРУСА КОРИ EDMONSTON-B	YESGVRIASL	2883	AAF85698.1
2001	ШТАММ ВИРУСА КОРИ EDMONSTON-B	YLKDKALA	2884	AAF85698.1
2002	ШТАММ ВИРУСА КОРИ EDMONSTON-B	YVYDHSGEAVK	2885	AAF85692.1
2003	ВИРУС КРАСНУХИ	ARVIDPAAQSFTGVV	2886	BAA28178.1
2004	ВИРУС КРАСНУХИ	SDRASARVIDPAAQS	2887	BAA28178.1
2005	ВИРУС КРАСНУХИ	VPPGKFVTAALLNTP	2888	BAA28178.1
2006	ВИРУС КРАСНУХИ	WVTPVIGSQARKCGL	2889	BAA28178.1
2007	РУБУЛАВИРУС ПАРОТИТА	GTYRLIPNARANLTA	400	AGC97176.1

I. Доставка праймерных редакторов

[0820] В другом аспекте настоящее изобретение предусматривает доставку праймерных редакторов *in vitro* и *in vivo* с помощью различных стратегий, включая отдельные векторы с использованием разделенных интеинов, а также стратегии

прямой доставки рибонуклеопротеинового комплекса (то есть праймерного редактора в виде комплекса с PEgRNA и/или нРНК для второго участка) с помощью таких способов как электропорация, применение катионных липид-опосредованных составов, и способов индуцированного эндоцитоза с использованием рецепторных лигандов, слитых с рибонуклеопротеиновыми комплексами. Здесь охватываются любые подобные способы.

Обзор вариантов способов доставки

[0821] В некоторых аспектах изобретение предусматривает способы, включающие доставку одного или нескольких кодирующих праймерный редактор полинуклеотидов, например одного или нескольких векторов, как описано здесь, кодирующих один или несколько компонентов системы праймерного редактирования, описанной здесь, одного или нескольких ее транскриптов, и/или одного или нескольких транскрибированных из нее белков в клетку-хозяин. В некоторых аспектах изобретение дополнительно предусматривает клетки, полученные с помощью таких способов, и организмы (такие как животные, растения, или грибы), состоящие или полученные из таких клеток. В некоторых вариантах осуществления праймерный редактор, как описано здесь, в комбинации (и, необязательно, в виде комплекса) с направляющей последовательностью доставляется в клетку. Для внесения нуклеиновых кислот в клетки млекопитающих или ткани-мишени могут применяться стандартные способы переноса генов, которые основаны или не основаны на вирусах. Такие способы могут применяться для введения нуклеиновых кислот, кодирующих компоненты праймерного редактора, в клетки в культуре или в организм-хозяин. Невирусные векторные системы доставки включают ДНК плазмиды, РНК (например описанный здесь транскрипт вектора), "голую" нуклеиновую кислоту, и нуклеиновую кислоту в виде комплекса со средством для доставки, например липосомой. Вирусные векторные системы доставки включают ДНК и РНК вирусы, которые имеют либо эписомальный, или встроенный геном после доставки в клетку. См. обзор процедур генной терапии в Anderson, Science 256:808-813 (1992); Nabel & Felgner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10):1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada et al., в Current

Topics in Microbiology and Immunology Doerfler and Bihm (eds) (1995); и Yu et al., Gene Therapy 1:13-26 (1994).

[0822] Способы невирусной доставки нуклеиновых кислот включают липофекцию, нуклеофекцию, микроинъекцию, биолистику, вирусомы, липосомы, иммунолипосомы, конъюгаты поликатион или липид:нуклеиновая кислота, голую ДНК, искусственные вирионы, и агент-энхансированный захват ДНК. Липофекция описана, например, в патентах США 5049386, 4946787, и 4897355), а реагенты для липофекции являются коммерчески доступными (например Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые являются подходящими для эффективной липофекции полинуклеотидов на основе распознавания рецептора, включают указанные у Feigner, WO 91/17424; WO 91/16024. Доставка может осуществляться в клетки (например при введении *in vitro* или *ex vivo*) или ткани-мишени (например при введении *in vivo*).

[0823] Получение комплексов липид:нуклеиновая кислота, включая липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области (см., например g., Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); патенты США 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028, и 4946787).

[0824] Применение систем на основе вирусной РНК или ДНК для доставки нуклеиновых кислот основано на продвинутых процессах нацеливания вируса на определенные клетки в организме и направлении полезной нагрузки вируса в ядро. Вирусные векторы могут быть введена напрямую пациентам (*in vivo*) или они могут быть использованы для лечения клеток *in vitro*, при этом модифицированные клетки могут необязательно быть введены пациентам (*ex vivo*). Стандартные системы на основе вирусов могут включать ретровирусные, лентивирусные, аденовирусные, адено-ассоциированные и основанные на вирусе простого герпеса векторы для переноса генов. Встраивание в геном хозяина возможно в случае способов на основе переноса генов с помощью ретровируса, лентивируса и адено-ассоциированного вируса, что зачастую приводит к долгосрочной экспрессии вставленного трансгена. Кроме того, высокие показатели эффективности трансдукции наблюдались в случае многих различных типов клеток и тканей-мишеней.

[0825] Тропизм вирусов может быть изменен за счет встраивания белков чужеродной оболочки, что обеспечивает увеличение потенциальной популяции клеток-мишеней. Лентивирусные векторы представляют собой ретровирусные векторы, которые способны трансдуцировать или инфицировать неделящиеся клетки и обычно продуцируют высокие вирусные титры. Таким образом, выбор ретровирусной системы для трансфера генов будет зависеть от ткани-мишени. Ретровирусные векторы состоят из cis-действующих длинных концевых повторов с показателем упаковывания до 6-10 kb чужеродной последовательности. Минимальные cis-действующие LTR являются достаточными для репликации и упаковывания векторов, которые затем используются для встраивания терапевтического гена в клетку-мишень для обеспечения перманентной экспрессии трансгена. Широко применяемые ретровирусные векторы включают основанные на вирусе лейкоза мышей (MuLV), вирусе лейкоза гиббонов (GaLV), вируес иммунодефицита обезьян (SIV), вирусе иммунодефицита человека (ВИЧ), и их комбинациях (см., например, Buchscher et al., J. Virol. 66:2731-2739 (1992); Johann et al., J. Virol. 66:1635-1640 (1992); Sommnerfelt et al., Virol. 176:58-59 (1990); Wilson et al., J. Virol. 63:2374-2378 (1989); Miller et al., J. Virol. 65:2220-2224 (1991); PCT/US94/05700). В рамках вариантах использования, где является предпочтительной транзиентная экспрессия, могут использоваться системы на основе аденовирусов. Векторы на основе аденовирусов способны обеспечивать очень высокую эффективность трансдукции во многих типах клеток и не требуют разделения клеток. С помощью подобных векторов был достигнут высокий титр и уровни экспрессии. Данный вектор может быть получен в больших количествах в рамках относительно простой системы. Векторы на основе адено-ассоциированного вируса ("AAV") также могут применяться для трансдукции клеток с использованием нуклеиновых кислот-мишеней, например при *in vitro* получении нуклеиновых кислот и пептидов и для *in vivo* и *ex vivo* процедур генной терапии (см., например West et al., Virology 160:38-47 (1987); патент США 4797368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). Конструирование рекомбинантных AAV векторов описано в ряде публикаций, включая патент США 5173414; Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); и Samulski et al., J. Virol. 63:03822-3828 (1989).

[0826] Упаковывающие клетки обычно используются для образования вирусных частиц, которые способны инфицировать клетку-хозяина. Такие клетки включают клетки 293, которые упаковывают аденовирус, и клетки ψ2 или клетки PA317, которые упаковывают ретровирус. Вирусные векторы. используемые в рамках генной терапии обычно получают путем продуцирования линии клеток, которые упаковывают вектор нуклеиновой кислоты в вирусную частицу. Векторы обычно содержат минимальные вирусные последовательности, требуемые для упаковывания и последующего встраивания в хозяина, при этом другие вирусные последовательности замещаются кассетой экспрессии для подвергаемого экспрессии полинуклеотида (полинуклеотидов). Отсутствующие вирусные функции обычно обеспечиваются in trans за счет линии упаковывающих клеток. Например, AAV векторы, используемые в рамках генной терапии, обычно обладают только ITR последовательностями из генома AAV, которые необходимы для упаковывания и встраивания в геном хозяина. Вирусная ДНК упаковывается в линии клеток, которая содержит хелперную плазмиду, кодирующую другие гены AAV, а именно rep и cap, но не имеет ITR последовательностей. Линия клеток также может быть инфицирована аденовирусом в качестве хелпера. Хелперный вирус промотирует репликацию AAV вектора и экспрессию генов AAV из хелперной плазмиды. Хелперная плазмида не упаковывается в значительных количествах вследствие отсутствия ITR последовательностей. Контаминация аденовирусом может быть снижена, например, с помощью тепловой обработки, к которой аденовирус является более чувствительным, чем AAV. Дополнительные способы доставки нуклеиновых кислот в клетки известны специалистам в данной области. См., например, публикацию US20030087817, приведенную здесь в качестве ссылки.

[0827] В различных вариантах осуществления конструкты PE (включая разделенные конструкты) могут быть модифицированы для доставки в одном или нескольких rAAV векторах. rAAV в рамках любого из предусмотренных здесь способов и композиций может иметь любой серотип, включая любой производный тип или псевдотип (например 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 2/1, 2/5, 2/8, 2/9, 3/1, 3/5, 3/8, или 3/9). rAAV может включать генетическую нагрузку (то есть вектор в виде рекомбинантной нуклеиновой кислоты, которая экспрессирует представляющий интерес ген, например целый или разделенный слитый белок PE, который переносится rAAV в клетку), которая предназначена для доставки в клетку. rAAV может являться химерным.

[0828] В соответствии с используемым здесь значением термин "серотип" rAAV обозначает серотип белков капсида рекомбинантного вируса. Неограничивающие примеры производных типов и псевдотипов включают rAAV2/1, rAAV2/5, rAAV2/8, rAAV2/9, гибрид AAV2-AAV3, AAVrh.10, AAVhu.14, AAV3a/3b, AAVrh32.33, AAV-HSC15, AAV-HSC17, AAVhu.37, AAVrh.8, CHt-P6, AAV2.5, AAV6.2, AAV2i8, AAV-HSC15/17, AAVM41, AAV9.45, AAV6(Y445F/Y731F), AAV2.5T, AAV-HAE1/2, клон AAV 32/83, AAVShH10, AAV2 (Y->F), AAV8 (Y733F), AAV2.15, AAV2.4, AAVM41, и AAVr3.45. Неограничивающий пример производных типов и псевдотипов, которые имеют химерные белки VP1, представляет собой rAAV2/5-1VP1u, который имеет геном AAV2, капсидный каркас AAV5 и VP1u AAV1. Другой неограничивающий пример производных типов и псевдотипов, которые имеют химерные белки VP1, представляет собой AAV2/5-8VP1u, rAAV2/9-1VP1u, и rAAV2/9-8VP1u.

[0829] Производные типа/псевдотипы AAV и способы получения таких производных типов/псевдотипов известны в данной области (см., например, Mol Ther. 2012 Apr;20(4):699-708. doi: 10.1038/mt.2011.287. Epub 2012 Jan 24. The AAV vector toolkit: poised at the clinical crossroads. Asokan A1, Schaffer DV, Samulski RJ.). Способы получения и применения псевдотипированных rAAV векторов известны в данной области (см., например, Duan et al., J. Virol., 75:7662-7671, 2001; Halbert et al., J. Virol., 74:1524-1532, 2000; Zolotukhin et al., Methods, 28:158-167, 2002; и Auricchio et al., Hum. Molec. Genet., 10:3075-3081, 2001).

[0830] Способы получения или упаковывания частиц rAAV известны в данной области, при этом соответствующие реагенты являются коммерчески доступными (см., например, Zolotukhin et al. Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. Methods 28 (2002) 158–167; и публикации патентов США US20070015238 и US20120322861, которые приведены здесь в качестве ссылки; при этом плазмиды и наборы доступны от ATCC и Cell Biolabs, Inc.). Например, плазмида, включающая представляющий интерес ген, может быть объединена с одной или несколькими хелперными плазмидами, например которые содержат ген rep (например кодирующий Rep78, Rep68, Rep52 и Rep40) и ген cap (кодирующий VP1, VP2, и VP3, включая модифицированный участок VP2, как описано здесь), и транфецированы в рекомбинантные клетки, так что частица rAAV может быть упакована и далее очищена.

[0831] Рекомбинантный AAV может включать вектор нуклеиновой кислоты, которая может включать, как минимум: (a) один или несколько гетерологичных участков нуклеиновой кислоты, включающих последовательность, кодирующую представляющий интерес белок или пептид или представляющую интерес РНК (например минтРНК или микроРНК), и (b) один или несколько участков, включающих подвергнутые инверсии последовательности концевого повтора (ITR) (например ITR последовательности дикого типа или модифицированные ITR последовательности), фланкирующие указанный один или несколько участков нуклеиновой кислоты (например гетерологичных участков нуклеиновой кислоты). При этом гетерологичные участки нуклеиновой кислоты, включающие последовательность, кодирующую представляющий интерес белок или представляющую интерес РНК, обозначаются как представляющие интерес гены.

[0832] Любая из предусмотренных здесь частиц rAAV может иметь капсидные белки, которые имеют аминокислоты различных серотипов вне участка VP1u. В некоторых вариантах осуществления серотип каркаса белка VP1 отличается от серотипа ITR и/или гена rep. В некоторых вариантах осуществления серотип каркаса капсидного белка VP1 частицы является таким же, что и серотип ITR. В некоторых вариантах осуществления серотип каркаса капсидного белка VP1 частицы является таким же, что и серотип гена rep. В некоторых вариантах осуществления капсидные белки частиц rAAV включают аминокислотные мутации, которые приводят к повышению эффективность трансдукции.

[0833] В некоторых вариантах осуществления вектор нуклеиновой кислоты включает один или несколько участков, включающих последовательность, которая обеспечивает экспрессию нуклеиновой кислоты (например гетерологичной нуклеиновой кислоты), например последовательности управления экспрессией, функционально связанные с нуклеиновой кислотой. Многочисленные подобные последовательности известны в данной области. Неограничивающие примеры последовательностей управления экспрессией включают промоторы, инсулаторы, сайленсеры, ответные элементы, интроны, энхансеры, участки инициации, сигналы терминации, и poly(A)-хвосты. Здесь охватывается любая комбинация подобных контрольных последовательностей (например промотор и энхансер).

[0834] Конечные конструкты AAV могут включать последовательность, кодирующую PEgRNA. В других вариантах осуществления конструкты AAV могут включать

последовательность, кодирующую направляющую РНК для никирования второго участка. В других вариантах осуществления конструкты AAV могут включать последовательность, кодирующую направляющую РНК для никирования второго участка, и последовательность, кодирующую PEgRNA.

[0835] В различных вариантах осуществления PEgRNA и направляющие РНК для никирования второго участка могут быть экспрессированы из подходящего промотора, например промотора U6 (hU6) человека, промотора U6 (mU6) мыши, или другого подходящего промотора. PEgRNA и направляющие РНК для никирования второго участка могут быть экспрессированы одними и теми же промоторами или различными промоторами.

[0836] В некоторых вариантах осуществления конструкты rAAV или указанные здесь композиции вводят субъекту энтерально. В некоторых вариантах осуществления конструкты rAAV или указанные здесь композиции вводят субъекту парентерально. В некоторых вариантах осуществления rAAV частицу или описанные здесь композиции вводят субъекту подкожно, внутриокулярно, интравитреально, субретинально, внутривенно (IV), интрацеребровентрикулярно, внутримышечно, интратекально (IT), внутрицистернально, внутрибрюшинно, посредством ингаляции, с помощью местного введения или путем прямой инъекции в одну или несколько клеток, тканей или органов. В некоторых вариантах осуществления rAAV частицу или описанные здесь композиции вводят субъекту с помощью инъекции в печеночную артерию или воротную вену.

Стратегии с разделенным PE на основе векторов

[0837] В данном аспекте праймерные редакторы могут быть разделены в участке разделения и поступать в виде двух половин целого/полного праймерного редактора. Две половины могут быть доставлены в клетки (например в виде экспрессированных белков или на отдельных векторах экспрессии), при этом при контактировании внутри клетки две половины образуют полный праймерный редактор за счет аутосплайсингового действия интеинов на каждой половине праймерного редактора. Последовательности разделенного интеина могут быть встроены в каждую из половин кодируемого праймерного редактора для обеспечения их транс-сплайсинга внутри клетки и сопутствующего восстановления полного функционируюоего PE.

[0838] Данные способы на основе разделенного интеина обеспечивают преодоление нескольких барьеров при *in vivo* доставке. Например, ДНК, кодирующая праймерные

редакторы, является более крупной, чем верхний предел упаковывания rAAV, и поэтому требует специальных решений. Одно подобное решение заключается в конструировании редактора, слитого с парами разделенного интеина, которые упакованы в две отдельные rAAV частицы, которые при совместной доставке в клетку обеспечивают восстановление функционального белка редактора. Описаны некоторые другие специальные факторы, которые следует учитывать в плане уникальных свойств праймерного редактирования, включая оптимизацию мишеней никирования второго участка и правильное упаковывание праймерных редакторов в вирусные векторы, включая лентивирусы и rAAV.

[0839] В данном аспекте праймерные редакторы могут быть разделены в участке разделения и поступать в виде двух половин целого/полного праймерного редактора. Две половины могут быть доставлены в клетки (например в виде экспрессированных белков или на отдельных векторах экспрессии), при этом при контактировании внутри клетки две половины образуют полный праймерный редактор за счет аутосплайсингового действия интеинов на каждой половине праймерного редактора. Последовательности разделенного интеина могут быть встроены в каждую из половин кодируемого праймерного редактора для обеспечения их транс-сплайсинга внутри клетки и сопутствующего восстановления полного функционируюоего PE.

[0840] На Фиг.66 изображен один вариант осуществления праймерного редактора, представленного в виде двух половин белка PE, которые соединяются в виде цельного праймерного редактора за счет аутосплайсингового действия половин разделенного интеина, расположенных в конце или начале каждой из половин белка праймерного редактора. В соответствии с используемым здесь значением термин "N-концевая половина PE" обозначает N-концевую половину полного праймерного редактора, которая включает "N-интеин" на C-конце N-концевой половины PE (то есть N-концевой экстеин) полного праймерного редактора. Термин "N-интеин" обозначает N-концевую половину полного полностью сформированного фрагмента в виде разделенного интеина. В соответствии с используемым здесь значением термин "C-концевая половина PE" обозначает C-концевую половину полного праймерного редактора, которая включает "C-интеин" на N-конце C-концевой половины (то есть C-концевой экстеин) полного праймерного редактора. Когда данные две половины белка, то есть N-концевая половина PE и C-концевая половина PE, контактируют друг с другом, например в клетке, N-интеин и C-интеин подвергаются одновременному

процессу самоэксцизии и образованию пептидной связи между С-концом N-концевой половины PE и N-концом С-концевой половины PE для образования полного белка праймерного редактора, включающего полный домен napDNAbp (например никазу Cas9) и домен RT. Хотя это не показано на фигурах, праймерный редактор также может включать дополнительные последовательности, включающие NLS на N-конце и/или С-конце, а также последовательности аминокислотных линкеров, присоединяющих каждый домен.

[0841] В различных вариантах осуществления праймерные редакторы могут быть сконструированы в виде двух половин белка (то есть N-концевой половины PE и С-концевой половины PE) путем "разделения" цельного праймерного редактора в "участке разделения". Термин "участок разделения" обозначает месторасположение вставки последовательностей разделенного интеина (то есть N-интеина и С-интеина) между двумя смежными аминокислотными остатками в рамках праймерного редактора. Более конкретно, термин "участок разделения" обозначает месторасположение разделения целого праймерного редактора на две отдельные половины, при этом каждая половина слита в участке разделения с мотивами N-интеина или С-интеина. Участок разделения может представлять собой любое подходящее месторасположение в рамках слитого белка праймерного редактора, однако участок разделения предпочтительно расположен в положении, которое обеспечивает образование двух половин белка, имеющих подходящий размер для доставки (например с помощью вектора экспрессии), и при этом интеины, которые слиты с каждой из половин белка на концах участка разделения, являются доступными для достаточного взаимодействия друг с другом, когда одна половина белка контактирует с другой половиной белка внутри клетки.

[0842] В некоторых вариантах осуществления участок разделения расположен в домене napDNAbp. В других вариантах осуществления участок разделения расположен в домене RT. В других вариантах осуществления участок разделения расположен в линкере, которые соединяет домен napDNAbp и домен RT.

[0843] В различных вариантах осуществления конструкция участка разделения требует нахождения участков для разделения и вставки N- и С-концевых интеинов, оба из которых являются структурно допустимыми для целей упаковывания двух половин в виде доменов праймерного редактора в два различных генома AAV. Кроме того, остатки интеина, необходимые для транс-сплайсинга, могут быть встроены за счет

мутирования остатков на N-конце C-концевого экстеина или вставки остатков, которые оставят интеиновый "шрам".

[0844] Примеры конфигураций разделенных праймерных редакторов, включающих никазу SpCas9 или никазу SaCas9, являются следующими:

PE S. PYOGENES, РАЗДЕЛЕН В ОБЛАСТИ ЛИНКЕРА, СТРУКТУРА N-КОНЦЕВОЙ ЧАСТИ: [N-ЭКСТЕИН]-[N-ИНТЕИН]
MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFD LAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHA ILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGD**SGGSSGGS**CLSYETEILTVEYGLLPIGKIVEKRIECT VYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPID EIFERELDLMRVDNLPNSGGSKRTADGSEFEPKKKRKV(SEQ ID NO: 443) ОБОЗНАЧЕНИЯ: NLS (SEQ ID NO: 124, 125) **CAS9** (SEQ ID NO: 445) **ЛИНКЕР** (SEQ ID NO: 446) ИНТЕИН NPUN (SEQ ID NO: 447)

PE S. PYOGENES, РАЗДЕЛЕН В ОБЛАСТИ ЛИНКЕРА, СТРУКТУРА C-КОНЦЕВОЙ ЧАСТИ: [C-ИНТЕИН]-[C-ЭКСТЕИН]
MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN***CFNSGS ETPGTSESATPESSGGSSGGSS***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGG*

*MGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLP VKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRL HPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILL QYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWL TEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDP VAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQ ALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTW YTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYT DSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP***SGGS***KRTADGSEFEPKKKRKV*(SEQ ID NO: 450)

ОБОЗНАЧЕНИЯ:
NLS (SEQ ID NO: 124, 125)
ЛИНКЕР 1 (SEQ ID NO: 453)
ЛИНКЕР 2 (SEQ ID NO: 174)

ИНТЕИН NPUC (SEQ ID NO: 452)
RT (SEQ ID NO: 454)

PE S. AUREUS, РАЗДЕЛЕН МЕЖДУ ОСТАТКАМИ 740/741, СТРУКТУРА N-КОНЦЕВОЙ ЧАСТИ: [N-ЭКСТЕИН]-[N-ИНТЕИН]

MKRTADGSEFESPKKKRKV**GKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFK EANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEA RVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKAL EEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFID TYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNE EDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDI QEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNR LKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIEL AREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEG KCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTP FQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINR NLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAE**CLSYETEILTVE YGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLIRATK DHKFMTVDGQMLPIDEIFERELDLMRVDNLPN***SGGS***KRTADGSEFEPKKKRKV*(SEQ ID NO: 458)

ОБОЗНАЧЕНИЯ:
NLS (SEQ ID NO: 124, 125)
CAS9 (SEQ ID NO: 460)
ЛИНКЕР (SEQ ID NO: 174)
ИНТЕИН NPUN (SEQ ID NO: 462)

PE S. AUREUS, РАЗДЕЛЕН МЕЖДУ ОСТАТКАМИ 740/741, СТРУКТУРА С-КОНЦЕВОЙ ЧАСТИ: [С-ИНТЕИН]-[С-ЭКСТЕИН]
MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN***CFNEIE TEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTL IVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKL SLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQA EFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHII KTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG*****SGGSSGGSSGSETPGTSESA TPESSGGSSGGSS***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDY RPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAF EWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAA TSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVM GQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQ ALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPC LRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLL QEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATA HIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMAD QAARKAAITETPDTSTLLIENSSP**SGGS**KRTADGSEFEPKKKRKV(SEQ ID NO: 465) ОБОЗНАЧЕНИЯ: NLS (SEQ ID NO: 124, 125) **CAS9** (SEQ ID NO: 467) **ЛИНКЕР 1** (SEQ ID NO: 127) **ЛИНКЕР 2** (SEQ ID NO: 174) *ИНТЕИН NPUC* (SEQ ID NO: 452) *RT* (SEQ ID NO: 471)

[0845] В различных вариантах осуществления с никазой SpCas9 в качестве примера (SEQ ID NO: 18, 1368 аминокислот) разделение может быть расположено между любыми двумя аминокислотами между 1 и 1368. Однако предпочтительно разделение будет расположено в центральном участке белка, например в диапазоне аминокислот 50-1250, или 100-1200, или 150-1150, или 200-1100, или 250-1050, или 300-1000, или 350-950, или 400-900, или 450-850, или 500-800, или 550-750, или 600-700 SEQ ID NO: 18. В определенных примерах вариантов осуществления участок разделения может быть расположен между 740/741, или 801/802, или 1010/1011, или 1041/1042. В других вариантах осуществления участок разделения может быть расположен между 1/2, 2/3, 3/4, 4/5, 5/6, 6/7, 7/8, 8/9, 9/10, 10/11, 12/13, 14/15, 15/16, 17/18, 19/20, 20/21, 21/22, 22/23, 23/24, 24/25, 25/26, 26/27, 27/28, 28/29, 29/30, 30/31, 31/32, 32/33, 33/34, 34/35, 35/36, 36/37, 38/39, 39/40, 41/42, 42/43, 43/44, 44/45, 45/46, 46/47, 47/48, 48/49, 49/50, 51/52, 52/53, 53/54, 54/55, 55/56, 56/57, 57/58, 58/59, 59/60, 61/62, 62/63, 63/64, 64/65,

65/66, 66/67, 67/68, 68/69, 69/70, 71/72, 72/73, 73/74, 74/75, 75/76, 76/77, 77/78, 78/79, 79/80, 81/82, 82/83, 83/84, 84/85, 85/86, 86/87, 87/88, 88/89, 89/90, или между любыми двумя парами смежных остатков между 90-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 450-500, 500-550, 550-600, 600-650, 650-700, 700-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, и 1350-1368, относительно SpCas9 согласно SEQ ID NO: 18, между любыми двумя соответствующими остатками в аминокислотной последовательности, имеющей, по меньшей мере, 80%, 85%, 90%, 95%, 98%, 99%, или 99,9% идентичность последовательности относительно SEQ ID NO: 18, или между любыми двумя соответствующими остатками в рамках варианта или эквивалента SpCas9 согласно любой из аминокислотных последовательностей SEQ ID NO. 19-88, или аминокислотной последовательности, имеющей, по меньшей мере, 80%, 85%, 90%, 95%, 98%, 99%, или 99,9% идентичность последовательности относительно любой из SEQ ID NO: 19-88.

[0846] В различных вариантах осуществления последовательности разделенного интеина могут быть сконструированы на основе следующих последовательностей интеина.

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
ИНТЕИН 2-4:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 472)
ИНТЕИН 3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDR VAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASM MGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLEIL MIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATS SRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRAL DKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHL YSMKYTNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDDK FLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHN C (SEQ ID NO: 473)

ИНТЕИН 30R3-1	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 474)
ИНТЕИН 30R3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 475)
ИНТЕИН 30R3-3	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC (SEQ ID NO: 476)
ИНТЕИН 37R3-1	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKG DRVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEA SMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWL EILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLL ATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH RALDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGM EHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADAL DDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVV VHNC ((SEQ ID NO: 477)
ИНТЕИН 37R3-2	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVV SWFDQGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH

	NC (SEQ ID NO: 478)
ИНТЕИН 37R3-3	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVS WFDQGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGD RVAGPGGSGNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEAS MMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLEI LMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLAT SSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRA LDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEH LYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVH NC (SEQ ID NO: 479)

[0847] В различных других вариантах осуществления последовательности разделенного интеина могут быть использованы следующим образом:

ИНТЕИН-N	ИНТЕИН-C
NPU-N CLSYETEILTVEYGLLPIGKIVEKRIEC TVYSVDNNGNIYTQPVAQWHDRGEQ EVFEYCLEDGSLIRATKDHKFMTVDG QMLPIDEIFERELDLMRVDNLPNSGG S (SEQ ID NO: 447)	NPU-C IKIATRKYLGKQNVYDIGVERDHNFALKNG FIASN (SEQ ID NO: 452)

[0848] В различных вариантах осуществления разделенные интеины могут применяться для отдельной доставки отдельных частей полного слитого белка PE в клетку, которые при экспрессии в клетке становятся восстановленными в виде полного слитого белка PE за счет транс-сплайсинга.

[0849] В некоторых вариантах осуществления изобретение предусматривает способ доставки слитого белка PE в клетку, включающий:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью разделенного интеина;

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностью разделенного интеина;

(c) доставку первого и второго вектора экспрессии в клетку,

при этом N-концевой и C-концевой фрагмент восстанавливаются в виде слитого белка PE в клетке в результате активности в виде транс-сплайсинга, вызывающей самоэксцизию первой и второй последовательности разделенного интеина.

[0850] В некоторых вариантах осуществления участок разделения может быть расположен где угодно в рамках продукта слияния праймерного редактора, включая домен napDNAbp, линкер, или домен обратной транскриптазы.

[0851] В других вариантах осуществления участок разделения расположен в домене napDNAbp.

[0852] В других вариантах осуществления участок разделения расположен в домене обратной транскриптазы или полимеразы.

[0853] В вариантах осуществления участок разделения расположен в линкере.

[0854] В различных вариантах осуществления настоящее изобретение предусматривает праймерные редакторы, включающие napDNAbp (например домен Cas9) и обратную транскриптазу, при этом один или оба из napDNAbp и/или обратной транскриптазы включают интеин, например лиганд-зависимый интеин. Как правило, интеин представляет собой лиганд-зависимый интеин, который демонстрирует отсутствие или минимальную активность в виде сплайсинга белка при отсутствии лиганда (например малых молекул, таких как 4-гидрокситамоксифен, пептиды, белки, полинуклеотиды, аминокислоты, и нуклеотиды). Лиганд-зависимые интеины являются известными и включают описанные в заявке на патент США 14/004280, опубликованной как 2014/0065711 A1, полное содержание которой приведено здесь в качестве ссылки. Кроме того, использование архитектуры в виде разделенного Cas9. В некоторых вариантах осуществления интеин включает аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 8-15, 447, 452, 462, и 472-479.

[0855] В различных вариантах осуществления домены napDNAbp являются доменами napDNAbp меньшего размера по сравнению со стандартным доменом SpCas9 согласно SEQ ID NO: 18.

[0856] Стандартный белок SpCas9 имеет длину 1368 аминокислот и имеет предполагаемую молекулярную массу 158 кДа. В соответствии с используемым здесь значением термин "вариант Cas9 малого размера" обозначает любой вариант Cas9 — встречающийся в природе, сконструированный или полученный иным образом — длиной менее чем приблизительно 1300 аминокислот, или менее чем 1290 аминокислот, или менее чем 1280 аминокислот, или менее чем 1270 аминокислот, или менее чем 1260 аминокислот, или менее чем 1250 аминокислот, или менее чем 1240 аминокислот, или менее чем 1230 аминокислот, или менее чем 1220 аминокислот, или менее чем 1210 аминокислот, или менее чем 1200 аминокислот, или менее чем 1190 аминокислот, или менее чем 1180 аминокислот, или менее чем 1170 аминокислот, или менее чем 1160 аминокислот, или менее чем 1150 аминокислот, или менее чем 1140 аминокислот, или менее чем 1130 аминокислот, или менее чем 1120 аминокислот, или

менее чем 1110 аминокислот, или менее чем 1100 аминокислот, или менее чем 1050 аминокислот, или менее чем 1000 аминокислот, или менее чем 950 аминокислот, или менее чем 900 аминокислот, или менее чем 850 аминокислот, или менее чем 800 аминокислот, или менее чем 750 аминокислот, или менее чем 700 аминокислот, или менее чем 650 аминокислот, или менее чем 600 аминокислот, или менее чем 550 аминокислот, или менее чем 500 аминокислот, но, по меньшей мере, более чем приблизительно 400 аминокислот, сохраняющий требуемые функции белка Cas9.

[0857] В одном варианте осуществления, как показано в Примере 20, описание охватывает следующие конструкты PE на основе разделенного интеина, которые разделены между остатками 1024 и 1025 стандартного SpCas9 (SEQ ID NO: 18) (или которые могут быть обозначены как остатки 1023 и 1024, соответственно, по сравнению с SEQ ID NO: 18 без Met).

[0858] Сначала показана аминокислотная последовательность SEQ ID NO: 18, которая является следующей, что свидетельствует о месторасположении участка разрыва между остатками 1024 ("K") и 1025 ("S"):

Описание	**Последовательность**	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* M1 Кат. # SwissProt Q99ZW2 Дикий тип	**M**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQL PGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK	SEQ ID NO: 18, с указанием участка разрыва 1024/1025 жирным шрифтом M в положении 1 не обязательно присутствует в слитом белке PE в определенных вариантах осуществления. Таким образом, нумерация участка разрыва соответствует

	LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA **KS**EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	1023/1024 в случае, если аминокислотная последовательность не включает Met в положении 1.

[0859] В этой конфигурации аминокислотная последовательность N-концевой половины (аминокислоты 1-1024) является следующей:

[0860] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH FLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKK IECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDK AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA**K** (SEQ ID NO: 3877).

[0861] В этой конфигурации аминокислотная последовательность N-концевой половины (аминокислоты 1-1023) (где белок не включает Met в положении 1) является следующей:

[0862] DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDK KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDE LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA**K** (SEQ ID NO: 3878).

[0863] В этой конфигурации аминокислотная последовательность С-концевой половины (аминокислоты 1024-1368 (или подсчитываемые как аминокислоты 1023-1367 в Cas9 без Met) является следующей:

[0864] **S**EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVK KDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GD (SEQ ID NO: 3879).

[0865] Как показано в Примере 20, конструкт PE2 (который основан на SpCas9 согласно SEQ ID NO: 18) был разделен в положении 1023/1024 (по сравнению с SEQ ID NO: 18 без Met) на два отдельных конструкта следующим образом:

[0866] SpPE2, разделенный в 1023/1024 N-концевой половины

[0867] MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAK*CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQW*
*HDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN***SGGS**
KRTADGSEFEPKKKRKV (SEQ ID NO: 3875)

[0868] Обозначения: NLS, Cas9, Мутированные остатки, **Линкер**, *интеин NpuN*, *интеин NpuC*, **RT**

[0869] SpPE2, разделенный в 1023/1024 С-концевой половины

MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN*CFNEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV
VAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF
ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD**SGGSSGGSS**
GSETPGTSESATPESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQA
WAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGIL
VPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPP
SHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN
SPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLG
NLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQL

REFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLT APALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAG WPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARM THYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTD QPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELI ALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEI LALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLI ENSSPSGGS*KRTADGSEFEPKKKRK*** (SEQ ID NO: 3876)**

[0870] Обозначения: NLS, Cas9, Мутированные остатки, **Линкер**, *интеин NpuN*, *интеин NpuC*, **RT**

[0871] Настоящее изобретение также охватывает способы доставки праймерных редакторов на основе разделенного интеина в клетки и/или обработку клеток праймерными редакторами на основе разделенного интеина.

[0872] В некоторых вариантах осуществления изобретение предусматривает способ доставки слитого белка PE в клетку, включающий:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью разделенного интеина;

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностью разделенного интеина;

(c) доставку первого и второго вектора экспрессии в клетку,

при этом N-концевой и C-концевой фрагмент восстанавливаются в виде слитого белка PE в клетке в результате активности в виде транс-сплайсинга, вызывающей самоэксцизию первой и второй последовательности разделенного интеина.

[0873] В определенных вариантах осуществления N-концевой фрагмент слитого белка PE, слитого с последовательностью первого разделенного интеина, представляет собой SEQ ID NO: 3875 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 98%, или, по меньшей мере, 99,9% идентичность последовательности относительно SEQ ID NO: 3875.

[0874] В других вариантах осуществления C-концевой фрагмент слитого белка PE, слитого с последовательностью первого разделенного интеина, представляет собой SEQ ID NO: 3876 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, по

меньшей мере, 98%, или, по меньшей мере, 99,9% идентичность последовательности относительно SEQ ID NO: 3876.

[0875] В других вариантах осуществления изобретение предусматривает способ редактирования последовательности-мишени ДНК в клетке, включающий:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью разделенного интеина;

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностью разделенного интеина;

(c) доставку первого и второго вектора экспрессии в клетку,

при этом N-концевой и C-концевой фрагмент восстанавливаются в виде слитого белка PE в клетке в результате активности в виде транс-сплайсинга, вызывающей самоэксцизию первой и второй последовательности разделенного интеина.

[0876] В определенных вариантах осуществления N-концевой фрагмент слитого белка PE, слитого с последовательностью первого разделенного интеина, представляет собой SEQ ID NO: 3875 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 98%, или, по меньшей мере, 99,9% идентичность последовательности относительно SEQ ID NO: 3875.

[0877] В других вариантах осуществления C-концевой фрагмент слитого белка PE, слитого с последовательностью первого разделенного интеина, представляет собой SEQ ID NO: 3876 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, по меньшей мере, 85%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 98%, или, по меньшей мере, 99,9% идентичность последовательности относительно SEQ ID NO: 3876.

Доставка рибонуклеопротеиновых комплексов PE

[0878] В данном аспекте праймерные редакторы могут быть доставлены с помощью стратегий невирусной доставки, включающих доставку праймерного редактора в виде комплекса с PEgRNA (то есть рибонуклеопротеинового комплекса PE) с помощью различных способов, включая электропорацию и липидные наночастицы. Способы невирусной доставки нуклеиновых кислот включают липофекцию, нуклеофекцию, микроинъекцию, биолистику, вирусомы, липосомы, иммунолипосомы, конъюгаты

поликатион или липид:нуклеиновая кислота, голую ДНК, искусственные вирионы, и агент-энхансированный захват ДНК. Липофекция описана, например, в патентах США 5049386, 4946787, и 4897355), а реагенты для липофекции являются коммерчески доступными (например Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые являются подходящими для эффективной липофекции полинуклеотидов на основе распознавания рецептора, включают указанные у Feigner, WO 91/17424; WO 91/16024. Доставка может осуществляться в клетки (например при введении *in vitro* или *ex vivo*) или ткани-мишени (например при введении *in vivo*).

[0879] Получение комплексов липид:нуклеиновая кислота, включая липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области (см., например, , Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); патенты США 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028, и 4946787).

[0880] Может быть сделана дополнительная ссылка на следующие ссылки, в рамках которых описаны походы для невирусной доставки рибонуклеопротеиновых комплексов, каждая из которых приведена здесь в качестве ссылки.

[0881] Chen, Sean, et al. "Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes." *Journal of Biological Chemistry* (2016): jbc-M116. PubMed

[0882] Zuris, John A., et al. "Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo." *Nature biotechnology* 33.1 (2015): 73. PubMed

[0883] Rouet, Romain, et al. "Receptor-Mediated Delivery of CRISPR-Cas9 Endonuclease for Cell-Type-Specific Gene Editing." *Journal of the American Chemical Society* 140.21 (2018): 6596-6603. PubMed.

[0884] На Фиг.68C приведены данные, которые показывают, что различные описанные рибонуклеопротеиновые PE комплексы (PE2 с высокой концентрацией, PE3 с высокой концентрацией и PE3 с низкой концентрацией) могут быть доставлены подобным образом.

Доставка PE с помощью иРНК

[0885] Другой способ, который может применяться для доставки праймерных редакторов и/или PEgRNA в клетки, в которых требуется осуществить геномное редактирование на основе праймерного редактирования, заключается в использовании способов и технологий доставки на основе информационной РНК (иРНК). Примеры способов доставки на основе иРНК и композиций, которые могут применяться в рамках настоящего изобретения, включают, например, PCT/US2014/028330, US8822663B2, NZ700688A, ES2740248T3, EP2755693A4, EP2755986A4, WO2014152940A1, EP3450553B1, BR112016030852A2, и EP3362461A1, каждый из которых полностью приведен здесь в качестве ссылки. См. дополнительное приводимое здесь в качестве ссылки описание в Kowalski et al., "Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery," *Mol Therap.*, 2019; 27(4): 710-728.

[0886] В отличие от ДНК вектора, кодирующего праймерные редакторы, применение РНК в качестве агента доставки для праймерных редакторов имеет преимущество, заключающееся в том, что генетический материал не обязательно должен проникать в ядро для реализации его функции. Доставленная иРНК может быть напрямую транслирована в цитоплазме в требуемый белок (например слитый белок праймерного редактора) и продукты в виде нуклеиновой кислоты (например PEgRNA). Однако для большей стабильности (например устойчивости к расщепляющим РНК ферментам в цитоплазме) в некоторых вариантах осуществления является необходимым стабилизировать иРНК для повышения эффективности доставки. Определенные носители для доставки, такие как катионные липиды или носители для полимерной доставки, также могут помогать обеспечивать защиту трансфицированной иРНК от ферментов в виде эндогенной РНКазы, которые в ином случае могут расщеплять терапевтическую иРНК, кодирующую требуемые слитые белки праймерного редактора. Кроме того, несмотря на повышенную стабильность модифицированной иРНК, доставка иРНК, в частности иРНК, кодирующей полноразмерный белок, в клетки in vivo таким образом, который обеспечивает терапевтические уровни продуцирования белка, остается трудной задачей.

[0887] С некоторыми исключениями, внутриклеточная доставка иРНК в целом является более трудной, чем доставка малых олигонуклеотидов, при этом она требует инкапсуляции в наночастицу для доставки, частично вследствие значительно более крупного размера молекул иРНК (300–5000 кДа, ~1–15 kb) по сравнению с другими

типами РНК (малыми интерферирующими РНК [минтРНК], ~14 кДа; антисмысловыми олигонуклеотидами [ASO], 4-10 кДа).

[0888] иРНК должны пересекать клеточную мембрану для достижения цитоплазмы. Клеточная мембрана представляет собой динамичный и трудный барьер для внутриклеточной доставки. Она в основном состоит из липидного двойного слоя цвиттерионных и отрицательно заряженных фосфолипидов, где полярные головки фосфолипидов обращены к водной среде, а гидрофобные хвосты образуют гидрофобное ядро.

[0889] В некоторых вариантах осуществления композиции на основе иРНК по изобретению включают иРНК (кодирующую праймерный редактор и/или PEgRNA), средство для транспорта, и необязательно агент, который обеспечивает контактирование с клеткой-мишенью и последующее трансфицирование.

[0890] В некоторых вариантах осуществления иРНК может включать одну или несколько модификаций, которые придают стабильность иРНК (например по сравнению с относящейся к дикому типу или нативной версией иРНК) и участвуют в ассоциированной анормальной экспрессии белка. Также может быть включена одна или несколько модификаций дикого типа, которые корректируют дефект. Например, нуклеиновые кислоты по изобретению могут включать модификации одного или обоих из 5' нетранслируемого участка или 3' нетранслируемого участка. Такие модификации могут предусматривать включение последовательностей, кодирующих частичную последовательность немедленного раннего гена 1 (IE1) цитомегаловируса (CMV), poly A-хвост, структуру Cap1, или гормон роста человека (hGH). В некоторых вариантах осуществления иРНК является модифицированной для снижения иммуногенности иРНК.

[0891] В одном варианте осуществления иРНК праймерного редактора в композиции по изобретению может быть сконструирована в рамках носителя для липосомного переноса для обеспечения доставки в клетки-мишени. Предполагаемые носители для переноса могут включать один или несколько катионных липидов, некатионных липидов, и/или ПЭГ-модифицированных липидов. Например, носитель для переноса может включать, по меньшей мере, один из следующих катионных липидов: C12-200, DLin-KC2-DMA, DODAP, HGT4003, ICE, HGT5000, или HGT5001. В некоторых вариантах осуществления носитель для переноса включает холестерин (chol) и/или модифицированные ПЭГ липиды. В некоторых вариантах осуществления носитель для

переноса включает DMG-PEG2K. В определенных вариантах осуществления носитель для переноса имеет следующий липидный состав: C12-200, DOPE, chol, DMG-PEG2K; DODAP, DOPE, холестерин, DMG-PEG2K; HGT5000, DOPE, chol, DMG-PEG2K, HGT5001, DOPE, chol, один из DMG-PEG2K.

[0892] Настоящее изобретение также предусматривает композиции и способы, которые могут применяться для обеспечения трансфекции клеток-мишеней одной или несколькими кодирующими РЕ молекулами иРНК. Например, композиции и способы по настоящему изобретению предусматривают применение нацеленных лигандов, которые могут повышать аффинность композиции в отношении одной или нескольких клеток-мишеней. В одном варианте осуществления нацеливающий лиганд представляет собой аполипопротеин В или аполипопротеин Е, при этом соответствующие клетки-мишени экспрессируют рецепторы липопротеина низкой плотности и, таким образом, промотируют распознавание нацеливающего лиганда. Огромное количество клеток-мишеней может предпочтительным образом становиться мишенью с помощью способов и композиций по настоящему изобретению. Например, предусмотренные клетки-мишени включают гепатоциты, эпителиальные клетки, гемопоэтические клетки, эпителиальные клетки, эндотелиальные клетки, клетки легкого, костные клетки, стволовые клетки, мезенхимальные клетки, нервные клетки, клетки сердца, адипоциты, клетки сосудов, клетки гладких мышц, кардиомиоциты, клетки скелетных мышц, бета-клетки, клетки гипофиза, клетки синовиальной оболочки, клетки яичника, клетки семенника, фибробласты, В-клетки, Т-клетки, ретикулоциты, лейкоциты, гранулоциты и опухолевые клетки. Однако они не ограничиваются ими.

[0893] В некоторых вариантах осуществления кодирующая РЕ иРНК может необязательно иметь химические или биологически модификации, которые, например, повышают стабильность и/или время полужизни подобной иРНК или которые повышают или иным образом обеспечивают продуцирование белка. При трансфекции природная иРНК в рамках композиций по изобретению может расщепляться со временем полужизни от 30 минут до нескольких дней. иРНК в композициях по изобретению могут сохранять, по меньшей мере, некоторую способность подвергаться трансляции, тем самым продуцируя функциональный белок или фермент. Соответственно, изобретение предусматривает композиции, которые включают стабилизированную иРНК, а также способы ее введения. В некоторых вариантах

осуществления активность иРНК пролонгируется в течение увеличенного периода времени. Например, активность иРНК может быть пролонгирована таким образом, что композиции по настоящему изобретению вводят субъекту на полуеженедельной или двухнедельной основе, или, более предпочтительно, на ежемесячной, двухмесячной, квартальной или ежегодной основе. Увеличенная или пролонгированная активность иРНК по настоящему изобретению напрямую связана с количеством белка или фермента, продуцируемого из такой иРНК. Аналогичным образом, активность композиций по настоящему изобретению может быть дополнительно увеличена или пролонгировано за счет модификаций, реализованный для улучшения или повышения трансляции иРНК. Более того, количество функционального белка или фермента, продуцируемого клеткой-мишенью, зависит от количества иРНК, доставляемой в клетки-мишени, и стабильности такой иРНК. В той степени, в которой стабильность иРНК по настоящему изобретению может быть улучшена или повышена, время полужизни, активность продуцируемого белка или фермента и частота введения композиции может быть также увеличена.

[0894] Соответственно, в некоторых вариантах осуществления иРНК в композициях по изобретению включает, по меньшей мере, одну модификацию, которая придает улучшенную или повышенную стабильность нуклеиновой кислоте, включая, например, улучшенную устойчивость в расщеплению нуклеазой *in vivo*. В соответствии с используемым здесь значением термины "модификация" и "модифицированная" при их использовании в отношении указанных здесь нуклеиновых кислот включают, по меньшей мере, одно изменение, которое предпочтительно повышает стабильность и делает иРНК более стабильной (например устойчивой к расщеплению нуклеазой), чем относящаяся к дикому типу или встречающаяся в природе версия иРНК. В соответствии с используемым здесь значением термины "стабильная" и "стабильность" при их использовании в отношении нуклеиновых кислот по настоящему изобретению и, в частности, в отношении иРНК обозначают повышенную или улучшенную устойчивость к расщеплению, например, нуклеазами (то есть эндонуклеазами или экзонуклеазами), которые обычно способны расщеплять такую иРНК. Повышенная стабильность может включать, например, меньшую чувствительность к гидролизу или другому разрушению эндогенными ферментами (например эндонуклеазами или экзонуклеазами) или условиям в клетке-мишени или ткани, тем самым повышая или стимулируя содержание такой иРНК в

клетке-мишени, ткани, субъекте и/или цитоплазме. Предусмотренные здесь молекулы стабилизированной иРНК демонстрируют более длительное время полужизни по сравнению с их встречающимися в природе немодифицированными аналогами (например версией иРНК дикого типа). Термины "модификация" и "модифицированная" при их использовании в отношении иРНК по настоящему изобретению также охватывают изменения, которые улучшают или повышают трансляцию нуклеиновых кислот иРНК, включая, например, включение последовательностей, которые участвуют в инициировании трансляции белка (например консенсусную последовательность Kozak). (Kozak, M., Nucleic Acids Res 15 (20): 8125-48 (1987)).

[0895] В некоторых вариантах осуществления иРНК, используемые в композициях по изобретению, подверглись химическому или биологическому модифицированию, чтобы сделать их более стабильными. Примеры модификаций иРНК включают устранение основания (например путем удаления или замещения одного нуклеотида на другой) или модификацию основания, например химическую модификацию основания. Термин "химические модификации" в соответствии с используемым здесь значением включает модификации, в рамках которых вносятся химические признаки, которые отличаются от тех, которые присущи встречающейся в природе иРНК, например ковалентные модификации, такие как внесение модифицированных нуклеотидов (например аналогов нуклеотидов, или внесение маятниковых групп, которые обычно не присущи таким молекулам иРНК).

[0896] Другие подходящие полинуклеотидные модификации, которые могут быть встроены в РЕ-кодирующую иРНК, используемую в композициях по настоящему изобретению, включают, но без ограничения, 4'-тио-модифицированные основания: 4'-тио-аденозин, 4'-тио-гуанозин, 4'-тио-цитидин, 4'-тио-уридин, 4'-тио-5-метилцитидин, 4'-тио-псевдоуридин и 4'- тио-2-тиоуридин, пиридин-4-он рибонуклеозид, 5-азауридин, 2-тио-5-азауридин, 2-тиоуридин, 4-тио-псевдоуридин, 2-тио-псевдоуридин, 5-гидроксиуридин, 3 -метилуридин, 5-карбоксиметил-уридин, 1-карбоксиметил-псевдоуридин, 5-пропинил-уридин, 1-пропинил-псевдоуридин, 5-тауринометилуридин, 1-тауринометил-псевдоуридин, 5-тауринометил-2-тио-уридин, 1-тауринометил-4-тио-уридин, 5-метил-уридин, 1-метил-псевдоуридин, 4-тио-1-метил-псевдоуридин, 2-тио-1-метил-псевдоуридин, 1-метил-1-деаза-псевдоуридин, 2 -тио-1-метил-1-деаза-псевдоуридин, дигидроуридин, дигидропсевдоуридин, 2-тио-дигидроуридин, 2-тио-

дигидропсевдоуридин, 2-метоксиуридин, 2-метокси-4-тио-уридин, 4-метокси-псевдоуридин, 4 -метокси-2-тио-псевдоуридин, 5-аза-цитидин, псевдоизоцитидин, 3-метил-цитидин, N4-ацетилцитидин, 5-формилцитидин, N4-метилцитидин, 5-гидроксиметилцитидин, 1-метил-псевдоизоцитидин, пирроло-цитидин, пирроло-псевдоизоцитидин, 2-тио-цитидин, 2-тио-5-метил-цитидин, 4-тио-псевдоизоцитидин, 4-тио-1-метил-псевдоизоцитидин, 4-тио-1-метил-1-деаза-псевдоизоцитидин, 1-метил-1-деаза-псевдоизоцитидин, зебуларин, 5-аза-зебуларин, 5-метил-зебуларин, 5-аза-2-тио-зебуларин, 2-тио-зебуларин, 2-метокси-цитидин, 2-метокси-5-метилцитидин, 4-метокси-псевдоизоцитидин, 4-метокси-1-метил-псевдоизоцитидин, 2-аминопурин, 2,6-диаминопурин, 7- деаза-аденин, 7-деаза-8-аза-аденин, 7-деаза-2-аминопурин, 7-деаза-8-аза-2-аминопурин, 7-деаза-2,6-диаминопурин, 7-деаза-8-аза-2,6-диаминопурин, 1-метиладенозин, N6-метиладенозин, N6-изопентениладенозин, N6-(цис-гидроксиизопентенил)аденозин, 2-метилтио-N6-(цис-гидроксиизопентенил)аденозин, N6-глицинилкарбамоиладенозин, N6-треонилкарбамоиладенозин, 2-метилтио-N6-треонилкарбамоиладенозин, N6,N6-диметиладенозин, 7-метиладенин, 2-метилтио-аденин и 2-метокси-аденин, инозин, 1-метил-инозин, виозин, вибутозин, 7-деаза-гуанозин, 7-деаза-8-аза-гуанозин, 6-тио-гуанозин, 6-тио-7-деаза-гуанозин, 6-тио-7-деаза-8-азагуанозин, 7-метил-гуанозин, 6-тио-7-метилгуанозин, 7-метилинозин, 6-метокси-гуанозин, 1-метилгуанозин, N2-метилгуанозин, N2,N2-диметилгуанозин, 8-оксо-гуанозин, 7-метил-8-оксо-гуанозин, 1-метил-6-тио-гуанозин, N2-метил-6-тио-гуанозин, и N2,N2-диметил-6-тио-гуанозин и их комбинации. Термин "модификация" также включает, например, встраивание ненуклеотидных связей или модифицированных нуклеотидов в последовательности иРНК по настоящему изобретению (например модификаций на одном или обоих из 3' и 5'-концов молекулы иРНК, кодирующей функциональны белок или фермент). Такие модификации включают добавление оснований в последовательность иРНК (например включение poly A-хвоста или более длинного poly A-хвоста), изменение 3' UTR или 5' UTR, образование комплексов иРНК с агентом (например белком или комплементарной молекулой нуклеиновой кислоты), и включение элементов, которые изменяют структуру молекулы иРНК (например которые образуют вторичные структуры).

[0897] В некоторых вариантах осуществления РЕ-кодирующие иРНК включают структуру в виде 5' кэпа. 5' кэп обычно добавляется следующим образом: сначала концевая фосфатаза РНК удаляет одну из концевых фосфатных групп из 5' нуклеотида,

оставляя два концевых фосфата; затем к концевым фосфатам добавляется гуанозинтрифосфат (GTP) с помощью гуанилилтрансферазы с получением 5'5'5 трифосфатной связи; и затем 7-азот гуанина метилируется с помощью метилтрансферазы. Примеры структур в виде кэпа включают, но без ограничения, m7G(5')ppp (5'(A,G(5')ppp(5')A и G(5')ppp(5')G. Встречающиеся в природе структуры в виде кэпа включают 7-метилгуанозин, который присоединен через трифосфатный мостик к 5'-концу первого транскрибированного нуклеотида, что приводит к образованию динуклеотидного кэпа в виде m7G(5')ppp(5')N, где N представляет собой любой нуклеозид. *In vivo* кэп добавляется ферментативно. Кэп добавляется в ядро и катализируется ферментом гуанилилтрансферазой. Добавление кэпа к 5'-концу РНК происходит непосредственно после инициирования транскрипции. концевой нуклеозид обычно представляет собой гуанозин, при этом он имеет обратную ориентацию относительно всех других нуклеотидов, то есть G(5')ppp(5')GpNpNp.

[0898] Дополнительные аналоги кэпа включают, но без ограничения, химические структуры, выбранные из группы, состоящей из m7GpppG, m7GpppA, m7GpppC; неметилированных аналогов кэпа (например GpppG); диметилированного аналога кэпа (например m2,7GpppG), триметилированного аналога кэпа (например m2,2,7GpppG), диметилированных симметричных аналогов кэпа (например m7Gpppm7G), или аналогов кэпа с правильной ориентацией (например ARCA; m7,2'OmeGpppG, m72'dGpppG, m7,3'OmeGpppG, m7,3'dGpppG и их тетрафосфатных производных) (см., например, Jemielity, J. et al., "Novel 'anti-reverse' cap analogs with superior translational properties", RNA, 9: 1108-1122 (2003)).

[0899] Как правило, наличие "хвоста" служит для защиты иРНК от расщепления экзонуклеазой. Предполагается, что poly A- или poly U-хвост стабилизирует природные информационные и синтетические смысловые РНК. Таким образом, в определенных вариантах осуществления длинный poly A- или poly U-хвост может быть добавлен в молекулу иРНК, что, таким образом, делает РНК более стабильной. Poly A- или poly U-хвосты могут быть добавлены с помощью ряда известных в данной области способов. Например, длинные poly A-хвосты могут быть добавлены в синтетическую или транскрибированную in vitro РНК с помощью poly A-полимеразы (Yokoe, et al. Nature Biotechnology. 1996; 14: 1252-1256). Вектор транскрипции также может кодировать длинные poly A-хвосты. Кроме того, poly A-хвосты могут быть добавлены путем транскрипции напрямую из продуктов ПЦР. Poly A также могут быть

лигированы к 3'-концу смысловой РНК с помощью РНК-лигазы (см., например, Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1991 edition)).

[0900] Как правило, длина poly A- или poly U-хвоста может составлять, по меньшей мере, 10, 50, 100, 200, 300, 400, или, по меньшей мере, 500 нуклеотидов. В некоторых вариантах осуществления poly A-хвост на 3'-конце иРНК обычно включает приблизительно от 10 до 300 аденозиновых нуклеотидов (например приблизительно от 10 до 200 аденозиновых нуклеотидов, приблизительно от 10 до 150 аденозиновых нуклеотидов, приблизительно от 10 до 100 аденозиновых нуклеотидов, приблизительно от 20 до 70 аденозиновых нуклеотидов, или приблизительно от 20 до 60 аденозиновых нуклеотидов). В некоторых вариантах осуществления иРНК включают структуру в виде 3' poly(C)-хвоста. Подходящий poly C-хвост на 3'-конце иРНК обычно включает приблизительно от 10 до 200 цитозиновых нуклеотидов (например приблизительно от 10 до 150 цитозиновых нуклеотидов, приблизительно от 10 до 100 цитозиновых нуклеотидов, приблизительно от 20 до 70 цитозиновых нуклеотидов, приблизительно от 20 до 60 цитозиновых нуклеотидов, или приблизительно от 10 до 40 цитозиновых нуклеотидов). Poly C-хвост может быть добавлен к poly A- или poly U-хвосту или может замещать poly A- или poly U-хвост.

[0901] PE-кодирующие иРНК в соответствии с настоящим изобретением могут быть синтезированы в соответствии с любым из ряда известных способов. Например, иРНК в соответствии с настоящим изобретением могут быть синтезированы с помощью *in vitro* транскрипции (IVT). Кратко, IVT обычно осуществляют с использованием линейной или циркулярной ДНК-матрицы, содержащей промотор, пул рибонуклеотидтрифосфатов, буферную систему, которая может включать ДТТ и ионы магния, и подходящую РНК-полимеразу (например РНК-полимеразу T3, T7 или SP6), ДНКазу I, пирофосфатазу, и/или ингибитор РНКазы. Конкретные условия будут варьироваться в соответствии с конкретным вариантом использования.

[0902] В рамках вариантов осуществления, включающих доставку иРНК, соотношение иРНК, кодирующей слитый белок PE, и PEgRNA может являться важным для эффективного редактирования. В определенных вариантах осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEgRNA составляет 1:1. В определенных других вариантах осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEgRNA составляет 2:1. В других вариантах осуществления

отношение массы иРНК (кодирующей слитый белок PE) к PEgRNA составляет 1:2. В других вариантах осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEgRNA выбрано из группы, состоящей из приблизительно 1:1000, 1:900; 1:800; 1:700; 1:600; 1:500; 1:400; 1:300; 1:200; 1:100; 1:90; 1:80; 1:70; 1:60; 1:50; 1:40; 1:30; 1:20; 1:10; и 1:1. В других вариантах осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEgRNA выбрано из группы, состоящей из приблизительно 1:1000, 1:900; 800:1; 700:1; 600:1; 500:1; 400:1; 300:1; 200:1; 100:1; 90:1; 80:1; 70:1; 60:1; 50:1; 40:1; 30:1; 20:1; 10:1; и 1:1.

J. Применение праймерного редактирования для идентификации нецелевого редактирования объективным образом

[0903] Подобно другим геномным редакторам, существует некоторый риск того, что PE может вносить свои запрограммированные генетические изменения в нежелательных участках вокруг генома, то есть "нецелевых" участках. Однако в настоящее время отсутствуют описанные способы детектирования нецелевого редактирования при использовании праймерных редакторов. Такие способы позволили бы идентифицировать потенциальные участки нецелевого редактирования при использовании праймерных редакторов.

[0904] Ключевая концепция данного аспекта заключается в идее применения праймерного редактирования для вставки одной и той же адаптерной последовательности и/или участка связывания праймера в целевых и нецелевых участках, матрицей для которых служит одна и та же PEgRNA, для обеспечения быстрой идентификации геномных участков нецелевой модификации napDNAbp нуклеаз или праймерных редакторов. Данный способ является отличным от других способов, которые обеспечивают идентификацию нуклеазных нецелевых участков, поскольку адаптер и/или последовательность связывания праймера вставляется одновременно со связыванием ДНК и никированием с помощью napDNAbp, что упрощает дальнейшее процессирование.

[0905] На Фиг.33 показан основной принцип идентификации нецелевого редактирования. Данная фигура представляет собой схему, демонстрирующую конструкцию PEgRNA для вставок последовательности связывания праймера и вставку для связывания праймера в геномную ДНК с помощью праймерного редактирования для определения нецелевого редактирования. В данном варианте осуществления праймерное редактирование осуществляется внутри живой клетки, ткани, или модели с

животными. На первом этапе конструируют подходящую PEgRNA. На верхней схеме показан пример PEgRNA, которая может использоваться в данном аспекте. Спейсер в PEgRNA (обозначенный как "протоспейсер") комплементарен одной из цепей геномной мишени. Комплекс PE:PEgRNA (то есть комплекс PE) обеспечивает внесение одноцепочечного 3′-концевого флэпа в участке разрыва, который содержит закодированную последовательность связывания праймера и участок гомологии (кодируемый плечом гомологии PEgRNA), который комплементарен участку непосредственно ниже участка разрыва (отмечен красным). За счет внесения флэпа и процессов репарации/репликации ДНК синтезированная цепь становится встроенной в ДНК, тем самым обеспечивая внесение участка связывания праймера. Данный процесс может осуществляться в требуемой геномной мишени, но также в других геномных участках, которые могут взаимодействовать с PEgRNA нецелевыми образом (то есть PEgRNA направляет комплекс PE к другим нецелевым участкам вследствие комплементарности спейсерного участка в отношении других геномных участков, которые не являются требуемым геномным участком). Таким образом, последовательность связывания праймера может быть внесена не только в требуемую геномную мишень, но также нецелевые геномные участки в других местах генома. Для детектирования вставки данных участков связывания праймера как в требуемых геномных участках-мишенях, так и нецелевых геномных участках, геномная ДНК (после PE) может быть выделена, фрагментирована, и лигирована с нуклеотидами адаптера (отмечены красным). Далее может быть выполнена ПЦР с использованием ПЦР олигонуклеотидов, которые отжигаются на адаптерах и вставленной последовательности связывания праймера, для амплификации целевых и нецелевых участков геномной ДНК, в которые с помощью PE был вставлен участок связывания праймера. Затем может быть выполнено высокопроизводительное секвенирование и выравнивание последовательностей для идентификации точек вставки вставленных с помощью PE последовательностей связывания праймера на целевом участке или нецелевых участках.

[0906] Таким образом, на Фиг.33 проиллюстрирован один аспект, относящийся к идентификации нецелевых участков редактирования при редактировании внутри живой клетки, в культуре ткани, или моделях с использованием животных. Для реализации данного способа получают PEgRNA, которая имеет спейсер, являющийся идентичным конечному требуемому праймерному редактору (и, если смотреть на

нецелевые участки праймерного редактирования, последовательность участка связывания праймера, которая идентична конечному требуемому редактору), но включает необходимые последовательности для встраивания адаптера или участка связывания праймера после обратной транскрипции в рамках праймерного редактирования. *In vivo* редактирование осуществляется с помощью праймерного редактора или RT-слитой нуклеазы и выделения геномной ДНК. Геномная ДНК фрагментируется ферментативным или механическим способом и обеспечивает присоединение другого адаптера к участкам фрагментирования ДНК. ПЦР используется для амплификации от одного адаптера к адаптеру, внесенному с помощью PEgRNA. Полученный продукт подвергается глубокому секвенированию для идентификации всех модифицированных участков.

[0907] В другом аспекте оценка нецелевого редактирования при PE может быть осуществлена *in vitro*. В данном аспекте PE может использоваться при *in vitro* модификации в рамках идентификации геномной ДНК участков нецелевого редактирования с использованием *in vitro* модификации геномной ДНК. Для реализации данного способа собирается рибонуклеопротеин (RNP) из очищенного слитого белка праймерного редактора и PEgRNA (то есть комплекс PE) для встраивания адаптера или последовательности связывания праймера в участке-мишени, но при этом в других аспектах она является такой же, что и представляющая интерес PEgRNA. Данный RNP (то есть комплекс PE) инкубируется с выделенной геномной ДНК до или после фрагментации ДНК. После фрагментации различные адаптерные последовательности лигируются к концам фрагментированной ДНК. Для амплификации геномных участков со вставленной адаптерной последовательностью (то есть вставленной с помощью PE) используется ПЦР, при этом адаптеры лигируются к концам фрагментов. Высокопроизводительное секвенирование со сравнением адаптерных последовательностей может обеспечивать идентификацию геномных участков с модификацией, которые являются целевыми и нецелевыми. Данный *in vitro* способ редактирования должен повышать чувствительность детектирования, поскольку репарация клеточной ДНК не будет устранять подвергнутый обратной транскрипции ДНК-адаптер, добавленный с помощью праймерного редактора.

[0908] Данные способы могут применяться для идентификации нецелевого редактирования для любого праймерного редактора или любого геномного редактора,

в рамках которого используется направляющая РНК для распознавания участка-мишени разрыва (как в случае большинства Cas нуклеаз).

[0909] Данные способы могут применяться для любых генетических заболеваний, для лечения которых могут использоваться геномные редакторы.

[0910] Примеры адаптеров и/или последовательностей связывания праймера, которые могут быть внесены с помощью PE, включают, но без ограничения:

АДАПТЕР 1	5′-CGGTGGACCGATGATCT-3′ (SEQ ID NO: 177)
АДАПТЕР 2	5′-GCCACCTGGCTACTAGA-3′ (SEQ ID NO: 178)
АДАПТЕР 3	5′-AGATCATCGGTCCACCG-3′ (SEQ ID NO: 179)
АДАПТЕР 4	5′-TCTAGTAGCCAGGTGGC-3′ (SEQ ID NO: 180)

[0911] Данные адаптеры и/или последовательности связывания праймера также могут применяться на этапе лигирования после фрагментирования геномной ДНК, как описано выше.

[0912] Примеры конструкций PEgRNA, которые иллюстрируют применение описанного здесь способа для оценки нецелевого редактирования, и их локус-мишень редактирования являются следующими:

ОПРЕДЕЛЕНИЕ НЕЦЕЛЕВОГО ТЕСТОВОГО ЛОКУСА HEK3:	*GGCCCAGACTGAGCACGTGA*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC**TCTGCCATCA***CGGTG GACCGATGATCT*CGTGCTCAGTCTG (SEQ ID NO: 181)
ОПРЕДЕЛЕНИЕ НЕЦЕЛЕВОГО ТЕСТОВОГО ЛОКУСА HEK4:	GGCACTGCGGCTGGAGGTGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC**ACACAGCACCAGAG TCTCCGCTTTAACCCCCA***GCCACCTGGCTACTAGA*CCTCCAGCC (SEQ ID NO: 182)
ОПРЕДЕЛЕНИЕ НЕЦЕЛЕВОЙ КОРРЕКЦИИ СЕРПОВИДНЫХ КЛЕТОК:	GCATGGTGCACCTGACTCCTGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC**AGACTTCTCTTCA G***GCCACCTGGCTACTAGA*GAGTCAGGTGCAC (SEQ ID NO: 183)
ОБОЗНАЧЕНИЯ	*PBS спейсер* Каркас sgРНК **Матрица для синтеза DNA** *Адаптер* PBS

K. Применение праймерного редактирования для вставки доменов индуцируемой димеризации

[0913] Описанные здесь праймерные редакторы также могут применяться для внесения доменов димеризации в один или несколько белков-мишеней. Домены димеризации могут обеспечивать индуцируемое регулирование активности, связанной с димеризацией одного или нескольких белков-мишеней, с помощью связывающего фрагмента (например малой молекулы, пептида или белка), который связывается биспецифическим образом. В различных аспектах каждый домен димеризации при внесении в различные белки (например белки одного типа или разных типов) связывается с одним и тем же биспецифическим фрагментом (например биспецифической малой молекулой, пептидом, или полипептидом, имеющим, по меньшей мере, два участка, которые по отдельности связываются с доменами димеризации), тем самым вызывая димеризацию белков за счет их общего взаимодействия с биспецифическим лигандом. Таким образом, биспецифический лиганд функционирует в качестве "индуктора" димеризации двух белков. В некоторых случаях биспецифический лиганд или соединение будет иметь два фрагмента-мишени, которые являются одинаковыми. В других вариантах осуществления биспецифический лиганд или соединение будет иметь фрагменты-мишени, каждый из которых отличается друг от друга. Биспецифический лиганд или соединение, имеющее два одинаковых фрагмента-мишени, будет способно нацеливаться на тот же домен димеризации, внесенный в различные белки-мишени. Биспецифический лиганд или соединение, имеющее различные фрагменты-мишени, будет способно нацеливаться на различные домены димеризации, внесенные в различные белки-мишени.

[0914] В соответствии с используемым здесь значением термин "домен димеризации" обозначает лиганд-связывающий домен, который связывается со связывающимся фрагментом биспецифического лиганда. "Первый" домен димеризации связывается с первым связывающимся фрагментом биспецифического лиганда, а "второй" домен димеризации связывается со вторым связывающимся фрагментом того же биспецифического лиганда. Когда первый домен димеризации подвергается слиянию с первым белком (например за счет PE, как описано здесь), а второй домен димеризации (например за счет PE, как описано здесь) подвергается слиянию со вторым белком, первый и второй белок димеризуются в присутствии биспецифического лиганда, при этом биспецифический лиганд имеет, по меньшей мере, один фрагмент, который

связывается с первым доменом димеризации, и, по меньшей мере, другой фрагмент, который связывается со вторым доменом димеризации.

[0915] В соответствии с используемым здесь значением, термин "биспецифический лиганд" или "биспецифический фрагмент" обозначает лиганд, который связываются с двумя различными лиганд-связывающими доменами. В различных вариантах осуществления биспецифический фрагмент сам по себе представляет собой димер двух одинаковых или двух различных химических фрагментов, при этом каждый фрагмент специфично и прочно связывается с доменом димеризации. В определенных вариантах осуществления лиганд представляет собой низкомолекулярное соединение, или пептид, или полипептид. В других вариантах осуществления лиганд-связывающий фрагмент представляет собой "домен димеризации", который может быть внесен в виде пептидной метки в белок. В различных вариантах осуществления два белка, каждый из которых включает один и тот же или различные домены димеризации, могут быть индуцированы для димеризации за счет связывания каждого домена димеризации с биспецифическим лигандом. Данные молекула также могут быть обозначены как "химические индукторы димеризации" или CID. Кроме того, биспецифические лиганды могут быть получены путем соединения (например с помощью стандартных химических связей) двух одинаковых фрагментов вместе или двух различных фрагментов вместе, при этом каждый фрагмент прочно и специфично связывается с доменом димеризации.

[0916] В различных аспектах домены димеризации, внесенные с помощью PE, могут быть одинаковыми или различными.

[0917] Например, домены димеризации могут представлять собой FKBP12, который имеет следующую аминокислотную последовательность:

FKBP12	MGVQVETISPGDGRTFPKRGQTCVVHYTG MLEDGKKFDSSRDRNKPFKFMLGKQEVIR GWEEGVAQMSVGQRAKLTISPDYAYGATG HPGIIPPHATLVFDVELLKLE (SEQ ID NO: 488)

[0918] В другом примере домен димеризации может представлять собой мутант FKBP12, обозначаемый как FKBP12-F36V, то есть мутант FKBP12 со встроенным отверстием, которое связывает синтетический имеющий "выступы" миметик FK506 (2, Фиг.3)[107]:

FKBP12-F36V	MGVQVETISPGDGRTFPKRGQTCVVHYTG MLEDGKKVDSSRDRNKPFKFMLGKQEVI

	RGWEEGVAQMSVGQRAKLTISPDYAYGAT GHPGIIPPHATLVFDVELLKLE (SEQ ID NO: 489)

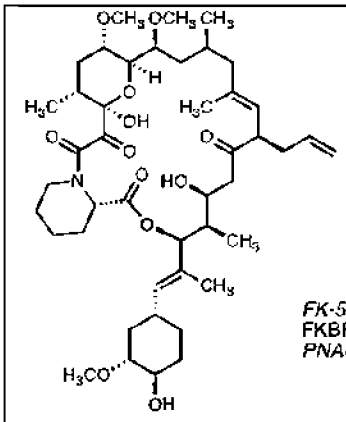
[0919] В другом примере домен димеризации может представлять собой циклофилин:

ЦИКЛОФИЛИН	MVNPTVFFDIAVDGEPLGRVSFELFADKVP KTAENFRALSTGEKGFGYKGSCFHRIIPGF MCQGGDFTRHNGTGGKSIYGEKFEDENFI LKHTGPGILSMANAGPNTNGSQFFICTAKT EWLDGKHVVFGKVKEGMNIVEAMERFGS RNGKTSKKITIADCGQLE (SEQ ID NO: 490)

[0920] В различных вариантах осуществления аминокислотные последовательности данных доменов димеризации могут быть изменены для оптимизации связывания или улучшения ортогональности связывания с нативными мишенями. Последовательности нуклеиновой кислоты генов, кодирующих низкомолекулярные связывающиеся белки, могут быть изменены для оптимизации эффективности процесса PE, например за счет снижения вторичной структуры PEgRNA.

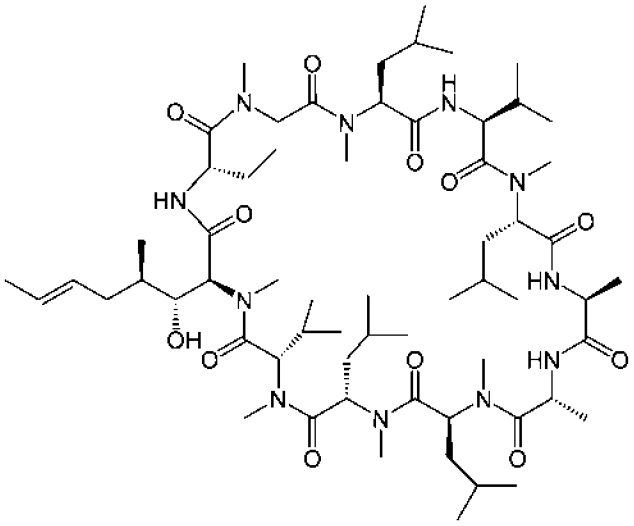
[0921] Другие примеры подходящих доменов димеризации и родственного низкомолекулярного соединения, которое связывается с ними, приведены ниже. Следует обратить внимание, что родственное низкомолекулярное соединение может быть соединено (например через химический линкер) со вторым низкомолекулярным соединением (которое представляет собой либо такое же соединение, либо другое соединение) для образования биспецифического лиганда, который может связывать два домена димеризации. В некоторых случаях, например при использовании FK506 и циклоспорина A, димеризация каждого из них (например FK506-FK506 или циклоспорин A-циклоспорин A) снижает или устраняет иммуносупрессивную активность мономерных соединений.

МАЛАЯ МОЛЕКУЛА – СВЯЗЫВАЕТСЯ С ДОМЕНОМ ДИМЕРИЗАЦИИ – ДИМЕР ДАННЫХ МОЛЕКУЛ БУДЕТ ЯВЛЯТЬСЯ БИСПЕЦИФИЧЕСКИМ ЛИГАНДОМ, КОТОРЫЙ БУДЕТ СВЯЗЫВАТЬСЯ С ДВУМЯ ДОМЕНАМИ ДИМЕРИЗАЦИИ	ДОМЕН(Ы) ДИМЕРИЗАЦИИ
FK506	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12: MGVQVETISPGDGRTFPKRGQTCVVH YTGMLEDGKKFDSSRDRNKPFKFMLG KQEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELLKL

<table>
<tr>
<td> 

МИШЕНИ: FKBP12 + КАЛЬЦИНЕВРИН</td>
<td>E (SEQ ID NO: 491)

КАЛЬЦИНЕВРИН:</td>
</tr>
<tr>
<td>ЦИКЛОСПОРИН А

МИШЕНИ: ЦИКЛОФИЛИН + КАЛЬЦИНЕВРИН</td>
<td>АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЦИКЛОСПОРИНА А ЧЕЛОВЕКА:

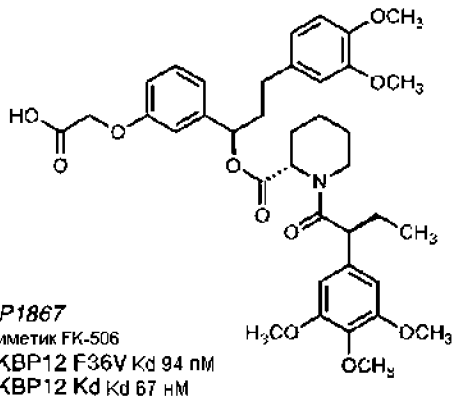
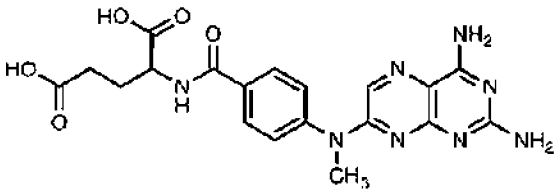
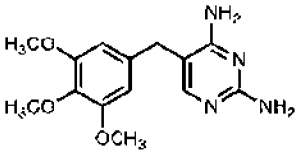
MVNPTVFFDIAVDGEPLGRVSFELF ADKVPKTAENFRALSTGEKGFGYK GSCFHRIIPGFMCQGGDFTRHNGTG GKSIYGEKFEDENFILKHTGPGILSM ANAGPNTNGSQFFICTAKTEWLDG KHVVFGKVKEGMNIVEAMERFGS RNGKTSKKITIADCGQLE (SEQ ID NO: 490)

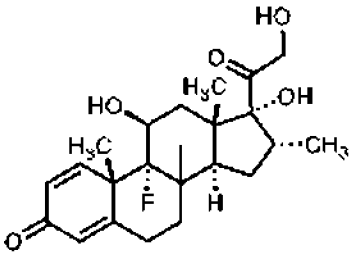
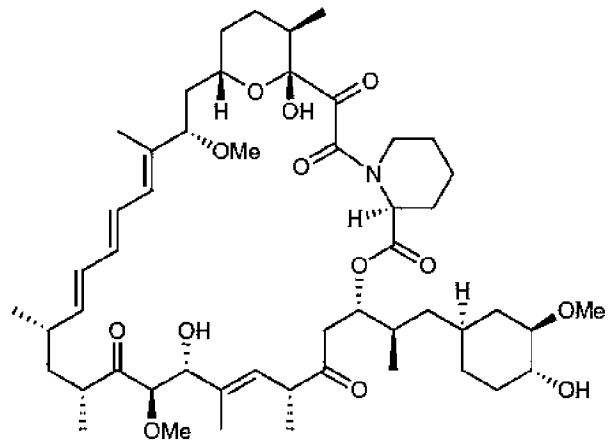
АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЦИКЛОСПОРИНА В ЧЕЛОВЕКА:

MLRLSERNMKVLLAAALIAGSVFF LLLPGPSAADEKKKGPKVTVKVYF DLRIGDEDVGRVIFGLFGKTVPKTV DNFVALATGEKGFGYKNSKFHRVIK DFMIQGGDFTRGDGTGGKSIYGER FPDENFKLKHYGPGWVSMANAGK DTNGSQFFITTVKTAWLDGKHVVF GKVLEGMEVVRKVESTKTDSRDKP LKDVIIADCGKIEVEKPFAIAKE (SEQ ID NO: 493)

АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЦИКЛОФИЛИНА С МЫШИ:

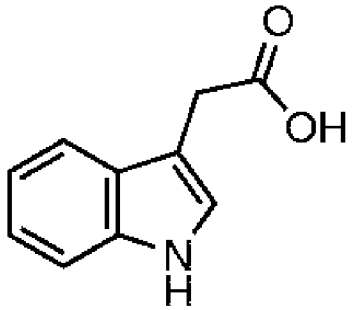
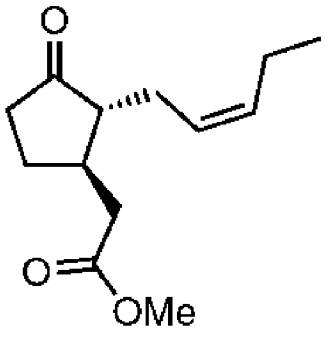
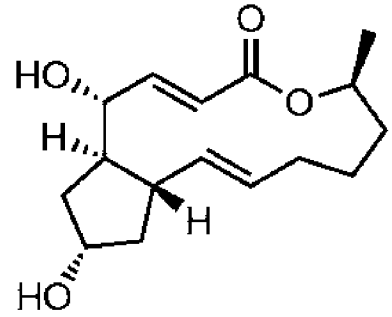
MSPGPRLLLPAVLCLGLGALVSSSG SSGVRKRGPSVTDKVFFDVRIGDK DVGRIVIGLFGNVVPKTVENFVALA</td>
</tr>
</table>

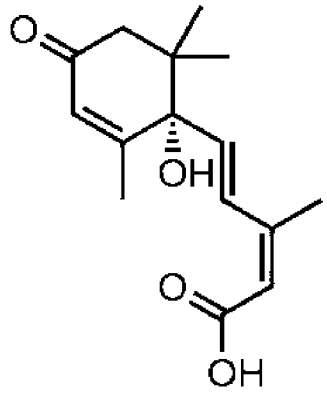
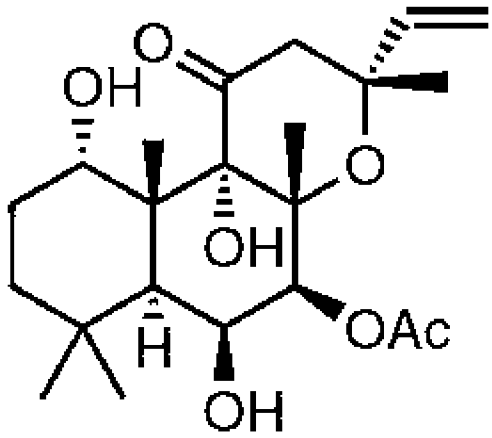
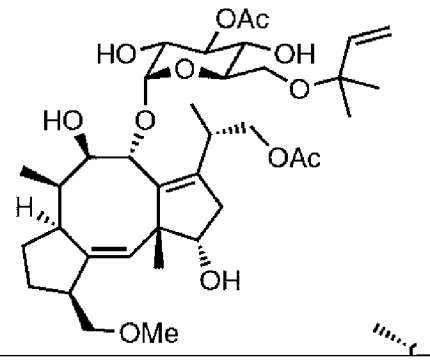
	TGEKGYGYKGSIFHRVIKDFMIQGG DFTARDGTGGMSIYGETFPDENFKL KHYGIGWVSMANAGPDTNGSQFFI TLTKPTWLDGKHVVFGKVLDGMT VVHSIELQATDGHDRPLTDCTIVNS GKIDVKTPFVVEVPDW (SEQ ID NO: 494)
AP1867  *AP1867* Миметик FK-506 FKBP12 F36V Kd 94 нМ FKBP12 Kd Kd 67 нМ *PNAS* **1998**, *95*, 10437. МИШЕНЬ (МИШЕНИ): FKBP12	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12-F36V MGVQVETISPGDGRTFPKRGQTCVVH YTGMLEDGKKVDSSRDRNKPFKFMLG KQEVIRGWEEGVAQMSVGQRAKLT ISPDYAYGATGHPGIIPPHATLVFDVELL KLE (SEQ ID NO: 495)
МЕТОТРЕКСАТ  *метотрексат* DHFR человека Kd < 10 нМ *J. Biol. Chem.* **1988**, *253*, 10304. *E. coli* DHFR Kd 9,5 нМ *PNAS* **2002**, *99*, 13481. МИШЕНЬ (МИШЕНИ): ДИГИДРОФОЛАТРЕДУКТАЗА	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ДИГИДРОФОЛАТРЕДУКТАЗЫ ЧЕЛОВЕКА MVGSLNCIVAVSQNMGIGKNGDLPWP PLRNEFRYFQRMTTTSSVEGKQNLVIM GKKTWFSIPEKNRPLKGRINLVLSRELK EPPQGAHFLSRSLDDALKLTEQPELAN KVDMVWIVGGSSVYKEAMNHPGHLK LFVTRIMQDFESDTFFPEIDLEKYKLLP EYPGVLSDVQEEKGIKYKFEVYEKND (SEQ ID NO: 496)
ТРИМЕТОПРИМ  *триметоприм* *E. coli* DHFR Kd 1,3 нМ *Biochemistry* **1982**, *21*, 5068. МИШЕНЬ (МИШЕНИ):	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ДИГИДРОФОЛАТРЕДУКТАЗЫ E. COLI MISLIAALAVDRVIGMENAMPWNLPAD LAWFKRNTLNKPVIMGRHTWESIGRPL PGRKNIILSSQPGTDDRVTWVKSVDEAI AACGDVPEIMVIGGGRVYEQFLPKAQ KLYLTHIDAEVEGDTHFPDYEPDDWES VFSEFHDADAQNSHSYCFEILERR (SEQ ID NO: 497)

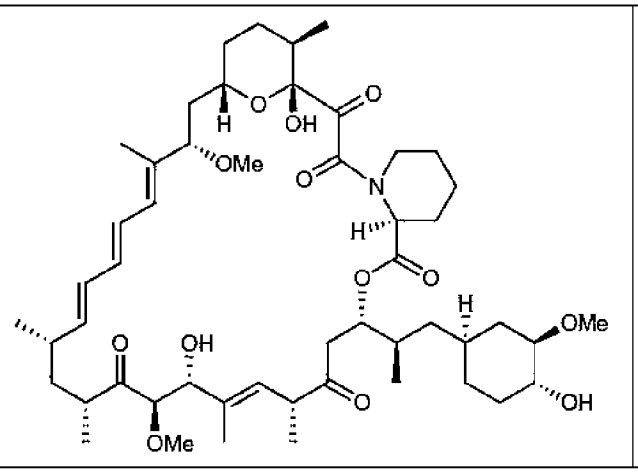
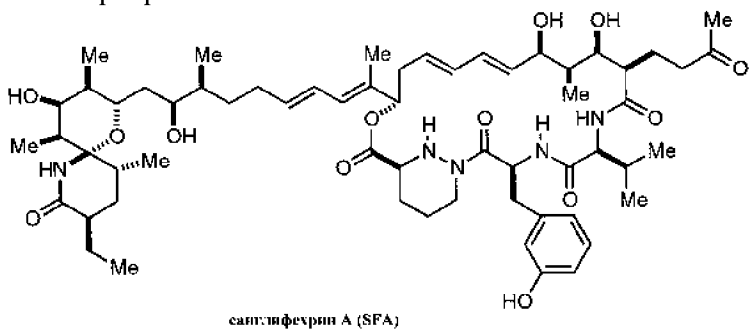
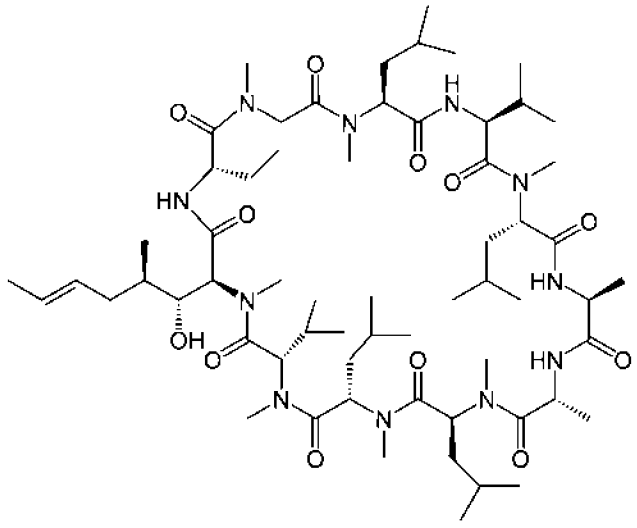
ДИГИДРОФОЛАТРЕДУКТАЗА	
ДЕКСАМЕТАЗОН  *дексаметазон* GR человека Kd 4,6 нМ *Mol. Endocrin.* **1999**, *13*, 1855.	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ГЛЮКОКОРТИКОИДНОГО РЕЦЕПТОРА ЧЕЛОВЕКА MDSKESLTPGREENPSSVLAQERGDV MDFYKTLRGGATVKVSASSPSLAVASQ SDSKQRRLLVDFPKGSVSNAQQPDLSK AVSLSMGLYMGETETKVMGNDLGFPQ QGQISLSSGETDLKLLEESIANLNRSTS VPENPKSSASTAVSAAPTEKEF PKTHSDVSSEQQHLKGQTGTNGGNVK LYTTDQSTFDILQDLEFSSGSPGKETNE SPWRSDLLIDENCLLSPLAGEDDSFLLE GNSNEDCKPLILPDTKPKIKDNGDLVLS SPSNVTLPQVKTEKEDFIELCTPGVIKQ EKLGTVYCQASFPGANIIGNKMSAISV HGVSTSGGQMYHYDMNTASLSQQQD QKPIFNVIPPIPVGSENWNRCQGSGDD NLTSLGTLNFPGRTVFSNGYSSPSMRPD VSSPPSSSSTATTGPPPKLCLVCSDEASG CHYGVLTCGSCKVFFKRAVEGQHNYL CAGRNDCIIDKIRRKNCPACRYRKCLQ AGMNLEARKTKKKIKGIQQATTGVSQ ETSENPGNKTIVPATLPQLTPTLVSLLEV IEPEVLYAGYDSSVPDSTWRIMTTLNM LGGRQVIAAVKWAKAIPGFRNLHLDD QMTLLQYSWMFLMAFALGWRSYRQS SANLLCFAPDLIINEQRMTLPCMYDQC KHMLYVSSELHRLQVSYEEYLCMKTL LLLSSVPKDGLKSQELFDEIRMTYIKEL GKAIVKREGNSSQNWQRFYQLTKLLD SMHEVVENLLNYCFQTFLDKTMSIEFP EMLAEIITNQIPKYSNGNIKKLLFHQK (SEQ ID NO: 498)
	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12: MGVQVETISPGDGRTFPKRGQTCVVH YTGMLEDGKKFDSSRDRNKPFKFMLG KQEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELLKL E (SEQ ID NO: 491)

РАПАМИЦИН	

[0922] Другие примеры встречающихся в природе бифункциональных молекул и их двойных рецепторов-мишеней являются следующими. Праймерное редактирование может применяться для внесения двойных рецепторов-мишеней в различные белки. После модификации различных белков с помощью PE в виде включения рецептора бифункциональной молекулы могут быть внесены бифункциональные молекулы, что вызывает димеризацию белков, которые были модифицированы для включения различных доменов димеризации. Примеры связывания (1) бифункциональных молекул и (2) их двойных рецепторов-мишеней являются следующими:

ВСТРЕЧАЮЩИЕСЯ В ПРИРОДЕ БИФУНКЦИОНАЛЬНЫЕ МОЛЕКУЛЫ	РЕЦЕПТОРЫ-МИШЕНИ БИФУНКЦИОНАЛЬНОЙ МОЛЕКУЛЫ
 ауксин	Рецептор-мишень 1: рецептор ауксина Рецептор-мишень 2: лигаза TIR1 E3
 метилжасмонат	Рецептор-мишень 1: рецептор JAZ Рецептор-мишень 2: лигаза Col1 E3
брефельдин А 	Рецептор-мишень 1: GBF1 Рецептор-мишень 2: GTPаза Arf1p

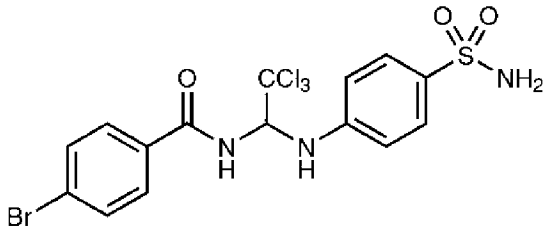
абсцизовая кислота  	Рецептор-мишень 1: рецептор PYR Рецептор-мишень 2: фосфопротеин фосфатаза 2C
форсколин  	Рецептор-мишень 1: мономеры аденилилциклазы Рецептор-мишень 2: мономеры аденилилциклазы
фузикокцин А  	Рецептор-мишень 1: белки 14-3-3 Рецептор-мишень 2: H^+-АТРаза
рапамицин	Рецептор-мишень 1: FKBP12 Рецептор-мишень 2: mTOR

	
санглифехрин A  санглифехрин A (SFA)	Рецептор-мишень 1: циклофилин Рецептор-мишень 2: IMP дегидрогеназа 2
циклоспорин A 	Рецептор-мишень 1: циклофилин Рецептор-мишень 2: кальциневрин

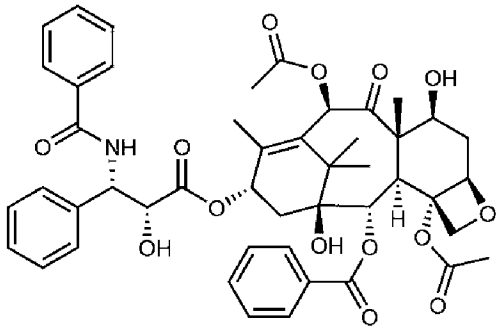
[0923]

[0924] Примеры других бифункциональных молекул, которые могут применяться в рамках данного аспекта праймерного редактирования, являются следующими:

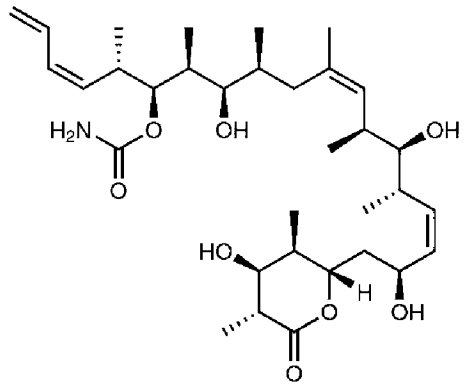
[0925] Синстаб А:



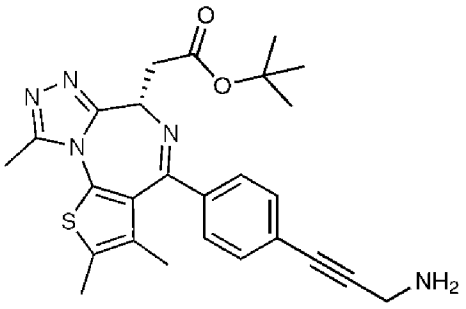
[0926] Паклитаксел:



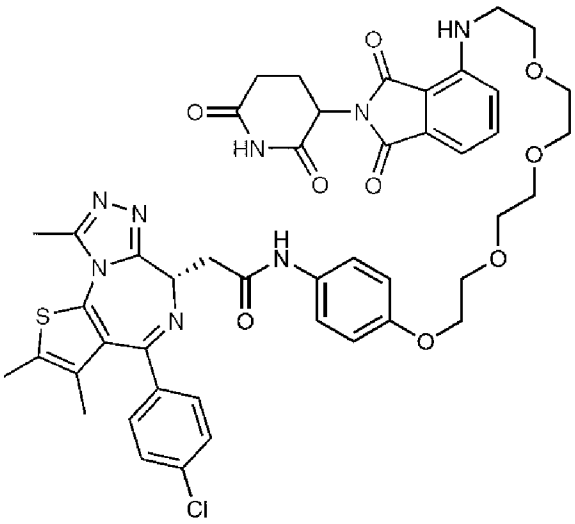
[0927] Дискодермолид:



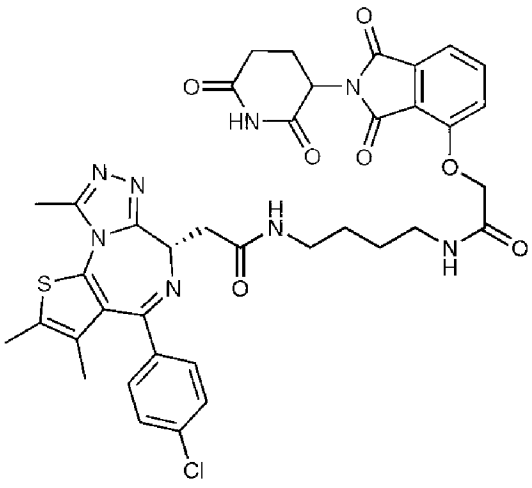
[0928] GNE-0011



[0929] ARV-825, и



[0930] dBET1



[0931] Синстаб A, паклитаксел, и дискодермолид представляют собой микротубулярные стабилизаторы. Таким образом, данные соединения могут применяться для димеризации белков, модифицированных с помощью PE, для

включения микротубулярных белков. GNE-0011, ARV-825, и dBET1 включают мотив связывания BRD4 и мотив связывания CRBN. Таким образом, данные соединения могут применяться для димеризации белков, модифицированных с помощью PE, для включения этих доменов-мишеней.

[0932] PEgRNA для внесения доменов димеризации могут включать следующие структуры (со ссылкой на Фиг.3D):

5′-[спейсер]-[ядро нРНК]-[плечо удлинения]-3′, где плечо удлинения включает 5′-[плечо гомологии]-[матрица для редактирования]-[участок связывания праймера]-3′; или

5′-[плечо удлинения]-[спейсер]-[ядро нРНК]-3′, где плечо удлинения включает 5′-[плечо гомологии]-[матрица для редактирования]-[участок связывания праймера]-3′, и где при любой конфигурации "матрица для редактирования" включает нуклеотидную последовательность домена димеризации.

[0933] В одном примере PEgRNA для вставки домена димеризации FKBP12 на C-конце рецептора инсулина человека является следующей (спейсер подчеркнут, ядро нРНК указано обычным шрифтом, флэп гомологии указан жирным шрифтом, вставка FKBP12 обозначена курсивом, участок отжига обозначен жирным курсивом):

PEGRNA ДЛЯ ВНЕСЕНИЯ FKBP12 В РЕЦЕПТОР ИНСУЛИНА ЧЕЛОВЕКА	CACGGUAGGCACUGUUAGGAGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC**UUGCCUCGGU CCAAUCCUUCC***GGAGUGCAGGUGGAAA CCAUCUCCCCAGGAGACGGGCGCACCUU CCCCAAGCGCGGCCAGACCUGCGUGGUG CACUACACCGGGAUGCUUGAAGAUGGAA AGAAAUUUGAUUCCUCCCGGGACAGAAA CAAGCCCUUUAAGUUUAUGCUAGGCAAG CAGGAGGUGAUCCGAGGCUGGGAAGAAG GGGUUGCCCAGAUGAGUGUGGGUCAGA GAGCCAAACUGACUAUAUCUCCAGAUUA UGCCUAUGGUGCCACUGGGCACCCAGGC AUCAUCCCACCACAUGCCACUCUCGUCU UCGAUGUGGAGCUUCUAAAACUGGAA**UA ACAGUGCCUACC*** (SEQ ID NO: 499)

[0934] В другом примере PEgRNA для вставки домена димеризации FKBP12 в локус HEK3 (для оптимизации) является следующей:

PEGRNA ДЛЯ ВНЕСЕНИЯ FKBP12 В	GGCCCAGACTGAGCACGTGAGTTTTAGA

HEK3	GCTAGAAATAGCAAGTTAAAATAAGGC TAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTGGAGGAAGCAGGG CTTCCTTTCCTCTGCCATCATTCCAGTTT TAGAAGCTCCACATCGAAGACGAGAGT GGCATGTGGTGGGATGATGCCTGGGTGC CCAGTGGCACCATAGGCATAATCTGGAG ATATAGTCAGTTTGGCTCTCTGACCCAC ACTCATCTGGGCAACCCCTTCTTCCCAG CCTCGGATCACCTCCTGCTTGCCTAGCA TAAACTTAAAGGGCTTGTTTCTGTCCCG GGAGGAATCAAATTTCTTTCCATCTTCA AGCATCCCGGTGTAGTGCACCACGCAGG TCTGGCCGCGCTTGGGGAAGGTGCGCCC GTCTCCTGGGGAGATGGTTTCCACCTGC ACTCCCGTGCTCAGTCTG (SEQ ID NO: 500)

[0935] Белки-мишени для внесения доменов димеризации не являются конкретно ограниченными; однако их димеризации (после модифицирования с помощью PE) в присутствии биспецифического лиганда продуцирует некоторый полезный биологический эффект, например путь передачи сигнала, сниженную иммунореактивность и так далее. В различных аспектах белки-мишени, подвергаемые димеризации с помощью PE-зависимого внесения доменов димеризации, могут представлять собой один и тот же белок или разные белки. Предпочтительно, белки при димеризации индуцируют один или несколько последующих биологических каскадов, например каскад трансдукции сигнала, фосфорилирование, и так далее. Примеры белков-мишеней, для которых может использоваться PE для внесения доменов димеризации, включают, но без ограничения:

СВЯЗАННЫЙ С МЕМБРАНОЙ РЕЦЕПТОР	КИНАЗНЫЙ ДОМЕН, СЛИТЫЙ С	ИСПОЛЬЗУЕМЫЙ CID	КАСКАД ПЕРЕДАЧИ СИГНАЛА	ССЫЛКА
Т-КЛЕТОЧНЫЙ РЕЦЕПТОР	FKBP12	FK1012 (ДИМЕР FK506)	СИГНАЛ Т-КЛЕТОЧНОГО РЕЦЕПТОРА	*SCIENCE* 262, 1019–24 (1993) *CHEM. BIOL.* 1, 163–172 (1994).
РЕЦЕПТОР FAS	ЦИКЛОФИЛИН С МЫШИ	ДИМЕР ЦИКЛОСПОРИНА А	ПУТЬ FAS ДЛЯ АПОПТОЗА	*CHEM. BIOL.* 3, 731–738 (1996).

РЕЦЕПТОР ИНСУЛИНА	FKBP12	FK1012 (ДИМЕР FK506)	СИГНАЛ ИНСУЛИНА	*CURR. BIOL.* 8, 11–18 (1998).
ПОЛУЧЕННЫЙ ИЗ ТРОМБОЦИТОВ ФАКТОР РОСТА БЕТА (PDGF)	FKBP12	FK1012 (ДИМЕР FK506)	СИГНАЛ ОБРАЗОВАНИЯ МЕЗОДЕРМЫ PDGF	*CURR. BIOL.* 8, 11–18 (1998).
РЕЦЕПТОР ЭРИТРОПОЭТИНА (EPOR)	FKBP12	FK1012 (ДИМЕР FK506)	EPOR-ОПОСРЕДОВАННЫЙ ПРОЛИФЕРАТИВНЫЙ СИГНАЛ	*PROC. NATL. ACAD. SCI.* 94, 3076–3081 (2002).

[0936] В одном аспекте описанные здесь праймерные редакторы могут применяться для внесения последовательностей, кодирующих домены димеризации, в один или несколько генов, кодирующих представляющие интерес белки-мишени в живой клетке или пациенте. Это может быть обозначено как "система праймерного редактирования–CID", где CID представляет собой биспецифический лиганд, который индуцирует димеризацию белков-мишеней, каждый из которых слит с доменом димеризации, внесенным с помощью PE. Данное изменение самостоятельно не должно иметь физиологического эффекта. При введении биспецифического лиганда, который обычно представляет собой димерную малую молекулу, которая может одновременно связываться с двумя доменами димеризации, каждый из которых слит с копией белка-мишени, биспецифический лиганд вызывает димеризацию белка-мишени. Данная димеризация белка-мишени затем вызывает биологический сигнал, например эритропоэз или сигнал инсулина. Здесь описан новый способ внесения индуцируемых димеризацией биологических процессов, например сигнала рецептора, под контролем удобного низкомолекулярного лекарства (то есть биспецифического лиганда) путем геномного встраивания генов, кодирующих связывающие малые молекулы белки (то есть домены димеризации), с помощью праймерного редактирования.

[0937] Димеризация белков представляет собой повсеместный биологический процесс. Следует отметить, что известно, что гомодимеризация многих связанных с мембраной рецепторов инициирует каскады передачи сигнала, зачастую со значительными биологическими последствиями. Ряд низкомолекулярных встречающихся в природе

продуктов, одобренных для применения в качестве лекарств, действуют как химические индукторы димеризации белков в виде части механизма их действия.[92] Например, FK506 прочно связывается с FKBP12, при этом образующийся комплекс малая молекула-белок затем связывается с фосфатазой кальциневрином, тем самым ингибируя один этап в рамках передачи сигнала Т-клеточного рецептора.[93] Аналогичным образом, циклоспорин А индуцирует димеризацию циклофилина и кальциневрина, а рапамицин индуцирует димеризацию FKBP и mTOR.[93,94]

[0938] В одном варианте осуществления на основе использования селективного высокоаффинного связывания типа малая молекула:белок в виде FK506:FKBP12 и циклоспорин А:циклофилин также были разработаны синтетические химические индукторы димеризации. В рамках одного примера было показано, что малая молекула, состоящая из двух единиц FK506, обозначенная как FK1012, оказывает влияние на трансдукцию сигнала, когда цитоплазматические домены сигнальных рецепторов были помечены FKBP12.[95] После этого химические индукторы димеризации (CID) стали использоваться для управления рядом путей передачи сигнала. [96-103]

[0939] Являясь полезными инструментами для исследования биологических процессов, одной из проблем, связанных с использованием CID в рамках терапевтических вариантов использования, является то, что внесение химер белка-мишени в виде FKBP12 (или циклофилина) пациентам является сложным.

[0940] Настоящее изобретение объединяет две концепции для создания ранее недоступного терапевтического процесса. Первая концепция представляет собой праймерное редактирование, которое описано здесь, которое обеспечивает возможность точного геномного редактирования, включая нацеленные вставки, в живых клетках. Вторая концепция представляет собой химически индуцируемую димеризацию, которая является мощным инструментом, обеспечивающим контроль процессов передачи сигнала и олигомеризации в клеточной культуре с помощью малых молекул.

[0941] Били идентифицированы конкретные случаи, при которых химический контроль над димеризацией белка может иметь полезный терапевтический эффект.

[0942] Рецептор инсулина представляет собой гетеротетрамерный трансмембранный белок, который отвечает за связывание инсулина с внеклеточным доменом за счет фосфорилирования цитоплазматического домена киназы.[104] Модифицированных

химерный белок, состоящий из компонента локализации на мембране, С-концевого киназного домена рецептора инсулина, и трех копий FKBP12, отвечает FK1012 и инициирует инсулиновый ответ в клеточной культуре.[99] Аналогичным образом, ожидается, что слияние FKBP12 с С-концом киназного домена нативного рецептора инсулина в клетках пациента должно обеспечивать FK1012-зависимое фосфорилирование и инициирование каскада передачи сигнала инсулина. Данная система должна заменять или дополнять использование инсулина у пациентов, которые не могут продуцировать инсулин (например при диабете 1 типа), или тех из них, которые имеют слабую реакцию на инсулин (например при диабете 2 типа).

[0943] Кроме того, эритропоэтин стимулирует пролиферацию эритроцитов за счет связывания с рецептором эритропоэтина (EpoR), индуцируя димеризацию или конформационное изменение в предварительно образованном димере рецептора, которая приводит к активации каскада передачи сигнала Jak/STAT.[105] Было продемонстрировано, что FK1012-индуцированная олигомеризация связанного с мембраной цитоплазматического домена EpoR, меченного FKBP12, является достаточной для инициирования каскада передачи сигнала Jak/STAT и промотирования клеточной пролиферации.[106] Ожидается, что слияние FKBP12 с нативным EpoR с помощью праймерного редактирования в клетках пациентов обеспечит FK1012-индуцированный контроль над пролиферацией эритроцитов (эритропоэз). Данная система может применяться для инициирования роста красных кровяных клеток у пациентов с анемией. FK1012-индуцируемый EpoR также может использоваться в качестве *in vivo* маркера отбора для кровяных клеток, которые подверглись *ex vivo* модифицированию.

[0944] В принципе, любая рецепторная тирозинкиназа может представлять собой жизнеспособную мишень для лекарства на основе праймерного редактирования-CID. Таблица ниже включает перечень всех рецепторных тирозинкиназ в геноме человека.[110]

Семейство	Рецептор	Синонимы	Кат. # NT	Кат. # PROT	Хромосома
Семейство ALK	ALK	Kl	NM_004304	NP_004295	2p23
	LTK	TYK1	NM_002344	NP_002335	15q15.1-q21.1
Семейство AXL	AXL	UFO, Tyro7(r) Ark(m)	NM_001699	NP_001690	19q13.1
	MER	MERTK, NYK, Eyk(ch)	NM_006343	NP_006334	2q14.1
	TYRO3	RSE, SKY, BRT, DTK, TIF	NM_006293	NP_006284	15q15.1-q21.1
Семейство DDR	DDR1	CAK, TRKE, NEP, NTRK4, EDDR1, PTK3	NM_013993	NP_001945	6p21.3
	DDR2	TKT, TYRO10, NTRKR3	NM_006182	NP_006173	1q21-q22
Семейство EGFR	EGFR	ERBB, ERBB1	NM_005228	NP_005219	7p12
	ERBB2	HER2, Neu(r), NGL	NM_004448	NP_004439	17q11.2-q12
	ERBB3	HER3	NM_001982	NP_001973	12q13
	ERBB4	HER4	NM_005235	NP_005226	2q33.3-q34
Семейство EPH	EPHA1	EPH, EPHT	NM_005232	NP_005223	7q32-q36
	EPHA2	ECK, Sek2(m), Myk2(m)	NM_004431	NP_004422	1p34
	EPHA3	HEK, ETK1, Tyro4(r), Mek4(m), Cek4(ch)	NM_005233	NP_005224	3p11.2
	EPHA4	HEK8, Tyro1(r), Sek1(m), Cek8(ch)	NM_004438	NP_004429	2q36-qter
	EPHA5	HEK7, Ehk1(r), Bsk(r), Cek7(ch)	L36644	P54756	
	EPHA6	DKFZp434C1418, Ehk2(r)	AL133666		
	EPHA7	HEK11, Mdk1(m), Ebk(m), Ehk3(r), Cek11(ch)	NM_004440	NP_004431	6q21
	EPHA8	HEK3, KIAA1459, Eek(r), Cek10(ch)	AB040892	CAB81612	1q23-q24
	EPHB1	NET, EPHT2, HEK6, Elk(r), Cek6(ch)	NM_004441	NP_004432	3q21-q23
	EPHB2	HEK5, ERK, DRT, EPHT3, Tyro5(r), Nuk(m), Sek3(m), Cek5(ch)	AF025304	AAB94602	1p36.1-p35
	EPHB3	HEK2, Tyro6, Mdk5(m), Sek4(m)	NM_004443	NP_004434	3q21-qter
	EPHB4	HTK, Tyro11(r), Mdk2(m), Myk1(m)	NM_004444	NP_004435	
	EPHB6	HEP, Mep(m), Cek1(ch)	NM_004445	NP_004436	7q33-q35
Семейство FGFR	FGFR1	FLT2, bFGFR, FLG, N-SAM	M34641	AAA35835	8p11.2
	FGFR2	KGFR, K-SAM, Bek(m), CFD1, JWS, Cek3(ch)	NM_000141	NP_000132	10q26
	FGFR3	HBGFR, ACH, Cek2(ch)	NM_000142	NP_000133	4p16.3
	FGFR4		NM_002011	NP_002002	5q35.1-qter
Семейство INSR	IGF1R	JTK13	NM_000875	NP_000866	15q25-q26
	INSR	IR	NM_000208	NP_000199	19p13.3-p13.2
	INSRR	IRR	J05046	AAC31759	1q21-q23
Семейство MET	MET	HGFR	NM_000245	NP_000236	7q31
	RON	MST1R, CDw136, Pv2(m), STK(m), SEA(ch)	NM_002447	NP_002438	3p21.3
Семейство MUSK	MUSK	Nsk2(m), Mik1(m), Mtk2(m)	NM_005592	NP_005583	9q31.3-q32
Семейство PDGFR	CSF1R	FMS, C-FMS, CD115	NM_005211	NP_005202	5q31-q32
	FLT3	FLK2, STK1, CD135	NM_004119	NP_0041110	13q12
	KIT	Sfr(m), CKIT	NM_000222	NP_000213	4q11-q12
	PDGFRA		NM_006206	NP_006197	4q11-q13
	PDGFRB	PDGFR, JTK12	NM_002609	NP_002600	5q31-q32
Семейство PTK7	PTK7	CCK4, KLG(ch)	NM_002821	NP_002812	6p21.1-p12.2
Семейство RET	RET	MEN2A/B, HSCR1, MTC1	X12949	P07949	10q11.2
Семейство ROR	ROR1	NTRKR1	NM_005012	NP_005003	1p32-p31
	ROR2	NTRKR2	NM_004560	NP_004551	
Семейство ROS	ROS1	MCF3	NM_002944	NP_002935	6q22
Семейство RYK	RYK	Vik(m), Mrk(m)	S59184	AAB26341	3q22
Семейство TIE	TEK	TIE2	NM_000459	NP_000450	9p21
	TIE	TIE1, JTK14	NM_005424	NP_005415	1p34-p33
Семейство TRK	NTRK1	TRK, TRKA	NM_002529	NP_002520	1q21-q22
	NTRK2	TRKB	NM_006180	NP_006171	9q22.1
	NTRK3	TRKC	NM_002530	NP_002521	15q25
Семейство VEGFR	VEGFR1	FLT1	NM_002019	NP_002010	13q12
	VEGFR2	KDR, FLK1		AAB88005	4q11-q12
	VEGFR3	FLT4, PCL	NM_002020	NP_002011	5q34-q35
Семейство AATYK	AATYK	AATK, KIAA0641	NM_004920	NP_004911	17q25.3
	AATYK2	KIAA1079	NM_014916	NP_055731	7q21-q22
	AATYK3				19q13.2-q13.3

[0945] Существует множество преимуществ системы праймерного редактирования–CID. Одно подобное преимущество заключается в том, что она может обеспечивать замену эндогенных лигандов, которые обычно представляют собой белки, вызывающие трудности в плане производства, стоимости, доставки, продуцирования, или хранения, подобными лекарствам малыми молекулами, которые могут быть

введены орально вместо введения в/в или с помощью инъекции, могут быть быстро получены из одобренных FDA лекарств (или сами по себе являются лекарствами), и не несут особых затрат в плане производства или хранения, которые обычно связаны с лекарствами на основе белков. Другое преимущество заключает в том, что изменение само по себе не должно иметь физиологического эффекта. Величина димеризации белка-мишени может контролироваться на основе дозировки низкомолекулярного CID. Кроме того, димеризация белка-мишени является легко и быстро обратимой путем добавления мономерной формы CID. Другое преимущество заключается в том, что в случаях, когда один лиганд нацелен на несколько рецепторов, может быть достигнута селективность путем праймерного редактирования только одного рецептора. Наконец, в зависимости от используемого способа доставки для праймерного редактирования, также является возможным ограничивать редактирование локализованной тканью или органом, что обеспечивает активацию индуцируемого рецептора только в определенных областях.

[0946] Если показатели эффективности редактирования являются достаточно высокими при праймерном редактировании, так что два отдельных события редактирования могут осуществляться с высокими уровнями, то является также возможным пометить два представляющих интерес белка различными связывающими малые молекулы доменами (например FKBP и цикрофилином) и индуцировать гетеродимеризацию с низкомолекулярными гетеродимерами (например димером FK506–циклоспорин A).

[0947] Было достигнуто слияние FKBP12 или других связывающих малые молекулы белков с нативными белками, в целом с помощью сверхэкспрессии из плазмиды в культуре ткани. Было продемонстрировано, что последующая химически индуцированная димеризация индуцирует фенотипические изменения в клетках, продуцирующих слитые белки.

[0948] Следующие ссылки приведены выше в Разделе G и приведены здесь в качестве ссылки.

1. Crabtree, G. R. & Schreiber, S. L. Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem. Sci.* 21, 418–22 (1996).

2. Liu, J. *et al.* Calcineurin Is a Common Target of A and FKBP-FK506 Complexes. *Cell* 66, 807–815 (1991).

3. Keith, C. T. *et al.* A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* 369, 756–758 (2003).

4. Spencer, D. M., Wandless, T. J., Schreiber, S. L. S. & Crabtree, G. R. Controlling signal transduction with synthetic ligands. *Science* 262, 1019–24 (1993).

5. Pruschy, M. N. *et al.* Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chem. Biol.* 1, 163–172 (1994).

6. Spencer, D. M. *et al.* Functional analysis of Fas signaling in vivo using synthetic inducers of dimerization. *Curr. Biol.* 6, 839–847 (1996).

7. Belshaw, P. J., Spencer, D. M., Crabtree, G. R. & Schreiber, S. L. Controlling programmed cell death with a cyclophilin-cyclosporin-based chemical inducer of dimerization. *Chem. Biol.* 3, 731–738 (1996).

8. Yang, J. X., Symes, K., Mercola, M. & Schreiber, S. L. Small-molecule control of insulin and PDGF receptor signaling and the role of membrane attachment. *Curr. Biol.* 8, 11–18 (1998).

9. Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc.Natl. Acad. Sci.* 93, 4604–4607 (2002).

10. Stockwell, B. R. & Schreiber, S. L. Probing the role of homomeric and heteromeric receptor interactions in TGF-β signaling using small molecule dimerizers. *Curr. Biol.* 8, 761–773 (2004).

11. Spencer, D. M., Graef, I., Austin, D. J., Schreiber, S. L. & Crabtree, G. R. A general strategy for producing conditional alleles of Src-like tyrosine kinases. *Proc. Natl. Acad. Sci.* 92, 9805–9809 (2006).

12. Holsinger, L. J., Spencer, D. M., Austin, D. J., Schreiber, S. L. & Crabtree, G. R. Signal transduction in T lymphocytes using a conditional allele of Sos. *Proc. Natl. Acad. Sci.* 92, 9810–9814 (2006).

13. Myers, M. G. Insulin Signal Transduction and the IRS Proteins. *Annu. Rev. Pharmacol. Toxicol.* 36, 615–658 (1996).

14. Watowich, S. S. The erythropoietin receptor: Molecular structure and hematopoietic signaling pathways. *J. Investig. Med.* 59, 1067–1072 (2011).

15. Blau, C. A., Peterson, K. R., Drachman, J. G. & Spencer, D. M. A proliferation switch for genetically modified cells. *Proc. Natl. Acad. Sci.* 94, 3076–3081 (2002).

16. Clackson, T. *et al.* Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc. Natl. Acad. Sci.* 95, 10437–10442 (1998).

17. Diver, S. T. & Schreiber, S. L. Single-step synthesis of cell-permeable protein dimerizers that activate signal transduction and gene expression. *J. Am. Chem. Soc.* 119, 5106–5109 (1997).

18. Guo, Z. F., Zhang, R. & Liang, F. Sen. Facile functionalization of FK506 for biological studies by the thiol-ene ‘click’ reaction. *RSC Adv.* 4, 11400–11403 (2014).

19. Robinson, D. R., Wu, Y.-M. & Lin, S.-F. The protein tyrosine kinase family of the human genome. *Oncogene* 19, 5548–5557 (2000).

L. <u>Применение праймерного редактирования для записи клеточных данных</u>

[0949] Праймерные редакторы и полученные и их помощью геномные модификации также могут применяться для исследования и записи клеточных процессов и развития. Например, описанные здесь праймерные редакторы могут применяться для записи наличия и продолжительности стимула в клетке за счет поступления в клетку первой последовательности нуклеиновой кислоты, которая кодирует слитый белок с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и обратную транскриптазу, и поступления в клетку, по меньшей мере, второй последовательности нуклеиновой кислоты, которая кодирует PEgRNA. Либо первая, либо вторая, или обе последовательности нуклеиновой кислоты являются функционально связанными с индуцируемыми промоторами, которые отвечают за стимул клетки, так что он индуцирует экспрессию слитого белка и/или PEgRNA, тем самым вызывая модифицирование последовательности-мишени в клетке.

[0950] Описанные здесь праймерные редакторы также могут применяться для мечения клеток и отслеживания происходления. Например, путем мечения каждой клетки уникальной геномной меткой праймерный редактор может помогать выявлять карту происхождения клеток за счет обеспечения конструирования филогенетических деревьев на основе модификаций, реализованных в одной или нескольких последовательностях-мишенях. Начиная с прогениторных клеток система праймерного редактора может обеспечивать построение карты происхождения клеток для отдельных клеток в рамках всего организма, которая может быть расшифрована путем анализа модификаций, реализованных в одной или нескольких последовательностях-мишенях. Способ отслеживания происхождения клеток может включать получение нуклеиновой кислоты, кодирующей слитый белок с napDNAbp и обратной транскриптазой, и получение, по меньшей мере, одной второй нуклеиновой кислоты, кодирующей PRнРНК. Уникальная клеточная метка может быть получена с использованием слитого белка и PEgRNA для создания одной или нескольких модификаций в одной или нескольких последовательностях-мишенях, тем самым обеспечивая отслеживание происхождения любой клетки, которая образуется из первой клетки, с помощью уникальной клеточной метки. Применение праймерных редакторов как для записи клеточных данных, так и отслеживания происхождение описано более подробно в Примере 13.

[0951] Праймерные редакторы могут обеспечивать как отслеживание происхождения, так и запись клеточных сигналов с помощью модифицирования геномных

последовательностей-мишеней или встроенных предварительно сконструированных последовательностей. В рамках праймерных редакторов используется синтетический слитый белок, включающий фрагменты никазы Cas9 (включая, но без ограничения, вариант H840A SpCas9) и домен обратной транскриптазы, наряду со сконструированной направляющей нРНК для праймерного редактирования (PEgRNA). Данные компоненты совместно нацеливаются на определенную геномную последовательность или встроенную предварительно сконструированную последовательность и обеспечивают внесение предварительного определенного изменения. Поскольку PEgRNA обеспечивает указание как геномной последовательности-мишени, так и результата редактирования, может быть достигнуто высокоспецифическое и контролируемое геномное модифицирование одновременно при использовании нескольких PEgRNA в рамках одной клетки. Вспомогательные геномные модификации включают все замещения одиночных нуклеотидов, вставки последовательностей с размером от малого до среднего, и делеции с размером от малого до среднего. Универсальность данной технологии геномного редактирования может обеспечивать возможность временно связанной сигнал-специфичной записи в клетках.

[0952] Применение праймерных редакторов для записи клеточных данных может включать композиции (например нуклеиновые кислоты), клетки, системы, наборы, и способы для записи силы и/или продолжительности эндогенных или экзогенных стимулов в течение срока жизни клетки. Система для записи клеточных данных может включать слитый белок, состоящий из napDNAbp (например домена Cas9) и обратной транскриптазы, функционально связанной с промотором, который индуцирует экспрессию слитого белка для индуцирования изменений за счет создание нацеленных и специфичных в плане последовательности геномных вставок, делеций или мутаций в ответ на стимул или изменение в клетке. В отличие от устройств на основе цифровой памяти, которые хранят информацию (например связанную с наличием или отсутствием стимула) в виде одного из двух отдельных состояний (то есть "on" или "off"), данные средства для записи клеточных данных могут индуцировать перманентные отметки в клеточной ДНК, которые соответствуют как силе (то есть амплитуде), так и продолжительности одного или нескольких стимулов. Таким образом, в некоторых аспектах системы для записи клеточных данных имеют способность одновременно записывать несколько состояний клетки, включая,

например, подвергание действию малой молекулы, белка, пептида, аминокислоты, метаболита, неорганической молекулы, органометаллической молекулы, органической молекулы, лекарства или лекарства-кандидата, сахара, липида, металла, нуклеиновой кислоты, молекулы, полученный при активации эндогенного или экзогенного каскада передачи сигнала, света, температуры, звука, давления, механического давления, давления вследствие сдвига, или вируса или другого микроорганизма. Данные средства для записи клеточных данных могут использовать технологии секвенирования (например высокопроизводительное секвенирование) для измерения считывания (например изменений в клеточной ДНК) и не зависят от крупных клеточных популяций как в плане записи стимула, так и считывания изменения (изменений) в клеточной ДНК, индуцированных стимулом.

[0953] В целом, предусмотренные здесь системы для записи клеточных данных для применения в клетке включают слитый белок, состоящий из napDNAbp и обратной транскриптазы, при этом последовательность нуклеиновой кислоты, кодирующая слитую плазмиду, функционально связана с промотором (например индуцируемым промотором или конститутивным промотором). При наличии стимула или возникновении изменения состояния клетки стимул индуцирует экспрессию слитого белка. В клетке также присутствуют одна или несколько нуклеиновых кислот, кодирующих, по меньшей мере, одну PEgRNA, которая связывается с napDNAbp и направляет napDNAbp или другой слитый белок к последовательности-мишени (то есть PEgRNA комплементарна последовательности-мишени). Нуклеиновая кислота, кодирующая PEgRNA, также может быть или может быть в качестве альтернативы функционально связана с промотором (например индуцируемым промотором или конститутивным промотором). При соответствующем стимуле или соответствующем наборе стимулов как слитый белок, так и PEgRNA экспрессируются в клетки, при этом PRнРНК связывается со слитым белком для его направления к последовательности-мишени. Данная последовательность-мишень записывает активность праймерного редактора, тем самым обеспечивая запись наличия стимула или набора стимулов, или изменения состояния клетки. В клетке может присутствовать более чем одна последовательность PEgRNA, при этом каждая из данных дополнительных последовательностей PEgRNA, которые могут направлять слитый белок к определенным последовательностям-мишеням, может быть функционально связана с промотором, который чувствует присутствие другого стимула, что обеспечивает

возможность конструирования сложных систем для записи клеточных данных для упорядоченной записи присутствия и продолжительности стимула или набора стимулов. В некоторых случаях один или несколько компонентов системы для записи клеточных данных (например слитый белок или PEgRNA) может быть конститутивно экспрессирован в клетке. Здесь описаны примеры компонентов системы для записи клеточных данных для применения с композициями. Дополнительные подходящие комбинации компонентов, которые описаны здесь, будут очевидны среднему специалисту в данной области на основе данного описания и знаний в этой области, при этом они, таким образом, входят в объем настоящего изобретения.

[0954] Повторная модификация ДНК-мишени, которая может быть подвергнута секвенированию с помощью нацеленного секвенирования ампликонов и/или секвенирования РНК (которое является особенно ценным для экспериментов в виде записи с использованием отдельных клеток), может применяться для записи множества важных биологических процессов, включая активацию каскадов передачи сигнала, метаболические состояния, и программы клеточной дифференциации. Соединение внутренних и внешних клеточных сигналов с модификацией последовательности в геноме является возможным для любого сигнала, для которого существует отвечающий на сигнал промотор. В некоторых вариантах осуществления промотор представляет собой промотор, подходящий для применения в рамках прокариотической системы (например бактериальный промотор). В некоторых вариантах осуществления промотор представляет собой промотор, подходящий для применения в рамках эукариотической системы (например эукариотический промотор). В некоторых вариантах осуществления промотор представляет собой промотор, подходящий для применения в рамках системы млекопитающего (например человека) (например промотор млекопитающего). В некоторых вариантах осуществления промотор индуцируется стимулом (то есть представляет собой индуцируемый промотор). В некоторых вариантах осуществления стимул представляет собой малую молекулу, белок, пептид, аминокислоту, метаболит, неорганическую молекулу, органометаллическую молекулу, органическую молекулу, лекарство или лекарство-кандидат, сахар, липид, металл, нуклеиновую кислоту, молекулу, полученную при активации эндогенного или экзогенного каскада передачи сигнала, свет, температуру, звук, давление, механическое давление, давление вследствие сдвига, или вирус или другой микроорганизм, изменение pH, или изменение в плане состояния

окисления/восстановления. В некоторых вариантах осуществления стимул представляет собой свет. В некоторых вариантах осуществления стимул представляет собой вирус. В некоторых вариантах осуществления стимул представляет собой малую молекулу. В некоторых вариантах осуществления стимул представляет собой антибиотик. В некоторых вариантах осуществления стимул представляет собой ангидротетрациклин или доксициклин. В некоторых вариантах осуществления стимул представляет собой сахар. В некоторых вариантах осуществления стимул представляет собой арабинозу, рамнозу, или IPTG. В некоторых вариантах осуществления стимул представляет собой сигнальную молекулу, полученную при активации каскада передачи сигнала (например бета-катенин, полученный при активации каскада передачи сигнала Wnt). Могут быть получены дополнительные промоторы, которые детектируют сигнальные молекулы, для индуцирования экспрессии последовательности нуклеиновой кислоты, функционально связанной с промотором, например промоторы, которые записывают эндогенный путь, включая иммунный ответ (промотор IL-2), отвечающий на cAMP элемент (CREB), сигнал NFκB, ответ на интерферон, P53 (повреждение ДНК), Sox2, сигнал TGF-ß (SMAD), Erk (например из активированного каскада Ras/Raf/Mek/Erk), PI3K/AKT (например из активированного каскада Ras/PI3K/Akt), тепловой шок, сигнал Notch, Oct4, рецептор арилуглеводородов, или фактор транскрипции AP-1. В некоторых вариантах осуществления промотор представляет собой конститутивный промотор. В некоторых вариантах осуществления промотор представляет собой промотор, указанный в Таблице 3. Дополнительные подходящие промоторы для применения в рамках как прокариотических, так и эукариотических систем будут очевидны средним специалистам в данной области на основе настоящего описания и знаний в данной области, при этом они входят в объем настоящего изобретения.

[0955] Праймерные редакторы также могут применяться для отслеживания происхождения клеток. Повторные модификации последовательности могут применяться для получения уникальных клеточных меток для отслеживания отдельных клеток. Массивы меток, их порядок и размер могут использоваться для определения происхождения клеток. Например, вставка последовательностей гомологии (например последовательностей 3' участка разрыва Cas9) и, в особенности, последовательностей гомологии с ассоциированными метками, является особенно полезной стратегией отслеживания происхождения в рамках праймерного

редактирования. Данные системы могут быть сконструированы таким образом, что последующие раунды редактирования приводят к вставке метки из кассеты PEgRNA, которая не может быть модифицирована в рамках других событий в виде редактирования с использованием PRнРНК в той же клетке. Система мечения может использовать несколько меток, которые могут быть ассоциированы с определенным стимулом. Данная система может сохранять большинство протоспейсеров-мишеней, но изменять основную последовательность, PAM, а также расположенные ниже смежные нуклеотиды. Это обеспечивает возможность соединения нескольких сигналов с одним локусом редактирования без значительного изменения используемых PEgRNA. Данная стратегия может обеспечивать возможность множественных вставок меток в ответ на большое количество клеточных стимулов (внутренних или внешних) в отдельном локусе. Она может обеспечивать запись интенсивности, продолжительности и порядка такого количества сигналов, которое соответствует количеству существующих уникальных меток (которые могут быть сконструированы с использованием нескольких N-нуклеотидов для получения 4^N возможных меток, например метка размером 5 nt обеспечит одновременную запись 4^5 или 1024 уникальных сигналов). Данная система может применяться как *in vitro*, так и *in vivo*.

M. Применение праймерного редактирования для модулирования активности биомолекулы

[0956] Применение описанных здесь праймерных редакторов также может использоваться для регулирования субклеточной локализации и связанных с модификацией состояний биомолекул, таких как ДНК, РНК, и белки. Определенные биологические функции, такие как транскрипционный контроль, клеточный метаболизм, и каскады передачи сигналов осторожно управляются в определенных месторасположениях в рамках клетки. Способность направлять белки к данным и другим уникальным клеточным компартментам может обеспечивать возможность изменять ряд биологических процессов.

[0957] Соответственно, праймерное редактирование может применяться для внесения генетически кодируемых средств, которые обеспечат возможность изменения связанных с модификацией состояний и субклеточного направления биомолекул с использованием генетически кодируемого сигнала (например белков, липидов, сахаров, и нуклеиновых кислот). В различных вариантах осуществления биомолекулы-мишени для лекарства на основе праймерных редакторов представляют

собой ДНК. Например, ДНК может быть модифицирована путем внесения ряда последовательностей ДНК, которые изменяют доступность локуса-мишени, что может приводить либо к повышенной, или пониженной транскрипции требуемых последовательностей. В других вариантах осуществления биомолекулы-мишени для лекарства на основе праймерных редакторов представляют собой РНК. Например, активность РНК может быть модифицирована путем изменения ее клеточной локализации, партнеров по взаимодействию, структурной динамики, или термодинамики фолдинга. В других вариантах осуществления биомолекулы-мишени для лекарства на основе праймерных редакторов представляют собой белок. Белки могут быть модифицированы для влияния на посттрансляционные модификации, при этом могут быть внесены мотивы белка для изменения субклеточной локализации белка или белки могут быть модифицированы для создания или устранения их способности существовать в рамках событий в виде образования комплексов белок-белок.

[0958] Данный вариант использования праймерного редактирования дополнительно описан в Примере 14.

Модификации ДНК

[0959] Одна биомолекула-мишень для PE-опосредованной модификации представляет собой ДНК. Модификации ДНК могут быть реализованы для внесения ряда последовательностей ДНК, которые изменяют доступность локуса-мишени. Доступность хроматина обеспечивает контроль выхода в виде транскрипции генов. Внесение отметок для рекрутирования компактирующих хроматин ферментов должно снижать транскрипционный выход соседних генов, в то время как внесение последовательностей, ассоциированных с открытием хроматина, должно делать участки более доступными и, в свою очередь, повышать транскрипцию. Внесение более сложных мотивов последовательности, которые отражают нативные регуляторные последовательности, должно обеспечивать более точный и биологически чувствительный контроль, чем доступные в настоящее время продукты слияния dCas9 с различным эпигенетическими ферментами для считывания, записи, или стирания — инструменты, которые обычно обеспечивают внесение большого количества одного типа отметок, которые могут не иметь конкретного биологического предка. Внесение последовательностей, которые обеспечат сближение двух локусов или приведут локусы в контакт с ядерной мембраной, должно также влиять на транскрипционных

выход данных локусов, как было продемонстрировано в рамках развивающейся области трехмерной геномной архитектуры.

Модификации РНК

[0960] Также могут быть реализованы модификации РНК для изменения их активности за счет изменения их клеточной локализации, партнеров по взаимодействию, структурной динамики, или термодинамики фолдинга. Внесение мотивов, которые будут вызывать приостановку трансляции или сдвиг рамки считывания, может изменять содержание иРНК за счет различных механизмов процессирования иРНК. Модифицирование консенсусных сплайсинговых последовательностей также будет изменять содержание и распространение различных видов РНК. Изменение соотношения различных сплайсинговых изоформ будет предсказуемо приводить к изменению соотношения продуктов трансляции белка, при этом это может применяться для изменения многих биологических путей. Например, сдвиг баланса белков репарации митохондриальной и ядерной ДНК будет изменять устойчивость различных видов рака в отношении химиотерапевтических реагентов. Более того, РНК могут быть модифицированы с использованием последовательностей, которые обеспечивают связывание с новыми белками-мишенями. Был разработан ряд РНК аптамеров, которые связываются с клеточными белками с высокой аффинностью. Внесение одного из данных аптамеров может применяться либо для отделения различных видов РНК за счет связывания с белком-мишенью, который будет предотвращать их трансляцию, биологическую активность, или для направления различных видов РНК в определенные субклеточные компартменты. Расщепление биомолекул является другим классом локализационной модификации.

[0961] Например, метилирование РНК применяется для регулирования РНК в клетке. Консенсусные мотивы для метилирования могут быть введены в РНК-мишени, кодирующие последовательности, с помощью РЕ. РНК также могут быть модифицированы для включения последовательностей, которые направляют механизм несмыслового расщепления или другие пути метаболизма нуклеиновой кислоты для расщепления вида-мишени РНК и изменяют пул РНК в клетке. Кроме того, определенные виды РНК могут быть модифицированы для изменения их состояния агрегации. Последовательности могут быть внесены в отдельный представляющий интерес вид РНК или несколько видов РНК для получения клубков РНК, которые будут делать их неэффективными субстратами для трансляции или передачи сигнала.

Модификации белка

[0962] Модификация белков с использованием посттрансляционной модификации (PTM) также является важным классом биомолекулярных манипуляций, которые могут быть реализованы с помощью РЕ. Как и в случае с отдельными видами РНК, изменение содержания белков в клетке является важной способностью РЕ. Редактирование может быть осуществлено для внесения стоп-кодонов в открытую рамку считывания — это устранит продуцирование полноразмерного продукта подвергнутой редактированию последовательностью ДНК. В качестве альтернативы могут быть внесены пептидные мотивы, которые вызывают изменение степени расщепления белка для белка-мишени. Внесение меток расщепления в тело гена может использоваться для изменения содержания белка в клетке. Более того, внесение дегронов, которые индуцируются малыми молекулами, может обеспечивать временно́й контроль над расщеплением белка. Это может являться важным как для исследований, так и для лечения, поскольку исследователи могут быстро определить являлось ли опосредованное малой молекулой расщепление терапевтического белка определенной мишени жизнеспособной терапевтической стратегией. Мотивы белка также могут быть внесены для изменения субклеточной локализации белка. Аминокислотные мотивы могут быть внесены для предпочтительного направления белков к ряду субклеточных компартментов, включая ядро, митохондрии, клеточную мембрану, пероксисому, лизосому, протеасому, экзосому, и другие.

[0963] Внесение или устранение мотивов, модифицированных с помощью механизма PTM, может изменять посттрансляционные модификации белка. Было идентифицировано фосфорилирование, убиквитинилирование, гликозилирование, липидирование (например фарнезилирование, миристоилирование, палмитоилирование, пренилирование, GPI якори), гидроксилирование, метилирование, ацетилирование, кротонилирование, SUMOилирование, образование дисульфидных связей, события в виде расщепления связей боковой цепи, события в виде расщепления полипептидного каркаса (протеолиз), и ряд других PTM белка. Данные PTM изменяют функцию белка, зачастую за счет изменения субклеточной локализации. Действительно, киназы зачастую активируют последующие каскады передачи сигнала за счет событий в виде фосфорилирования. Удаление фосфозита-мишени будет предотвращать трансдукцию сигнала. Способность специфическим в отношении участка образом устранять или вносить любой PTM мотив при сохранении экспрессии

полноразмерного белка будет важным достижением как для исследований, так и для терапевтического использования. Объем вносимых последовательностей и целевое окно PE делают его хорошо подходящим для широкого диапазона PTM модификаций.

[0964] Удаление участков липидирования должно предотвращать направление белков к клеточным мембранам. Заметным ограничением текущих видов лечения, нацеленных на процессы посттрансляционного модифицирования, является их специфичность. Ингибиторы фарнезилтрансферазы интенсивно тестировались на предмет их способности устранять локализацию KRAS на клеточных мембранах. К сожалению, глобальное ингибирование фарнезилирования связано с множеством нецелевых эффектов, которые препятствуют широкому использованию данных малых молекул. Аналогичным образом, специфичное ингибирование протеинкиназ с использованием малых молекул может быть очень затруднено вследствие большого размера генома человека и сходства между различными киназами. PE предлагает потенциальное решение данной проблемы, связанной со специфичностью, поскольку оно обеспечивает ингибирование модификации белка-мишени путем устранения участка модифицирования вместо глобального ингибирования фермента. Например, удаление лапидированного пептидного мотива в KRAS будет представлять собой целевой подход, который может использоваться вместо ингибирования фарнезилтрансферазы. Данный подход является функционально обратным ингибированию активности белка-мишени за счет внесения нацеленного на липид мотива в белок, который не предназначен для связывания с мембраной.

[0965] PE также может применяться для стимулирования событий в виде образования комплексов белок-белок. Белки зачастую функционируют в рамках комплексов для реализации их биологической активности. PE может использоваться для создания или устранения способности белков существовать в рамках таких комплексов. Для устранения событий в виде образования комплексов могут быть внесены аминокислотные замещения или вставки в рамках интерфейса белок:белок для исключения образования комплексов. SSX18 представляет собой белковый компонент комплекса BAF, который является важным комплексом, обеспечивающим ремоделирование гистонов. Мутации в SSX18 вызывают синовиальные саркомы. PE может использоваться для внесения боковых цепей, которые предотвращают связывание SSX18 с его белками-партнерами в рамках комплекса, для предотвращения его онкогенной активности. PE также может применяться для удаления патогенных

мутаций для восстановления активности данного белка, соответствующей дикому типу. В качестве альтернативы, PE может применяться для сохранения белков в рамках их нативного комплекса или для их привлечения для участия во взаимодействиях, которые не относятся к их нативной активности, для ингибирования их активности. Образующиеся комплексы, которые сохраняют одно состояние взаимодействия вместо другого, могут обеспечивать важную терапевтическую модальность. Изменение интерфейсов белок:белок для снижения Kd взаимодействия будет обеспечивать более длительное связывание данных белков. Белковые комплексы могут иметь несколько комплексов передачи сигнала, например как n-myc, запускающий каскады передачи сигнала нейробластомы при заболевании, но в иных случаях принимающий участие в соответствующем здоровому состоянию транскрипционном контроле в других клетках. PE может быть использовано для внесения мутаций, которые запускают ассоциацию n-myc со соответствующими здоровому состоянию партнерами по взаимодействию и снижают его аффинность в отношении онкогенных партнеров по взаимодействию.

Ссылки, приведенные в Разделе I

Каждая из следующих ссылок приведена здесь в качестве ссылки.

1. Selective Target Protein Degradation via Phthalimide Conjugation. Winter et al. Science. Author manuscript; available in PMC 2016 Jul 8.

2. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. Kadoch and Crabtree. Cell. 2013 Mar 28;153(1):71-85. doi: 10.1016/j.cell.2013.02.036.

3. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. Atkins et al. Nucleic Acids Research, Volume 44, Issue 15, 6 September 2016, Pages 7007–7078.

4. Transcriptional Regulation and its Misregulation in Disease. Lee and Young. Cell. Author manuscript; available in PMC 2014 Mar 14.

5. Protein localization in disease and therapy. Mien-Chie Hung, Wolfgang Link Journal of Cell Science 2011 124: 3381-3392.

6. Loss of post-translational modification sites in disease. Li et al. Pac Symp Biocomput. 2010:337-47. PTMD: A Database of Human Disease-associated Post-translational Modifications. Xu et al. Genomics Proteomics Bioinformatics. 2018 Aug;16(4):244-251. Epub 2018 Sep 21.

7. Post-transcriptional gene regulation by mRNA modifications. Zhao et al. Nature Reviews Molecular Cell Biology volume18, pages31–42 (2017).

N. Аспекты усовершенствованной конструкции PEgRNA для праймерного редактирования

[0966] В других вариантах осуществления система праймерного редактирования может включать применение конструкций PEgRNA и стратегий, которые могут повышать эффективность праймерного редактирования. Целью данных стратегий является преодоление некоторых существующих трудностей вследствие многоэтапного процесса, требуемого для праймерного редактирования. Например, нежелательные структуры РНК, которые могут образовываться в PEgRNA, могут приводить к ингибированию изменений ДНК при копировании из PEgRNA в геномный локус. Данные ограничения могут быть преодолены за счет изменения конструкции и модифицирования компонента в виде PEgRNA. Данные изменения конструкции могут повышать эффективность праймерного редактора и могут обеспечивать внесение более длинных вставляемых последовательностей в геном.

[0967] Соответственно, в различных вариантах осуществления конструкции PEgRNA могут приводить к более длинным PEgRNA за счет обеспечения эффективной экспрессии функциональных PEgRNA из не относящихся к полимеразе III (pol III) промоторов, устраняя обременительные требования в плане последовательности. В других вариантах осуществления ядро, представляющее собой связывающий Cas9 каркас PEgRNA, может быть усовершенствовано для повышения эффективности системы. В других вариантах осуществления могут быть реализованы модификации PEgRNA для повышения процессирования обратной транскриптазы (RT), что будет обеспечивать вставку более длинных последовательностей в геномных локусах-мишенях. В других вариантах осуществления могут быть добавлены мотивы РНК к 5′ и/или 3′-концам PEgRNA для повышения стабильности, увеличения процессирования RT, предотвращения неправильного фолдинга и/или рекрутирования дополнительных факторов, являющихся важными для геномного редактирования. В другом варианте осуществления предусмотрена платформа для эволюции PEgRNA для определенной последовательности-мишени, которая может обеспечивать усовершенствование каркаса PEgRNA и повышать эффективность праймерного редактора. Данные конструкции могут применяться для усовершенствования любой PEgRNA, распознаваемой любой Cas9 или ее вариантом.

[0968] Данный вариант использования праймерного редактирования дополнительно описан в Примере 15.

[0969] PEgRNA могут включать дополнительные усовершенствования конструкции, связанные с модификацией свойств и/или характеристик PEgRNA, тем самым повышая эффективность праймерного редактирования. В различных вариантах осуществления данные усовершенствования могут относится к одной или нескольким из ряда различных категорий, включая, но без ограничения: (1) конструкции для обеспечения эффективной экспрессии функциональных PEgRNA из не относящихся к полимеразе III (pol III) промоторов, которые будут обеспечивать экспрессию более длинных PEgRNA без обременительных требований в плане последовательности; (2) усовершенствования ядра, Cas9-связывающего каркаса PEgRNA, которые могут повысить эффективность; (3) модификации PEgRNA для повышения процессирования RT, обеспечивающие вставку более длинных последовательностей в геномных локусах-мишенях; и (4) добавление мотивов РНК на 5' или 3'-концы PEgRNA, которые повышают стабильность PEgRNA, стимулируют процессирование RT, предотвращают неправильный фолдинг PEgRNA, или рекрутируют дополнительные факторы, являющиеся важными для геномного редактирования.

[0970] В одном варианте осуществления PEgRNA может быть сконструирована с pol III промоторами для повышения экспрессии более длинной PEgRNA с более крупными плечами удлинения. sgРНК обычно экспрессируются из промотора мяРНК U6. Данный промотор рекрутирует pol III для экспрессии ассоциированной РНК и может применяться для экспрессии коротких РНК, которые остаются в ядре. Однако pol III не является высокопроцессивной и не способна экспрессировать РНК с длиной более чем несколько сотен нуклеотидов на уровнях, требуемых для эффективного геномного редактирования. Кроме того, pol III может вызывать остановку или терминирование на длинных участках в виде U, что потенциально ограничивает разнообразие цепей, которые могут быть вставлены с помощью PEgRNA. Были проанализированы другие промоторы, которые рекрутируют полимеразу II (например pCMV) или полимеразу I (например промотор мяРНК U1) на их способность экспрессировать более длинные sgРНК. Однако данные промоторы обычно частично транскрибируются, что будет приводить к дополнительной последовательности 5' спейсера в экспрессируемой PEgRNA, что, как было показано, приводит к заметному снижению активности Cas9:sgРНК в зависимости от участка. Кроме того, в то время как транскрибированные

из pol III PEgRNA могут просто приводить к терминации на участке из 6-7 U, PEgRNA, транскрибированные из pol II или pol I, будут требовать другого сигнала терминации. Зачастую такие сигналы также приводят к полиаденилированию, которое приведет к нежелательному транспорту PEgRNA из ядра. Аналогичным образом, РНК, экспрессируемые из pol II промоторов, например pCMV, обычно являются 5'-кэпированными, что также приводит к их экспорту из ядра.

[0971] Ранее Rinn и коллеги провели скрининг ряда платформ экспрессии на предмет продуцирования sgРНК, меченных длинными некодирующими РНК (днкРНК)[183]. Данные платформы включают РНК, экспрессированные из pCMV, которые приводят к терминации в элементе ENE из нкРНК MALAT1 человека[184], элементе PAN ENE из KSHV[185], или 3' боксе из sgРНК U1[186]. Следует отметить, что нкРНК MALAT1 и PAN ENE образуют тройные спирали, защищающие polyA-хвост [184, 187]. Данные конструкты также могут повышать стабильность РНК. Предполагается, что данные системы экспрессии также обеспечат возможность экспрессии более длинных PEgRNA.

[0972] Кроме того, был разработан ряд способов для расщепления части pol II промотора, которая будет транскрибирована в виде части PEgRNA, обеспечивая добавление либо саморасщепляемого рибозима, например рибозимов типа "головка молотка" (hammerhead)[188], "пистолет" (pistol)189, "топор" (hatchet)189, "шпилька" (hairpin)[190], VS[191], "твистер"192, или "твистер-систер" (twister sister)192, или других саморасщепляемых элементов для процессирования транскрибированной направляющей, или шпильки, которая распознается Csy4[193] и также ведет к процессированию направляющей. Кроме того, существует гипотеза, что встраивание нескольких мотивов ENE может приводить к повышенной экспрессии PEgRNA и стабильности, как ранее было продемонстрировано для РНК KSHV PAN и элемента[185]. Также ожидается, что циркуляризация PEgRNA в форме циркулярной интронной РНК (циРНК) также может приводить к повышенной экспрессии РНК и стабильности, а также ядерной локализации[194].

[0973] В различных вариантах осуществления PEgRNA может включать различные указанные выше элементы, как показано на примере следующей последовательности.

[0974] Неограничивающий пример 1 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и MALAT1 ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC

ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTT
CTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGAT
GCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCT
TTGCTTTGACT (SEQ ID NO: 501)

[0975] Неограничивающий пример 2 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и PAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA
TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 502)

[0976] Неограничивающий пример 3 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3xPAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA

CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA
TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAAACACACTGTT
TTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGG
TTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCT
AGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGG
TCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAATCTCTCTGTTTTGGCTGGGTT
TTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTATCCCAG
TGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAAT
TAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGGTCAATGTTTAAT
CCATAAAAAAAAAAAAAAAAAA (SEQ ID NO: 503)

[0977] Неограничивающий пример 4 – платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3′ бокса

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTGGTTT
GTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 504)

[0978] Неограничивающий пример 5 - платформа экспрессии PEgRNA, состоящая из pU1, Csy4 шпильки, PEgRNA, и 3′ бокса

CTAAGGACCAGCTTCTTTGGGAGAGAACAGACGCAGGGGCGGGAGGGAAAAAG
GGAGAGGCAGACGTCACTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCGGTTGA
GTGGCAGAAAGGCAGACGGGGACTGGGCAAGGCACTGTCGGTGACATCACGGAC
AGGGCGACTTCTATGTAGATGAGGCAGCGCAGAGGCTGCTGCTTCGCCACTTGCT
GCTTCACCACGAAGGAGTTCCCGTGCCCTGGGAGCGGGTTCAGGACCGCTGATCG
GAAGTGAGAATCCCAGCTGTGTGTCAGGGCTGGAAAGGGCTCGGGAGTGCGCGG
GGCAAGTGACCGTGTGTGTAAAGAGTGAGGCGTATGAGGCTGTGTCGGGGCAGA
GGCCCAAGATCTCAGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG
TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAA
AGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTCAGCAAG
TTCAGAGAAATCTGAACTTGCTGGATTTTTGGAGCAGGGAGATGGAATAGGAGCT
TGCTCCGTCCACTCCACGCATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGC

ACCCACTTTCTGGAGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTT GTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 505)

[0979] В различных других вариантах осуществления PEgRNA может быть усовершенствована путем введения усовершенствований в последовательности каркаса или ядра. Это может быть осуществлено путем внесения известных

[0980] Ядро, представляющее собой Cas9-связывающий каркас PEgRNA, может быть вероятно усовершенствовано для повышения активности PE. Некоторые подобные подходы уже были продемонстрированы. Например, первый элемент связывания каркаса (P1) содержит элемент связывания GTTTT-AAAAC. Было показано, что такие повторы T приводят к остановке pol III и преждевременной терминации РНК транскрипта. Было показано, что разумная мутация одной из пар T-A на пару G-C в этой части P1 повышает активность sgРНК, что свидетельствует о том, что данный подход также может применяться для PEgRNA[195]. Кроме того, также было показано, что повышение длины P1 стимулирует фолдинг sgРНК и ведет к повышенной активности[195], что свидетельствует о том, что оно является другой возможностью повысить активность PEgRNA. Примеры усовершенствований ядра могут включать:

[0981] PEgRNA, содержащую удлинение P1 длиной 6 nt

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGCTCATGAAAATGAGCTAGCAAG TTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTC TGCCATCAAAGCGTGCTCAGTCTGTTTTTTT (SEQ ID NO: 228)

[0982] PEgRNA, содержащую мутацию T-A на G-C в P1

GGCCCAGACTGAGCACGTGAGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGC TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC GTGCTCAGTCTGTTTTTTT (SEQ ID NO: 229)

[0983] В различных других вариантах осуществления PEgRNA может быть усовершенствована путем внесения модификаций в участок матрицы для редактирования. С увеличением размера вставки, матрицей которого служит PEgRNA, она с большей вероятностью будет расщепляться эндонуклеазами, подвергаться спонтанному гидролизу или фолдингу во вторичные структуры, не способные подвергаться обратной транскрипции за счет RT или нарушающие фолдинг каркаса PEgRNA и последующее связывание Cas9-RT. Соответственно, является вероятным, что для реализации крупных вставок может быть необходимым модифицирование матрицы PEgRNA, например для вставок целых генов. Некоторые стратегии для этого включают встраивание модифицированных нуклеотидов в синтетическую или

полусинтетическую PEgRNA, которые делают РНК более устойчивой к расщеплению или гидролизу или менее подверженной включению ингибиторных вторичных структур[196]. Подобные модификации могут включать 8-аза-7-деазагуанозин, который будет снижать содержание вторичной структуры РНК в последовательностях с высоким содержанием G; "заблокированные" нуклеиновые кислоты (LNA), которые снижают расщепление и повышают определенные виды вторичной структуры РНК; модификации в виде 2'-O-метил, 2'-фтор, или 2'-O-метоксиэтокси, которые повышают стабильность РНК. Такие модификации также могут быть включены в другие части PEgRNA для повышения стабильности и активности. В качестве альтернативы или дополнительно, матрица PEgRNA может быть сконструирована таким образом, что она как кодирует требуемый продукт в виде белка, так и с большей вероятностью включает простые вторичные структуры, которые способны подвергаться обратному фолдингу за счет RT. Такие простые структуры будут действовать в качестве термодинамического стока, делая менее вероятным появление более сложных структур, предотвращающих обратную транскрипцию. Наконец, также является возможным разделить матрицу на две отдельные PEgRNA. При такой конструкции PE будет использоваться для инициирования транскрипции, а также рекрутирования отдельной матричной РНК к являющемуся мишенью участку с помощью РНК-связывающего белка, слитого с Cas9, или элемента распознавания РНК на самой PEgRNA, например аптамера MS2. RT может либо напрямую связываться с данной отдельной матричной РНК, или инициировать обратную транскрипцию на исходной PEgRNA перед заменой на вторую матрицу. Такой подход может обеспечивать возможность осуществления длинных вставок за счет как предотвращения неправильного фолдинга PEgRNA при добавлении длинной матрицы, так и за счет отсутствия необходимости в диссоциации Cas9 из генома для реализации длинных вставок, что вероятно может ингибировать длинные вставки на основе PE.

[0984] В других вариантах осуществления PEgRNA может быть усовершенствована путем внесения дополнительных мотивов РНК на 5′ и 3′-концах PEgRNA. Некоторые подобные мотивы, например PEN ENE из KSHV и ENE из MALAT1, были описаны выше в качестве возможных способов терминации экспрессии длинных PEgRNA из не относящихся к pol III промоторов. Эти элементы образуют тройные спирали РНК, которые окружают polyA-хвост, что приводит к их удержанию в ядре[184,187]. Однако за счет образования сложных структур на 3′-конце PEgRNA, которые закрывают

концевой нуклеотид, данные структуры также вероятно будут помогать предотвращать опосредованное экзонуклеазой расщепление PEgRNA.

[0985] Другие структурные элементы, вставленные на 3′-конце, также могут повышать стабильность РНК, хотя и без обеспечения терминации из не относящихся к pol III промоторов. Такие мотивы могут включать шпильки или квадруплексы РНК, которые будут закрывать 3′-конец[197], или саморасщепляющиеся рибозимы, например HDV, которые будут приводить к образованию 2’-3′-циклического фосфата на 3′-конце, а также потенциально делать PEgRNA менее подверженной расщеплению экзонуклеазами[198]. Индуцирование циклизации PEgRNA за счет неполного сплайсинга – для образования циРНК – также может повышать стабильность PEgRNA и приводить к удержанию PEgRNA в ядре[194].

[0986] Дополнительные мотивы РНК также могут повышать процессирование RT или повышать активность PEgRNA за счет повышения связывания RT с дуплексом ДНК-РНК. Добавление нативной последовательности, связанной RT в ее родственном ретровирусном геноме, может повышать активность RT[199]. Это может включать нативный участок связывания праймера (PBS), полипуриновый тракт (PPT), или целующиеся петли, принимающие участие в димеризациии ретровирусного генома и инициировании транскрипции[199].

[0987] Добавление мотивов димеризации, таких как целующиеся петли или пары тетрапетля GNRA/рецептор тетрапетли [200], на 5′ и 3′-концах PEgRNA также может приводить к эффективной циркуляризации PEgRNA, что повышает стабильность. Кроме того, предполагается, что добавление данных мотивов может обеспечивать физическое разделение спейсера и праймера PEgRNA, предотвращать закрытие спейсера, ухудшающее активность PE. Короткие 5′ удлинения или 3′ удлинения PEgRNA, которые образуют малую опорную шпильку в спейсерном участке, также могут благоприятным образом конкурировать против участка отжига PEgRNA, связывающего спейсер. Наконец, целующиеся петли также могут использоваться для рекрутирования других матричных РНК в участок генома и обеспечения обмена RT активностью от одной РНК к другой. Примеры усовершенствований включают, но без ограничения:

[0988] Продукт слияния PEgRNA-HDV

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC

GTGCTCAGTCTGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAA
CATGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 230)

[0989] Целующуюся цепь PEgRNA-MMLV

GGTGGGAGACGTCCCACCGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAAT
AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCG
GTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCT
TTTTTT (SEQ ID NO: 231)

[0990] Целующуюся цепь PEgRNA-VS рибозим

GAGCAGCATGGCGTCGCTGCTCACGGCCCAGACTGAGCACGTGAGTTTTAGAGCT
AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACC
GAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTCCATCAGTTGACA
CCCTGAGGTTTTTTT (SEQ ID NO: 232)

[0991] PEgRNA-GNRA тетрапетля/рецептор тетрапетли

GCAGACCTAAGTGGUGACATATGGTCTGGGCCCAGACTGAGCACGTGAGTTTTAG
AGCTAUACGTAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTUACGAAGTGG
GACCGAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGCATGCGATT
AGAAATAATCGCATGTTTTTTT (SEQ ID NO: 233)

[0992] PEgRNA матрица с заменой вторичного продукта слияния в виде РНК-HDV

TCTGCCATCAAAGCTGCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCGGCCG
GCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCG
AATGGGACTTTTTTT (SEQ ID NO: 234)

[0993] Каркас PEgRNA может быть дополнительно усовершенствован с помощью направленной эволюции аналогично тому, как был усовершенствован SpCas9 и праймерный редактор (PE). Направленная эволюция может повышать степень распознавания PEgRNA Cas9 или подвергнутыми эволюции вариантами Cas9. Кроме того, является возможным, что различные последовательности каркаса PEgRNA будут оптимальными в рамках различных геномных локусов, обеспечивая повышение активности PE в рассматриваемом участке, снижая нецелевые виды активности, или осуществляя оба из данных действий. Наконец, эволюция каркасов PEgRNA, к которым были добавлены другие мотивы РНК, будет практически наверняка повышать активность слитой PEgRNA по сравнению с не подвергавшейся эволюции слитой РНК. Например, эволюция аллостерических рибозимов, состоящих из аптамеров c-di-GMP-I и рибозимов типа "головка молотка" (hammerhead), приводила к значительному повышению активности[202], что свидетельствует о том, что эволюция также будет повышать активность PEgRNA типа "головка молотка" (hammerhead). Кроме того, в то время как Cas9 в настоящее время в целом не совместим с 5′ удлинением sgРНК, направленная эволюция вероятно обеспечит реализацию мутаций, которые устранят

данную несовместимость, обеспечивая возможность использования дополнительных мотивов РНК.

[0994] Настоящее изобретение охватывает любые подобные подходы к дальнейшему усовершенствованию эффективности описанных здесь систем праймерного редактирования.

О. Применение праймерного редактирования с расширенной областью нацеливания

[0995] Праймерное редактирование (PE) с использованием Cas9 (SpCas9) *Streptococcus pyogenes* может обеспечивать эффективное внесение всех замещений отдельных оснований, вставок, делеций, и их комбинаций в геномных локусах с подходящим образом расположенным смежным с протоспейсером мотивом NGG (PAM), который может быть эффективно связан SpCas9. Однако в другом аспекте описанные здесь способы расширяют способность нацеливания PE за счет увеличения доступных PAM и, таким образом, геномных локусов-мишеней, доступных для эффективного PE. Праймерные редакторы на основе использования РНК-направляемых связывающих ДНК белков, отличных от SpCas9, обеспечивают расширение области нацеливания в плане геномных локусов за счет предоставления доступа к различным PAM. Кроме того, использование РНК-направляемых связывающих ДНК белков с размером менее чем у SpCas9 также обеспечивает более эффективную вирусную доставку. PE с белками Cas или другими РНК-направляемыми связывающими ДНК белками помимо SpCas9 будет обеспечивать высокоэффективные терапевтические изменения, которые являлись недоступными или неэффективными при использовании PE на основе SpCas9.

[0996] Ожидается, что это будет использоваться в ситуациях, где PE на основе SpCas9 является либо неэффективным вследствие неидеального расположения изменения относительно PAM NGG, или общий размер конструкта на основе SpCas9 препятствует клеточной экспрессии и/или доставке. Определенные связанные с заболеваниями локусы, такие как ген хантингтина, который имеет редкие и неудачно расположенные PAM NGG для SpCas9 рядом с участком-мишенью, могут быть легко нацелены с использованием различных белков Cas в рамках системы PE, например SpCas9-VRQR, который распознает PAM NGA. Белки Cas меньшего размера будут использоваться для получения конструктов PE меньшего размера, которые могут быть упакованы в AAV векторы более эффективным образом, обеспечивая улучшенную доставку в

ткани-мишени. На Фиг.61 показан переход к практике праймерного редактирования с использованием CRISPR-Cas на основе *Staphylococcus aureus* в качестве РНК-направляемого связывающего ДНК белка. NT представляет собой контроль без обработки.

[0997] На Фиг.62A-62B продемонстрирована важность протоспейсера для эффективного внесения требуемого изменения в точном месторасположении с помощью праймерного редактирования. Это подчеркивает важность альтернативных PAM и протоспейсеров в качестве новых свойств данной технологии. "н.д." на Фиг.62A означает "не детектирован".

[0998] На Фиг.63 показан переход к практике PE и использованием SpCas9(H840A)-VRQR и SpCas9(H840A)-VRER в качестве РНК-направляемого связывающего ДНК белка в рамках системы праймерного редактора. napDNAbp SpCas9(H840A)-VRQR описан здесь как SEQ ID NO: 87. napDNAbp SpCas9(H840A)-VRER описан здесь как SEQ ID NO: 88. Слитый белок SpCas9(H840A)-VRER-RT MMLV описан здесь как SEQ ID NO: 516, в то время как RT MMLV включает D200N, L603W, T330P, T306K, и W313F замещения по сравнению с RT MMLV дикого типа. Слитый белок SpCas9(H840A)-VRQR-RT MMLV описан здесь как SEQ ID NO: 515, в то время как RT MMLV включает D200N, L603W, T330P, T306K, и W313F замещения по сравнению с RT MMLV дикого типа. Мишенями являются семь различных локусов в геноме человека: 4 для системы праймерного редактора SpCas9(H840A)-VRQR- RT MMLV и 3 для системы SpCas9(H840A)-VRER-RT MMLV. Аминокислотные последовательности тестируемых конструктов являются следующими:

ПРАЙМЕРНЫЙ РЕДАКТОР SACAS9- RT M-MLV	MKRTADGSEFESPKKKRKVGKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK LLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLA KRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLE RLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFID TYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPE ELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYH DIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKL VPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYG LPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGK ENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHII PRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETF KKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVD TRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFK

	KERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMF EEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKK PNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLIN KSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGN YLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE EAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIE VNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNL YEVKSKKHPQIIKKGSGGSSGGSSGSETPGTSESATPESSGGSSG GSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGL AVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGIL VPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR DPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDL ILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQ ICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRR PVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVV ALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAEL IALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSE GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFEPKKKR KV (SEQ ID NO: 514)
ПРАЙМЕРНЫЙ РЕДАКТОР SPCAS9(H840A)-VRQR-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA

	VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 515)
ПРАЙМЕРНЫЙ РЕДАКТОР SPCAS9(H840A)-VRER-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA

	VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 516)

[0999] Как показано на Фиг.63, SpCas9(H840A)-VRQR-RT MMLV являлся функциональным в участках PAM, которые включали "AGAG" и "GGAG", и имел некоторую связанную с редактированием активность в последовательностях PAM "GGAT" и "AGAT". SpCas9(H840A)-VRER-RT MMLV являлся функциональным в участках PAM, которые включали "AGCG" и "GGCG", и имел некоторую связанную редактированием активность в "TGCG".

[1000] Эти данные демонстрируют, что праймерное редактирование может быть реализовано с помощью napDNAbp с различной специфичностью в отношении PAM, например описанных здесь вариантов Cas9.

[1001] В различных вариантах осуществления napDNAbp (например Cas9) с измененной специфичностью в отношении PAM
включают комбинацию мутаций, которые демонстрируют активность на последовательности-мишени, включающей последовательность PAM 5′-NAA-3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 1. В некоторых вариантах

осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 1. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 1.

[1002] Таблица 1: Клоны PAM NAA

Мутации на основе SpCas9 дикого типа (например SEQ ID NO: 18)
D177N, K218R, D614N, D1135N, P1137S, E1219V, A1320V, A1323D, R1333K
D177N, K218R, D614N, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, G715C, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A367T, K710E, R1114G, D1135N, P1137S, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D861N, D1135N, K1188R, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, V743I, R753G, E762G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, S1274R, A1320V, R1333K
A10T, I322V, S409I, E427G, A589S, R753G, D1135N, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, E757K, G865G, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, R753G, E757K, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, K599R, M631A, R654L, K673E, V743I, R753G, N758H, E762G, D1135N, D1180G, E1219V, Q1221H, Q1256R, H1264Y, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N869S, N1054D, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, L727I, V743I, R753G, E762G, R859S, N946D, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, N1317T, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, K1151E, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1080S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, E630K, R654L, K673E, V743I, R753G, E762G, Q768H,

N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, Q768H, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, G1223S, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1801S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, M673I, N803S, N869S, G1077D, R1114G, D1135N, V1139A, D1180G, E1219V, Q1221H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, R1114G, D1135N, E1219V, Q1221H, A1320V, R1333K

[1003] В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 80% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 1. В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 92%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на 99%, или, по меньшей мере, на 99,5% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 1.

[1004] В некоторых вариантах осуществления белок Cas9 демонстрирует повышенную активность на последовательности-мишени, которая не включает стандартный PAM (5′-NGG-3′) на ее 3′-конце в отличие от Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, имеющей 3′-конец, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 5-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 10-кратно, по меньшей мере, 50-кратно, по меньшей мере, 100-кратно, по меньшей мере, 500-кратно, по меньшей мере, 1000-кратно, по меньшей мере, 5000-кратно, по меньшей мере, 10000-кратно, по меньшей

мере, 50000-кратно, по меньшей мере, 100000-кратно, по меньшей мере, 500000-кратно, или, по меньшей мере, 1000000-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 2 на той же последовательности-мишени. В некоторых вариантах осуществления 3′-конец последовательности-мишени является непосредственно смежным с последовательностью AAA, GAA, CAA, или TAA. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций, которые демонстрируют активность на последовательности-мишени, включающей последовательность PAM 5′-NAC -3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 2. В некоторых вариантах осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 2. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 2.

[1005] Таблица 2: Клоны PAM NAC

МУТАЦИИ НА ОСНОВЕ SPCAS9 ДИКОГО ТИПА (НАПРИМЕР SEQ ID NO: 18)
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
D1135N, E1219V, D1332N, R1335Q, T1337N
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
T472I, R753G, Q771H, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, K1156E, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, V647I, R753G, N803S, K959N, G1030R, I1055E, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, E630G, T638P, V647A, G687R, N767D, N803S, K959N, R1114G, D1135N, E1219V, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, I1057T, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, M631I, T638P, R753G, N803S, K959N, Y1036H, R1114G, D1135N, E1219V, D1251G, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, V875I, K959N, Y1016C, R1114G, D1135N, E1219V,

D1251G, D1332G, R1335Q, T1337N, I1348V
K608R, E627K, T638P, V647I, R654L, R753G, N803S, T804A, K848N, V922A, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R753G, N803S, V922A, K959N, K1014N, V1015A, R1114G, D1135N, K1156N, E1219V, N1252D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, A711T, R753G, K775R, K789E, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
I670S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, K797N, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, G752R, R753G, K797N, N803S, K948E, K959N, V1015A, Y1016S, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
I570T, A589V, K608R, E627K, T638P, V647I, R654L, Q716R, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1207G, E1219V, N1234D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, R654L, Q740R, R753G, N803S, K959N, N990S, T995S, V1015A, Y1036D, R1114G, D1135N, E1207G, E1219V, N1234D, N1266H, D1332N, R1335Q, T1337N
I562F, V565D, I570T, K608R, L625S, E627K, T638P, V647I, R654I, G752R, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, N1177S, N1234D, D1332N, R1335Q, T1337N
I562F, I570T, K608R, E627K, T638P, V647I, R753G, E790A, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, D1180E, A1184T, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1015A, R1114G, D1127A, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, L625S, E627K, T638P, V647I, R654I, T703P, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1016A, R1114G, D1135N, E1219V, K1246E, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R654L, K673E, R753G, E790A, N803S, K948E, K959N, R1114G, D1127G, D1135N, D1180E, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, L625S, E627K, T638P, V647I, R654I, I670T, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
E627K, M631V, T638P, V647I, K710E, R753G, N803S, N808D, K948E, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N, S1338T, H1349R

[1006] В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 80% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 2. В некоторых вариантах осуществления белок Cas9 включает аминокислотную последовательность, которая, по меньшей мере, на 85%, по меньшей мере, на 90%, по меньшей мере, на 92%, по меньшей мере, на 95%, по меньшей мере, на 96%, по меньшей мере, на 97%, по меньшей мере, на 98%, по меньшей мере, на

99%, или, по меньшей мере, на 99,5% идентична аминокислотной последовательности белка Cas9, представленной в рамках любого одного из вариантов в Таблице 2.

[1007] В некоторых вариантах осуществления белок Cas9 демонстрирует повышенную активность на последовательности-мишени, которая не включает стандартный PAM (5′-NGG-3′) на ее 3′-конце в отличие от Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, имеющей 3′-конец, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 5-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления белок Cas9 демонстрирует активность на последовательности-мишени, которая не является непосредственно смежной со стандартной последовательностью PAM (5′-NGG-3′), которая является, по меньшей мере, 10-кратно, по меньшей мере, 50-кратно, по меньшей мере, 100-кратно, по меньшей мере, 500-кратно, по меньшей мере, 1000-кратно, по меньшей мере, 5000-кратно, по меньшей мере, 10000-кратно, по меньшей мере, 50000-кратно, по меньшей мере, 100000-кратно, по меньшей мере, 500000-кратно, или, по меньшей мере, 1000000-кратно повышенной по сравнению с активностью Cas9 *Streptococcus pyogenes* согласно SEQ ID NO: 18 на той же последовательности-мишени. В некоторых вариантах осуществления 3′-конец последовательности-мишени является непосредственно смежным с последовательностью AAC, GAC, CAC, или TAC.

[1008] В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций, которые демонстрируют активность на последовательности-мишени, включающей последовательность PAM 5′- NAT -3′ на ее 3′-конце. В некоторых вариантах осуществления комбинация мутаций присутствует в любом одном из клонов, перечисленных в Таблице 3. В некоторых вариантах осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 3. В некоторых вариантах осуществления белок Cas9 включает комбинацию мутаций любого одного из клонов Cas9, перечисленных в Таблице 3.

[1009] Таблица 3: Клоны PAM NAT

МУТАЦИИ НА ОСНОВЕ SPCAS9 ДИКОГО ТИПА (НАПРИМЕР SEQ ID NO: 18)

K961E, H985Y, D1135N, K1191N, E1219V, Q1221H, A1320A, P1321S, R1335L
D1135N, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
V743I, R753G, E790A, D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, V748I, V743I, R753G, D853E, V922A, R1114G D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1286K, P1321S, D1322G, R1335L
M631L, R654L, R753G, K797E, D853E, V922A, D1012A, R1114G D1135N, G1218S, E1219V, Q1221H, P1249S, N1317K, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, D596Y, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, Q1256R, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, K710E, V750A, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, K649R, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, K1156E, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, I1057G, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1308D, P1321S, D1322G, R1335L
M631L, R654L, R753G, D853E, V922A, R1114G, Y1131C, D1135N, E1150V, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, D853E, I1057V, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, I1057V, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L

[1010] Любой из указанных выше вариантов Cas9, демонстрирующих различную специфичность в отношении PAM в отличие от стандартного SpCas9, может применяться в рамках описанных здесь праймерных редакторов.

P. Применение праймерного редактирования для вставки участков-мишеней рекомбиназы

[1011] В другом аспекте праймерное редактирование может применяться для вставки участков-мишеней рекомбиназы (или "последовательностей распознавания рекомбиназой") в требуемый участок генома. Вставка участков-мишеней рекомбиназы обеспечивает программируемое месторасположение для эффективных специфических в отношении участка генетических изменений в геноме. Такие генетические изменения могут включать, например, встраивание в геном плазмиды, геномную делецию или вставку, хромосомные транслокации и обмен кассетами, среди других

генетических изменений. Данные примеры типов генетических изменений показаны на Фиг.64(b)-(f). Внесенные последовательности распознавания рекомбиназой затем могут использоваться для осуществления специфичной в отношении участка рекомбинации в данном участке для получения ряда результатов рекомбинации, например эксцизии, встраивания, инверсии, или замены фрагментов ДНК. Например, на Фиг.65 показано внесение участка-мишени рекомбиназы, который затем может применяться для встраивания донорской матрицы ДНК, включающей маркер экспрессии GFP. Клетки, содержащие встроенную систему экспрессии GFP в участке рекомбиназы, будут флуоресцировать.

[1012] Механизм внесения участка-мишени рекомбиназы в геном аналогичен внесению других последовательностей, например пептидных/белковых и РНК-меток. Схема, являющаяся примером внесения последовательности-мишени рекомбиназы, показана на Фиг.64(a). Процесс начинается с выбора требуемого локуса-мишени, в который будет внесена последовательность-мишень рекомбиназы. Далее предоставляет продукт слияния в виде праймерного редактора ("RT-Cas9:нРНК"). При этом "нРНК" обозначает PEgRNA, которая может быть сконструирована с применением описанных здесь способов. В различных вариантах осуществления PEgRNA будет включать архитектуру, соответствующую Фиг.3D (5′-[~20 nt спейсер]-[ядро нРНК]-[плечо удлинения]-3′, где плечо удлинения включает в направлении от 3′ к 5′ участок связывания праймера ("A"), матрицу для редактирования ("B"), и плечо гомологии ("C"). Матрица для редактирования ("B") будет включать последовательность, включающую участок-мишень рекомбиназы, то есть одноцепочечную РНК PEgRNA, которая кодирует комплементарную одноцепочечную ДНК, которая представляет собой смысловую или антисмысловую цепь участка-мишени рекомбиназы и которая встраивается в локус-мишень геномной ДНК за счет процесса праймерного редактирования.

[1013] В различных аспектах настоящее изобретение предусматривает применение PE для внесения последовательностей распознавания рекомбиназой в имеющие высокую ценность локусы генома человека или другие геномы, которые после подвергания действию специфичной в отношении участка рекомбиназы (рекомбиназ) будут обеспечивать точные и эффективные геномные модификации (Фиг.64). В различных вариантах осуществления, показанных на Фиг.64, PE может применяться для (b) вставки отдельной мишени SSR для использования в качестве участка для геномного

встраивания донорской матрицы ДНК. (c) показано, как тандемная вставка участков-мишеней SSR может использоваться для делеции части генома. (d) показано, как тандемная вставка участков-мишеней SSR может использоваться для инверсии части генома. (e) показано, как вставка двух участков-мишеней SSR в двух удаленных хромосомных участках может приводить к хромосомной транслокации. (f) показано, как вставка двух различных участков-мишеней SSR в геноме может использоваться для обмена кассетами из донорской матрицы ДНК. Каждый из данных типов геномных модификаций охватывается при применении PE для вставки мишеней SSR, однако данный перечень также не является ограничивающим.

[1014] PE-опосредованное внесение последовательностей распознавания рекомбиназой может быть особенно полезным при лечении генетических заболеваний, которые вызываются крупными геномными дефектами, такими как потеря генов, инверсия, или дупликация, или хромосомная транслокация[1-7] (Таблица 6). Например, синдром Вильямса-Бойрена представляет собой нарушение развития, вызванное делецией 24 в хромосоме 721. В настоящее время не существует технологии для эффективной и нацеленной вставки нескольких целых генов в живые клетки (потенциал PE в плане реализации подобной вставки полноразмерных генов в настоящее время оценивается, но еще не был подтвержден); однако опосредуемое рекомбиназой встраивание в мишени, вставленной с помощью PE, является одним из возможных подходов к перманентному излечению данного и других заболеваний. Кроме того, нацеленное внесение последовательностей распознавания рекомбиназой может эффективно обеспечивать различные варианты применения, включая получения трансгенных растений, исследовательских моделей на основе животных, биопроизводственных линий клеток, или других линий модифицированных эукариотических клеток. Например, опосредуемое рекомбиназой геномное модифицирование в трансгенных растениях в PE-специфичных мишенях может обеспечивать преодоление одного из препятствий в плане получения сельскохозяйственных растений с улучшенными свойствами[8,9].

[1015] Таблица 6. Примеры генетических заболеваний, связанных с крупными геномными модификациями, которые могут быть скорректированы с помощью внесения последовательностей распознавания рекомбиназой на основе PE.

ЗАБОЛЕВАНИЕ	ПРИЧИНА
ТРИСОМИЯ 17P	ДУПЛИКАЦИЯ ГЕНОВ

БОЛЕЗНЬ ШАРКО-МАРИ-ТУТА ТИП I	ДУПЛИКАЦИЯ ГЕНОВ
СИНДРОМ СМИТА-МАГЕНИСА	ДЕЛЕЦИЯ ГЕНОВ
СИНДРОМ ВИЛЬЯМСА-БОЙРЕНА	ДЕЛЕЦИЯ ГЕНОВ
СИНДРОМ ДЕ ЛА ШАПЕЛЬ	ХРОМОСОМНАЯ ТРАНСЛОКАЦИЯ
СИНДРОМ ДАУНА (НЕКОТОРЫЕ ФОРМЫ)	ХРОМОСОМНАЯ ТРАНСЛОКАЦИЯ
ГЕМОФИЛИЯ А	ИНВЕРСИЯ ГЕНОВ
СИНДРОМ ХАНТЕРА	ИНВЕРСИЯ ГЕНОВ

[1016] Был охарактеризован ряд членов семейства SSR, а их последовательности распознавания рекомбиназой описаны, включая природные и модифицированные тирозиновые рекомбиназы (Таблица 7), крупные сериновые интегразы (Таблица 8), сериновые резольвазы (Таблица 9), и тирозиновые интегразы (Таблица 10). Модифицированные последовательности-мишени, которые демонстрируют повышенные показатели геномного встраивания, также были описаны для нескольких SSR[22-30]. Помимо природных рекомбиназ, были разработаны программируемые рекомбиназы с различной специфичностью[31-40]. С помощью PE одна или нескольких из данных последовательностей распознавания может быть внесена в геном в определенном месте, например в локус "безопасная гавань"[41-43], в зависимости от требуемого варианта использования.

[1017] Например, внесение одной последовательности распознавания рекомбиназой в геном с помощью праймерного редактирования будет приводить к интегративной рекомбинации на основе донорской матрицы ДНК (Фиг.64b). Сериновые интегразы, которые функционируют повсеместно в рамках клеток человека, могут являться особенно подходящими для генного встраивания[44,45].

[1018] Кроме того, внесение двух последовательностей распознавания рекомбиназой может приводить к делеции промежуточной последовательности, инверсии промежуточной последовательности, хромосомной транслокации, или обмену кассетами, в зависимости от вида и ориентации мишеней (Фиг.64c-f). За счет выбора эндогенных последовательностей, которые уже близко напоминают мишени рекомбиназы, объем редактирования, необходимого для внесения полной мишени рекомбиназы, будет снижен.

[1019] Наконец, было продемонстрировано, что несколько рекомбиназ встраиваются в геном человека или эукариотические геномы в нативно расположенных

псевдосайтах[46-64]. PE может применяться для модификации данных локусов для повышения показателей интеграции в данных природных псевдосайтах или, в качестве альтернативы, для устранения псевдосайтов, которые могут служить нежелательными нецелевыми последовательностями.

[1020] В рамках данного описания описана общая методология внесения последовательностей-мишеней рекомбиназы в эукариотические геномы с помощью PE, варианты использования которого являются практически неограниченными. Реакции геномного редактирования предназначены для применения с "праймерным редактором", который представляет собой химерный продукт слияния CRISPR/белка Cas9 и домена обратной транскриптазы, который использует модифицированную направляющую РНК для праймерного редактирования (PEgRNA). За счет удлинения инструменты на основе Cas9 и пути направленной гомологичной репарации (HDR) также могут использоваться для внесения последовательностей распознавания рекомбиназой с помощью ДНК-матриц путем понижения показателей инделей с помощью ряда способов[65-67]. Эксперимент для подтверждения данной концепции в рамках культуры клеток человека показан на Фиг.65.

[1021] В приведенном выше описании указаны следующие несколько таблиц, относящихся к PE-опосредованному внесению последовательностей распознавания рекомбиназой, в которых указан перечень примеров рекомбиназ, которые могут применяться, а также соответствующие им последовательности распознавания рекомбиназой, которые могут быть внесены с помощью PE.

Таблица 7. Тирозиновые рекомбиназы и последовательности мишеней SSR.

Рекомбиназа	**Последовательность распознавания рекомбиназой**	**Название**
Cre	ATAACTTCGTATAGCATACATTATACGAAGTTAT (SEQ ID NO: 517)	*loxP*
Dre	TAACTTTAAATAATGCCAATTATTTAAAGTTA (SEQ ID NO: 518)	*rox*
VCre	TCAATTTCTGAGAACTGTCATTCTCGGAAATTGA (SEQ ID NO: 519)	*loxV*
SCre	CTCGTGTCCGATAACTGTAATTATCGGACATGAT (SEQ ID NO: 520)	*loxS*
Flp	GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC (SEQ ID NO: 521)	*FRT*
B2	GAGTTTCATTAAGGAATAACTAATTCCCTAATGAA ACTC (SEQ ID NO: 522)	*loxB*

Рекомбиназа	Последовательность распознавания рекомбиназой	Название
B3	GGTTGCTTAAGAATAAGTAATTCTTAAGCAACC (SEQ ID NO: 523)	*loxB3*
Kw	ACGAAAAATGGTAAGGAATAGACCATTCCTTACC ATTTTTGGT (SEQ ID NO: 524)	
R	TTGATGAAAGAATAACGTATTCTTTCATCAA (SEQ ID NO: 525)	*RS*
TD1-40	GTGCGTCAAATAATAACGTATTATTTGACACTT (SEQ ID NO: 526)	*TDRS*
Vika	AATAGGTCTGAGAACGCCCATTCTCAGACGTATT (SEQ ID NO: 527)	*vox*
Nigri	TGAATGTCCTATAATTACACTTATAGGACATTCA (SEQ ID NO: 528)	*nox*
Panto	GAAACTTTAAATAATAAGTCTTATTTAAAGTTTC (SEQ ID NO: 529)	*pox*
Kd	AAACGATATCAGACATTTGTCTGATAATGCTTCA TTATCAGACAAATGTCTGATATCGTTT (SEQ ID NO: 530)	*loxK*
Fre	ATATATACGTATATAGACATATATACGTATATAT (SEQ ID NO: 531)	*loxH*
CreALSHG	ATAACTCTATATAATGTATGCTATATAGAGTTAT (SEQ ID NO: 532)	*loxM7*
Tre	ACAACATCCTATTACACCCTATATGCCAACATGG (SEQ ID NO: 533)	*loxLTR*
Brec1	AACCCACTGCTTAAGCCTCAATAAAGCTTGCCTT (SEQ ID NO: 534)	*loxBTR*
Cre-R3M3	GATACAACGTATATACCTTTCTATACGTTGTTTA (SEQ ID NO: 535)	*loxK2*

Таблица 8. Крупные сериновые интегразы и последовательности мишеней SSR.

Рекомбиназа	Последовательность распознавания рекомбиназой (левая)	Последовательность распознавания рекомбиназой (правая)
Bxb1	GGTTTGTCTGGTCAACCACCGC GGTCTCAGTGGTGTACGGTACA AACC (SEQ ID NO: 536)	GGCTTGTCGACGACGGCGGTCTCC GTCGTCAGGATCAT (SEQ ID NO: 537)
phiC31	GTGCCCCAACTGGGGTAACCTT TGAGTTCTCTCAGTTGGGGG (SEQ ID NO: 538)	TGCGGGTGCCAGGGCGTGCCCTTG GGCTCCCCGGGCGCGTACTCC (SEQ ID NO: 539)
R4	TGTTCCCCAAAGCGATACCACT TGAAGCAGTGGTACTGCTTGTG	GCATGTTCCCCAAAGCGATACCAC TTGAAGCAGTGGTACTGCTTGTGG

Рекомбиназа	Последовательность распознавания рекомбиназой (левая)	Последовательность распознавания рекомбиназой (правая)
	GGTACA (SEQ ID NO: 540)	GTACACTCTGCGGGTG (SEQ ID NO: 541)
phiBT1	GGTGCTGGGTTGTTGTCTCTGGACAGTGATCCATGGGAAACTACTCAGCACC (SEQ ID NO: 542)	CAGGTTTTTGACGAAAGTGATCCAGATGATCCAG (SEQ ID NO: 543)
MJ1 (phiFC1)	ATTTTAGGTATATGATTTTGTTTATTAGTGTAAATAACACTATGTACCTAAAAT (SEQ ID NO: 544)	CAAAGGATCACTGAATCAAAAGTATTGCTCATCCACGCGAAA (SEQ ID NO: 545)
MR11	TTTGTGCGGAACTACGAACAGTTCATTAATACGAAGTGTACAAACTTCCATACAA (SEQ ID NO: 546)	CGAAAATGTATGGAGGCACTTGTATCAATATAGGATGTATACCTTCGAAGACACTT (SEQ ID NO: 547)
TP901-1	GAGTTTTTATTTCGTTTATTTCAATTAAGGTAACTAAAAAACTCCTTTTAAGG (SEQ ID NO: 548)	ATGCCAACACAATTAACATCTCAATCAAGGTAAATGCTTTTTGCTTTTTTTGC (SEQ ID NO: 549)
A118	TTCCTCGTTTTCTCTCGTTGGAAGAAGAAGAAACGAGAAA (SEQ ID NO: 550)	TTTCGGATCAAGCTATGAAGGACGCAAAGAGGGAACTAAA(SEQ ID NO: 551)
U153	TTCCTCGTTTTCTCTCGTTGGACGGAAACGAATCGAGAAA (SEQ ID NO: 552)	TTTCGGATCAAGCTATGAAGGACGCAAAGAGGGAACTAAA (SEQ ID NO: 553)
phiRV1	GTAGTGTATCTCACAGGTCCACGGTTGGCCGTGGACTGCTGAAGAACATTCC (SEQ ID NO: 554)	GAAGGTGTTGGTGCGGGGTTGGCCGTGGTCGAGGTGGGGT (SEQ ID NO: 555)
phi370.1	AAAAAAATACAGCGTTTTTCATGTACAACTATACTAGTTGTAGTGCCTAAAA (SEQ ID NO: 556)	TTGTAAAGGAGACTGATAATGGCATGTACAACTATACTCGTCGGTAAAAAGGCA (SEQ ID NO: 557)
TG1	TCCAGCCCAACAGTGTTAGTCTTTGCTCTTACCCAGTTGGGCGGGA (SEQ ID NO: 558)	GATCAGCTCCGCGGGCAAGACCTTTCTCCTTCACGGGGTGGAAGGTC (SEQ ID NO: 559)
WB	CTAGTTTTAAAGTTGGTTATTAGTTACTGTGATATTTATCACGGTACCCAATAACCAATGAAT (SEQ ID NO: 560)	CGGAAGGTAGCGTCAACGATAGGTGTAACTGTCGTGTTTGTAACGGTACTTCCAACAGCTGGCGCCGCCAC (SEQ ID NO: 561)
BL3	CAATGAAAAACTAGGCATGTAGAAGTTGTTTGT (SEQ ID NO: 562)	TTTCCACAGACAACTCACGTGGAGGTAGTCAC (SEQ ID NO: 563)
SprA	TGTAGTAAGTATCTTAATATACAGCTTTATCTGTTTTTTAAGATACTTACTACTTT (SEQ ID NO: 564)	CACCCATTGTGTTCACAGGAGATACAGCTTTATCTGTACTGATATTAATGACATGCTG (SEQ ID NO: 565)

Рекомбиназа	Последовательность распознавания рекомбиназой (левая)	Последовательность распознавания рекомбиназой (правая)
phiJoe	AGTTGTGGCCATGTGTCCATCTGGGGGGCAGATGGAGACGGGGTCACA (SEQ ID NO: 566)	ATCTGGATGTGGGTGTCCATCTGCGGGCAGACGCCGCAGTCGAAGCACGG (SEQ ID NO: 567)
--	--	ACCTTGATCTCGGTGTCCATCGCCGGGCAGACGCCGCAGTCGAAGCACGG (SEQ ID NO: 568)
phiK38	CCCTAATACGCAAGTCGATAACTCTCCTGGGAGCGTTGACAACTTGCGCACCCTGATCTG (SEQ ID NO: 569)	GAGCGCCGGATCAGGGAGTGGACGGCCTGGGAGCGCTACACGCTGTGGCTGCGGTCGGTGC (SEQ ID NO: 570)
Int2	GCTCATGTATGTGTCTACGCGAGATTCTCGCCCGAGAACTTCTGCAAGGCACTGCTCTTGGCT (SEQ ID NO: 571)	GGACGGCGCAGAAGGGGAGTAGCTCTTCGCCGGACCGTCGACATACTGCTCAGCTCGTC (SEQ ID NO: 572)
Int3	ATGGATAAAAAAATACAGCGTTTTTCATGTACAACTATACTAGTTGTAGTGCCTAAATAATGCTT (SEQ ID NO: 573)	GTTTGTAAAGGAGACTGATAATGGCATGTACAACTATACTCGTCGGTAAAAAGGCATCTTAT (SEQ ID NO: 574)
Int4	AAAAATTACAAAGTTTTCAACCCTTGATTTGAATTAGCGGTCAAATAATTTGTAATTCGTTT (SEQ ID NO: 575)	TTCCAAAGAGCGCCCAACGCGACCTGAAATTTGAATAAGACTGCTGCTTGTGTAAAGGCGATGATT (SEQ ID NO: 576)
Int7	GTGTTATAAACCTGTGTGAGAGTTAAGTTTACATGCCTAACCTTAACTTTTACGCAGGTTCAGCTT (SEQ ID NO: 577)	AGACGAGAAACGTTCCGTCCGTCTGGGTCAGTTGGGCAAAGTTGATGACCGGGTCGTCCGTT (SEQ ID NO: 578)
Int8	TTAATAAACTATGGAAGTATGTACAGTCTTGCAATGTTGAGTGAACAAACTTCCATAATAAAAT (SEQ ID NO: 579)	CAATCATCAGATAACTATGGCGGCACGTGCATTAACCACGGTTGTATCCCGTCTAAAGTACTCGT (SEQ ID NO: 580)
Int9	GTGGTTGTTTTTGTTGGAAGTGTGTATCAGGTATCTGCATAGTTATTCCGAACTTCCAATTA (SEQ ID NO: 581)	TTTATATTGCGAAAAATAATTGGCGAACGAGGTAACTGGATACCTCATCCGCCAATTAAAATTTG (SEQ ID NO: 582)
Int10	GGAAAATATAAATAATTTTAGTAACCTACATCTCAATCAAGGATAGTAAAACTCTCACTCTT (SEQ ID NO: 583)	AGCACGCTGATAATCAGCAAGACCACCAACATTTCCACCAATGTAAAAGCTTTAACCTTAGC (SEQ ID NO: 584)
Int11	GTTTATATGTTTACTAATAAGACGCTCTCAACCCATAAAGTCTTATTAGTAAACATATTTCAACT	ATGGATTTTGCAGATTCCCAGATGCCCCTACAGAAAGAGGTACAAAACATTTATTGGAATTAATT (SEQ ID NO:

Рекомбиназа	Последовательность распознавания рекомбиназой (левая)	Последовательность распознавания рекомбиназой (правая)
	(SEQ ID NO: 585)	586)
Int12	TTTTTGTATGTTAGTTGTGTCA CTGGGTAGACCTAAATAGTGA CACAACTGCTATTAAAATTTAA (SEQ ID NO: 587)	GTTCGTGGTAACTATGGGTGGTAC AGGTGCCACATTAGTTGTACCATTT ATGTTTATGTGGTTAAC (SEQ ID NO: 588)
Int13	CAATAACGGTTGTATTTGTAGA ACTTGACCAGTTGTTTTAGTAA CATAAATACAACTCCGAATA (SEQ ID NO: 589)	GCATACATTGTTGTTGTTTTTCCAG ATCCAGTTGGTCCTGTAAATATAA GCAATCCATGTGAGT (SEQ ID NO: 590)
LI	GTTTAGTATCTCGTTATCTCTC GTTGGAGGGAGAAGAAACGGG ATACCAAAA (SEQ ID NO: 591)	TAACTTTTTCGGATCGAGTTATGAT GGACGTAAAGAGGGAACAAAGCA TCTA (SEQ ID NO: 592)
Peaches	TAGTTTCCAATGTTACAGGAAC TGCTGGCAGAATCCAACACATT GGAAGTCG (SEQ ID NO: 593)	CGGTCTCCATCGGGATCTGCTGATC GAGCAGCATGCCGACCA (SEQ ID NO: 594)
Bxz2	TAACCGCAAGTGTACATCCCTC GGCTGGCCGAGACAAGTACAG TTGCGACAG (SEQ ID NO: 595)	CGGTCTCCATCGGGATCTGCTGATC GAGCAGCATGCCGACCA (SEQ ID NO: 596)
SV1	ATGTGGTCCTTTAGATCCACTG ACGTGGGTCAGTGTCTCTAAAG GACTCGCG (SEQ ID NO: 597)	CATCAGGGCGGTCAGGCCGTAGAT GTGGAAGAAACGGCAGCACGGCG AGGACG (SEQ ID NO: 598)

Таблица 9. Сериновые резольвазы и последовательности мишеней SSR.

Резольваза	Последовательность распознавания рекомбиназой (левая)	Последовательность распознавания рекомбиназой (правая)
Gin	CGTTTCCTGTAAACCGAGGTTTTG GATAAACA (SEQ ID NO: 599)	CGTTTCCTGTAAACCGAGGTTTTG GATAATGG (SEQ ID NO: 600)
Cin	GAGTTCTCTTAAACCAAGGTTTA GGATTGAAA (SEQ ID NO: 601)	GAGTTCTCTTAAACCAAGGTATT GGATAACAG (SEQ ID NO: 602)
Hin	TGGTTCTTGAAAACCAAGGTTTTT GATAAAGC (SEQ ID NO: 603)	AAATTTTCCTTTTTGGAAGGTTTT TGATAACCA (SEQ ID NO: 604)
Min	GCCTTCCCCTAAACCAACGTTTTT ATGCCGCC (SEQ ID NO: 605)	GCCTTCCCCCAAACCAAGGTAAT CAAGAACGC (SEQ ID NO: 606)
Sin	TTGTGAAATTTGGGTACACCCTA ATCATACAA (SEQ ID NO: 607)	CGTATGATTAGGGTGTATATTAA TTT (SEQ ID NO: 608)

Таблица 10. Тирозиновые интегразы и последовательности мишеней.

Интеграза	*attP*	*attB*
HK022	CAAATGATTTTATTTTGACTAATAA TGACCTACTTACATTAATTTACTGAT AATTAAAGAGATTTTAAATATACAA CTTATTCACCTAAAGGATGACAAAA (SEQ ID NO: 609)	GCACTTTAGGTGAAAAAGGT T (SEQ ID NO: 610)
	TAACATTAATCACTTAAAAATCATC GCATTACACTAATCTGTGGTTAAAT GATAGACTACATAATGCGACAAAA CGCAACATATCCAGTCACTATGAAT CAACTACTTAGATAGTATTAGTGAC CT (SEQ ID NO: 611)	
P22	CTAAGTGGTTTGGGACAAAAATGGG ACATACAAATCTTTGCATCGGTTTG CAAGGCTTTGCATGTCTTTCGAAGA TGGGACGTGTGAGCGCAGGTATGAC GTGGTATGTGTTGACTTAAAAGGTA GTTCTTATAATTCGTAATGCGAAGG TCGTAGGTTCGACTCCTATTATCGG CACCAGTTAAATCAAATACTTACGT ATTATTCGTGCCTTCCTTATTTTTAC TGTGGGACATATTTGGGACAGAAGT ACCAAAAA (SEQ ID NO: 612)	GCAGCGCATTCGTAATGCGA AGGTCGT (SEQ ID NO: 613)
L5	GCGATCCCCATCCGCGACGTGCCAA CTAGGTCTCCTCTCGTCGTGAACAA GGCTACCGGGTTGCAACTCCTGTGC AACTCTCAGGCTTCAACGCGCTTCT ACGACCTGCAATTTCTTTCCACTTA GAGGATGCAGCCGAGAGGGGTAAA AACCTATCTTGACCGGCCCATATGT GGTCGGCAGACACCCATTCTTCCAA ACTAGCTACGCGGGTTCGATTCCCG TCGCCCGCTCCGCTGGTCAGAGGGT GTTTTCGCCCTCTGGCCATTTTTCTT TCCAGGGGTCTGCAACTCTTGTGCG ACTCTTCTGACCTGGGCATACGCGG TTGCAACGCATCCCTGATCTGGCTA CTTTCGATGCTGACAAACGAATAGA GCCCCCCGCCTGCGCGAACAGACG AGGGGCATTCACA (SEQ ID NO: 614)	GAGCGGGCGACGGGAATCG AACCCGCGTAGCTAGTTTGG AAGA (SEQ ID NO: 615)

[1022] В различных других аспектах настоящее изобретение относится к способам применения PE для внесения одной или нескольких последовательностей распознавания рекомбиназой и их применению для специфичной в отношении участка рекомбинации.

[1023] В некоторых вариантах осуществления специфичная в отношении участка рекомбинация может обеспечивать получение ряда результатов рекомбинации, например эксцизии, встраивания, инверсии, или замены фрагментов ДНК.

[1024] В некоторых вариантах осуществления данные способы могут применяться для внесения рекомбинации одного или нескольких участков или между одним или несколькими участками одной или нескольких молекул нуклеиновых кислот (например ДНК). В других вариантах осуществления данные способы могут применяться для внесения рекомбинации одного или нескольких участков или между одним или несколькими участками в одной молекуле нуклеиновой кислоты (например ДНК).

[1025] В некоторых вариантах осуществления изобретение предусматривает способ встраивания донорской матрицы ДНК с помощью специфической в отношении участка рекомбинации, включающий: (a) внесение последовательности распознавания рекомбиназой в геномный локус с помощью праймерного редактирования; (b) контактирование геномного локуса с донорской матрицей ДНК, которая также включает последовательность распознавания рекомбиназой, в присутствии рекомбиназы.

[1026] В других вариантах осуществления изобретение предусматривает способ удаления участка генома с помощью специфической в отношении участка рекомбинации, включающий: (a) внесение пары последовательностей распознавания рекомбиназой в геномный локус с помощью праймерного редактирования; (b) контактирование геномного локуса с рекомбиназой, тем самым обеспечивая катализирование удаления участка генома между парой последовательностей распознавания рекомбиназой.

[1027] В других вариантах осуществления изобретение предусматривает способ инверсии участка генома с помощью специфической в отношении участка рекомбинации, включающий: (a) внесение пары последовательностей распознавания рекомбиназой в геномный локус с помощью праймерного редактирования; (b) контактирование геномного локуса с рекомбиназой, тем самым обеспечивая катализирование инверсии участка генома между парой последовательностей распознавания рекомбиназой.

[1028] В других вариантах осуществления изобретение предусматривает способ индуцирования хромосомной транслокации между первым участком генома и вторым участком генома, включающий: (a) внесение первой последовательности

распознавания рекомбиназой в первый геномный локус с помощью праймерного редактирования; (b) внесение второй последовательности распознавания рекомбиназой во второй геномный локус с помощью праймерного редактирования; (c) контактирование первого и второго геномного локуса с рекомбиназой, тем самым обеспечивая хромосомную транслокацию первого и второго геномного локуса

[1029] В других вариантах осуществления изобретение предусматривает способ индуцирования обмена кассетами между геномным локусом и донорской ДНК, включающей кассету, включающий: (a) внесение первой последовательности распознавания рекомбиназой в первый геномный локус с помощью праймерного редактирования; (b) внесение второй последовательности распознавания рекомбиназой во второй геномный локус с помощью праймерного редактирования; (c) контактирование первого и второго геномного локуса с донорской ДНК, включающей кассету, которая фланкируется первой и второй последовательностью распознавания рекомбиназой, и рекомбиназой, тем самым обеспечивая катализирование обмена фланкированного геномного локуса и кассеты в донорской ДНК.

[1030] В различных вариантах осуществления, включающих вставку более чем одной последовательности распознавания рекомбиназой в геном, последовательности распознавания рекомбиназой могут являться одинаковыми или различающимися. В некоторых вариантах осуществления последовательности распознавания рекомбиназой являются одинаковыми. В других вариантах осуществления данные последовательности распознавания рекомбиназой являются различающимися.

[1031] В различных вариантах осуществления рекомбиназа может представлять собой тирозиновую рекомбиназу, например Cre, Dre, Vcre, Scre, Flp, B2, B3, Kw, R, TD1-40, Vika, Nigri, Panto, Kd, Fre, Cre(ALSHG), Tre, Brec1, или Cre-R3M3, как показано в Таблице 7. В таких вариантах осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 7, которая соответствует используемой рекомбиназе.

[1032] В различных других вариантах осуществления рекомбиназа может представлять собой крупную сериновую рекомбиназу, например Bxb1, PhiC31, R4, phiBT1, MJ1, MR11, TP901-1, A118, V153, phiRV1, phi370.1, TG1, WB, BL3, SprA, phiJoe, phiK38, Int2, Int3, Int4, Int7, Int8, Int9, Int10, Int11, Int12, Int13, L1, peaches, Bxz2, или SV1, как показано в Таблице 8. В таких вариантах осуществления последовательность

распознавания рекомбиназой может представлять собой RRS из Таблицы 8, которая соответствует используемой рекомбиназе.

[1033] В различных других вариантах осуществления рекомбиназа может представлять собой сериновую рекомбиназу, например Bxb1, PhiC31, R4, phiBT1, MJ1, MR11, TP901-1, A118, V153, phiRV1, phi370.1, TG1, WB, BL3, SprA, phiJoe, phiK38, Int2, Int3, Int4, Int7, Int8, Int9, Int10, Int11, Int12, Int13, L1, peaches, Bxz2, или SV1, как показано в Таблице 8. В таких вариантах осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 8, которая соответствует используемой рекомбиназе.

[1034] В других вариантах осуществления рекомбиназа может представлять собой сериновую резольвазу, например Gin, Cin, Hin, Min, или Sin, как показано в Таблице 9. В таких вариантах осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 9, которая соответствует используемой рекомбиназе.

[1035] В различных других вариантах осуществления рекомбиназа может представлять собой тирозиновую интегразу, например HK022, P22, или L5, как показано в Таблице 10. В таких вариантах осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 10, которая соответствует используемой рекомбиназе.

[1036] В некоторых вариантах осуществления любой из способов специфической в отношении участка рекомбинации с использованием PE может быть реализован *in vivo* или *in vitro*. В некоторых вариантах осуществления любой из способов специфической в отношении участка рекомбинации реализуется в клетке (например рекомбинация геномной ДНК в клетке). Клетка может являться прокариотической или эукариотической. Клетка, такая как эукариотическая клетка, может находится в субъекте, таком как описанный здесь субъект (например представляющий собой человека субъект). Описанные здесь способы могут применяться для генетического модифицирования клеток *in vitro* и *in vivo*, например в контексте получения трансгенных клеток, линий клеток, или животных, или при изменении геномной последовательности, например коррекции генетического дефекта в клетке в субъекте.

Ссылки, приведенные в Разделе I

[1037] Каждая из следующих ссылок приведена в Примере 17, каждая из которых приведена здесь в качестве ссылки.

1. Feuk, L. Inversion variants in the human genome: role in disease and genome architecture. Genome Med 2, 11 (2010).

2. Zhang, F., Gu, W., Hurles, M.E. & Lupski, J.R. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet 10, 451-481 (2009).

3. Shaw, C.J. & Lupski, J.R. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. Hum Mol Genet 13 Spec No 1, R57-64 (2004).

4. Carvalho, C.M., Zhang, F. & Lupski, J.R. Evolution in health and medicine Sackler colloquium: Genomic disorders: a window into human gene and genome evolution. Proc Natl Acad Sci U S A 107 Suppl 1, 1765-1771 (2010).

5. Rowley, J.D. Chromosome translocations: dangerous liaisons revisited. Nat Rev Cancer 1, 245-250 (2001).

6. Aplan, P.D. Causes of oncogenic chromosomal translocation. Trends Genet 22, 46-55 (2006).

7. McCarroll, S.A. & Altshuler, D.M. Copy-number variation and association studies of human disease. Nat Genet 39, S37-42 (2007).

[8] Способы лечения

[1038] Настоящее изобретение предусматривает способы лечения субъекта, у которого диагностировано заболевание, связанной или вызванное точечной мутацией или другими видами мутаций (например делецией, вставкой, инверсией, дупликацией и так далее), которые могут быть скорректированы с помощью описанной здесь системы праймерного редактирования, например, но без ограничения, прионное заболевание (например Пример 5 здесь), связанное экспансией тринуклеотидных повторов заболевание (например Пример 3 здесь), или связанное с дефицитом CDKL5 заболевание (CDD) (например Пример 23 здесь).

[1039] Практически любой вызывающий заболевание генетический дефект может быть скорректирован с помощью праймерного редактирования, которое включает выбор подходящего слитого белка праймерного редактора (включая napDNAbp и полимеразу (например обратную транскриптазу)) и конструирование подходящей PEgRNA, сконструированной для (a) нацеливания на подходящую ДНК, содержащую участок редактирования, и (b) обеспечения матрицы для синтеза одиночной цепи ДНК от 3'-конца участка разрыва, которая включает требуемое изменение, которое вытесняет и

замещает эндогенную цепь непосредственно ниже участка разрыва. Праймерное редактирование может применяться, без ограничения, для (a) внесения корректирующих мутации изменений в нуклеотидную последовательность, (b) внесения белковых и РНК меток, (c) внесения иммуноэпитопов в представляющие интерес белки, (d) внесения доменов индуцируемой димеризации в белки, (e) внесения или удаления последовательностей для изменения активности биомолекулы, (f) внесения участков-мишеней рекомбиназы для реализации определенных генетических изменения, и (g) мутагенеза последовательности-мишени за счет использования подверженной ошибкам RT.

[1040] Способ лечения заболевания может включать в качестве первого этапа конструирование подходящей PEgRNA и слитого белка праймерного редактора в соответствии с описанными здесь способами, которое включает ряд условий, которые могут быть приняты во внимание, таких как:

(a) последовательность-мишень, то есть нуклеотидная последовательность, в которую требуется внести одну или несколько модификаций азотистых оснований с помощью праймерного редактора;

(b) расположение участка разреза в рамках последовательности-мишени, то есть конкретное положение азотистого основания, в котором праймерный редактор индуцирует одноцепочечный разрыв для создания 3′-концевой праймерной последовательности RT на одной стороне разрыва и 5′-концевого эндогенного флэпа на другой стороне разрыва (который в конечном итоге удаляется FEN1 или ее эквивалентом и замещается 3′ оцДНК-флэпом. Участок разрыва создает 3′-концевую праймерную последовательность, которая становится удлиненной за счет полимеразы слитого белка PE (например фермента RT) при РНК-зависимой полимеризации ДНК для создания 3′ оцДНК-флэпа, содержащего требуемое изменение, который затем замещает 5′ эндогенный ДНК-флэп в последовательности-мишени.

(c) доступность последовательностей PAM (включая стандартные участки PAM SpCas9, а также нестандартные участки PAM, распознаваемые вариантами и эквивалентами Cas9 с расширенной или иной специфичностью в отношении PAM).

(d) промежуток между доступными последовательностями PAM и расположением участка разрыва в цепи PAM;

(e) конкретного доступного для использования Cas9, варианта Cas9, или эквивалента Cas9 праймерного редактора (который частично определяется на основе доступного PAM);

(f) последовательность и длина участка связывания праймера;

(g) последовательность и длина матрицы для редактирования;

(h) последовательность и длина плеча гомологии;

(i) спейсерная последовательность и длина; и

(j) последовательность ядра нРНК.

[1041] Подходящая PEgRNA и необязательно направляющая для конструирования никирующей sgРНК для никирования второго участка могут быть сконструированы на основе следующего являющегося примером пошагового набора инструкций, который учитывает одно или несколько из приведенных выше условий. В рамках этапов приведены ссылки на примеры, показанные на Фиг.70A-70I.

1. **Определение последовательности-мишени и изменения.** Получить последовательность участка-мишени ДНК (~200bp) с центром вокруг месторасположения требуемого изменения (точечной мутации, вставки, делеции, или их комбинации). См. Фиг.70A.

2. **Поиск PAM-мишеней.** Идентифицировать PAM поблизости от требуемого месторасположения изменения. PAM могут быть идентифицированы на любой цепи ДНК поблизости от требуемого месторасположения изменения. В то время как являются предпочтительными PAM рядом с положением изменения (то есть где участок разрыва расположен менее чем в 30 nt от положения изменения или менее чем в 29 nt, 28 nt, 27 nt, 26 nt, 25 nt, 24 nt, 23 nt, 22 nt, 21 nt, 20 nt, 19 nt, 18 nt, 17 nt, 16 nt, 15 nt, 14 nt, 13 nt, 12 nt, 11 nt, 10 nt, 9 nt, 8 nt, 7 nt, 6 nt, 5 nt, 4 nt, 3 nt, или 2 nt от положения изменения), является возможным вносить изменения с помощью протоспейсеров и PAM, которые обеспечивают расположение разрыва на расстоянии ≥ 30 nt от положения изменения. См. Фиг.70B.

3. **Поиск участков разрыва.** Для каждого рассматриваемого PAM идентифицировать соответствующий участок разрыва и цепь. Для никазы Sp Cas9 H840A расщепление осуществляется в PAM-содержащей цепи между 3-м и 4-м основаниями 5′ относительно PAM NGG. Все подвергнутые редактированию нуклеотиды должны быть расположены 3′ относительно участка разрыва, так что подходящие PAM должны обеспечивать расположение разрыва 5′ относительно

изменения-мишени на PAM-содержащей цепи. В показанном ниже примере присутствуют два возможных PAM. Для простоты остальные этапы продемонстрируют конструирование PEgRNA с использованием только PAM 1. См. Фиг.70C.

4. **Конструирование спейсерной последовательности.** Протоспейсер Sp SpCas9 соответствует 20 нуклеотидам 5′ относительно PAM NGG на PAM-содержащей цепи. Эффективное инициирование транскрипции Pol III требует, чтобы G был первым транскрибированным нуклеотидом. Если первый нуклеотид протоспейсера представляет собой G, то спейсерная последовательность для PEgRNA будет просто представлять собой протоспейсерную последовательность. Если первый нуклеотид протоспейсера не представляет собой G, то спейсерная последовательность PEgRNA представляет собой G и далее протоспейсерную последовательность. См. Фиг.70D.

5. **Конструирование участка связывания праймера (PBS).** Идентифицировать праймер ДНК на PAM-содержащей цепи с помощью последовательности начального аллеля. 3′-конец праймера ДНК представляет собой нуклеотид, расположенный непосредственно выше участка разрыва (то есть 4-го основания 5′ относительно PAM NGG для Sp Cas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера (PBS) PEgRNA, содержащий от 12 до 13 нуклеотидов комплементарности относительно праймера ДНК, может использоваться для последовательностей, которые содержат ~40-60% содержимого в виде GC. Для последовательностей с низким содержанием GC должны быть протестированы более длинные PBS (от 14 до 15 nt). Для последовательностей с высоким содержанием GC должны быть протестированы более короткие PBS (от 8 до 11 nt). Оптимальные последовательности PBS должны быть определены эмпирически независимо от содержания GC. Для конструирования PBS последовательности длиной p следует взять обратный комплемент первых p нуклеотидов 5′ участка разрыва в PAM-содержащей цепи с использованием последовательности *начального* аллеля. См. Фиг.70E.

6. **Конструирование RT матрицы (или матрицы для синтеза ДНК).** RT матрица (или матрица для синтеза ДНК, где полимераза не является обратной транскриптазой) кодирует требуемое изменение и гомологию относительно последовательности, являющейся смежной относительно изменения. В одном варианте осуществления данные участки соответствуют матрице для синтеза ДНК согласно Фиг.3D и Фиг.3E, где матрица для синтеза ДНК включает "матрицу для редактирования" и "плечо гомологии". Оптимальные величины длины RT матрицы

варьируются в зависимости от участка-мишени. Для коротких изменений (положения от +1 до +6) рекомендуется тестировать короткую (от 9 до 12 nt), среднюю (от 13 до 16 nt), и длинную (от 17 до 20 nt) RT матрицу. Для длинных изменений (положения +7 и более) рекомендуется использовать RT матрицы, которые имеют длину, по меньше мере, на 5 nt превышающую (предпочтительно на 10 nt или более) положение изменения для обеспечения достаточной гомологии 3′ ДНК-флэпа. Для длинных изменений несколько RT матриц должны быть подвергнуты скринингу для идентификации функциональных конструкций. Для более крупных вставок и делеций (≥5 nt) рекомендуется встраивание большей 3′ гомологии (~20 nt или более) в RT матрицу. Эффективность редактирования обычно ухудшается, когда RT матрица кодирует синтез G в качестве последнего нуклеотида в обратно транскрибированном ДНК продукте (соответствует C в RT матрице PEgRNA). Поскольку многие RT матрицы поддерживают эффективное праймерное редактирование, при конструировании RT матриц рекомендуется не использовать G в качестве последнего синтезируемого нуклеотида Для конструирования последовательности RT матрицы длиной r следует использовать последовательность требуемого аллеля и взять обратный комплемент первых r нуклеотидов 3′ участка разрыва в цепи, которая первоначально содержала PAM. Следует обратить внимание, что по сравнению с изменениями SNP, изменения в виде вставки или делеции с использованием RT матриц той же длины не будут содержать идентичную гомологию. См. Фиг.70F.

7. **Сборка полной последовательности PEgRNA.** Следует конкатенировать компоненты PEgRNA в следующем порядке (от 5′ к 3′): спейсер, каркас, RT матрица и PBS. См. Фиг.70G.

8. **Конструирование никирующих sgРНК для PE3.** Идентифицировать PAM на не подвергаемой редактированию цепи выше и ниже изменения. Оптимальные положения никирования имеют высокую зависимость от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии от 40 до 90 нуклеотидов 5′ относительно положения PEgRNA-индуцированного разрыва, ведут к более высоким выходам редактирования и меньшему количеству инделей. Никирующая sgРНК имеет спейсерную последовательность, которая совпадает с протоспейсером длиной 20 nt в *начальном* аллеле, с добавлением 5′ G, если протоспейсер не начинается на G. См. Фиг.70H.

9. **Конструирование никирующих sgРНК PE3b** Если в комплементарной цепи присутствует PAM и его соответствующий протоспейсер накладывается на последовательность, являющуюся мишенью редактирования, то такое изменение может являться кандидатом для системы PE3b. В системе PE3b спейсерная последовательность никирующей sgРНК совпадает с последовательностью требуемого подвергаемого редактированию аллеля, но не начального аллеля. Система PE3b работает эффективно, когда подвергаемый редактированию нуклеотид (нуклеотиды) попадают в участок затравки (~10 nt от PAM) протоспейсера никирующей sgРНК. Это предотвращает никирование комплементарной цепи до готовности подвергнутой редактированию цепи, предотвращая конкурирование между PEgRNA и sgРНК за связывание с ДНК-мишенью. PE3b также обеспечивает избежание образования одновременных разрывов на обеих цепях, тем самым значительно снижая образование инделей, в то же время сохраняя высокую эффективность редактирования. sgРНК PE3b должны иметь спейсерную последовательность, которая совпадает с протоспейсером длиной 20 nt в *требуемом* аллеле, с добавлением 5′ G, если это необходимо. См. Фиг.70I.

[1042] Указанный выше пошаговый процесс конструирования подходящей PEgRNA и sgРНК для никирования второго участка не является ограничивающим каким-либо образом. Изобретение охватывает вариации описанного выше пошагового процесса, которые будут очевидны среднему специалисту в данной области.

[1043] После выбора/конструирования подходящей PEgRNA и слитого белка PE они могут быть введены с помощью подходящих способов, например трансфекции на основе векторов (при которой один или несколько векторов включают ДНК, кодирующую PEgRNA и слитый белок PE, которые экспрессируются в клетке при трансфекции с использованием векторов), прямой доставки слитого белка PE с PEgRNA (например доставки на основе RNP) в формате для доставки (например в виде липидных частиц, наночастиц), или с помощью системы доставки на основе иРНК. Подобные способы описаны здесь, при этом может использоваться любой известный способ.

[1044] PEgRNA и слитый белок PE (или совместно обозначаемые как PE комплекс) могут быть доставлены в клетку в терапевтически эффективном количестве, так что при контактировании с представляющей интерес ДНК-мишенью в нее будет вноситься требуемое изменение.

[1045] Любое заболевание является предположительно подлежащим лечению с помощью подобных способов при условии, что доставка в подходящие клетки является реализуемой. Средний специалист в данной области будет способен выбрать и/или определить способ доставки PE, соответствующий предполагаемой цели и предполагаемым клеткам-мишеням.

[1046] Например, в некоторых вариантах осуществления предусмотрен способ, который включает введение субъекту, имеющему такое заболевание, например рак, ассоциированный с точечной мутацией, как описано выше, эффективного количества описанной здесь системы для праймерного редактирования, которая корректирует точечную мутацию или вносит деактивирующую мутацию в ассоциированный с заболеванием ген на основе определяемой гомологией репарации в присутствии донорской молекулы ДНК, включающей требуемое генетическое изменение. В некоторых вариантах осуществления предусмотрен способ, который включает введение субъекту, имеющему такое заболевание, например рак, ассоциированный с точечной мутацией, как описано выше, эффективного количества описанной здесь системы для праймерного редактирования, которая корректирует точечную мутацию или вносит деактивирующую мутацию в ассоциированный с заболеванием ген. В некоторых вариантах осуществления заболевание представляет собой пролиферативное заболевание. В некоторых вариантах осуществления заболевание представляет собой генетическое заболевание. В некоторых вариантах осуществления заболевание представляет собой неопластическое заболевание. В некоторых вариантах осуществления заболевание представляет собой метаболическое заболевание. В некоторых вариантах осуществления заболевание представляет собой лизосомное заболевание накопления. Другие заболевания, которые могут подвергаться лечению путем корректирования точечной мутации или введения деактивирующей мутации в ассоциированный с заболеванием ген, будут известны специалистам в данной области, при этом изобретение не является ограниченным в этом отношении.

[1047] Настоящее изобретение предусматривает способы лечения дополнительных заболеваний или расстройств, например заболеваний или расстройств, которые ассоциированы или вызваны точечной мутацией, которая может быть скорректирована с помощью TPRT-опосредованного редактирования генов. Некоторые подобные заболевания описаны здесь, при этом дополнительные подходящие заболевания, которые могут подвергаться лечению с использованием описанных здесь стратегий и

слитых белков, будут очевидны специалистам в данной области на основе настоящего описания. Примеры подходящих заболеваний и расстройств перечислены ниже. Очевидно, что нумерация конкретных положений или остатков в соответствующих последовательностях зависит от конкретного используемого белка и схемы нумерации. Нумерация может различаться, например в предшественниках зрелого белка и самом зрелом белке, при этом различия в последовательностях от одного вида к другому могут влиять на нумерацию. Специалист в данной области будет способен идентифицировать соответствующий остаток в любом гомологичном белке и в соответствующей кодирующей нуклеиновой кислоте с помощью способов, которые хорошо известны в данной области, например путем выравнивания последовательностей и определения гомологичных остатков. Примеры подходящих заболеваний и расстройств включают, без ограничения, такие заболевания и расстройства как: 2-метил-3-гидроксимасляная ацидурия; дефицит 3 бета-гидроксистероиддегидрогеназы; 3-метилглутаконовая ацидурия; дефицит 3-оксо-5-альфа-стероид-дельта-4-дегидрогеназы; 46, реверсия пола XY, тип 1, 3 и 5; дефицит 5-оксопролиназы; дефицит 6-пирувоил-тетрагидроптеринсинтазы; синдром Аарског; синдром Аазе; ахондрогенез 2 типа; ахроматопсия 2 и 7; приобретенный синдром удлиненного интервала QT; акрокаллозальный синдром по типу Шинцеля; акрокапитофеморальная дисплазия; акродизостоз 2 с гормональной резистентностью или без нее; акроэритрокератодермия; акромикрическая дисплазия; Acth-независимая макронодулярная гиперплазия надпочечников 2; синдром активированного PI3K-дельта; острая перемежающаяся порфирия; дефицит семейства ацил-КоА дегидрогеназ, член 9; синдром Адамса-Оливера 5 и 6; дефицит аденинфосфорибозилтрансферазы; дефицит аденилаткиназы; гемолитическая анемия вследствие дефицита аденилосукцинатлиазы; подростковый нефронофтиз; почечно-печеночно-панкреатическая дисплазия; синдром Меккеля тип 7; адренолейкодистрофия; буллезный соединительный эпидермолиз у взрослых; буллезный эпидермолиз, соединительный, локализованный вариант; цероидный липофусциноз нейронов взрослых; цероидный липофусциноз нейронов взрослых; атаксия у взрослых с глазодвигательной апраксией; синдром ADULT; афибриногенемия и врожденная афибриногенемия; аутосомно-рецессивная агаммаглобулинемия 2; возрастная дегенерация желтого пятна 3, 6, 11 и 12; синдромы Айкарди-Гутьереса 1, 4 и 5; хилбеновая волчанка 1; синдромы Алажилля 1 и 2; болезнь Александра;

алькаптонурия; синдром Аллана-Херндона-Дадли; врожденная общая алопеция; энцефалопатия Альперса; дефицит альфа-1-антитрипсина; аутосомно-доминантный, аутосомно-рецессивный и Х-сцепленный рецессивный синдром Альпорта; болезнь Альцгеймера, семейная, 3, со спастическим парапарезом и апраксией; болезнь Альцгеймера, типы 1, 3 и 4; IIA1 несовершенный амелогенез, тип гипокальцификации и тип гипоматурации; дефицит аминоацилазы 1; синдром детской эпилепсии амишей; амилоидогенный транстиретиновый амилоидоз; амилоидная кардиомиопатия, связанная с транстиретином; кардиомиопатия; боковой амиотрофический склероз типов 1, 6, 15 (с лобно-височной деменцией или без нее), 22 (с лобно-височной деменцией или без нее) и 10; лобно-височная деменция с включениями TDP43, связанная с TARDBP; синдром Андерманна; синдром Андерсена-Тавиля; врожденный синдром удлиненного интервала QT; анемия, несфероцитарная гемолитическая, вследствие дефицита G6PD; синдром Ангельмана; тяжелая неонатальная энцефалопатия с микроцефалией; предрасположенность к аутизму, Х-сцепленный 3; ангиопатия, наследственная, с нефропатией, аневризмами и мышечными спазмами; ангиотензин-превращающий фермент, доброкачественное повышение в сыворотке крови; аниридия, мозжечковая атаксия и умственная отсталость; анонихия; дефицит антитромбина III; синдром Антли-Бикслера с генитальными аномалиями и нарушенным стероидогенезом; аневризм аорты, семейная торакальная 4, 6 и 9; аневризмы грудной аорты и расслоение аорты; синдром мультисистемной дисфункции гладких мышц; болезнь моямоя 5; апластическая анемия; очевидный избыток минералокортикоидов; дефицит аргиназы; дефицит аргининосукцинатлиазы; дефицит ароматазы; аритмогенная кардиомиопатия правого желудочка 5, 8 и 10 типов; первичная семейная гипертрофическая кардиомиопатия; врожденный множественный артрогрипоз, дистальный, Х-сцепленный; холестазный синдром на фоне артрогрипоза и почечной дисфункции ; артрогрипоз, нарушение функции почек и холестаз 2; дефицит аспарагинсинтетазы; нарушение миграции нейронов; атаксия при дефиците витамина Е; атаксия сенсорная, аутосомно-доминантная; синдром атаксии-телеангиэктазии; наследственный предрасполагающий к раку синдром; атрансферринемия; фибрилляция предсердий, семейная, 11, 12, 13 и 16; дефекты межпредсердной перегородки 2, 4 и 7 (с дефектами атриовентрикулярной проводимости или без них); остановка предсердий 2; дефект атриовентрикулярной перегородки 4; наследственная луковичная атрофия; синдром ATR-X; аурикуло-

мыщелковый синдром 2; аутоиммунное заболевание, мультисистемное, с младенческим началом ; аутоиммунный лимфопролиферативный синдром, тип 1a; аутосомно-доминантная гипогидротическая эктодермальная дисплазия; аутосомно-доминантная прогрессирующая наружная офтальмоплегия с делециями 1 и 3 митохондриальной ДНК; аутосомно-доминантная торсионная дистония 4; аутосомно-рецессивная центроядерная миопатия; аутосомно-рецессивный врожденный ихтиоз 1, 2, 3, 4A и 4B; аутосомно-рецессивная кутислакса типа IA и 1B; синдром аутосомно-рецессивной гипогидротической эктодермальной дисплазии; эктодермальная дисплазия 11b; гипогидротическая / волос / зубов, аутосомно-рецессивная; аутосомно-рецессивная гипофосфатемическая болезнь костей; синдром Аксенфельда-Ригера тип 3; синдром Бейнбриджа-Роперса; синдром Баннаяна-Райли-Рувалькабы; синдром опухоли гамартомы PTEN; синдромы Барайцера-Винтера 1 и 2; синдром Бараката; синдромы Барде-Бидля 1, 11, 16 и 19; синдром голых лимфоцитов тип 2, группа комплементации E; синдром Барттера антенатальный тип 2; синдром Барттера 3, 3 типа с гипокальциурией и 4 типа; кальцификация базальных ганглиев, идиопатическая, 4; монилетрикс; доброкачественная семейная гематурия; доброкачественные семейные неонатальные судороги 1 и 2; судороги, доброкачественные семейные неонатальные 1, и/или миокимия; судороги, ранняя детская эпилептическая энцефалопатия 7; доброкачественные семейные неонатально-младенческие судороги; доброкачественная наследственная хорея; доброкачественная лопаточно-перонеальная мышечная дистрофия с кардиомиопатией; синдром Бернара-Сулье, типы A1 и A2 (аутосомно-доминантный); бестрофинопатия, аутосомно-рецессивная; бета-талассемия; миопатия Бетлема и миопатия Бетлема 2; кристаллическая корнеоретинальная дистрофия Бьетти; нарушение синтеза желчной кислоты, врожденный, 2; дефицит биотинидазы; синдром дисморфизма и умственной отсталости Бирка-Барела; блефарофимоз, птоз и обратный эпикантус; синдром Блума; синдром Борхесона-Форссмана-Лемана; синдром Буше-Нейхаузера; брахидактилия, типы A1 и A2; брахидактилия при гипертонии; заболевание мелких сосудов головного мозга с кровотечением; дефицит киназы кетоациддегидрогеназы с разветвленной цепью; жаберные синдромы 2 и 3; рак груди с ранним началом; рак груди-яичников, семейный 1, 2 и 4; синдром хрупкой роговицы 2; миопатия Броди ; бронхоэктазия с повышенным содержанием хлорида пота или без него 3; синдром Брауна-Виалетто-Ван Лаэра и синдром Брауна-Виалетто-Ван Лаэра 2; синдром Бругада; синдром Бругада 1; фибрилляция желудочков; пароксизмальная

семейная фибрилляция желудочков; синдром Бругада и синдром Бругада 4; синдром удлиненного интервала QT; внезапная сердечная смерть; макулярная дистрофия ("бычий глаз"); болезнь Штаргардта 4; конус-стержневая дистрофия 12; буллезная ихтиозиформная эритродермия; синдром Бёрна-Маккеуна; кандидоз, семейный, 2, 5, 6 и 8; гликопротеиновый синдром дефицита углеводов I и II типа; дефицит карбоангидразы VA, гипераммониемия вследствие нее; рак толстой кишки; аритмия сердца; синдром удлиненного интервала QT, подтип LQT1; кардиоэнцефаломиопатия со смертельным исходом из-за недостаточности цитохром-с-оксидазы; кардиофациально-кожный синдром; кардиомиопатия; болезнь Данона; гипертрофическая кардиомиопатия; кардиомиопатия без компактации левого желудочка; синдром Карневале; комплекс Карни, тип 1; дефицит карнитин-ацилкарнитинтранслоказы; дефицит карнитин-пальмитоилтрансферазы I, II, II (позднее начало) и II (младенческий); катаракта 1, 4, аутосомно-доминантная, аутосомно-доминантная, множественных типов, с микрокорнеей, коппокоподобная, ювенильная, с микрокорнеей и глюкозурией, и ядерная диффузная непрогрессирующая; катехоламинергическая полиморфная желудочковая тахикардия; синдром каудальной регрессии; дефицит Cd8, семейный; заболевание центрального ядра; центромерная нестабильность хромосом 1,9 и 16 и иммунодефицит; мозжечковая атаксия в младенчестве с прогрессирующей внешней офтальмоплегией и мозжечковой атаксией, умственной отсталостью и синдромом дисбаланса 2; церебральная амилоидная ангиопатия, связанная с APP; церебральная аутосомно-доминантная и рецессивная артериопатия с подкорковыми инфарктами и лейкоэнцефалопатией; кавернозные мальформации головного мозга 2; цереброокулофациоскелетный синдром 2; церебро-окуло-фациально-скелетный синдром; цереброретинальная микроангиопатия с кальцификациями и кистами; цероидный липофусциноз нейронов 2, 6, 7 и 10; синдром Чедиака-Хигаши, взрослый тип; болезнь Шарко-Мари-Тута типов 1B, 2B2, 2C, 2F, 2I, 2U (аксональная), 1C (демиелинизирующая), доминантная промежуточная C, рецессивная промежуточная A, 2A2, 4C, 4D, 4H, IF, IVF и X; спинальная мышечная атрофия лопатки и надкостницы; дистальная спинномозговая мышечная атрофия, врожденная непрогрессирующая; спинальная мышечная атрофия дистальная, аутосомно-рецессивная, 5 ; CHARGE-ассоциация; гипофосфатазия в детстве; гипофосфатазия у взрослых; холецистит; прогрессирующий семейный внутрипеченочный холестаз 3; внутрипеченочный холестаз при беременности 3;

болезнь накопления холестанола; дефицит холестеринмонооксигеназы (расщепление боковой цепи); хондродисплазия типа Бломстранда; хондродисплазия точечная 1, Х-сцепленная рецессивная и 2 Х-сцепленная доминантная; синдром CHOPS; хроническая гранулематозная болезнь, аутосомно-рецессивный цитохром b-положительный, типы 1 и 2; синдром Чадли-Маккалоу; цилиарная дискинезия, первичная, 7, 11, 15, 20 и 22; цитруллинемия I типа; цитруллинемия I и II типа; клейдокраниальный дизостоз; c-подобный синдром; синдром Кокейна типа A,; дефицит коэнзима Q10, первичный 1, 4 и 7; синдром Коффина-Сириса/умственная отсталость; синдром Коффина-Лоури; синдром Коэна; синдром холодового потоотделения 1; синдром Коула-Карпентера 2; комбинированные клеточные и гуморальные иммунные дефекты с гранулемами; комбинированная d-2- и l-2-гидроксиглутаровая ацидурия; комбинированная малоновая и метилмалоновая ацидурия; недостатки комбинированного окислительного фосфорилирования 1, 3, 4, 12, 15 и 25; комбинированный частичный и полный дефицит 17-альфа-гидроксилазы/17,20-лиазы; общий вариабельный иммунодефицит 9; компонент 4 комплемента, частичный дефицит из-за дисфункционального ингибитора c1; дефицит фактора комплемента B; конусный монохроматизм; дистрофия палочек и колбочек 2 и 6; дистрофия палочек и колбочек, несовершенный амелогенез; врожденная гиперплазия надпочечников и Врожденная гипоплазия надпочечников, Х-сцепленная; врожденная амегакариоцитарная тромбоцитопения; врожденная аниридия; врожденная центральная гиповентиляция; болезнь Гиршпрунга 3; врожденная контрактурная арахнодактилия; врожденные контрактуры конечностей и лица, гипотония и задержка развития; врожденное нарушение гликозилирования типов 1B, 1D, 1G, 1H, 1J, 1K, 1N, 1P, 2C, 2J, 2K, IIm; врожденная дизеритропоэтическая анемия I и II типа; врожденная эктодермальная дисплазия лица; врожденная эритропоэтическая порфирия; врожденная генерализованная липодистрофия 2 типа; врожденный порок сердца, несколько типов, 2; врожденный порок сердца; прерванная дуга аорты; врожденный избыточный липоматозный рост, сосудистые аномалии и эпидермальные невусы; немелкоклеточный рак легкого; новообразование яичника; нарушение сердечной проводимости неспецифическое; врожденная микроворсинчатая атрофия; врожденная мышечная дистрофия; врожденная мышечная дистрофия из-за частичного дефицита LAMA2; врожденная мышечная дистрофия-дистрогликанопатия с аномалиями мозга и глаз, типы A2, A7, A8, A11 и A14; врожденная мышечная дистрофия-дистрогликанопатия с умственной отсталостью, типы B2, B3, B5 и B15;

врожденная мышечная дистрофия-дистрогликанопатия без умственной отсталости, тип B5; врожденный мышечно-гипертрофический церебральный синдром; врожденный миастенический синдром, чувствительный к ацетазоламиду; врожденная миопатия с диспропорцией типов волокон; врожденная колобома глаза; врожденная стационарная куриная слепота типа 1A, 1B, 1C, 1E, 1F и 2A; копропорфирия; плоская роговица 2; дистрофия роговицы, эндотелий Фукса, 4; эндотелиальная дистрофия роговицы 2 типа; хрупкость роговицы, кератоглобус, синие склеры и гипермобильность суставов; синдромы Корнелии де Ланге 1 и 5; ишемическая болезнь сердца аутосомно-доминантная 2; ишемическая болезнь сердца; гипераль фалипопротеинемия 2; корковая дисплазия, сложная, с другими пороками развития головного мозга 5 и 6; корковые аномалии, затылочные; дефицит кортикостероид-связывающего глобулина; дефицит кортикостеронметилоксидазы 2 типа; синдром Костелло; синдром Каудена 1; болезнь Пертеса; краниодиафизарная дисплазия, аутосомно-доминантная; краниосиностоз 1 и 4; краниосиностоз и стоматологические аномалии; дефицит креатина, X-связанный; синдром Крузона; синдром криптофтальма; крипторхизм, односторонний или двусторонний; симфалангизм Кушинга; кожная злокачественная меланома 1; дряблая кожа при остеодистрофии и тяжелых патологиях легких, желудочно-кишечного тракта и мочевыводящих путей; цианоз, преходящий неонатальный и атипичный нефропатический; муковисцидоз; цистинурия; дефицит цитохром-с-оксидазы I; дефицит цитохром-с оксидазы; D-2-гидроксиглутаровая ацидурия 2; болезнь Дарье, сегментарная; глухота с лабиринтной аплазией, микротией и микродонтией (LAMM) ; глухота, аутосомно-доминантная 3a, 4, 12, 13, 15, аутосомно-доминантная несиндромная нейросенсорная 17, 20 и 65; глухота, аутосомно-рецессивная 1A, 2, 3, 6, 8, 9, 12, 15, 16, 18b, 22, 28, 31, 44, 49, 63, 77, 86 и 89; глухота, кохлеарная, с миопией и интеллектуальными нарушениями, без вестибулярного поражения, аутосомно-доминантная, X-сцепленная 2; дефицит 2-метилбутирил-КоА дегидрогеназы; дефицит 3-гидроксиацил-КоА дегидрогеназы; дефицит альфа-маннозидазы; дефицит декарбоксилазы ароматических L-аминокислот; дефицит бисфосфоглицератмутазы; дефицит бутирил-КоА дегидрогеназы; дефицит ферроксидазы; дефицит галактокиназы; дефицит гуанидиноацетатметилтрансферазы; дефицит гиалуроноглюкозаминидазы; дефицит рибозо-5-фосфатизомеразы; дефицит стероид-11-бета-монооксигеназы; дефицит УДФ-глюкозо-гексозо-1-фосфатуридилилтрансферазы; дефицит ксантиноксидазы; болезнь Дежерина-Соттаса;

болезнь Шарко-Мари-Тута, типы ID и IVF; синдром Дежерина-Соттаса, аутосомно-доминантный; дефицит дендритных клеток, моноцитов, B-лимфоцитов и естественных лимфоцитов-киллеров; дисплазия Десбуки 2; синдром Десбуки; DFNA 2 несиндромная потеря слуха; сахарный и несахарный диабет с атрофией зрительного нерва и глухотой; сахарный диабет 2 типа и инсулинозависимый, 20; анемия Даймонда-Блэкфана 1, 5, 8 и 10; диарея 3 (секреторная, натриевая, врожденная, синдромальная) и 5 (с тафтинговой энтеропатией, врожденная); дикарбоновая аминоацидурия; диффузная ладонно-подошвенная кератодермия ботнического типа; дигитореноцеребральный синдром; дефицит дигидроптеридинредуктазы; дилатационная кардиомиопатия 1A, 1AA, 1C, 1G, 1BB, 1DD, 1FF, 1HH, 1I, 1KK, 1N, 1S, 1Y и 3B; некомпакция левого желудочка 3; нарушение стероидогенеза из-за дефицита оксидоредуктазы цитохрома p450; дистальный артрогрипоз типа 2B; дистальная наследственная мотонейронопатия типа 2B; дистальная миопатия по типу Маркесбери-Григгса; дистальная спинномозговая мышечная атрофия, X-сцепленная 3; дистихиазно-лимфедемный синдром; доминирующий дистрофический буллезный эпидермолиз с отсутствием кожи; доминирующая наследственная атрофия зрительного нерва; синдром Доннаи-Барроу; дефицит дофамин-бета-гидроксилазы; рецептор дофамина d2, снижение плотности в мозге; болезнь Даулинга-Дего 4; сотовая дистрофия сетчатки Дойна; Malattia leventinese; синдром Дуэйна 2 типа; синдром Дубина-Джонсона; мышечная дистрофия Дюшенна; мышечная дистрофия Беккера; дисфибриногенемия; врожденный дискератоз аутосомно-доминантный и аутосомно-доминантный, 3; врожденный дискератоз, аутосомно-рецессивный, 1, 3, 4 и 5; врожденный дискератоз X-сцепленный; дискинезия, семейная, с миокимией лица; дисплазминогенемия; дистония 2 (скрученность, аутосомно-рецессивная), 3 (скрученность, X-сцепленная), 5 (Dopa-чувствительный тип), 10, 12, 16, 25, 26 (миоклоническая); судороги, доброкачественные семейные детские, 2; ранняя детская эпилептическая энцефалопатия 2, 4, 7, 9, 10, 11, 13 и 14; атипичный синдром Ретта; ранний T-клеточный предшественник острого лимфобластного лейкоза; синдром хрупкости кожи при эктодермальной дисплазии; синдром эктодермальной дисплазии-синдактилии 1; Ectopia lentis, изолированная аутосомно-рецессивная и доминантная; эктродактилия, эктодермальная дисплазия и синдром расщелины губы/неба 3; синдром Элерса-Данлоса тип 7 (аутосомно-рецессивный), классический тип, тип 2 (прогероидный), с дефицитом гидроксилизина, тип 4, вариант типа 4, и вследствие

дефицита тенасцина-X; врожденная мышечная дистрофия типа Эйхсфельда; эндокринная цереброостеодисплазия; синдром повышенных s-колбочек; синдром расширенного вестибулярного акведука; недостаточность энтерокиназы; бородавчатая эпидермодисплазия; простая буллезная эпидермолиза и мышечная дистрофия пояса конечностей, симплекс с пятнистой пигментацией, симплекс с атрезией привратника, симплекс, аутосомно-рецессивный и с атрезией привратника; эпидермолитическая ладонно-подошвенная кератодермия; семейные фебрильные судороги 8; эпилепсия, отсутствие в детстве 2, 12 (идиопатическая генерализованная, предрасположенность) 5 (ночная лобная доля), ночная лобная доля 1 типа, частичная, с переменными очагами, прогрессирующая миоклоническая 3 и X-сцепленная, с различными нарушениями обучаемости и поведенческими расстройствами; эпилептическая энцефалопатия, в детстве, в раннем младенчестве, 1, 19, 23, 25, 30 и 32; эпифизарная дисплазия, множественная, с миопией и кондуктивной глухотой; эпизодическая атаксия 2 типа; эпизодический болевой синдром, семейный, 3; синдром Эпштейна; синдром Фехтнера; эритропоэтическая протопорфирия; устойчивость к эстрогенам; экссудативная витреоретинопатия 6; болезнь Фабри и болезнь Фабри, кардиальный вариант; фактор H, VII, X, v и фактор viii, комбинированный дефицит 2, xiii, субъединица, дефицит; семейный аденоматозный полипоз 1 и 3; семейная амилоидная нефропатия с крапивницей и глухотой; семейная холодовая крапивница; семейная аплазия дермы; семейная доброкачественная пузырчатка; семейный рак груди; восприимчивость к раку груди; остеосаркома; рак поджелудочной железы 3; семейная кардиомиопатия; семейный аутовоспалительный холодовой синдром 2 ; семейный колоректальный рак; семейная экссудативная витреоретинопатия, X-сцепленная; семейная гемиплегическая мигрень 1 и 2 типа; семейная гиперхолестеринемия; семейная гипертрофическая кардиомиопатия 1, 2, 3, 4, 7, 10, 23 и 24; семейная гипокалиемия-гипомагниемия; семейная гипоплазия, гломеруло-кистозная почка; семейная инфантильная миастения; семейная юношеская подагра; семейная средиземноморская лихорадка и Семейная средиземноморская лихорадка, аутосомно-доминантная; семейная порэнцефалия; семейная поздняя кожная порфирия; семейный гемангиоматоз легочных капилляров; семейная почечная глюкозурия; семейная почечная гипоурикемия; семейная рестриктивная кардиомиопатия 1; семейная гиперлипопротеинемия 1 и 3 типа; анемия Фанкони, группа комплементации E, I, N и O; синдром Фанкони-Бикеля; восприимчивость к фавизму; фебрильные судороги, семейные, 11; синдром

Фейнгольда 1; количественный признак гемоглобина плода, локус 1; синдром FG и синдром FG 4; фиброз экстраокулярных мышц, врожденный, 1, 2, 3a (с поражением глазного яблока или без него), 3b; болезнь рыбьего глаза; дистрофия роговицы Флека; синдром плавающей гавани; фокальная эпилепсия с нарушением речи с умственной отсталостью или без нее; фокальный сегментарный гломерулосклероз 5; дефекты фронтальной части мозга; синдром Франка-Тер-Хаара; синдром Борроне ди Рокко Кровато; синдром Фрейзера; опухоль Вильмса 1; синдром Фримена-Шелдона; лобно-метафизарная дисплазия 1 и 3; лобно-височная деменция; лобно-височная деменция и/или боковой амиотрофический склероз 3 и 4; лобно-височная деменция, связанная с хромосомой 3, и лобно-височная деменция, убиквитин-положительная; дефицит фруктозобифосфатазы; синдром Фурмана; дефицит трансаминазы гамма-аминомасляной кислоты; синдром Гамсторпа-Вольфарта; болезнь Гоше 1 типа и подострая нейронопатия; паралич взгляда, семейный горизонтальный, с прогрессирующим сколиозом; генерализованный доминантный дистрофический буллезный эпидермолиз; генерализованная эпилепсия с фебрильными судорогами плюс 3, тип 1, тип 2; эпилептическая энцефалопатия типа Леннокса-Гасто; гигантская аксональная нейропатия; тромбастения Гланцмана; глаукома 1, открытый угол, e, F и G; глаукома 3, первичная врожденная, d; глаукома, врожденная и глаукома, врожденная, колобома; глаукома, первично открытоугольная, ювенильное начало; восприимчивость к глиоме 1; синдром дефицита переносчика глюкозы 1 типа; нарушение транспорта глюкозо-6-фосфата; синдром дефицита GLUT1 2; эпилепсия, генерализованная идиопатическая, подверженность, 12; дефицит глутаматформиминотрансферазы; глутаровая ацидемия IIA и IIB; глутаровая ацидурия, тип 1; дефицит глутатионсинтетазы; болезнь накопления гликогена 0 (мышцы), II (взрослая форма), IXa2, IXc, тип 1A; тип II, тип IV, IV (комбинированный печеночный и миопатический), тип V и тип VI; синдром Гольдмана-Фавра; синдром Гордона; синдром Горлина; последовательность голопрозэнцефалии; голопрозэнцефалия 7; гранулематозная болезнь, хроническая, X-сцепленная, вариант; гранулезно-клеточная опухоль яичника; синдром серых тромбоцитов; синдром Грищелли тип 3; дистрофия роговицы Греноува типа I; задержка роста и умственного развития, нижнечелюстной дизостоз, микроцефалия и волчья пасть; дефицит гормона роста при аномалиях гипофиза; нечувствительность к гормону роста при иммунодефиците; дефицит GTP циклогидролазы I; синдром Хайду-Чейни; синдром матки, кисти, и стопы; нарушение

слуха; гемангиома, инфантильная капиллярная; гематологическое новообразование; гемохроматоз типа 1, 2B и 3; микрососудистые осложнения диабета 7; количественный признак уровня трансферрина в сыворотке, локус 2; болезнь гемоглобина H неделеционная; гемолитическая анемия, несфероцитарная, вследствие дефицита глюкозо-фосфат-изомеразы; гемофагоцитарный лимфогистиоцитоз, семейный, 2; гемофагоцитарный лимфогистиоцитоз, семейный, 3; дефицит кофактора гепарина II; наследственный энтеропатический акродерматит; синдром наследственного рака груди и яичников; расстройство, подобное атаксии-телеангиэктазии; наследственный диффузный рак желудка; наследственная диффузная лейкоэнцефалопатия со сфероидами; наследственные факторы дефицита II, IX, VIII.; наследственная геморрагическая телеангиэктазия 2 типа ; наследственная нечувствительность к боли при ангидрозе; наследственная лимфедема I типа; наследственная моторная и сенсорная нейропатия с атрофией зрительного нерва; наследственная миопатия с ранней дыхательной недостаточностью; наследственная невралгическая амиотрофия; наследственные неполипозные колоректальные новообразования; синдром Линча I и II; наследственный панкреатит; предрасположенность к хроническому панкреатиту; наследственная сенсорная и вегетативная нейропатия типа IIB и IIA; наследственная сидеробластная анемия; синдром Германского-Пудлака 1, 3, 4 и 6; гетеротаксия, висцеральная, 2, 4 и 6, аутосомная; гетеротаксия, висцеральная, X-сцепленная; гетеротопия; гистиоцитарный медуллярный ретикулез; синдром гистиоцитоза-лимфаденопатии плюс; дефицит синтетазы холокарбоксилазы; голопрозэнцефалия 2, 3,7 и 9; синдром Холта-Орама; гомоцистеинемия из-за дефицита MTHFR, дефицита CBS и гомоцистинурии, чувствительная к пиридоксину; гомоцистинурия-мегалобластная анемия из-за нарушения метаболизма кобаламина, тип комплементации cblE; синдром Хауэла-Эванса; синдром Гурлера; синдром Хатчинсона-Гилфорда; гидроцефалия; гипераммонемия, тип III; гиперхолестеринемия и гиперхолестеринемия, аутосомно-рецессивные; гиперэкплексия 2 и наследственная гиперэкплексия; синдром гиперферритинемии катаракты; гиперглицинурия; гипериммуноглобулин D при периодической лихорадке; мевалоновая ацидурия; синдром гипериммуноглобулина E; гиперинсулинемическая гипогликемия семейная 3, 4 и 5; синдром гиперинсулинизма-гипераммониемии; гиперлизинемия; гиперманганеземия с дистонией, полицитемией и циррозом печени; синдром гиперорнитинемии-гипераммониемии-гомоцитруллинурии; гиперпаратиреоз 1 и 2;

гиперпаратиреоз тяжелой степени у новорожденных; гиперфенилаланинемия, с дефицитом bh4, a, из-за частичного дефицита pts, с дефицитом BH4, D и без pku; гиперфосфатазия с синдромом умственной отсталости 2, 3 и 4; гипертрихотическая остеохондродисплазия; гипобеталипопротеинемия, семейная, связанная с apob32; гипокальциемия аутосомно-доминантная 1; гипокальциурическая гиперкальциемия, семейная, 1 и 3 типы; гипохондрогенез; гипохромная микроцитарная анемия с перегрузкой железом; гипогликемия с дефицитом гликогенсинтетазы в печени; гипогонадотропный гипогонадизм 11 с аносмией или без нее; гипогидротическая эктодермальная дисплазия с иммунодефицитом; гипогидротическая X-сцепленная эктодермальная дисплазия; гипокалиемический периодический паралич 1 и 2; гипомагниемия 1, кишечная; гипомагниемия, судороги и умственная отсталость; гипомиелинизирующая лейкодистрофия 7; синдром гипоплазии левых отделов сердца; дефект атриовентрикулярной перегородки и общий атриовентрикулярный переход; гипоспадия 1 и 2, X-сцепленная; гипотироидизм, врожденный, незобный, 1; гипотрихоз 8 и 12; синдром гипотрихоз-лимфедема-телеангиэктазия; I группа крови; буллезный ихтиоз Сименса; ихтиоз отшелушивающий; ихтиозный синдром при недоношенности; идиопатическая кальцификация базальных ганглиев 5; идиопатический фиброзирующий альвеолит, хроническая форма; врожденный дискератоз, аутосомно-доминантный, 2 и 5; идиопатическая гиперкальциемия у младенцев; иммунная дисфункция с инактивацией Т-клеток из-за дефекта поступления кальция 2; иммунодефицит 15, 16, 19, 30, 31C, 38, 40, 8 из-за дефекта в cd3-zeta, с гипер-IgM типа 1 и 2, и X-сцепленный с дефектом магния, инфекцией вируса Эпштейна-Барра и неоплазией; иммунодефицит-центромерная нестабильность-синдром лицевых аномалий 2; миопатия с тельцами включения 2 и 3; миопатия Нонака; детские судороги и пароксизмальный хореоатетоз, семейный; детский кортикальный гиперостоз; детский ганглиозидоз GM1; детская гипофосфатазия; детский нефронофтиз; детский нистагм, X-сцепленный; детский паркинсонизм-дистония; бесплодие, связанное с многохвостыми сперматозоидами и избытком ДНК; резистентность к инсулину; инсулинорезистентный сахарный диабет и черный акантоз; синдром секреторной диареи инсулинозависимого сахарного диабета; интерстициальный нефрит, кариомегалия; задержка внутриутробного развития, метафизарная дисплазия, врожденная гипоплазия надпочечников и генитальные аномалии; дефект йодтирозильной связи; дефицит IRAK4; доминантный тип иридогониодисгенеза и тип

1; накопление железа в головном мозге; ишиопателлярная дисплазия; гиперплазия островковых клеток; изолированный дефицит 17,20-лиазы; изолированный дефицит лутропина; дефицит изовалерил-КоА дегидрогеназы; синдром Янковича-Риверы; синдром Джервелла и Ланге-Нильсена 2 ; синдром Жубера 1, 6, 7, 9/15 (дигенный), 14, 16 и 17 и орофациальный синдром xiv; узловой буллезный эпидермолиз Херлитца; юношеский GM>1<ганглиозидоз; синдром юношеского полипоза; юношеский полипоз/синдром наследственной геморрагической телеангиэктазии; юношеский ретиношизис; синдром Кабуки; синдром Каллмана 1, 2 и 6; задержка полового созревания; болезнь Канзаки; синдром Карак; синдром Картагенера; синдром Кенни-Каффи 2 типа; синдром Кеппена-Любинского; кератоконус 1; фолликулярный кератоз; кератоз ладонно-подошвенный полосатый 1; синдром Киндлера; L-2-гидроксиглутаровая ацидурия; синдром Ларсена, доминантный тип; решеточная дистрофия роговицы III типа; амавроз Лебера; синдром Зеллвегера; нарушения биогенеза пероксисом; спектр синдрома Зеллвегера; врожденный амавроз Лебера 11, 12, 13, 16, 4, 7 и 9; атрофия зрительного нерва Лебера; глухота, вызванная аминогликозидами; глухота, несиндромная нейросенсорная, митохондриальная; некомпакция левого желудочка 5; пороки развития левой-правой оси; болезнь Ли; дефицит митохондриальной короткоцепочечной эноил-КоА гидратазы 1; синдром Ли из-за дефицита митохондриального комплекса I; болезнь Лейнера; дисхондростеоз Лери-Вейля; синдром летальной врожденной контрактуры 6; нарушение адгезии лейкоцитов I и III типа; лейкодистрофия, гипомиелинизирующая, 11 и 6; лейкоэнцефалопатия с атаксией, поражением ствола и спинного мозга и повышением уровня лактата, с исчезновением белого вещества, и прогрессирующая, с недостаточностью яичников; тотальная лейконихия; деменция с тельцами Леви; синдром Лихтенштейна-Норра; синдром Ли-Фраумени 1; синдром Lig4; конечностно-поясная мышечная дистрофия, тип 1B, 2A, 2B, 2D, C1, C5, C9, C14; врожденная мышечная дистрофия-дистрогликанопатия с аномалиями мозга и глаз, тип A14 и B14; комбинированная недостаточность липазы; липидный протеиноз; липодистрофия, семейная парциальная, тип 2 и 3; лиссэнцефалия 1, 2 (X-сцепленная), 3, 6 (с микроцефалией), X-сцепленная; подкорковая ламинарная гетеротопия, X-сцепленная; печеночная недостаточность острая младенческая; синдром Лойса-Дитца 1, 2, 3; синдром удлиненного интервала QT 1, 2, 2/9, 2/5, (дигенный), 3, 5 и 5, приобретенный, восприимчивость; рак легкого; лимфедема, наследственная, id; лимфедема первичная с

миелодисплазией; лимфопролиферативный синдром 1, 1 (Х-сцепленный) и 2; дефицит липазы лизосомальной кислоты; макроцефалия, макросомия, синдром лицевого дисморфизма; макулярная дистрофия, желточная форма, у взрослых; восприимчивость к злокачественной гипертермии 1 типа; злокачественная неходжкинская лимфома; злокачественная меланома; злокачественная опухоль простаты; мандибулоакральный дизостоз; мандибулоакральная дисплазия с липодистрофией типа А или В, атипичная; мандибуло-лицевой дизостоз по типу Тричера-Коллинза, аутосомно-рецессивный; дефицит маннозо-связывающего белка; заболевание мочи типа "кленовый сироп" 1А и 3 типа; синдром Мардена-Уокера; синдром Марфана; синдром Маринеско-Съёргена; синдром Марцольфа; диабет зрелого возраста у молодых людей типа 1, типа 2, типа 11, типа 3 и типа 9; аномалия Мэй-Хегглина; расстройства, связанные с MYH9; синдром Себастьяна; синдром МакКьюна-Олбрайта; соматотропная аденома; опухоль полового канатика и стромы гонад; синдром Кушинга; синдром МакКьюсика-Кауфмана; синдром нейроакантоцитоза Маклеода; синдром Меккеля-Грубера; дефицит ацил-кофермента А-дегидрогеназы со средней длиной цепи; медуллобластома; мегаленцефальная лейкоэнцефалопатия с подкорковыми кистами 1 и 2а; мегалэнцефалия мраморной кожи телеангиэктатическая врожденная; связанный с PIK3CA спектр чрезмерного роста; синдром мегалэнцефалии-полимикрогирии-полидактилии-гидроцефалии 2; мегалобластная анемия, тиамин-чувствительная, с сахарным диабетом и нейросенсорной глухотой; синдромы Мейера-Горлина 1 и 4; синдром Мельника-Нидлза; менингиома; умственная отсталость, Х-сцепленная, 3, 21, 30 и 72; умственная отсталость и микроцефалия с гипоплазией моста и мозжечка; умственная отсталость Х-сцепленная синдромная 5; умственная отсталость, протрузия верхней челюсти и косоглазие ; умственная отсталость, аутосомно-доминантный тип 12, 13, 15, 24, 3, 30, 4, 5, 6 и 9; умственная отсталость, аутосомно-рецессивный 15, 44, 46 и 5; умственная отсталость, стереотипные движения, эпилепсия и/или пороки развития головного мозга; умственная отсталость, синдромальная, типа Класа-Йенсена, Х-сцепленная; умственная отсталость, Х-сцепленная, неспецифическая, синдромная, по типу Хедера и синдромальная, тип wu; врожденная мышечная дистрофия с недостаточностью мерозина; метахроматическая лейкодистрофия ювенильного, позднего инфантильного и взрослого типов; метахроматическая лейкодистрофия; метатрофическая дисплазия; метгемоглобинемия I и 2 типа; дефицит метионинаденозилтрансферазы, аутосомно-доминантный; метилмалоновая ацидемия с

гомоцистинурией,; метилмалоновая ацидурия типа cblB ; метилмалоновая ацидурия из-за дефицита метилмалонил-КоА мутазы; метилмалоновая ацидурия, тип mut (0; микроцефальная остеодиспластическая примордиальная карликовость 2 типа; микроцефалия с хориоретинопатией, лимфедемой или умственной отсталостью или без них; микроцефалия, грыжа пищеводного отверстия диафрагмы и нефротический синдром; микроцефалия; гипоплазия мозолистого тела; спастическая параплегия 50, аутосомно-рецессивная; глобальная задержка развития; гипомиелинизация ЦНС; атрофия головного мозга; микроцефалия, нормальный интеллект и иммунодефицит; синдром микроцефалии-капиллярной мальформации; микроцитарная анемия; синдромный микрофтальм 5, 7 и 9; микрофтальм, изолированный 3, 5, 6, 8 и с колобомой 6; микросферофакия; мигрень, семейная базилярная; синдром Миллера; миопатия миопатия с наружной офтальмоплегией; миопатия врожденная с стержнями; синдром Митчелла-Райли; митохондриальная недостаточность 3-гидрокси-3-метилглутарил-КоА-синтазы; дефицит митохондриального комплекса I, II, III, III (ядерный тип 2, 4 или 8); синдром истощения митохондриальной ДНК 11, 12 (кардиомиопатический тип), 2, 4B (тип MNGIE), 8B (тип MNGIE); синдром истощения митохондриальной ДНК 3 и 7, гепатоцеребральные типы и 13 (энцефаломиопатический тип); дефицит митохондриального переносчика фосфата и переносчика пирувата; дефицит митохондриального трифункционального белка; дефицит длинноцепочечной 3-гидроксиацил-КоА дегидрогеназы; мышечная дистрофия Миёси 1; миопатия, дистальная, с началом в передней большеберцовой кости; синдром Мора-Транебьерга; дефицит кофактора молибдена, группа комплементации A; синдром Мовата-Вильсона; муколипидоз III гамма; мукополисахаридоз типа VI, типа VI (тяжелый) и типа VII; мукополисахаридоз, MPS-I-H / S, MPS-II, MPS-III-A, MPS-III-B, MPS-III-C, MPS-IV-A, MPS-IV-B; пигментный ретинит 73; ганглиозидоз GM1 типа 1 (с поражением сердца) 3; многоцентровая остеолизная нефропатия; мультицентрический остеолиз, нодулез и артропатия; множественные врожденные аномалии; дефект межпредсердной перегородки 2; синдром множественных врожденных аномалий-гипотонии-судорог 3; множественные кожные и венозные пороки слизистых оболочек; множественная эндокринная неоплазия 1 и 4 типа; множественная эпифизарная дисплазия 5 или доминантная; множественные атрезии желудочно-кишечного тракта; синдром множественного птеригиума типа Эскобара; множественный дефицит сульфатазы; синдром

множественных синостозов 3; дефицит гуаниноксидазы AMP в мышцах; заболевание мышц глаз и головного мозга; мышечная дистрофия, врожденная, мегакониального типа; миастения, семейная инфантильная, 1; миастенический синдром, врожденный, 11, связанный с дефицитом рецепторов ацетилхолина; миастенический синдром, врожденный, 17, 2A (медленный канал), 4B (быстрый канал) и без тубулярных агрегатов; дефицит миелопероксидазы; MYH-ассоциированный полипоз; карцинома эндометрия; инфаркт миокарда 1; миоклоническая дистония; миоклонико-атоническая эпилепсия; миоклонус с эпилепсией с рваными красными волокнами; миофибриллярная миопатия 1 и связанная с ZASP; миоглобинурия, острая рецидивирующая, аутосомно-рецессивная; синдром мионевральной желудочно-кишечной энцефалопатии; мозжечковая атаксия младенческого возраста с прогрессирующей внешней офтальмоплегией; синдром истощения митохондриальной ДНК 4B, тип MNGIE; миопатия, центральноядерная, 1, врожденная, с избытком мышечных веретен, дистальная, 1, лактоацидоз и сидеробластная анемия 1, митохондриально-прогрессирующая анемия с врожденной катарактой, потерей слуха и задержкой развития и тубулярным агрегатом, 2; близорукость 6; миосклероз аутосомно-рецессивный; миотония врожденная; врожденная миотония, аутосомно-доминантная и рецессивная формы; синдром ногтей-надколенника; синдром Нэнси-Хорана; нанофтальм 2; нейрогепатопатия Навахо; немалиновая миопатия 3 и 9; неонатальная гипотония; интеллектуальная инвалидность; судороги; задержка речи и языкового развития; умственная отсталость, аутосомно-доминантная 31; неонатальный внутрипеченочный холестаз, вызванный дефицитом цитрина; нефрогенный несахарный диабет, Нефрогенный несахарный диабет, X-сцепленный; нефролитиаз/остеопороз, гипофосфатемический, 2 ; нефронофтис 13, 15 и 4; бесплодие; церебелло-окуло-почечный синдром (нефронофтиз, глазодвигательная апраксия и аномалии мозжечка); нефротический синдром, тип 3, тип 5, с глазными аномалиями или без них, тип 7 и тип 9; синдром прогерии Нестора-Гильермо; синдром Ней-Лаксова 1; нейродегенерация с накоплением железа в головном мозге 4 и 6; нейроферритинопатия; нейрофиброматоз 1 и 2 типа; нейрофибросаркома; нейрогипофизарный несахарный диабет; невропатия, наследственная сенсорная, тип IC; дефект транспорта нейтральной 1 аминокислоты; болезнь накопления нейтральных липидов с миопатией; синдром нейтрофильного иммунодефицита; синдром Николаидеса-Барайцера; болезнь Ниманна-Пика типа C1, C2, типа A и типа C1,

взрослая форма; некетотическая гиперглицинемия; синдром Нунана 1 и 4, синдром LEOPARD 1; заболевание, подобное синдрому Нунана, с ювенильным миеломоноцитарным лейкозом или без него; нормокалиемический периодический паралич, чувствительный к калию; болезнь Норума; эпилепсия, потеря слуха и синдром задержки умственного развития; умственная отсталость, Х-сцепленная 102 и синдромальная 13; ожирение; глазной альбинизм, тип I; глазокожный альбинизм типа 1B, типа 3 и типа 4; окулодентодигитальная дисплазия; одонтогипофосфатазия; одонтотрихомелический синдром; болезнь Огучи; синдром олигодонтии-колоректального рака; синдром Опица G/BBB; атрофия зрительного нерва 9; орально-лицевой-дигитальный синдром; дефицит орнитинаминотрансферазы; орофациальная расщелина 11 и 7, синдром расщелины губы / неба и эктодермальной дисплазии; синдром Орставика-Линдеманна-Сольберга; остеоартроз с легкой хондродисплазией; рассекающий остеохондрит; несовершенный остеогенез тип 12, тип 5, тип 7, тип 8, тип I, тип III, с нормальными склерами, доминантная форма, рецессивный перинатальный летальный; полосчатая остеопатия с краниальный склерозом; остеопетроз аутосомно-доминантный тип 1 и 2, рецессивный 4, рецессивный 1, рецессивный 6; остеопороз с псевдоглиомой; ото-небно-дигитальный синдром, типы I и II; дисгенезия яичников 1; овариолейкодистрофия; врожденная пахионихия 4 и 2 типа; болезнь Педжета костей, семейная; синдром Паллистера-Холла; ладонно-подошвенная кератодермия, неэпидермолитическая, очаговая или диффузная; агенезия поджелудочной железы и врожденный порок сердца; синдром Папийона-Лефевре; параганглиомы 3; врожденная парамиотония фон Эйленбурга; карцинома паращитовидной железы; болезнь Паркинсона 14, 15, 19 (ювенильное начало), 2, 20 (раннее начало), 6 (аутосомно-рецессивное раннее начало) и 9; частичный альбинизм; частичный дефицит гипоксантин-гуанинфосфорибозилтрансферазы; узорчатая дистрофия пигментного эпителия сетчатки; PC-K6a; болезнь Пелицея-Мерцбахера; синдром Пендреда; периферическая демиелинизирующая нейропатия, центральная дисмиелинизация; болезнь Гиршпрунга; постоянный неонатальный сахарный диабет; сахарный диабет, перманентный неонатальный, с неврологическими особенностями; неонатальный инсулинозависимый сахарный диабет; диабет зрелости у молодых, тип 2; нарушение биогенеза пероксисом 14B, 2A, 4A, 5B, 6A, 7A и 7B; синдром Перро 4; синдром Перри; стойкая гиперинсулинемическая гипогликемия младенческого возраста; семейный гиперинсулинизм; фенотипы; фенилкетонурия; феохромоцитома; наследственные

синдромы параганглиомы-феохромоцитомы; параганглиомы 1; карциноидная опухоль кишечника; синдром Каудена 3; дефицит фосфоглицератдегидрогеназы; дефицит фосфоглицераткиназы 1; фоточувствительная трихотиодистрофия; болезнь накопления фитановой кислоты; болезнь Пика; синдром Пирсона; пигментная дистрофия сетчатки; пигментная узловая болезнь коры надпочечников, первичная, 1; пиломатриксома; синдром Питта-Хопкинса; гипофизарный гиперкортизолизм; дефицит гормона гипофиза, комбинированный 1, 2, 3 и 4; дефицит ингибитора активатора плазминогена типа 1; дефицит плазминогена, тип I; расстройство свертываемости тромбоцитов 15 и 8; пойкилодермия, наследственное фиброзирование с контрактурами сухожилий, миопатией и фиброзом легких; поликистоз почек 2, взрослый тип и детский тип; поликистозная липомембранозная остеодисплазия со склерозирующей лейкоэнцефалопатией; миопатия с полиглюкозановыми тельцами 1 с иммунодефицитом или без него ; полимикрогирия, асимметричная, двусторонняя лобно-теменная; полинейропатия, потеря слуха, атаксия, пигментный ретинит и катаракта; понтоцеребеллярная гипоплазия 4 типа; синдром подколенного птеригиума; порэнцефалия 2; порокератоз 8, диссеминированный поверхностный актинический тип; дефицит порфобилиногенсинтазы; поздняя кожная порфирия; атаксия задней колонны с пигментным ретинитом; задняя полярная катаракта 2 типа; синдром Прадера-Вилли; преждевременная недостаточность яичников 4, 5, 7 и 9; первичная аутосомно-рецессивная микроцефалия 10, 2, 3 и 5; первичная цилиарная дискинезия 24; первичная дилатационная кардиомиопатия; некомпакция левого желудочка 6; 4, некомпакция левого желудочка 10; пароксизмальная фибрилляция предсердий; первичная гипероксалурия, тип I, тип и тип III; первичная гипертрофическая остеоартропатия, аутосомно-рецессивная 2; первичная гипомагниемия; первичная открытоугольная глаукома с юношеским дебютом 1; первичная легочная гипертензия; синдром первоцвета; прогрессирующая семейная блокада сердца типа 1B; прогрессирующий семейный внутрипеченочный холестаз 2 и 3; прогрессирующий внутрипеченочный холестаз; прогрессирующая миоклоническая эпилепсия с атаксией; прогрессирующая ревматоидная дисплазия; прогрессирующая склерозирующая полиодистрофия; дефицит пролидазы; дефицит пролиндегидрогеназы; шизофрения 4; дефицит пропердина, X-сцепленный; пропионовая академия; дефицит пропротеинконвертазы 1/3; рак простаты, наследственный, 2; дефект протана; протеинурия; финский врожденный нефротический синдром; синдром Протея;

аденокарцинома груди; синдром псевдоахондропластической спондилоэпифизарной дисплазии; псевдогипоальдостеронизм, аутосомно-доминантный и рецессивный тип 1 и тип 2; псевдогипопаратиреоз типа 1A, псевдопсевдогипопаратиреоз; псевдонеонатальная адренолейкодистрофия; псевдопервичный гиперальдостеронизм; эластическая псевдоксантома; обобщенный кальциноз артерий младенческого возраста 2; псевдоксантомное эластикоподобное расстройство с недостаточностью множественных факторов свертывания крови; восприимчивость к псориазу 2; синдром опухоли гамартомы PTEN; легочная артериальная гипертензия, связанная с наследственной геморрагической телеангиэктазией; фиброз легких и/или недостаточность костного мозга, связанные с теломерами, 1 и 3; легочная гипертензия, первичная, 1, с наследственной геморрагической телеангиэктазией; дефицит пуриннуклеозидфосфорилазы; дефицит пируваткарбоксилазы; дефицит пируватдегидрогеназы E1-альфа; дефицит пируваткиназы эритроцитов; синдром Рейна; разопатия; рецессивный дистрофический буллезный эпидермолиз; поражение ногтей несиндромное врожденное, 8; синдром Рейфенштейна; почечная адисплазия; почечный дефект транспорта карнитина; синдром почечной колобомы; почечная дисплазия; дисплазия почек, пигментная дистрофия сетчатки, мозжечковая атаксия и дисплазия скелета; почечный канальцевый ацидоз, дистальный, аутосомно-рецессивный, с поздней нейросенсорной тугоухостью или с гемолитической анемией; почечный канальцевый ацидоз, проксимальный, с глазными аномалиями и умственной отсталостью; дистрофия конуса сетчатки 3B; пигментный ретинит; пигментный ретинит 10, 11, 12, 14, 15, 17 и 19; пигментный ретинит 2, 20, 25, 35, 36, 38, 39, 4, 40, 43, 45, 48, 66, 7, 70, 72; ретинобластома; расстройство Ретта; синдром предрасположенности к рабдоидным опухолям 2; регматогенная отслойка сетчатки по аутосомно-доминантному типу; точечная ризомелическая хондродисплазия 2 и 3 типа; синдром фокомелии Робертса-SC; синдром Робиноу-Сорауфа; синдром Робиноу, аутосомно-рецессивный, аутосомно-рецессивный, с брахи-син-полидактилией; синдром Ротмунда-Томсона; синдром Рападилино; митохондриальное заболевание, связанное с RRM2B; синдром Рубинштейна-Тайби; болезнь Салла; болезнь Сандхоффа, взрослые и младенческие типы; саркоидоз, раннее начало; синдром Блау; болезнь Шиндлера, тип 1; шизэнцефалия; шизофрения 15; дисплазия Шнекенбекена; шванноматоз 2; синдром Шварца-Джампеля тип 1; склерорнеа аутосомно-рецессивная; склеростеоз; вторичный гипотиреоз ; синдром Сегавы, аутосомно-рецессивный;

синдром Сеньора-Локена 4 и 5,; сенсорная атаксическая нейропатия, дизартрия и офтальмопарез; дефицит сепиаптеринредуктазы; синдром SeSAME; тяжелый комбинированный иммунодефицит из-за дефицита ADA, с микроцефалией, задержкой роста и чувствительностью к ионизирующему излучению, атипичный, аутосомно-рецессивный, Т-клеточно-отрицательный, В-клеточный, NK-отрицательный или NK-положительный; тяжелая врожденная нейтропения; тяжелая врожденная нейтропения 3, аутосомно-рецессивная или доминантная; тяжелая врожденная нейтропения и 6, аутосомно-рецессивная; тяжелая миоклоническая эпилепсия в младенчестве; генерализованная эпилепсия с фебрильными судорогами плюс, типы 1 и 2; тяжелая Х-сцепленная миотубулярная миопатия; синдром короткого интервала QT 3; низкий рост с неспецифическими аномалиями скелета; низкий рост, атрезия слухового прохода, гипоплазия нижней челюсти, аномалии скелета; низкий рост, ониходисплазия, дисморфизм лица и гипотрихоз; врожденная карликовость; дисплазия грудного отдела коротких ребер 11 или 3 с полидактилией или без нее; сиалидоз I и II типа; синдром серебряной спастической параплегии; замедленная скорость нервной проводимости, аутосомно-доминантная; синдром Смита-Лемли-Опица; синдром Снайдера-Робинсона; соматотропная аденома; пролактинома; семейная предрасположенность к аденоме гипофиза; синдром Сотоса 1 или 2; спастическая атаксия 5, аутосомно-рецессивная, тип Шарлевуа-Сагеней, 1,10 или 11, аутосомно-рецессивная; боковой амиотрофический склероз 5 типа; спастическая параплегия 15, 2, 3, 35, 39, 4, аутосомно-доминантная, 55, аутосомно-рецессивная и 5А; нарушение синтеза желчных кислот, врожденное, 3; сперматогенная недостаточность 11, 3 и 8; сфероцитоз 4 и 5 типов; миопатия сфероидного тела; спинальная мышечная атрофия, с преобладанием нижних конечностей 2, аутосомно-доминантная; спинальная мышечная атрофия II типа; спиноцеребеллярная атаксия 14, 21, 35, 40 и 6; аутосомно-рецессивная спиноцеребеллярная атаксия 1 и 16; гипоплазия селезенки; синдром спондилокарпотарзального синостоза; спондилохейродисплазия, подобная синдрому Элерса-Данлоса, с нарушением иммунной регуляции, тип Аггрекана, с врожденными вывихами суставов, тип короткой конечности-руки, тип Седагата, с дистрофией палочек и колбочек, тип Козловски; парастремматическая карликовость; болезнь Штаргардта 1; дистрофия колбочек и палочек 3; синдром Стиклера тип 1; дисплазия Кнеста; синдром Стиклера, типы 1 (несиндромный глазной) и 4; васкулопатия, связанная с укусом, младенческое начало; синдром Сторморкена; синдром Стерджа-

Вебера, пороки развития капилляров, врожденные, 1; дефицит сукцинил-КоА-ацетоацетаттрансферазы; дефицит сахаразы-изомальтазы; синдром внезапной детской смерти; дефицит сульфитоксидазы, изолированный; суправальварный стеноз аорты; дисфункция метаболизма ПАВ, легочная, 2 и 3; симфалангизм проксимальный, 1b; синдактилия типа Ченани-Ленца; синдактилия 3-го типа; синдромная Х-сцепленная умственная отсталость 16; эквиноварусная косолапость; болезнь Танжера; TARP-синдром; болезнь Тея-Сакса, вариант B1, Gm2-ганглиозидоз (у взрослых), Gm2-ганглиозидоз (у взрослых); синдром Темтами; синдром Тенорио; терминальная костная дисплазия; недостаток тестостерон-17-бета-дегидрогеназы; тетраамелия, аутосомно-рецессивная; тетралогия Фалло; синдром гипоплазии левых отделов сердца 2; артериальный ствол; пороки развития сердца и магистральных сосудов; дефект межжелудочковой перегородки 1; дистрофия роговицы Тиля-Бенке; аневризмы грудной аорты и расслоение аорты; марфановский синдром; синдром 3M 2; тромбоцитопения, дисфункция тромбоцитов, гемолиз и несбалансированный синтез глобина; Х-сцепленная тромбоцитопения; тромбофилия, наследственная, из-за дефицита протеина С, аутосомно-доминантная и рецессивная; агенезия щитовидной железы; рак щитовидной железы, фолликулярный; метаболизм гормонов щитовидной железы, анормальный; резистентность к гормонам щитовидной железы, генерализованная, аутосомно-доминантная; тиреотоксический периодический паралич и Тиреотоксический периодический паралич 2; общая резистентность к высвобождающему тиротропин гормону; синдром Тимоти; синдром периодической лихорадки, связанный с рецепторами TNF (TRAPS); агенезия зубов, выборочная, 3 и 4; пируэтная тахикардия; синдром Таунса-Брокса брангиоретенального типа; транзиторный буллезный дермолиз новорожденного; синдром Тричера-Коллинза 1; трихомегалия с умственной отсталостью, карликовостью и пигментной дегенерацией сетчатки; трихоринофаланговая дисплазия I типа; трихоринофалангеальный синдром 3 типа; триметиламинурия; синдром туберозного склероза; лимфангиомиоматоз ; туберозный склероз 1 и 2; тирозиназа-отрицательный окулокожный альбинизм; тирозиназа-положительный окулокожный альбинизм; тирозинемия I типа; дефицит UDP-глюкозо-4-эпимеразы; врожденная мышечная дистрофия Ульриха; отсутствие локтевой и малоберцовой костей при тяжелой недостаточности конечностей; синдром Апшоу-Шульман; дефицит уроканатгидратазы; синдром Ашера, типы 1, 1B, 1D, 1G, 2A, 2C и 2D; пигментный ретинит 39; уФ-чувствительный синдром; синдром Ван дер

Вуде; синдром Ван Мальдергема 2; синдром лимфангиэктазии-лимфедемы Хеннекама 2; вариегатная порфирия; вентрикуломегалия при кистозной болезни почек; синдром Верхей; дефицит ацил-КоА-дегидрогеназы с очень длинной цепью; пузырно-мочеточниковый рефлюкс 8; висцеральная гетеротаксия 5, аутосомная; висцеральная миопатия; витамин D-зависимый рахит 1 и 2 типа; желточная дистрофия; болезнь фон Виллебранда типа 2M и типа 3; синдром Ваарденбурга типа 1, 4C и 2E (с неврологическим поражением); синдром Клейна-Ваарденберга; врожденная мышечная дистрофия Уокера-Варбурга; микросиндром Варбурга 2 и 4; бородавки, гипогаммаглобулинемия, инфекции и миелокатексис; синдром Уивера; синдром Вейля-Марчезани 1 и 3; синдром Вейля-Марчезани; синдром Вайссенбахера-Цвеймюллера; болезнь Верднига-Гофмана; болезнь Шарко-Мари-Тута; синдром Вернера; заболевания, связанные с WFS1; синдром Видеманна-Штайнера; болезнь Вильсона; вольфрамоподобный синдром, аутосомно-доминантный; болезнь Ворта; болезнь Ван Бухема 2 типа; Xeroderma pigmentosum, группа комплементации b, группа D, группа E и группа G; X-сцепленная агаммаглобулинемия; X-сцепленная наследственная моторная и сенсорная нейропатия; X-сцепленный ихтиоз с дефицитом стерилсульфатазы; X-сцепленная перивентрикулярная гетеротопия; ото-небно-дигитальный синдром, тип I; X-сцепленный тяжелый комбинированный иммунодефицит; синдром Циммермана-Лабанда и синдром Циммермана-Лабанда 2; и зонулярная пылевидная катаракта 3.

[1048] Нуклеотидная последовательность-мишень может включать последовательность-мишень (например точечную мутацию), ассоциированную с заболеванием, расстройством, или состоянием. Последовательность-мишень может включать точечную мутацию T на C (или A на G), ассоциированную с заболеванием, расстройством, или состоянием, и при этом дезаминирование мутантного основания C приводит к опосредованной репарацией несовпадения коррекции в последовательности, которая не ассоциирована с заболеванием, расстройством, или состоянием. Последовательность-мишень может включать точечную мутацию G на A (или C на T), ассоциированную с заболеванием, расстройством, или состоянием, и при этом дезаминирование мутантного основания A приводит к опосредованной репарацией несовпадения коррекции в последовательности, которая не ассоциирована с заболеванием, расстройством, или состоянием. Последовательность-мишень может кодировать белок, и при этом точечная мутация располагается в кодоне и приводит к

изменению аминокислоты, кодируемой мутантным кодоном, по сравнению с кодоном дикого типа. Последовательность-мишень также может быть расположена в участке сплайсинга, при этом точечная мутация приводит к изменению в сплайсинге иРНК транскрипта по сравнению с транскриптом дикого типа. Кроме того, мишень может быть расположена в некодирующей последовательности гена, например в промоторе, при этом точечная мутация приводит к повышенной или пониженной экспрессии гена.

[1049] Таким образом, в некоторых аспектах дезаминирование мутантного С приводит к изменению аминокислоты, кодируемой мутантным кодоном, что в некоторых случаях может приводить к экспрессии аминокислоты дикого типа. В других аспектах дезаминирование мутантного А приводит к изменению аминокислоты, кодируемой мутантным кодоном, что в некоторых случаях может приводить к экспрессии аминокислоты дикого типа.

[1050] Описанные здесь способы, включающие контактирование клетки с композицией или rAAV частицей, могут осуществляться *in vitro, ex vivo,* или *in vivo*. В определенных вариантах осуществления этап контактирования осуществляется в субъекте. В определенных вариантах осуществления у субъекта было диагностировано заболевание, расстройство, или состояние.

[1051] В некоторых вариантах осуществления описанные здесь способы включают контактирование клетки млекопитающего с композицией или rAAV частицей. В определенных вариантах осуществления способы включают контактирования клетки сетчатки, кортикальной клетки или церебеллярной клетки.

[1052] Разделенный белок Cas9 или разделенный праймерный редактор, доставляемый с помощью описанных здесь способов, предпочтительно имеет сравнимую активность по сравнению с исходным белком Cas9 или праймерным редактором (то есть неразделенным белком, доставляемым в клетку или экспрессируемым в клетке в целом). Например, разделенный белок Cas9 или разделенный праймерный редактор сохраняет, по меньшей мере, 50% (например, по меньшей мере, 50%, по меньшей мере, 60%, по меньшей мере, 70%, по меньшей мере, 80%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 98%, по меньшей мере, 99%, или 100%) активности исходного белка Cas9 или праймерного редактора. В некоторых вариантах осуществления разделенный белок Cas9 или разделенный праймерный редактор является более активным (например, 2-кратно, 5- кратно, 10- кратно, 100- кратно, 1000-кратно, или более), чем исходный белок Cas9 или праймерный редактор.

[1053] Описанные здесь композиции могут быть введены нуждающемуся в этом субъекту в терапевтически эффективном количестве для лечения и/или предотвращения заболевания или расстройства, которым страдает субъект. Любое заболевание или расстройство, которое может быть подвергнуто лечению и/или предотвращено с помощью технологии геномного редактирования на основе CRISPR/Cas9, может быть подвергнуто лечению с помощью описанного здесь разделенного белка Cas9 или разделенного праймерного редактора. Следует понимать, что если нуклеотидные последовательности, кодирующие разделенный белок Cas9 или праймерный редактор, далее не кодируют нРНК, то может быть введен отдельный вектор нуклеиновой кислоты, кодирующей нРНК, наряду с описанными здесь композициями.

[1054] Примеры подходящих заболеваний, расстройств или состояний включают, без ограничения, заболевание или расстройство, выбранное из группы, состоящей из: муковисцидоза, фенилкетонурии, эпидермолитического гиперкератоза (ЕНК), хронической обструктивной болезни легких (ХОБЛ), болезни Шарко-Мари-Тута типа 4J, нейробластомы (NB), болезни фон Виллебранда (vWD), врожденной миотонии, наследственного амилоидоза почек, дилатационной кардиомиопатии, наследственной лимфедемы, семейной болезни Альцгеймера, прионной болезни, хронического детского неврологического кожно-суставного синдрома (CINCA), врожденной глухоты, болезни Ниманна-Пика типа С (NPC), и миопатии, связанной с десмином (DRM). В определенных вариантах осуществления заболевание или состояние представляет собой болезнь Ниманна-Пика типа С (NPC).

[1055] В некоторых вариантах осуществления заболевание, расстройство, или состояние ассоциировано с точечной мутацией в гене *NPC*, гене *DNMT1*, гене *PCSK9*, или гене *TMC1*. В определенных вариантах осуществления точечная мутация представляет собой мутацию T3182C в *NPC*, которая приводит к аминокислотному замещению I1061T.

[1056] В определенных вариантах осуществления точечная мутация представляет собой мутацию A545G в *TMC1*, которая приводит к аминокислотному замещению Y182C. *TMC1* кодирует белок, который образует механочувствительные ионные каналы в сенсорных волосковых клетках внутреннего уха и необходим для нормального функционирования слуха. Аминокислотное замещение Y182C ассоциировано с врожденной глухотой.

[1057] В некоторых вариантах осуществления заболевание, расстройство или состояние ассоциировано с точечной мутацией, которая приводит к образованию стоп-кодона, например преждевременного стоп-кодона в кодирующем участке гена.

[1058] Дополнительные примеры заболеваний, расстройств и состояний включают муковисцидоз (см., например, Schwank *et al.*, Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. Cell stem cell. 2013; 13: 653-658; и Wu et. al. *Cell stem cell.* **2013**; 13: 653-658; и Wu et. al., Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell stem cell.* **2013;** 13: 659-662, при этом ни в одном из них не используется дезаминазный слитый белок для корректирования генетического дефекта); фенилкетонурию – например мутацию фенилаланин на серин в положении 835 (мышь) или 240 (человек) или гомологичном остатке в гене фенилаланингидроксилазы (мутация T>C) – см., например, McDonald *et al.*, *Genomics.* **1997;** 39:402-405; синдром Бернара-Сулье (BSS) – например мутацию фенилаланин на серин в положении 55 или гомологичном остатке, или цистеин на аргинин в положении 24 или гомологичном остатке в гликопротеине мембраны тромбоцитов IX (мутация T>C) – см., например, Noris *et al.*, *British Journal of Haematology.* **1997;** 97: 312-320, и Ali *et al.*, *Hematol.* **2014;** 93: 381-384; эпидермолитический гиперкератоз (EHK) – например мутацию лейцин на пролин в положении 160 или 161 (если считать начальный метионин) или гомологичном остатке в кератине 1 (мутация T>C) – см., например, Chipev *et al.*, *Cell.* **1992;** 70: 821-828, также см. кат. # P04264 в базе данных UNIPROT на www[dot]uniprot[dot]org; хроническую обструктивную болезнь легких (ХОБЛ) – например мутацию лейцин на пролин в положении 54 или 55 (если считать начальный метионин) или гомологичном остатке в процессированной форме α_1-антитрипсина или остатке 78 в непроцессированной форме или гомологичном остатке (мутация T>C) – см., например, Poller *et al.*, *Genomics.* **1993;** 17: 740-743, см. кат. # P01011 в базе данных UNIPROT; болезнь Шарко-Мари-Тута типа 4J – например мутацию изолейцин на треонин в положении 41 или гомологичном остатке в FIG4 (мутация T>C) – см., например, Lenk *et al.*, PLoS Genetics. 2011; 7: e1002104; нейробластому (NB) – например мутацию в положении 197 или гомологичном остатке в каспазе-9 (мутация T>C) – см., например, Kundu *et al.*, *3 Biotech.* **2013, 3:225-234**; болезнь фон Виллебранда (vWD) – например мутацию цистеин на аргинин в положении 509 или гомологичном остатке в процессированной форме фактора Виллебранда, или в положении 1272 или гомологичном остатке в непроцессированной

форме фактора Виллебранда (мутация T>C) – см., например, Lavergne *et al.*, *Br. J. Haematol.* **1992**, также см. кат. # P04275 в базе данных UNIPROT**;** 82: 66-72; врожденную миотонию – например мутацию цистеин на аргинин в положении 277 или гомологичном остатке в гене мышечного хлоридного канала CLCN1 (мутация T>C mutation) – см., например, Weinberger *et al.*, *The J. of Physiology.* **2012;** 590: 3449-3464; наследственный амилоидоз почек – например мутацию стоп-кодон на аргинин в положении 78 или гомологичном остатке в процессированной форме аполипопротеина AII или в положении 101 или гомологичном остатке в непроцессированной форме (мутация T>C) – см., например, Yazaki *et al.*, *Kidney Int.* **2003;** 64: 11-16; дилатационную кардиомиопатию (DCM) – например мутацию триптофан на аргинин в положении 148 или гомологичном остатке в гене FOXD4 (мутация T>C), см., например, Minoretti et. al., *Int. J. of Mol. Med.* **2007;** 19: 369-372; наследственную лимфедему – например мутацию гистидин на аргинин в положении 1035 или гомологичном остатке в VEGFR3 тирозинкиназе (мутация A>G), см., например, Irrthum *et al.*, *Am. J. Hum. Genet.* **2000;** 67: 295-301; фамильную болезнь Альцгеймера – например мутацию изолейцин на валин в положении 143 или гомологичном остатке в презенилине 1 (мутация A>G), см., например, Gallo et. al., *J. Alzheimer's disease.* **2011;** 25: 425-431; прионную болезнь – например мутацию метионин на валин в положении 129 или гомологичном остатке в прионном белке (мутация A>G) – см., например, Lewis et. al., *J. of General Virology.* **2006;** 87: 2443-2449; хронический детский неврологический кожно-суставной синдром (CINCA) – например мутацию тирозин на цистеин в положении 570 или гомологичном остатке в криопирине (мутация A>G) – см., например, Fujisawa et. al. *Blood.* **2007;** 109: 2903-2911; и связанную с десмином миопатию (DRM) – например мутацию аргинин на глицин в положении 120 или гомологичном остатке в кристаллине αβ (мутация A>G mutation) – см., например, Kumar *et al.*, *J. Biol. Chem.* **1999;** 274: 24137-24141. Полное содержание всех ссылок и записей в базах данных приведено здесь в качестве ссылки.

Заболевание, связанное с экспансией тринуклеотидных повторов

[1059] Экспансия тринуклеотидных повторов связана с рядом заболеваний человека, включая болезнь Хантингтона, синдром ломкой X-хромосомы и атаксию Фридрейха. Наиболее часто встречающийся тринуклеотидный повтор содержит триплеты CAG, хотя также встречаются триплеты GAA (атаксия Фридрейха), и триплеты CGG (синдром ломкой X-хромосомы). Унаследование предрасположенности к экспансии

или приобретение уже экспансированного родительского аллеля повышает вероятность возникновения заболевания. Патогенные экспансии тринуклеотидных повторов гипотетически могут быть скорректированы с помощью праймерного редактирования.

[1060] Участок выше участка повтора может быть никирован с помощью РНК-направляемой нуклеазы и затем использован для праймирования синтеза новой цепи ДНК, которая содержит здоровое количество повторов (которое зависит от конкретного гена и заболевания) в соответствии с общим механизмом, показанным на Фиг.1G или Фиг.22. После последовательности повтора добавляется короткий фрагмент гомологии, который соответствует последовательности, примыкающей к другому концу повтора (красная цепь). Внесение новой синтезированной цепи с помощью TPRT системы и последующее замещение эндогенной ДНК новым синтезированным флэпом ведет к сокращению повторяющегося аллеля. Термин "сокращенный" обозначает укорачивание длины участка нуклеотидного повтора, что тем самым приводит к репарации участка тринуклеотидного повтора.

[1061] Описанная здесь система праймерного редактирования (PE) может применяться для сокращения мутаций в виде тринуклеотидных повторов (или "заболеваний, связанных с экспансией тринуклеотидных повторов") для лечения состояний, таких как болезнь Хантингтона и другие заболевания, связанные с тринуклеотидными повторами). Заболевания, связанные с экспансией тринуклеотидных повторов, представляют собой комплексные, прогрессирующие заболевания, которые затрагивают связанные с развитием нейробиологические процессы и зачастую влияют на когнитивные способности, а также сенсорно-моторные функции. Данные заболевания демонстрируют наличие генетической связи (то есть повышение тяжести с каждым поколением). Экспансии или сокращения ДНК обычно происходят при мейозе (то есть во время гаметогенеза или при раннем эмбриональном развитии) и зачастую связаны с полом, что свидетельствует о том, что некоторые гены подвергаются экспансии только при наследовании от матери, а другие – только от отца. У человека связанные с тринуклеотидными повторами заболевания могут вызывать молчание генов либо на транскрипционном, или трансляционном уровне, что по существу блокирует функцию гена. В качестве альтернативы, связанные с тринуклеотидными повторами заболевания могут вызывать образование измененных белков, которые образуются за счет крупных аминокислотных последовательностей с повторами,

которые нарушают или изменяют функцию белка, зачастую доминантно-отрицательным образом (например при полиглутаминовых заболеваниях).

[1062] Без связи с теорией, экспансия триплетов вызывается проскальзыванием при репликации ДНК или при репаративном синтезе ДНК. Поскольку тандемные повторы имеют идентичную последовательность относительно друг друга, спаривание оснований между двумя цепями ДНК может происходить в нескольких точках в рамках последовательности. Это может приводить к образованию "выступающих в виде петель" структур при репликации ДНК или репаративном синтезе ДНК. Это может приводить к повторяющемуся копированию повторенной последовательности, что увеличивает количество повторов. Были предложены дополнительные механизмы, связанные с промежуточными продуктами в виде РНК:ДНК гибридов. Праймерное редактирование может применяться для снижения или устранения данных участков экспансии триплетов за счет делеции одного или нескольких триплетов в виде нежелательных повторяющихся кодонов. В рамках варианта осуществления, соответствующего данному варианту использования, на Фиг.23 приведена схема конструкции PEgRNA для сокращения или уменьшения последовательностей тринуклеотидных повторов с помощью праймерного редактирования.

[1063] Праймерное редактирование может быть реализовано для сокращения участков экспансии триплетов за счет никирования участка выше участка в виде триплетного повтора с помощью праймерного редактора, включающего PEgRNA, которая подходящим образом нацелена на участок разрыва. Затем праймерный редактор обеспечивает синтез новой цепи ДНК (оцДНК-флэпа) на основе PEgRNA в качестве матрицы (то есть матрицы для ее редактирования), которая кодирует здоровое количество триплетных повторов (которое зависит от конкретного гена и заболевания). Новая синтезированная цепь оцДНК, включающая здоровую последовательность триплетных повторов, также синтезируется таким образом, чтобы включать короткий участок гомологии (то есть плечо гомологии), который соответствует последовательности, являющейся смежной относительного другого конца повтора (красная цепь). Внесение новой синтезированной цепи и последующее замещение эндогенной ДНК новым синтезированным оцДНК-флэпом ведет к сокращению повторяющегося аллеля.

[1064] В зависимости от конкретного связанного с экспансией тринуклеотидов заболевания, дефектные триплетные экспансии могут располагаться в "связанных с

экспансией тринуклеотидных повторов белках". Связанные с экспансией тринуклеотидных повторов белки представляют собой разнообразный набор белков, связанных с подверженностью к развитию заболевания, связанного с экспансией тринуклеотидных повторов, наличием заболевания, связанного с экспансией тринуклеотидных повторов, тяжестью заболевания, связанного с экспансией тринуклеотидных повторов, или любой их комбинацией. Заболевания, связанные с экспансией тринуклеотидных повторов, разделяются на две категории, определяемые на основе типа повтора. Наиболее частый повтор представляет собой повтор CAG, который при его наличии в кодирующем участке гена кодирует аминокислоту глутамин (Q). Таким образом, данные заболевания обозначаются как полиглутаминовые (polyQ) заболевания и включают следующие заболевания: болезнь Хантингтона (HD); спинальную и бульбарную мышечную атрофию (SBMA); спинально-церебеллярную атаксию (SCA) (SCA типов 1, 2, 3, 6, 7 и 17); и дентато-рубро-паллидо-льюисову атрофию (Dentatorubro-Pallidoluysian Atrophy, DRPLA). Остальные заболевания, связанные с экспансией тринуклеотидных повторов, либо не включают триплет CAG, или триплет CAG не располагается в кодирующем участке гена и, таким образом, они обозначаются как неполиглутаминовые заболевания. Неполиглутаминовые заболевания включают синдром ломкой X-хромосомы (FRAXA), ломкую XE MR (FRAXE), атаксию Фрейдрейха (FRDA), миотоническую дистрофию (DM), и спиноцеребеллярные атаксии (SCA типов 8 и 12).

[1065] Белки, ассоциированные с заболеваниями, связанными с экспансией тринуклеотидных повторов, могут быть выбраны на основе экспериментального ассоциирования белка, ассоциированного с заболеванием, связанным с экспансией тринуклеотидных повторов, с заболеванием, связанным с экспансией тринуклеотидных повторов. Например, скорость продуцирования или концентрация в крови белка, ассоциированного с заболеванием, связанным с экспансией тринуклеотидных повторов, может быть повышена или понижена у группы, имеющей заболевание, связанное с экспансией тринуклеотидных повторов, по сравнению с группой без заболевания, связанного с экспансией тринуклеотидных повторов. Различия в плане уровней белка могут быть проанализированы с помощью протеомных способов, включая, но без ограничения, вестерн-блот, иммуногистохимическое окрашивание, иммуноферментный анализ (ELISA), и масс-спектометрию. В качестве альтернативы, белки, ассоциированные с заболеваниями, связанными с экспансией тринуклеотидных

повторов, могут быть идентифицированы путем получения профилей экспрессии генов, кодирующих данные белки, с помощью геномных способов, включая, но без ограничения, анализы на основе ДНК-микрочипов, последовательные анализы экспрессии генов (SAGE), и количественную полимеразную цепную реакцию в реальном времени (кПЦР).

[1066] Неограничивающие примеры белков, ассоциированных с заболеваниями, связанными с экспансией тринуклеотидных повторов, которые могут быть скорректированы с помощью праймерного редактирования, включают AR (андрогенный рецептор), FMR1 (белок 1, ассоциированный с умственной отсталостью, сцепленной с ломкой X-хромосомой), HTT (гентингтин), DMPK (протеинкиназу, ассоциированную с миотонической дистрофией), FXN (фратаксин), ATXN2 (атаксин 2), ATN1 (атрофии 1), FEN1 (структуроспецифичную флэп-эндонуклеазу 1), TNRC6A (белок, кодируемый геном 6A, содержащим тринуклеотидные повторы), PABPN1 (ядерный поли(A)-связывающий белок 1), JPH3 (юнктофилин 3), MED15 (субъединицу 15 медиаторного комплекса), ATXN1 (атаксин 1), ATXN3 (атаксин 3), TBP (TATA-бокс-связывающий белок), CACNA1A (альфа-1A-субъединицу потенциал-зависимого кальциевого канала P/Q-типа), ATXN80S (белок, синтезируемый с противоположной нити ATXN8 (не кодирующей белок)), PPP2R2B (бета-изоформу регуляторной субъединицы B протеинфосфатазы 2), ATXN7 (атаксин 7), TNRC6B (белок, кодируемый геном 6B, содержащим тринуклеотидные повторы), TNRC6C (белок, кодируемый геном 6C, содержащим тринуклеотидные повторы), CELF3 (CUGBP, член 3 семейства Elav-подобных белков), MAB21L1 (mab-21-подобный белок 1 (C.elegans)), MSH2 (гомолог 2 mutS, ассоциированный с неполипозным колоректальным раком типа 1 (E. coli)), TMEM185A (трансмембранный белок 185A), SIX5 (белок, кодируемый гомеобоксом 5 SIX), CNPY3 (гомолог Canopy 3 (данио-рерио)), FRAXE (белок, ассоциированный с "редким" ломким сайтом, проявляющимся при недостатке фолиевой кислоты, fra(X)(q28) E), GNB2 (бета полипептид 2 белка, связывающего гуаниновые нуклеотиды (G-белка)), RPL14 (рибосомный белок L14), ATXN8 (атаксин 8), INSR (инсулиновый рецептор), TTR (транстиретин), EP400 (E1A-связывающий белок p400), GIGYF2 (белок GYF 2, взаимодействующий с GRB10), OGG1 (8-оксогуанин-ДНК-гликозилазу), STC1 (станниокальцин 1), CNDP1 (карнозиндипептидазу 1 (металлопептидазу семейства M20)), C10orf2 (белок, кодируемый открытой рамкой считывания 2 хромосомы 10), MAML3 (mastermind-

подобный белок 3 (Drosophila)), DKC1 (белок 1, ассоциированный с врожденным дискератозом, дискерин), PAXIP1 (белок 1, взаимодействующий с PAX (с доменом активации транскрипции)), CaSK (кальций/кальмодулин-зависимую сериновую протеинкиназу (семейства MAGUK)), MAPT (белок tau, ассоциированный с микротрубочками), SP1 (фактор транскрипции Sp1), POLG (полимеразу гамма (ДНК-направленную)), AFF2 (член 2 семейства AF4/FMR2), THBS1 (тромбоспондин 1), TP53 (опухолевый белок p53), ESR1 (эстрогеновый рецептор 1), CGGBP1 (белок 1, связывающий триплетный повтор CGG), ABT1 (активатор 1 базальной транскрипции), KLK3 (родственную калликреину пептидазу 3), PRNP (белок приона), JUN (онкоген jun), KCNN3 (кальций-активируемый калиевый канал средней/малой проводимости, представитель 3 подсемейства N), BAX (BCL2-ассоциированный белок X), FRAXA (белок, ассоциированный с "редким" ломким сайтом, проявляющимся при недостатке фолиевой кислоты, fra(X)(q27.3) A (макроорхидизм, умственная отсталость)), KBTBD10 (белок 10, содержащий повтор Kelch и домен BTB (POZ)), MBNL1 (muscleblind-подобный белок (Drosophila)), RAD51 (гомолог RAD51 (гомолог RecA, E. coli) (S. cerevisiae)), NCOA3 (коактиватор 3 ядерных рецепторов), ERDA1 (белок с экспансией повторяющихся доменов, CAG/CTG 1), TSC1 (белок 1, ассоциированный с туберозным склерозом), COMP (олигомерный матриксный белок хряща), GCLC (каталитическую субъединицу глутаматцистеинлигазы), RRAD (Ras-родственный белок, ассоциированный с сахарным диабетом), MSH3 (гомолог 3 mutS (E. coli)), DRD2 (дофаминовый рецептор D2), CD44 (молекулу CD44 (система групп крови Indian)), CTCF (CCCTC-связывающий фактор (белок с "цинковыми пальцами")), CCND1 (циклин D1), CLSPN (гомолог класпина (Xenopus laevis)), MEF2A (энхансерный фактор 2A миоцитов), PTPRU (протеинтирозинфосфатазу рецепторного типа U), GAPDH (глицеральдегид-3-фосфатдегидрогеназу), TRIM22 (белок 22, содержащий тройной мотив), WT1 (белок 1 опухоли Вильмса), AHR (арил-углеводородный рецептор), GPX1 (глутатионпероксидазу 1), TPMT (тиопурин-S-метилтрансферазу), NDP (белок, ассоциированный с болезнью Норри (псевдоглиомой)), ARX (белок, кодируемый гомеобоксом гена, родственного aristaless), MUS81 (гомолог эндонуклеазы MUS81 (S. cerevisiae)), TYR (тирозиназу (глазокожный альбинизм IA)), EGR1 (белок 1 раннего ростового ответа), UNG (урацил-ДНК-гликозилазу), NUMBL (белок, подобный гомологу numb (Drosophila)), FABP2 (белок 2, связывающий жирные кислоты в кишечнике), EN2 (белок, кодируемый гомеобоксом engrailed 2), CRYGC

(гамма-С-кристаллин), SRP14 (гомологичный РНК-связывающий белок Alu размером 14 кДа из частицы узнавания сигнала), CRYGB (гамма-В-кристаллин), PDCD1 (белок 1 запрограммированной гибели клеток), HOXA1 (белок, кодируемый гомеобоксом А1), ATXN2L (атаксин-2-подобный белок), PMS2 (PMS2, белок 2, противодействующий повышению уровня постмейотической сегрегации (S. cerevisiae)), GLA (альфа-галактозидазу), CBL (белок, кодируемый последовательностью, трансформирующей с экотропным ретровирусом Cas-Br-M (мышей)), FTH1 (полипептид 1 тяжелой субъединицы ферритина), IL12RB2 (бета-2-субъединицу рецептора интерлейкина 12), OTX2 (белок, кодируемый гомеобоксом orthodenticle 2), HOXA5 (белок, кодируемый гомеобоксом А5), POLG2 (вспомогательную гамма-2-субъединицу полимеразы (ДНК-направленной)), DLX2 (белок, кодируемый гомеобоксом distal-less 2), SIRPA (сигнально-регуляторный белок альфа), OTX1 (белок, кодируемый гомеобоксом orthodenticle 1), AHRR (репрессор арил-углеводородного рецептора), MANF (мезэнцефальный нейротрофический фактор, происходящий из астроцитов), TMEM158 (трансмембранный белок 158 (ген/псевдоген)) и ENSG00000078687.

[1067] В определенном аспекте настоящее изобретение предусматривает способы на основе TPRT для лечения субъекта, у которого диагностировано заболевание, связанное экспансией повторов (также известное как связанное с экспансией повторов заболевание или связанное с тринуклеотидными повторами заболевание). Связанные с экспансией повторов заболевания возникают когда микросателлитные повторы подвергаются экспансии сверх пороговой длины. В нестоящее время считывается, что, по меньшей мере 30 генетических заболеваний вызываются экспансиями повторов. Научное понимание данной разнообразной группы заболеваний возникло в ранних 1990-х годах с открытием того, что тринуклеотидные повторы связаны с несколькими заметными наследственными состояниями, включая синдром ломкой Х-хромосомы, спинальную и бульбарную мышечную атрофию, миотоническую дистрофию и болезнь Хантингтона (Nelson et al, “The unstable repeats – three evolving faces of neurological disease,” *Neuron*, March 6, 2013, Vol.77; 825-843, который полностью приведен здесь в качестве ссылки), а также синдром Хау-Ривера, синдром Якобсена, дентато-рубро-паллидо-льюисову атрофию (DRPLA), болезнь Мачадо-Джозефа, синполидактилию (SPD II), генитальный синдром кисти-стопы (HFGS), клидокраниальную дисплазию (CCD), голопрозэнцефалическое расстройство (HPE), врожденный синдром центральной гиповентиляции (CCHS), ARX-несиндромную Х-сцепленную умственную

отсталость (XLMR) и окулофарингеальную мышечную дистрофию (OPMD) (см.). Было обнаружено, что нестабильность микросателлитных повторов является признаком данных состояний, как и предполагалось, то есть явлением, при котором экспансия повторов может возникать с каждым последующим поколением, что приводит к более тяжелому фенотипу и более раннему возрасту возникновения заболевания у потомства. Считается, что экспансии повторов вызывают заболевания за счет нескольких различных механизмов. Более конкретно, экспансии могут нарушать функцию клеток на уровне гена, иРНК транскрипта, и/или кодируемого белка. При некоторых состояниях мутации действуют за счет механизма потери функции путем сайленсинга содержащих повторы генов. При других из них заболевание возникает вследствие механизмов приобретения функции, при которых иРНК транскрипт или белок приобретает новые аберрантные функции.

[1068] В рамках одного варианта осуществления на Фиг.23 изображен способ лечения заболевания, связанного с тринуклеотидными повторами. В целом, данный подход включает применение геномного редактирования на основе TPRT (то есть праймерного редактирования) в комбинации с удлиненной нРНК, которая включает участок, который кодирует требуемую и здоровую замещающую последовательность тринуклеотидного повтора, которая должна заместить эндогенную связанную с заболеванием последовательность тринуклеотидного повтора за счет механизма процесса праймерного редактирования. Схема примера конструкции нРНК для сокращения последовательностей тринуклеотидных повторов и сокращение тринуклеотидных повторов с помощью редактирования генома на основе TPRT (то есть праймерного редактирования) показаны на Фиг.23.

Прионное заболевание

[1069] Праймерное редактирование также может применяться для предотвращения или остановки прогрессирования прионного заболевания за счет внесения одной или нескольких защитных мутаций в прионные белки (PRNP), которые подвергаются неправильному фолдингу при заболевании. Прионные болезни или трансмиссивные губчатые энцефалопатии (TSE) представляют собой семейство редких прогрессирующих нейродегенеративных заболеваний, которыми страдают как люди, так и животные. Они характеризуются длительными инкубационными периодами, характеризуются губчатыми изменениями, связанными с потерей нейронов, и неспособностью индуцировать воспалительный ответ.

[1070] У человека прионное заболевание включает болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, и куру. У животных прионное заболевание включает коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных. Праймерное редактирование может применяться для внесения защитных точечных мутаций в прионный белок для предотвращения или остановки прогрессирования любого из данных прионных заболеваний.

[1071] Классическая CJD представляет собой прионное заболевание человека. Она представляет собой нейродегенеративное заболевание с характерными клиническими и диагностическими признаками. Данное заболевание быстро прогрессирует и всегда является смертельным. Заражение данным заболеванием приводит к смерти обычно в течение 1 года после возникновения заболевания. CJD является быстро прогрессирующим неизменно смертельным нейродегенеративным заболеванием, которое, как считается, вызывается анормальной изформой клеточного гликопротеина, известного как прионный белок. CJD возникает по всему миру, при этом сообщается, что оцениваемая ежегодная частотность во многих странах, включая Соединенные Штаты, составляет около одного случая на один миллион населения. Подавляющее большинство пациентов с CJD обычно умирают в течение 1 года после возникновения заболевания. CJD классифицируется как трансмиссивная губчатая энцефалопатия (TSE) наряду с другими прионными заболеваниями, которые возникают у человека и животных. У приблизительно 80% пациентов CJD возникает в виде случайного заболевания без определенного паттерна передачи. У меньшей части пациентов (от 5 до 15%) CJD развивается вследствие наследственных мутаций гена прионного белка. Данные наследственные формы включают синдром Герстмана-Штрауслера-Шейнкера и фатальную семейную бессонницу. Способы лечения CJD в настоящее время неизвестны.

[1072] Вариант болезни Крейтцфельдта-Якоба (vCJD) представляет собой прионное заболевание, которое было впервые описано в 1996 в Соединенном Королевстве. В настоящее время существует сильное научное свидетельство того, что агент, отвечающий за вспышку прионного заболевания у коров, коровьей губчатой энцефалопатии (BSE или "коровьего бешенства"), является тем же агентом,

ответственным за вспышку vCJD у людей. Вариант CJD (vCJD) не является тем же заболеванием, что и классическая CJD (которую часто называют просто CJD). Она имеет клинические и патологические характеристики, отличающиеся от классической CJD. Каждое заболевание также обладает определенным генетическим профилем гена прионного белка. Оба заболевания являются необратимо смертельными заболеваниями мозга с необычно длительными инкубационными периодами, измеряемыми годами, и вызываются нестандартным переносимым агентом, называемым прионом. Способы лечения vCJD в настоящее время неизвестны.

[1073] BSE (коровья губчатая энцефалопатия или "коровье бешенство") представляет собой прогрессирующее неврологическое заболевание крупного рогатого скота, которое возникает вследствие инфицирования необычным переносимым агентом, называемым прионом. Природа данного переносимого агента не до конца понятна. В настоящее время наиболее распространенная теория заключается в том, что данный агент представляет собой модифицированную форму нормального белка, известную как прионный белок. Вследствие неизвестных в настоящее время причин нормальный прионный белок превращается в патогенную (вредоносную) форму, которая затем повреждает центральную нервную систему крупного рогатого скота. Существует значительное свидетельство того, что существуют различные виды BSE: типичный или классический вид BSE, ответственный за вспышку в Соединенном Королевстве, и два атипичных вида (виды H и L). Способы лечения BSE в настоящее время неизвестны.

[1074] Хроническое истощающее заболевание (CWD) представляет собой прионное заболевание, которое поражает оленей, настоящих оленей, северных оленей, пятнистых оленей и лосей. Оно было обнаружено в некоторых областях Северной Америки, включая Канаду и Соединенные Штаты, Норвегию и Южную Корею. Может пройти более года прежде чем у инфицированного животного появятся симптомы, которые могут включать значительную потерю веса (истощение), спотыкание, вялость и другие неврологические симптомы. CWD может поражать животных всех возрастов, при этом некоторые инфицированные животные могут умирать даже без развития заболевания. CWD является смертельным для животных, при этом способы его лечения или вакцины отсутствуют.

[1075] Считается, что вызывающими TSE агентами являются прионы. Термин "прионы" обозначает анормальные патогенные агенты, которые являются трансмиссивными и способными индуцировать анормальный фолдинг определенных

нормальных клеточных белков, называемых прионными белками и находящихся в основном в мозге. Функции данных нормальных прионных белков все еще не полностью известны. Анормальный фолдинг прионных белков приводит к повреждению мозга и характерным признакам и симптомам заболваниея. Прионные заболевания обычно быстро прогрессируют и всегда являются смертельными.

[1076] В соответствии с используемым здесь значением, термин "прион" означает инфекционную частицу, в отношении которой известно, что она вызывает заболевания (губчатую энцефалопатию) у человека и животных. Термин "прион" является сокращением слов "белок" (протеин) и "инфекция", при этом данные частицы в основном (если не полностью) состоят из молекул PRNPSc, кодируемых геном PRNP, который экспрессирует PRNPC, который изменяет конформацию с образованием PRNPSc. Прионы являются отличными от бактерий, вирусов и вироидов. Известные прионы включают те из них, которые инфицируют животных и вызывают скрепи, представляющее собой трансмиссивное дегенеративное заболевание нервной системы овец и коз, а также коровью губчатую энцефалопатию (BSE) или коровье бешенство и губчатую энцефалопатию кошек. Известно, что четырьмя описанными выше прионными заболеваниями, которые поражают человека, являются (1) куру, (2) болезнь Крейтцфельдта-Якоба (CJD), (3) болезнь Герстмана-Штрауслера-Шейнкера (GSS), и (4) фатальная семейная бессонница (FFI). В соответствии с используемым здесь значением термин "прион" включает все формы прионов, вызывающие все или любые из данных заболеваний, или другие из них у любых животных, в частности у человека и домашних сельскохозяйственных животных.

[1077] В целом, и без связи с теорией, прионные заболевания вызываются неправильным фолдингом прионных белков. При таких заболеваниях, которые часто называются здесь связанными с отложением заболеваниями, неправильный фолдинг прионных белков может осуществляться следующим образом. Если A представляет собой нормально синтезированный генный продукт, который реализует требуемую физиологическую роль в мономерном или олигомерном состоянии, то A* представляет собой конформационно активированную форму A, которая способна подвергаться значительному конформационному изменению, B представляет собой конформационно измененное состояние, для которого являются предпочтительными мультимерные сборки (то есть подвергнутая неправильному фолдингу форма, которая образует отложения), а B_n представляет собой мультимерный материал, который

является патогенным и относительно сложным в плане рециклизации. Для прионных заболеваний $PRNP^{C}$ и $PRNP^{Sc}$ соответствуют состояниям A и B_n, где A является по большей части спиральным и мономерным, а B_n имеет высокое содержание β и является мультимерным.

[1078] Известно, что определенные мутации в прионных белках могут быть ассоциированы с повышенным риском прионного заболевания. С другой стороны, определенные мутации в прионных белках могут являться защитными по своей природе. См. Bagynszky et al., "Characterization of mutations in PRNP (prion) gene and their possible roles in neurodegenerative diseases," *Neuropsychiatr Dis Treat.*, 2018; 14: 2067-2085, содержание которого приведено здесь в качестве ссылки.

[1079] PRNP (NCBI RefSeq No. NP_000302.1 (SEQ ID NO: 291)) — прионный белок человека — кодируется геном длиной 16 kb, расположенным на хромосоме 20 (4686151-4701588). Он содержит два экзона, при этом экзон 2 несет открытую рамку считывания, которая кодирует белок PrP длиной 253 аминокислоты (AA). Экзон 1 представляет собой некодирующий экзон, который может служить в качестве участка инициации транскрипции. Посттрансляционные модификации приводят к удалению N-концевого фрагмента (NTF) в виде первых 22 AA и C-концевого фрагмента (CTF) в виде последних 23 AA. NTF отщепляется после транспорта PrP в эндоплазматическую сеть (ER), в то время как CTF отщепляется GPI-якорем (сигнальный пептид [GPI-SP] гликозилфосфатидилинозитол [GPI]). GPI-якорь может принимать участие в транспорте белка PrP. Он также может принимать участие в присоединении прионного белка к внешней поверхности мембраны клетки. Нормальный PrP состоит из длинной N-концевой петли (которая содержит участок октапептидного повтора), двух коротких β-листов, трех α-спиралей, и C-концевого участка (который содержит GPI-якорь). Расщепление PrP приводит к получению гликопротеина длиной 208 AA, присоединенного к клеточной мембране.

[1080] Аминокислотная последовательность 253 PRNP (NP_000302.1) является следующей:

MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGG
GWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKT
NMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQV
YYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQ
YERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 291).

[1081] Аминокислотная последовательность 253 PRNP (NP_000302.1) кодируется нуклеотидной последовательностью (NCBI Ref. Seq No. NM_000311.5, "прионный

белок homo sapiens (PRNP), транскриптный вариант 1, иРНК)", которая является следующей:

GCGAACCTTGGCTGCTGGATGCTGGTTCTCTTTGTGGCCACATGGAGTGACCTGG
GCCTCTGCAAGAAGCGCCCGAAGCCTGGAGGATGGAACACTGGGGGCAGCCGAT
ACCCGGGGCAGGGCAGCCCTGGAGGCAACCGCTACCCACCTCAGGGCGGTGGTG
GCTGGGGGCAGCCTCATGGTGGTGGCTGGGGGCAGCCTCATGGTGGTGGCTGGG
GGCAGCCCCATGGTGGTGGCTGGGGACAGCCTCATGGTGGTGGCTGGGGTCAAG
GAGGTGGCACCCACAGTCAGTGGAACAAGCCGAGTAAGCCAAAAACCAACATGA
AGCACATGGCTGGTGCTGCAGCAGCTGGGGCAGTGGTGGGGGGCCTTGGCGGCT
ACATGCTGGGAAGTGCCATGAGCAGGCCCATCATACATTTCGGCAGTGACTATGAG
GACCGTTACTATCGTGAAAACATGCACCGTTACCCCAACCAAGTGTACTACAGGCC
CATGGATGAGTACAGCAACCAGAACAACTTTGTGCACGACTGCGTCAATATCACA
ATCAAGCAGCACACGGTCACCACAACCACCAAGGGGGAGAACTTCACCGAGACC
GACGTTAAGATGATGGAGCGCGTGGTTGAGCAGATGTGTATCACCCAGTACGAGA
GGGAATCTCAGGCCTATTACCAGAGAGGATCGAGCATGGTCCTCTTCTCCTCTCCA
CCTGTGATCCTCCTGATCTCTTTCCTCATCTTCCTGATAGTGGGATGAGGAAGGTCT
TCCTGTTTTCACCATCTTTCTAATCTTTTTCCAGCTTGAGGGAGGCGGTATCCACCT
GCAGCCCTTTTAGTGGTGGTGTCTCACTCTTTCTTCTCTCTTTGTCCCGGATAGGCT
AATCAATACCCTTGGCACTGATGGGCACTGGAAAACATAGAGTAGACCTGAGATG
CTGGTCAAGCCCCCTTTGATTGAGTTCATCATGAGCCGTTGCTAATGCCAGGCCAG
TAAAAGTATAACAGCAAATAACCATTGGTTAATCTGGACTTATTTTTGGACTTAGTG
CAACAGGTTGAGGCTAAAACAAATCTCAGAACAGTCTGAAATACCTTTGCCTGGA
TACCTCTGGCTCCTTCAGCAGCTAGAGCTCAGTATACTAATGCCCTATCTTAGTAGA
GATTTCATAGCTATTTAGAGATATTTTCCATTTTAAGAAAACCCGACAACATTTCTG
CCAGGTTTGTTAGGAGGCCACATGATACTTATTCAAAAAAATCCTAGAGATTCTTA
GCTCTTGGGATGCAGGCTCAGCCCGCTGGAGCATGAGCTCTGTGTGTACCGAGAA
CTGGGGTGATGTTTTACTTTTCACAGTATGGGCTACACAGCAGCTGTTCAACAAGA
GTAAATATTGTCACAACACTGAACCTCTGGCTAGAGGACATATTCACAGTGAACAT
AACTGTAACATATATGAAAGGCTTCTGGGACTTGAAATCAAATGTTTGGGAATGGT
GCCCTTGGAGGCAACCTCCCATTTTAGATGTTTAAAGGACCCTATATGTGGCATTCC
TTTCTTTAAACTATAGGTAATTAAGGCAGCTGAAAAGTAAATTGCCTTCTAGACAC
TGAAGGCAAATCTCCTTTGTCCATTTACCTGGAAACCAGAATGATTTTGACATACA
GGAGAGCTGCAGTTGTGAAAGCACCATCATCATAGAGGATGATGTAATTAAAAAAT
GGTCAGTGTGCAAAGAAAAGAACTGCTTGCATTTCTTTATTTCTGTCTCATAATTG
TCAAAAACCAGAATTAGGTCAAGTTCATAGTTTCTGTAATTGGCTTTTGAATCAAA
GAATAGGGAGACAATCTAAAAAATATCTTAGGTTGGAGATGACAGAAATATGATTG
ATTTGAAGTGGAAAAAGAAATTCTGTTAATGTTAATTAAAGTAAAATTATTCCCTG
AATTGTTTGATATTGTCACCTAGCAGATATGTATTACTTTTCTGCAATGTTATTATTG
GCTTGCACTTTGTGAGTATTCTATGTAAAAATATATATGTATATAAAATATATATTGCA
TAGGACAGACTTAGGAGTTTTGTTTAGAGCAGTTAACATCTGAAGTGTCTAATGCA
TTAACTTTTGTAAGGTACTGAATACTTAATATGTGGGAAACCCTTTTGCGTGGTCCT
TAGGCTTACAATGTGCACTGAATCGTTTCATGTAAGAATCCAAAGTGGACACCATT
AACAGGTCTTTGAAATATGCATGTACTTTATATTTTCTATATTTGTAACTTTGCATGT
TCTTGTTTTGTTATATAAAAAAATTGTAAATGTTTAATATCTGACTGAAATTAAACG
AGCGAAGATGAGCACCA (SEQ ID NO: 292)

[1082] Сообщается, что участки мутирования, которые относятся к PRNP (NP_000302.1), которые связаны с CJD и FFI, являются следующими. Данные мутации могут быть удалены или внесены с помощью описанных здесь праймерных редакторов.

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННОГО С ПРИОННЫМ ЗАБОЛЕВАНИЕМ CJD (СМ. ТАБЛИЦУ 1 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 PRNP NP_000302.1)
D178N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHNCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 293)
T188K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHKVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 294)
E196K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGKNFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 295)
E196A	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGANFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 296)
E200K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTKTDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 297)
E200G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTGTDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 298)
V203I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF

	GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDIKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 299)
R208H	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMEHVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 300)
V210I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVIEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 301)
E211Q	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVQQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 302)
M232R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSRVLFSSPPV (SEQ ID NO: 303)

[1083] Сообщается, что участки мутирования, которые относятся к PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с GSS, являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННОГО С ПРИОННЫМ ЗАБОЛЕВАНИЕМ GSS (СМ. ТАБЛИЦУ 2 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 PRNP NP_000302.1)
P102L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKLSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 304)
P105L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKLKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 305)
A117V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAVAGAVVGGLGGYMLGSAMSRPIIHF

	GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 306)
G131V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLVSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 307)
V176G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFGHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 308)
H187R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQRTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 309)
	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 291)
F198S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENSTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 311)
D202N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETNVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 312)
Q212P	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEPMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 313)
Q217R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT

	TTTKGENFTETDVKMMERVVEQMCITRYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 314)
M232T	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSTVLFSSPPV (SEQ ID NO: 315)

[1084] Участки мутирования, которые относятся к PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с возможной защитной природой против прионного заболевания, являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТНОГО PRNP, СВЯЗАННАЯ С ЗАЩИТНОЙ ПРИРОДОЙ ПРОТИВ ПРИОННОГО ЗАБОЛЕВАНИЯ (СМ. ТАБЛИЦУ 4 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 PRNP NP_000302.1)
G127S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGSYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 316)
G127V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGVYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 317)
M129V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYVLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 318)
D167G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMGEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 319)
D167N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMNEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 320)
N171S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG

	GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSSQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 321)
E219K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYKRESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 322)
P238S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNR YPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGG GTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHF GSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVT TTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSSPV (SEQ ID NO: 323)

[1085] Таким образом, в различных вариантах осуществления праймерное редактирование может применяться для удаления мутации в PRNP, которая связана с прионным заболеванием, или внесения мутации в PRNP, которая считается являющейся защитной против прионного заболевания. Например, праймерное редактирование может применяться для удаления или восстановления мутации D178N, V180I, T188K, E196K, E196A, E200K, E200G, V203I, R208H, V210I, E211Q, I215V, или M232R в белке PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). Например, праймерное редактирование может применяться для удаления или восстановления мутации P102L, P105L, A117V, G131V, V176G, H187R, F198S, D202N, Q212P, Q217R, или M232T в белке PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). Путем удаления или корректирования наличия подобных мутаций в PRNP с помощью праймерного редактирования риск развития прионного заболевания может быть снижен или устранен.

[1086] В других вариантах осуществления праймерное редактирование может применяться для внесения защитной мутации в PRNP, которая связана с защитным эффектом против одного или нескольких прионных заболеваний. Например, праймерное редактирование может применяться для внесения защитной мутации G127S, G127V, M129V, D167G, D167N, N171S, E219K, или P238S в PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291). В других вариантах осуществления защитная мутация может представлять собой любую альтернативную аминокислоту, внесенную в G127, G127, M129, D167, D167, N171, E219, или P238 в PRNP (относительно PRNP NP_000302.1) (SEQ ID NO: 291).

[1087] В определенных вариантах осуществления праймерное редактирование может применяться для внесения защитной мутации G127V в PRNP, как показано на Фиг.27 и описано в Примере 5.

[1088] В другом варианте осуществления праймерное редактирование может применяться для внесения защитной мутации E219K в PRNP.

[1089] Белок *PRNP* и участок защитной мутации являются консервативными у млекопитающих, поэтому помимо лечения заболевания у человека он также может использоваться для получения коров и овец с иммунитетом против прионного заболевания или даже осуществления лечения диких популяций животных, которые страдают прионным заболеванием. Праймерное редактирование уже использовалось для достижения ~25% внесения встречающегося в природе защитного аллеля в клетки человека, при этом предшествующие эксперименты на мышах показывают, что данный уровень внесения является достаточным для возникновения иммунитета против большинства прионных заболеваний. Данный способ является первым и потенциально единственным доступным в настоящее время способом внесения данного аллеля с настолько высокой эффективностью в большинство типов клеток. Другая возможная стратегия лечения заключается в применении праймерного редактирования для снижения или устранения экспрессии *PRNP* путем внесения в ген раннего стоп-кодона.

[1090] С помощью описанных здесь способов конструирования PEgRNA могут быть сконструированы походящие PEgRNA для внесения требуемых защитных мутаций или для удаления ассоциированных с прионным заболеванием мутаций из PRNP. Например, приведенный ниже перечень PEgRNA может применяться для внесения защитного аллеля G127V и защитного аллеля E219K в PRNP человека, а также защитного аллеля G127V в PRNP различных животных.

[9] Фармацевтические композиции

[1091] Другие аспекты настоящего изобретения относятся к фармацевтическим композициям, включающим любые из различных компонентов описанной здесь системы праймерного редактирования (например включая, но без ограничения, napDNAbp, обратные транскриптазы, слитые белки (например включающие napDNAbp и обратные транскриптазы), удлиненные направляющие РНК, и комплексы, включающие слитые белки и удлиненные направляющие РНК, а также вспомогательные элементы, такие как компоненты для никирования второй цепи и нуклеазы для удаления 5' эндогенного ДНК-флэпа для реализации процесса

праймерного редактирования с образованием подвергнутого редактированию продукта).

[1092] Термин "фармацевтическая композиция" в соответствии с используемым здесь значением обозначает композицию, составленную для фармацевтического применения. В некоторых вариантах осуществления фармацевтическая композиция дополнительно включает фармацевтически приемлемый носитель. В некоторых вариантах осуществления фармацевтическая композиция включает дополнительные агенты (например для специфичной доставки, увеличения периода полужизни, или другие терапевтические соединения).

[1093] В соответствии с используемым здесь значением термин "фармацевтически приемлемый носитель" означает фармацевтически приемлемый материал, композицию или носитель, такой как жидкий или твердый наполнитель, разбавитель, инертный наполнитель, вспомогательное средство для производства (например лубрикант, магниевый тальк, стеарат кальция или цинка, или стеариновую кислоту), или материал для инкапсуляции разбавителя, принимающий участие в переносе или транспортировке соединения из одного участка организма (например участка доставки) в другой участок (например орган, ткань, или часть тела). Фармацевтически приемлемый носитель является "приемлемым" в плане того, что он является совместимым с другими ингредиентами состава и не наносит вреда ткани субъекта (например является физиологически совместимым, стерильным, имеет физиологический pH и так далее). Некоторые примеры материалов, которые могут служить в качестве фармацевтически приемлемых носителей, включают: (1) сахара, такие как лактоза, глюкоза и сахароза; (2) крахмалы, такие как кукурузный крахмал и картофельный крахмал; (3) целлюлоза и ее производные, такие как карбоксиметилцеллюлоза натрия, метилцеллюлоза, этилцеллюлоза, микрокристаллическая целлюлоза и ацетат целлюлозы; (4) порошкообразный трагакант; (5) солод; (6) желатин; (7) смазывающие агенты, такие как стеарат магния, лаурилсульфат натрия и тальк; (8) наполнители, такие как масло какао и воски для суппозиториев; (9) масла, такие как арахисовое масло, хлопковое масло, сафлоровое масло, кунжутное масло, оливковое масло, кукурузное масло и соевое масло; (10) гликоли, такие как пропиленгликоль; (11) полиолы, такие как глицерин, сорбит, маннит и полиэтиленгликоль (ПЭГ); (12) сложные эфиры, такие как этилолеат и этиллаурат; (13) агар; (14) буферные агенты, такие как гидроксид магния и гидроксид

алюминия; (15) альгиновую кислоту; (16) апирогенную воду; (17) изотонический физиологический раствор; (18) раствор Рингера; (19) этиловый спирт; (20) pH-буферные растворы; (21) сложные полиэфиры, поликарбонаты и/или полиангидриды; (22) наполнители, такие как полипептиды и аминокислоты; (23) компонент сыворотки, такой как сывороточный альбумин, HDL и LDL; (22) C2-C12 спирты, такие как этанол; и (23) другие нетоксичные совместимые вещества, используемые в фармацевтических составах. В составе также могут присутствовать смачивающие агенты, красители, агенты для высвобождения, агенты для покрытия, подсластители, ароматизаторы, отдушки, консерванты и антиоксиданты. Такие термины как "инертный наполнитель", "носитель", "фармацевтически приемлемый носитель" или им подобные используются здесь взаимозаменяемым образом.

[1094] В некоторых вариантах осуществления фармацевтическая композиция составлена для доставки субъекту, например для редактирования генов. Подходящие способы введения описанной здесь фармацевтической композиции включают, без ограничения: местное, подкожное, трансдермальное, внутрикожное, внутриочаговое, внутрисуставное, внутрибрюшинное, внутрипузырное, чрезслизистое, десневое, внутризубное, интракохлеарное, транстимпаническое, внутриорганное, эпидуральное, интратекальное, внутримышечное, внутривенное, внутрисосудистое, внутрикостное, периокулярное, внутриопухолевое, внутримозговое и интрацереброventрикулярное введение.

[1095] В некоторых вариантах осуществления описанную здесь фармацевтическую композицию вводят местно в участок заболевания (например участок с опухолью). В некоторых вариантах осуществления описанную здесь фармацевтическую композицию вводят субъекту с помощью инъекции, с помощью катетера, с помощью суппозитория, или с помощью имплантата, при этом имплантат является пористым, непористым, или изготовленным из мягкого материала, включая мембрану, например силиконовую мембрану или волокно.

[1096] В других вариантах осуществления описанная здесь фармацевтическая композиция доставляется в рамках системы с контролируемым высвобождением. В одном варианте осуществления может использоваться помпа (см., например, Langer, 1990, Science 249:1527-1533; Sefton, 1989, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald et al., 1980, Surgery 88:507; Saudek et al., 1989, N. Engl. J. Med. 321:574). В другом варианте осуществления могут использоваться полимерные материалы (см., например,

Medical Applications of Controlled Release (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1974); Controlled Drug Bioavailability, Drug Product Design and Performance (Smolen and Ball eds., Wiley, New York, 1984); Ranger and Peppas, 1983, Macromol. Sci. Rev. Macromol. Chem. 23:61. Также см. Levy et al., 1985, Science 228:190; During et al., 1989, Ann. Neurol. 25:351; Howard et al., 1989, J. Neurosurg. 71:105). Также описаны другие системы с контролируемым высвобождением, например см. Langer выше.

[1097] В некоторых вариантах осуществления фармацевтическая композиция составлена в соответствии со стандартными процедурами в виде композиции, адаптированной для внутривенного или подкожного введения субъекту, например человеку. В некоторых вариантах осуществления фармацевтическая композиция для введения с помощью инъекции представляет собой растворы в стерильном изотоническом водном буфере. При необходимости фармацевтическая композиция также может включать солюбилизирующий агент и местный анестетик, например лигнокаин, для облегчения боли в месте инъекции. В целом ингредиенты присутствуют либо по отдельности, либо смешиваются вместе в виде единичной лекарственной формы, например в виде сухого лиофилизированного порошка или безводного концентрата в герметично запечатанном контейнере, например ампуле или саше, с указанием количества активного агента. При введении фармацевтической композиции с помощью инфузии она может быть разведена с использованием бутылки для инфузий, содержащей стерильную подходящую для фармацевтического применения воду или солевой раствор. При введении фармацевтической композиции с помощью инъекции может быть добавлена ампула со стерильной водой или солевым раствором для инъекций, так что ингредиенты могут быть смешаны перед введением.

[1098] Фармацевтическая композиция для системного введения может представлять собой жидкость, например стерильный солевой раствор, лактатный раствор Рингера или Хэнка. Кроме того, фармацевтическая композиция может быть представлена в твердых формах и повторно растворена или суспендирована непосредственно перед применением. Также охватываются лиофилизированные формы.

[1099] Фармацевтическая композиция может содержаться в липидной частице или пузырьке, например липосоме или микрокристалле, который также является подходящим для парентерального введения. Частицы могут иметь любую подходящую структуру, например однослойную или многослойную, при условии, что в них содержатся композиции. Соединения могут быть заключены в "стабилизированных

плазмид-липидных частицах" (SPLP), содержащих фузогенный липид диолеоилфосфатидилэтаноламин (DOPE), низкие уровни (5-10 мол.%) катионного липида, и стабилизированы за счет полиэтиленгликолевого (ПЭГ) покрытия (Zhang Y. P. *et al.*, *Gene Ther.* 1999, 6:1438-47). Положительно заряженные липиды, такие как N-[1-(2,3-диолеоилокси)пропил]-N,N,N-триметил-аммониймутилсульфат, или "DOTAP", являются особенно предпочтительными для таких частиц и пузырьков. Получение подобных липидных частиц является хорошо известным. См., например, патенты США 4880635; 4906477; 4911928; 4917951; 4920016; и 4921757; каждый из которых приведен здесь в качестве ссылки.

[1100] Описанная здесь фармацевтическая композиция может быть, например, введена или упакована в виде единичной дозы. Термин "единичная доза" при использовании в отношении фармацевтической композиции по настоящему изобретению обозначает физически отдельные единицы, являющиеся подходящими в качестве единичной дозы для субъекта, при этом каждая единица содержит предварительно определенное количество активного материала, рассчитанное для продуцирования требуемого терапевтического эффекта, совместно с требуемым разбавителем; то есть носителем или индифферентным веществом.

[1101] Кроме того, фармацевтическая композиция может быть представлена в виде фармацевтического набора, включающего (a) контейнер, содержащий соединение по изобретению в лиофилизированной форме, и (b) второй контейнер, содержащий фармацевтически приемлемый разбавитель (например стерильную воду) для инъекции. Фармацевтически приемлемый разбавитель может применяться для восстановления или разбавления лиофилизированного соединения по изобретению. С таким контейнером (контейнерами) может быть необязательно связана этикетка в форме, соответствующей требованиям государственного учреждения, регулирующего производство, применение или продажу лекарственный средств или биологических продуктов, на которой отражено одобрение данного учреждения, регулирующего производство, применение, или продажу, для введения человеку.

[1102] В другом аспекте включено изделие, содержащее материалы, которые могут применяться для лечения описанных выше заболеваний. В некоторых вариантах осуществления изделие включает контейнер и этикетку. Подходящие контейнеры включают, например, бутылки, флаконы, шприцы и пробирки для тестов. Контейнеры могут быть изготовлены из различных материалов, таких как стекло или пластик. В

некоторых вариантах осуществления контейнер содержит композицию, которая является эффективной для лечения описанного здесь заболевания, и может иметь стерильное входное отверстие. Например, контейнер может представлять собой пакет с внутривенным раствором или флакон, имеющий пробку, протыкаемую иглой для подкожных инъекций. Активный агент в композиции представляет собой соединение по изобретению. В некоторых вариантах осуществления этикетка, которая расположена на контейнере или связана с контейнером, указывает, что композиция применяется для лечения выбранного заболевания. Изделие может дополнительно включать второй контейнер, включающий фармацевтически приемлемый буфер, например фосфатно-солевой буфер, раствор Рингера, или раствор декстрозы. Он может дополнительно включать другие материалы, являющиеся желательными с коммерческой и пользовательской точки зрения, включая другие буферы, разбавители, фильтры, иглы, шприцы, и вкладыши в упаковку с инструкциями по применению.

[10] Наборы, клетки, векторы, и доставка

Наборы

[1103] Композиции по настоящему изобретению могут быть собраны в наборы. В некоторых вариантах осуществления набор включает векторы нуклеиновой кислоты для экспрессии описанных здесь праймерных редакторов. В других вариантах осуществления набор дополнительно включает подходящие направляющие нуклеотидные последовательности (например PEgRNA и нРНК для второго участка) или векторы нуклеиновой кислоты для экспрессии таких направляющих нуклеотидных последовательностей для нацеливания белка Cas9 или праймерного редактора на требуемую последовательность-мишень.

[1104] Описанный здесь набор может включать один или несколько контейнеров, включающих компоненты для осуществления описанных здесь способов и необязательно инструкции по применению. Любой из описанных здесь наборов может дополнительно включать компоненты, требуемые для осуществления способов проведения анализов. Каждый компонент наборов где это возможно может быть представлен в жидкой форме (например в виде раствора) или в твердой форме (например в виде сухого порошка). В определенных случаях некоторые компоненты могут подвергаться восстановлению или быть обработаны иным образом (например до активной формы), например путем добавления подходящего растворителя или другого компонента (например воды), который может входить или не входить в состав набора.

[1105] В некоторых вариантах осуществления наборы могут необязательно включать инструкции и/или материалы относительно применения представленных компонентов. В соответствии с используемым здесь значением термин "инструкции" может определять компонент в виде инструкции и/или материала и обычно включает письменные инструкции, расположенные на упаковке или связанные с упаковкой. Инструкции также могут включать любые устные или электронные инструкции, предоставленные любым образом, так что пользователю станет быстро очевидно, что инструкции следует ассоциировать с набором, например аудиовизуальные (например в виде видеокассеты, DVD и так далее), основанные на интернете и/или получаемые через веб, и так далее. Письменные инструкции могут быть в любой форме, соответствующей требованиям государственного учреждения, регулирующего производство, применение или продажу лекарственный средств или биологических продуктов, которые также могут отражать одобрение данного учреждения, регулирующего производство, применение, или продажу, для введения человеку. В соответствии с используемым здесь значением термин "материалы" включает все способы ведения бизнеса, включая способы образовательного, больничного, и другого клинического инструктирования, научные запросы, разработку и открытие лекарств, академические исследования, активность в рамках фармацевтической промышленности, включая продажи лекарственных средств, и любые виды рекламы и другую связанную с продвижением активность, включая письменные, устные и электронные коммуникации в любой форме, связанные с данным изобретением. Кроме того, наборы могут включать другие компоненты в зависимости от определенного вида использования, как описано здесь.

[1106] Наборы могут содержать любой один или несколько описанных здесь компонентов в одном или нескольких контейнерах. Компоненты могут быть получены в стерильном виде, упакованы в шприц и доставлены в охлажденном виде. В качестве альтернативы, они могут быть заключены во флаконе или другом контейнере для хранения. Второй контейнер может включать другие компоненты, полученные в стерильном виде. В качестве альтернативы, наборы могут включать активные агенты, предварительно смешанные и доставленные во флаконе, пробирке, или другом контейнере.

[1107] Наборы могут иметь различные формы, такие как блистерный пакет, пакет, упакованный в термоусадочную пленку, пакет с вакуумным запечатыванием,

запечатываемый термоформованный лоток или аналогичный пакет или лоток, с аксессуарами, свободно упакованными внутри пакета, одной или несколькими пробирками, контейнерами, коробкой или пакетом. Наборы могут быть стерилизованы после добавления аксессуаров, что позволяет не запаковывать отдельные аксессуары в контейнере. Наборы могут быть простерилизованы с помощью любых подходящих способов стерилизации, таких как стерилизация радиацией, тепловая стерилизация, или другие способы стерилизации, известные в данной области. Наборы также могут включать другие компоненты в зависимости от определенного вида использования, например контейнеры, питательную среду для клеток, соли, буферы, реагенты, шприцы, иглы ткань, такую как марля, для нанесения или удаления дезинфицирующего средства, одноразовые перчатки, подставку для агентов перед введением и так далее. Некоторые аспекты настоящего изобретения относятся к наборам, включающим конструкт нуклеиновой кислоты, включающий нуклеотидную последовательность, кодирующую различные компоненты описанной здесь системы праймерного редактирования (например включая, но без ограничения, napDNAbp, обратные транскриптазы, полимеразы, слитые белки (например включающие napDNAbp и обратные транскриптазы (или, более широко, полимеразы), удлиненные направляющие РНК, и комплексы, включающие слитые белки и удлиненные направляющие РНК, а также вспомогательные элементы, такие как компоненты для никирования второй цепи (например нРНК для никирования второй цепи) и нуклеазы для удаления 5' эндогенного ДНК-флэпа для реализации процесса праймерного редактирования с образованием подвергнутого редактированию продукта). В некоторых вариантах осуществления нуклеотидная последовательность (последовательности) включает гетерологичный промотор (или более одного промотора), который обеспечивает экспрессию компонентов системы праймерного редактирования.

[1108] Другие аспекты данного изобретения предусматривают наборы, включающие один или несколько конструктов нуклеиновой кислоты, кодирующей различные компоненты описанной здесь системы праймерного редактирования, например включающие нуклеотидную последовательность, кодирующую компоненты системы праймерного редактирования, способные модифицировать последовательность-мишень ДНК. В некоторых вариантах осуществления нуклеотидная

последовательность включает гетерологичный промотор, который обеспечивает экспрессию компонентов системы праймерного редактирования.

[1109] Некоторые аспекты данного изобретения предусматривают наборы, включающие конструкт нуклеиновой кислоты, включающий (a) нуклеотидную последовательность, кодирующую napDNAbp (например домен Cas9), слитый с обратной транскриптазой, и (b) гетерологичный промотор, который обеспечивает экспрессию последовательности (a).

Клетки

[1110] Клетки, которые могут содержать любую из описанных здесь композиций, включают прокариотические клетки и эукариотические клетки. Описанные здесь способы применяются для доставки белка Cas9 или праймерного редактора в эукариотическую клетку (например клетку млекопитающего, такую как клетка человека). В некоторых вариантах осуществления клетка представлена *in vitro* (например, является культивируемой клеткой). В некоторых вариантах осуществления клетка представлена *in vivo* (например располагается в субъекте, таком как представляющий собой человека субъект). В некоторых вариантах осуществления клетка представлена *ex vivo* (например выделена из субъекта и может быть введена обратно тому же или другому субъекту).

[1111] Клетки млекопитающих по настоящему изобретению включают клетки человека, клетки приматов (например клетки Vero), клетки крыс (например клетки GH3, клетки OC23) или клетки мышей (например клетки MC3T3). Существует множество линий клеток человека, включая, без ограничения, эмбриональные клетки почки человека (HEK), клетки HeLa, раковые клетки из 60 линий раковых клеток Национального института рака (National Cancer Institute) (NCI60), клетки DU145 (рак простаты), клетки Lncap (рак простаты), клетки MCF-7 (рак груди), клетки MDA-MB-438 (рак груди), клетки PC3 (рак простаты), клетки T47D (рак груди), клетки THP-1 (острый миелоидный лейкоз), клетки U87 (глиобластома), клетки нейробластомы человека SHSY5Y (клонированные из миеломы) и клетки Saos-2 (рак кости). В некоторых вариантах осуществления rAAV векторы доставляют в эмбриональные клетки почки человека (HEK) (например клетки HEK 293 или HEK 293T). В некоторых вариантах осуществления векторы rAAV доставляются в стволовые клетки (например стволовые клетки человека), такие как, например, плюрипотентные стволовые клетки (например плюрипотентные стволовые клетки человека, включая индуцированные

плюрипотентные стволовые клетки человека (hiPSC)). Термин "стволовая клетка" обозначает клетку со способностью делиться в течение бесконечного времени в культуре и порождать специализированные клетки. Термин "плюрипотентная стволовая клетка" обозначает тип стволовой клетки, которая способна дифференцироваться во все ткани организма, но не способна поддерживать полное развитие организма самостоятельным образом. Термин "индуцированная плюрипотентная стволовая клетка человека" обозначает соматическую (например зрелую или взрослую) клетку, которая была перепрограммирована в соответствующее эмбриональной стволовой клетке состояние за счет индуцирования экспрессии генов и факторов, являющихся важными для поддержания основных свойств эмбриональных стволовых клеток (см., например, Takahashi and Yamanaka, *Cell* 126 (4): 663–76, 2006, который приведен здесь в качестве ссылки). Индуцированные плюрипотентные стволовые клетки человека экспрессируют маркеры стволовых клеток и способны продуцировать клетки, характерные для всех трех зародышевых листков (эктодермы, энтодермы, мезодермы).

[1112] Дополнительные неограничивающие примеры линий клеток, которые могут применяться в соответствии с настоящим изобретением, включают 293-T, 293-T, 3T3, 4T1, 721, 9L, A-549, A172, A20, A253, A2780, A2780ADR, A2780cis, A431, ALC, B16, B35, BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C2C12, C3H-10T1/2, C6, C6/36, Cal-27, CGR8, CHO, CML T1, CMT, COR-L23, COR-L23/5010, COR-L23/CPR, COR-L23/R23, COS-7, COV-434, CT26, D17, DH82, DU145, DuCaP, E14Tg2a, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, Hepa1c1c7, клетки High Five, HL-60, HMEC, HT-29, HUVEC, клетки J558L, Jurkat, клетки JY, клетки K562, KCL22, KG1, Ku812, KYO1, LNCap, Ma-Mel 1, 2, 3....48, MC-38, MCF-10A, MCF-7, MDA-MB-231, MDA-MB-435, MDA-MB-468, MDCK II, MG63, MONO-MAC 6, MOR/0.2R, MRC5, MTD-1A, MyEnd, NALM-1, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NW-145, OPCN/OPCT Peer, PNT-1A/PNT 2, PTK2, Raji, клетки RBL, RenCa, RIN-5F, RMA/RMAS, S2, клетки Saos-2, Sf21, Sf9, SiHa, SKBR3, SKOV-3, T-47D, T2, T84, THP1, U373, U87, U937, VCaP, WM39, WT-49, X63, YAC-1 и клетки YAR.

[1113] Некоторые аспекты данного изобретения предусматривают клетки, включающие любой из описанных здесь конструктов. В некоторых вариантах осуществления клетка-хозяин транзиентно или нетранзиентно трансфицируется одним

или несколькими описанными здесь векторами. В некоторых вариантах осуществления клетка трансфицируется в естественном виде, в котором она находится в субъекте. В некоторых вариантах осуществления клетку, которая подвергается трансфекции, забирают от субъекта. В некоторых вариантах осуществления клетку получают из клеток, взятых от субъекта, например из линии клеток. В данной области известно большое количество линий клеток для культивирования тканей. Примеры линий клеток включают, но без ограничения: C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, эпителиальные клетки BS-C-1 почки обезьяны, эмбриональный фибробласт мыши BALB/3T3, 3T3 Swiss, 3T3-L1, фибробласты зародышей человека 132-d5; фибробласты мыши 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A 172, A20, A253, A431, A-549, ALC, B16, B35, клетки BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293. BxPC3. C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr −/−, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY cells, клетки K562, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK 11, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, линии клеток OPCN/OPCT, Peer, PNT-1A/PNT 2, RenCa, RIN-5F, RMA/RMAS, клетки Saos-2, Sf-9, SkBr3, T2, T-47D, T84, линию клеток THP1, U373, U87, U937, VCaP, клетки Vero, WM39, WT-49, X63, YAC-1, YAR, и их трансгенные варианты.

[1114] Линии клеток доступны из ряда источников, известных специалистам в данной области (см., например, Американскую коллекцию типовых культур (American Type Culture Collection, ATCC) (Manassus, Va.)). В некоторых вариантах осуществления клетка, трансфицированная одним или несколькими описанными здесь векторами, применяется для получения новой линии клеток, включающих одну или несколько последовательностей из вектора. В некоторых вариантах осуществления клетка,

транзиентно трансфицированная компонентами системы CRISPR, как описано здесь, (например в виде транзиентной трансфекции одним или несколькими векторами, или трансфекции РНК) и модифицированная за счет активности комплекса CRISPR, применяется для получения новой линии клеток, включающих модификацию, но не имеющих какой-либо другой экзогенной последовательности. В некоторых вариантах осуществления клетки транзиентно или нетранзиентно трансфицируют одним или несколькими описанными здесь векторами или линии клеток, полученные из подобных клеток, применяют при анализе одного или нескольких тестируемых соединений.

Векторы

[1115] Некоторые аспекты настоящего изобретения относятся к применению рекомбинантных вирусных векторов (например адено-ассоциированных вирусных векторов, аденовирусных векторов, или векторов на основе вируса простого герпеса) для доставки описанных здесь праймерных редакторов или их компонентов, например разделенного белка Cas9 или разделенных праймерных редакторов, в клетку. В случае подхода на основе разделенного PE N-концевая часть слитого белка PE и C-концевая часть продукта слияния PE доставляются с помощью отдельных рекомбинантных вирусных векторов (например адено-ассоциированных вирусных векторов, аденовирусных векторов, или векторов на основе вируса простого герпеса) в одну клетку, поскольку полноразмерный белок Cas9 или праймерные редакторы превышают лимит упаковывания различных вирусных векторов, например rAAV (~4,9 kb).

[1116] Таким образом, в одном варианте осуществления изобретение охватывает векторы, способные обеспечивать доставку слитых белков разделенного праймерного редактора или его разделенных компонентов. В некоторых вариантах осуществления предусмотрена композиция для доставки разделенного белка Cas9 или разделенного праймерного редактора в клетку (например клетку млекопитающего, клетку человека). В некоторых вариантах осуществления композиция по настоящему изобретению включает: (i) частицу первого рекомбинантного адено-ассоциированного вируса (rAAV), включающую первую нуклеотидную последовательность, кодирующую N-концевую часть белка Cas9 или праймерный редактор, слитый на его C-конце с интеином-N; и (ii) вторую частицу рекомбинантного адено-ассоциированного вируса (rAAV), включающую вторую нуклеотидную последовательность, кодирующую интеин-C, слитый с N-концом C-концевой части белка Cas9 или праймерного

редактора. rAAV частицы по настоящему изобретению включают rAAV вектор (то есть рекомбинантный геном rAAV), инкапсулированный в белках капсида вируса.

[1117] В некоторых вариантах осуществления rAAV вектор включает: (1) гетерологичный участок нуклеиновой кислоты, включающий первую или вторую нуклеотидную последовательность, кодирующую N-концевую часть или C-концевую часть разделенного белка Cas9 или разделенного праймерного редактора в любой форме, как описано здесь, (2) одну или несколько нуклеотидных последовательностей, включающих последовательность, которая обеспечивает экспрессию гетерологичного участка нуклеиновой кислоты (например промотор), и (3) один или несколько участков нуклеиновой кислоты, включающих последовательность, которая обеспечивает встраивание гетерологичного участка нуклеиновой кислоты (необязательно с одним или несколькими участками нуклеиновой кислоты, включающими последовательность, которая обеспечивает экспрессию), в геном клетки. В некоторых вариантах осуществления вирусные последовательности, которые обеспечивают встраивание, включают последовательности инвертированных концевых повторов (ITR). В некоторых вариантах осуществления первая или вторая нуклеотидная последовательность, кодирующая N-концевую часть или C-концевую часть разделенного белка Cas9 или разделенного праймерного редактора, фланкируется на каждой стороне последовательностью ITR. В некоторых вариантах осуществления вектор нуклеиновой кислоты дополнительно включает участок, кодирующий белок AAV Rep, как описано здесь, который содержится либо в участке, фланкируемом ITR, либо вне данного участка. Последовательности ITR могут быть получены из любого серотипа AAV (например 1, 2, 3, 4, 5, 6, 7, 8, 9, или 10) или могут быть получены из более чем одного серотипа. В некоторых вариантах осуществления последовательности ITR получены из AAV2 или AAV6.

[1118] Таким образом, в некоторых вариантах осуществления описанные здесь rAAV частицы включают, по меньшей мере, одну rAAV2 частицу, rAAV6 частицу, rAAV8 частицу, rPHP.B частицу, rPHP.eB частицу, или rAAV9 частицу, или их вариант. В определенных вариантах осуществления описанные rAAV частицы представляют собой rPHP.B частицы, rPHP.eB частицы, rAAV9 частицы.

[1119] Последовательности ITR и плазмиды, содержащие последовательности ITR, хорошо известны в данной области и являются коммерчески доступными (см., например, продукты и услуги, предоставляемые Vector Biolabs, Philadelphia, PA;

Cellbiolabs, San Diego, CA; Agilent Technologies, Santa Clara, Ca; и Addgene, Cambridge, MA; и Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein. Kessler PD, Podsakoff GM, Chen X, McQuiston SA, Colosi PC, Matelis LA, Kurtzman GJ, Byrne BJ. Proc Natl Acad Sci USA. 1996 Nov 26;93(24):14082-7; и Curtis A. Machida. Methods in Molecular Medicine™. Viral Vectors for Gene Therapy Methods and Protocols. 10.1385/1-59259-304-6:201 © Humana Press Inc. 2003. Chapter 10. Targeted Integration by Adeno-Associated Virus. Matthew D. Weitzman, Samuel M. Young Jr., Toni Cathomen and Richard Jude Samulski; патенты США 5139941 и 5962313, все из которых приведены здесь в качестве ссылки).

[1120] В некоторых вариантах осуществления rAAV вектор по настоящему изобретению включает один или несколько регуляторных элементов для контролирования экспрессии гетерологичного участка нуклеиновой кислоты (например промоторы, терминаторы транскрипции, и/или другие регуляторные элементы). В некоторых вариантах осуществления первая и/или вторая нуклеотидная последовательность функционально связана с одной или несколькими (например 1, 2, 3, 4, 5 или более) терминаторами транскрипции. Неограничивающие примеры терминаторов транскрипции, которые могут применяться в соответствии с настоящим изобретением, включают терминаторы транскрипции гена гормона роста коровы (bGH), гена гормона роста человека (hGH), SV40, CW3, ϕ, или их комбинации. Была протестирована эффективность нескольких терминаторов транскрипции для определения их соответствующих эффектов на уровень экспрессии разделенного белка Cas9 или разделенного праймерного редактора. В некоторых вариантах осуществления терминатор транскрипции, используемый в настоящем изобретении, представляет собой терминатор транскрипции bGH. В некоторых вариантах осуществления rAAV вектор дополнительно включает посттранскрипционный регуляторный элемент вируса гепатита сурка (WPRE). В определенных вариантах осуществления WPRE представляет собой усеченную последовательность WPRE, например "W3". В некоторых вариантах осуществления WPRE вставлен 5' относительно терминатора транскрипции. При транскрипции такие последовательности создают третичную структуру, которая повышает экспрессию, особенно из вирусных векторов.

[1121] В некоторых вариантах осуществления используемые здесь векторы кодируют белки слитого PE или любые из его компонентов (например napDNAbp, линкеры, или полимеразы). Кроме того, применяемые здесь векторы могут кодировать PEgRNA

и/или вспомогательную нРНК для никирования второй цепи. Векторы могут быть способны обеспечивать экспрессию одной или нескольких кодирующих последовательностей в клетке. В некоторых вариантах осуществления клетка может представлять собой прокариотическую клетку, например бактериальную клетку. В некоторых вариантах осуществления клетка может представлять собой эукариотическую клетку, например клетку дрожжей, растения, насекомого или млекопитающего. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку млекопитающего. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку грызуна. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку человека. Подходящие промоторы для обеспечения экспрессии в различных типах клеток известны в данной области. В некоторых вариантах осуществления промотор может относиться к дикому типу. В других вариантах осуществления промотор может быть модифицирован для более эффективной экспрессии. В других вариантах осуществления промотор может быть усечен, но все же сохранять свою функцию. Например, промотор может иметь нормальный размер или уменьшенный размер, являющийся подходящим для правильного упаковывания вектора в вирус.

[1122] В некоторых вариантах осуществления промоторы, которые могут применяться в рамках векторов праймерного редактора, могут являться конститутивными, индуцируемыми, или специфичными в отношении ткани. В некоторых вариантах осуществления промоторы могут представлять собой конститутивные промоторы. Неограничивающие примеры конститутивных промоторов включают немедленный ранний промотор цитомегаловируса (CMV), промотор вируса обезьяны (SV40), большой поздний промотор аденовируса (MLP), промотор вируса саркомы Рауса (RSV), промотор вируса опухоли молочной железы мыши (MMTV), промотор фосфоглицераткиназы (PGK), промотор фактора-альфа элонгации (EFla), промоторы убиквитина, промоторы актина, промоторы тубулина, промоторы иммуноглобулина, их функциональный фрагмент или комбинацию любых из них. В некоторых вариантах осуществления промотор может представлять собой промотор CMV. В некоторых вариантах осуществления промотор может представлять собой усеченный промотор CMV. В других вариантах осуществления промотор может представлять собой промотор EFla. В некоторых вариантах осуществления промотор может представлять

собой индуцируемый промотор. Неограничивающие примеры индуцируемых промоторов включают промоторы, которые индуцируются температурным шоком, светом, химическими веществами, пептидами, металлами, стероидами, антибиотиками, или спиртом. В некоторых вариантах осуществления индуцируемый промотор может представлять собой промотор, имеющий низкий базальный (неиндуцированный) уровень экспрессии, например промотор Tet-On® (Clontech). В некоторых вариантах осуществления промотор может представлять собой специфический в отношении ткани промотор. В некоторых вариантах осуществления специфический в отношении ткани промотор исключительно или преимущественно экспрессируется в ткани печени. Неограничивающие примеры специфических в отношении ткани промоторов включают промотор B29, промотор CD14, промотор CD43, промотор CD45, промотор CD68, промотор десмина, промотор эластазы-1, промотор эндоглина, промотор фибронектина, промотор Flt-1, промотор GFAP, промотор GPIIb, промотор ICAM- 2, промотор INF-β, промотор Mb, промотор Nphsl, промотор OG-2, промотор SP-B, промотор SYN1, и промотор WASP.

[1123] В некоторых вариантах осуществления векторы праймерного редактора (например включая любые векторы, кодирующие слитый белок праймерного редактора и/или PEgRNA и/или вспомогательные нРНК для никирования второй цепи) могут включать индуцируемые промоторы для начала экспрессии только после его доставки в клетку-мишень. Неограничивающие примеры индуцируемых промоторов включают промоторы, которые индуцируются температурным шоком, светом, химическими веществами, пептидами, металлами, стероидами, антибиотиками, или спиртом. В некоторых вариантах осуществления индуцируемый промотор может представлять собой промотор, имеющий низкий базальный (неиндуцированный) уровень экспрессии, например промотор Tet-On® (Clontech).

[1124] В дополнительных вариантах осуществления векторы праймерного редактора (например включая любые векторы, кодирующие слитый белок праймерного редактора и/или PEgRNA и/или вспомогательные нРНК для никирования второй цепи) могут включать специфические в отношении ткани промоторы для начала экспрессии только после его доставки в определенную ткань. Неограничивающие примеры специфических в отношении ткани промоторов включают промотор B29, промотор CD14, промотор CD43, промотор CD45, промотор CD68, промотор десмина, промотор эластазы-1, промотор эндоглина, промотор фибронектина, промотор Flt-1, промотор

GFAP, промотор GPIIb, промотор ICAM- 2, промотор INF-β, промотор Mb, промотор Nphsl, промотор OG-2, промотор SP-B, промотор SYN1, и промотор WASP.

[1125] В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая PEgRNA (или любые направляющие РНК, используемые в рамках праймерного редактирования) может быть функционально связана, по меньшей мере, с одной последовательностью контроля транскрипции или трансляции. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая направляющую РНК, может быть функционально связана, по меньшей мере, с одним промотором. В некоторых вариантах осуществления промотор может распознаваться РНК-полимеразой III (Pol III). Неограничивающие примеры Pol III промоторов включают промоторы U6, HI, и тРНК. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая направляющую РНК, может быть функционально связана с промотором U6 мыши или человека. В других вариантах осуществления нуклеотидная последовательность, кодирующая направляющую РНК, может быть функционально связана с промотором HI мыши или человека. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая направляющую РНК, может быть функционально связана с тРНК промотором мыши или человека. В вариантах осуществления с более чем одной направляющей РНК промоторы, применяемые для обеспечения экспрессии, могут являться одинаковыми или различающимися. В некоторых вариантах осуществления нуклеотид, кодирующий crРНК направляющей РНК, и нуклеотид, кодирующей tracrРНК направляющей РНК, могут быть внесены на одном векторе. В некоторых вариантах осуществления нуклеотид, кодирующий crРНК, и нуклеотид, кодирующей tracrРНК, могут быть внесены с помощью одного промотора. В некоторых вариантах осуществления crРНК и tracrРНК могут быть транскрибированы в один транскрипт. Например, crРНК и tracrРНК могут быть процессированы из одного транскрипта для образования направляющей РНК в виде двойной молекулы. В качестве альтернативы, crРНК и tracrРНК могут быть транскрибированы в направляющую РНК в виде одной молекулы.

[1126] В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая направляющую РНК, может быть расположена на одном векторе, включающем нуклеотидную последовательность, кодирующую слитый белок PE. В некоторых вариантах осуществления экспрессия направляющей РНК и слитого белка PE может быть обеспечена за счет их соответствующих промоторов. В некоторых

вариантах осуществления экспрессия направляющей РНК может быть обеспечена за счет того же промотора, который обеспечивает экспрессию слитого белка PE. В некоторых вариантах осуществления направляющая РНК и транскрипт слитого белка PE могут содержаться в одном транскрипте. Например, направляющая РНК может располагаться в нетранслируемом участке (UTR) транскрипта белка Cas9. В некоторых вариантах осуществления направляющая РНК может располагаться в 5' UTR транскрипта слитого белка PE. В других вариантах осуществления направляющая РНК может располагаться в 3' UTR транскрипта слитого белка PE. В некоторых вариантах осуществления внутриклеточный период полужизни транскрипта слитого белка PE может быть снижен за счет включения направляющей РНК в его 3' UTR с укорачиванием, тем самым, длины его 3' UTR. В дополнительных вариантах осуществления направляющая РНК может располагаться в интроне транскрипта слитого белка PE. В некоторых вариантах осуществления в интрон, в котором располагается направляющая РНК, могут быть добавлены подходящие участки сплайсинга, так что направляющая РНК подвергается сплайсингу из транскрипта правильным образом. В некоторых вариантах осуществления экспрессия белка Cas9 и направляющей РНК в непосредственной близости на одном векторе может обеспечивать более эффективное образование комплекса CRISPR.

[1127] Система праймерного редактирования на основе векторов может включать один вектор, или два вектора, или три вектора, или четыре вектора, или пять векторов, или более. В некоторых вариантах осуществления векторная система может включать один единственный вектор, который кодирует ка слитый белок PE, так и PEgRNA. В других вариантах осуществления векторная система может включать два вектора, при этом один вектор кодирует слитый белок PE, а другой кодирует PEgRNA. В дополнительных вариантах осуществления векторная система может включать три вектора, при этом третий вектор кодирует нРНК для никирования второй цепи, применяемую в рамках описанных здесь способов.

[1128] В некоторых вариантах осуществления композиция, включающая rAAV частицу (в любой предусмотренной здесь форме), дополнительно включает фармацевтически приемлемый носитель. В некоторых вариантах осуществления композиция включает подходящие фармацевтические носители для введения субъектам, представляющим собой человека или животных.

[1129] Некоторые примеры материалов, которые могут служить в качестве фармацевтически приемлемых носителей, включают: (1) сахара, такие как лактоза, глюкоза и сахароза; (2) крахмалы, такие как кукурузный крахмал и картофельный крахмал; (3) целлюлоза и ее производные, такие как карбоксиметилцеллюлоза натрия, метилцеллюлоза, этилцеллюлоза, микрокристаллическая целлюлоза и ацетат целлюлозы; (4) порошкообразный трагакант; (5) солод; (6) желатин; (7) смазывающие агенты, такие как стеарат магния, лаурилсульфат натрия и тальк; (8) наполнители, такие как масло какао и воски для суппозиториев; (9) масла, такие как арахисовое масло, хлопковое масло, сафлоровое масло, кунжутное масло, оливковое масло, кукурузное масло и соевое масло; (10) гликоли, такие как пропиленгликоль; (11) полиолы, такие как глицерин, сорбит, маннит и полиэтиленгликоль (ПЭГ); (12) сложные эфиры, такие как этилолеат и этиллаурат; (13) агар; (14) буферные агенты, такие как гидроксид магния и гидроксид алюминия; (15) альгиновую кислоту; (16) апирогенную воду; (17) изотонический физиологический раствор; (18) раствор Рингера; (19) этиловый спирт; (20) pH-буферные растворы; (21) сложные полиэфиры, поликарбонаты и/или полиангидриды; (22) наполнители, такие как полипептиды и аминокислоты; (23) компонент сыворотки, такой как сывороточный альбумин, HDL и LDL; (22) C2-C12 спирты, такие как этанол; и (23) другие нетоксичные совместимые вещества, используемые в фармацевтических составах. В составе также могут присутствовать смачивающие агенты, красители, агенты для высвобождения, агенты для покрытия, подсластители, ароматизаторы, отдушки, консерванты и антиоксиданты. Такие термины как "инертный наполнитель", "носитель", "фармацевтически приемлемый носитель" или им подобные используются здесь взаимозаменяемым образом.

Способы доставки

[1130] В некоторых аспектах изобретение предусматривает способы, включающие доставку одного или нескольких полинуклеотидов, например одного или нескольких векторов, как описано здесь, одного или нескольких их транскриптов, и/или одного или нескольких транскрибированных из них белков в клетку-хозяин. В некоторых аспектах изобретение дополнительно предусматривает клетки, полученные с помощью таких способов, и организмы (такие как животные, растения, или грибы), состоящие или полученные из таких клеток. В некоторых вариантах осуществления редактор

оснований, как описано здесь, в комбинации (и, необязательно, в виде комплекса) с направляющей последовательностью доставляется в клетку.

[1131] Примеры стратегий доставки описаны в других разделах здесь и включают стратегии на основе векторов, доставку на основе рибонуклеопротеинового комплекса РЕ, и доставку РЕ с помощью способов на основе иРНК.

[1132] В некоторых вариантах осуществления предусмотренный способ доставки включает нуклеофекцию, микроинъекцию, биолистику, вирусомы, липосомы, иммунолипосомы, конъюгаты поликатион или липид:нуклеиновая кислота, голую ДНК, искусственные вирионы, и агент-энхансированный захват ДНК.

[1133] Примеры способов доставки нуклеиновых кислот включают липофекцию, нуклеофекцию, электропорацию, стабильную геномную интеграцию (например piggyback), микроинъекцию, биолистику, вирусомы, липосомы, иммунолипосомы, конъюгаты поликатион или липид:нуклеиновая кислота, голую ДНК, искусственные вирионы, и агент-энхансированный захват ДНК. Липофекция описана, например, в патентах США 5049386, 4946787, и 4897355), а реагенты для липофекции являются коммерчески доступными (например Transfectam™, Lipofectin™ и SF Cell Line 4D-Nucleofector X Kit™ (Lonza)). Катионные и нейтральные липиды, которые являются подходящими для эффективной липофекции полинуклеотидов на основе распознавания рецептора, включают указанные у Feigner, WO 91/17424; WO 91/16024. Доставка может осуществляться в клетки (например при введении *in vitro* или *ex vivo*) или ткани-мишени (например при введении *in vivo*). Доставка может быть достигнута за счет использования RNP комплексов.

[1134] Получение комплексов липид:нуклеиновая кислота, включая липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области (см., например, Crystal, Science 270:404-410 (1995); Blaese *et al., Cancer Gene Ther.* 2:291-297 (1995); Behr *et al., Bioconjugate Chem.* 5:382-389 (1994); Remy *et al., Bioconjugate Chem.* 5:647-654 (1994); Gao *et al., Gene Therapy* 2:710-722 (1995); Ahmad *et al., Cancer Res.* 52:4817-4820 (1992); патенты США 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028, и 4946787).

[1135] В других вариантах осуществления способ доставки и предусмотренный здесь вектор представляет собой RNP комплекс. RNP доставка слитых белков заметно повышает специфичность в плане ДНК при праймерном редактировании. RNP доставка слитых белков ведет к устранению одновременного целевого и нецелевого

редактирования ДНК. RNP доставка обеспечивает устранение нецелевого редактирования в участках без повторов, в то же время обеспечивая целевое редактирование, сравнимое с плазмидной доставкой, и значительно снижает нецелевое редактирование ДНК даже участке 2 VEGFA с высоким содержанием повторов. См. Rees, H.A. *et al.,* Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery, *Nat. Commun.* 8, 15790 (2017), патент США 9526784, выданный 27 декабря 27 2016 г., и патент США 9737604, выданный 22 августа 2017 г., каждый из которых приведен здесь в качестве ссылки.

[1136] Дополнительные способы доставки нуклеиновых кислот в клетки известны специалистам в данной области. См., например, публикацию US 2003/0087817, приведенную здесь в качестве ссылки.

[1137] Другие способы настоящего изобретения предусматривают способы доставки конструктов праймерного редактора в клетку для образования полного и функционального праймерного редактора в клетке. Например, в некоторых вариантах осуществления клетка контактирует с описанной здесь композицией (например композициями, включающими нуклеотидные последовательности, кодирующие разделенный Cas9 или разделенный праймерный редактор или AAV частицы, содержащие векторы нуклеиновой кислоты, включающие такие нуклеотидные последовательности). В некоторых вариантах осуществления контактирование приводит к доставке таких нуклеотидных последовательностей в клетку, при этом N-концевая часть белка Cas9 или праймерного редактора и C-концевая часть белка Cas9 или праймерного редактора экспрессируются в клетке и соединяются для образования полного белка Cas9 или полного праймерного редактора.

[1138] Является очевидным, что любая предусмотренная здесь rAAV частица, молекула нуклеиновой кислоты или композиция может быть внесена в клетку любым подходящим способом, либо стабильно, или транзиентно. В некоторых вариантах осуществления описанные белки могут быть трансфицированы в клетку. В некоторых вариантах осуществления клетка может быть трансдуцирована или трансфицирована молекулой нуклеиновой кислоты. Например, клетка может быть трансдуцирована (например вирусом, кодирующим разделенный белок) или трансфицирована (например плазмидой, кодирующей разделенный белок) молекулой нуклеиновой кислоты, которая кодирует разделенный белок, или rAAV частицей, содержащей вирусный геном, кодирующий одну или несколько молекул нуклеиновой кислоты. Такая

трансдукция может представлять собой стабильную или транзиентную трансдукцию. В некоторых вариантах осуществления клетки, экспрессирующие разделенный белок или содержащие разделенный белок, могут быть трансдуцированы или трансфицированы одной или несколькими последовательностями направляющей РНК, например при доставке разделенного белка Cas9 (например nCas9). В некоторых вариантах осуществления плазмида, экспрессирующая разделенный белок, может быть внесена в клетки с помощью электропорации, транзиентной (например с помощью липофекции) и стабильной геномной интеграции (например с помощью piggybac), и вирусной трансдукции, или других способов, известных специалистам в данной области.

[1139] В определенных вариантах осуществления предусмотренные здесь композиции включают липид и/или полимер. В определенных вариантах осуществления липид и/или полимер является катионным. Получение подобных липидных частиц является хорошо известным. См., например, патенты США 4880635; 4906477; 4911928; 4917951; 4920016; 4921757; и 9737604, каждый из которых приведен здесь в качестве ссылки.

[1140] Последовательность направляющей РНК может иметь длину 15-100 нуклеотидов и включать последовательность из, по меньшей мере, 10, по меньшей мере, 15, или, по меньшей мере, 20 последовательных нуклеотидов, которая комплементарна нуклеотидной последовательности-мишени. Направляющая РНК может включать последовательность из 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, или 40 последовательных нуклеотидов, которая является комплементарной нуклеотидной последовательности-мишени. Направляющая РНК может иметь длину 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, или 50 нуклеотидов.

[1141] В некоторых вариантах осуществления нуклеотидная последовательность-мишень представляет собой последовательность ДНК в геноме, например эукариотическом геноме. В определенных вариантах осуществления нуклеотидная последовательность-мишень расположена в геноме млекопитающего (например человека).

[1142] Композиции по данному изобретению могут быть введены или упакованы в виде, например, единичной дозы. Термин "единичная доза" при использовании в

отношении фармацевтической композиции по настоящему изобретению обозначает физически отдельные единицы, являющиеся подходящими в качестве единичной дозы для субъекта, при этом каждая единица содержит предварительно определенное количество активного материала, рассчитанное для продуцирования требуемого терапевтического эффекта, совместно с требуемым разбавителем; то есть носителем или индифферентным веществом.

[1143] Лечение заболевания или расстройства включает замедление развития или прогрессирования заболевания, или снижение тяжести заболевания. Лечение заболевания не обязательно требует результатов в виде излечения.

[1144] В соответствии с используемым здесь значением, "замедление" развития заболевания означает задерживание, сдерживание, замедление, притормаживание, стабилизацию, и/или отсрочку прогрессирования заболевания. Данное замедление может иметь различную длительность в зависимости от истории заболевания и/или подвергаемых лечению субъектов. Способ, который обеспечивает "замедление" или облегчение развития заболевания или задержку развития заболевания, представляет собой способ, который снижает вероятность развития одного или нескольких симптомов заболевания в рамках определенного промежутка времени и/или снижение тяжести симптомов в рамках определенного промежутка времени при сравнении с отсутствием применения данного способа. Такие сравнения обычно основаны на клинических исследованиях, использовании ряда субъектов, являющихся достаточными для получения статистически значимого результата.

[1145] "Развитие" или "прогрессирование" заболевания означает начальные проявления и/или наступающее прогрессирование заболевания. Развитие заболевания может являться детектируемым и может быть проанализировано с помощью стандартных клинических способов, которые хорошо известны в данной области. Однако термин "развитие" также обозначает прогрессирование, которое может являться недетектируемым. Для целей данного описания термины "развитие" или "прогрессирование" обозначают биологическое течение симптомов. Термин "развитие" включает возникновение, повторное возникновение, и развитие.

[1146] В соответствии с используемым здесь значением "развитие" или "возникновение" заболевания включают первоначальное развитие и/или повторное возникновение. Стандартные способы, известные средним специалистам в области медицины, могут применяться для введения выделенного полипептида или

фармацевтической композиции субъекту в зависимости от типа подвергаемого лечению заболевания или участка заболевания.

[1147] Предполагается, что специалист в данной области может полностью реализовать настоящее изобретение на основе приведенного выше описания без дополнительных уточнений. Приведенные ниже конкретные варианты осуществления, таким образом, являются лишь иллюстративными и не ограничивают остальной объем описания каким-либо образом. Все упомянутые здесь публикации приведены в качестве ссылок для указанных здесь целей или описываемого предмета.

ПРИМЕРЫ

ПРИМЕР 1. ПРАЙМЕРНОЕ РЕДАКТИРОВАНИЕ (PE) ДЛЯ ВНЕСЕНИЯ ТОЧНЫХ НУКЛЕОТИДНЫХ ИЗМЕНЕНИЙ В ГЕНОМ

[1148] Целью является разработать трансформативную технологию геномного редактирования для точного и универсального внесения изменений отдельных нуклеотидов в геномах млекопитающих. Данная технология позволила бы ученым исследовать эффекты вариаций отдельных нуклеотидов практически в любом гене млекопитающих и потенциально обеспечила бы возможность терапевтического вмешательства для корректирования патогенных точечных мутаций у представляющих собой людей пациентов.

[1149] Использование системы на основе коротких регулярно расположенных группами палиндромных повторов (CRISPR) для геномного редактирования стало революцией в науках о жизни[1-3]. Несмотря на то, что модификация генов с помощью CRISPR теперь является стандартной практикой, точное внесение изменений отдельных нуклеотидов остается трудной задачей, несмотря на то, что оно необходимо для исследования или корректирования большого количества вызывающих заболевания мутаций. Направленная гомологичная репарация (HDR) способна обеспечивать реализацию подобных изменений, но связана с низкой эффективностью (зачастую <5%), потребностью в донорных матрицах для репарации ДНК, и негативными эффектами образования двуцепочечных разрывов ДНК (DSB). Недавно в лаборатории профессора Liu было разработано редактирование оснований, которое обеспечивает достижение эффективного редактирования отдельных нуклеотидов без DSB. В рамках редакторов оснований (BE) система CRISPR комбинируется с ферментами в виде модифицирующей основания дезаминазы для конвертации пар-

мишеней оснований C•G или A•T в A•T или G•C, соответственно[4-6]. Несмотря на то, что они уже широко используются исследователями по всему миру (в виде > 5000 BE конструктов лаборатории профессора Liu, поставленных Addgene), текущие BE обеспечивают конверсию лишь четырех из двенадцати возможных пар оснований и не способны корректировать малые вставки или делеции. Более того, объем нацеливания при редактировании оснований ограничен редактированием не являющихся мишенями оснований C или A, являющихся смежными с основанием-мишенью ("побочное редактирование"), а также необходимостью наличия последовательности PAM на расстоянии 15±2 bp от основания-мишени. Таким образом, преодоление данных ограничений значительно расширит варианты основного исследовательского и терапевтического использования геномного редактирования.

[1150] Здесь предлагается разработать новый подход точного редактирования, который обеспечивает множество преимуществ редактирования оснований, а именно устранение двуцепочечных разрывов и потребности в донорных матрицах для репарации ДНК, в то же время обеспечивая преодоление его основных ограничений. Для достижения данной амбициозной цели предлагается напрямую вносить подвергнутые редактированию цепи ДНК в геномные участки-мишени с помощью мишень-инициированной обратной транскрипции (TPRT). В рамках описываемой здесь конструкции направляющая РНК CRISPR (нРНК) будет сконструирована таким образом, чтобы нести матрицу, кодирующую синтез мутагенной цепи ДНК, для выполнения ассоциированным ферментом в виде обратной транскриптазы (RT). Никированный нуклеазой (Cas9) CRISPR участок-мишень ДНК будет служить в качестве праймера для обратной транскрипции, обеспечивая прямое встраивание любого требуемого нуклеотидного изменения.

Раздел 1

[1151] Реализация обратной транскрипции мутагенных цепей ДНК с направляющей РНК в качестве матрицы. Предшествующие исследования показали, что после расщепления ДНК, но до диссоциации комплекса Cas9 высвобождает не являющуюся мишенью цепь ДНК для высвобождения свободного 3′-конца. Предполагается, что данная цепь ДНК является доступной для удлинения ферментами в виде полимеразы и что нРНК могут быть сконструированы путем удлинения их 5′ или 3′-конца для использования в качестве матриц для синтеза ДНК. В предварительных *in vitro* исследованиях было определено, что никированные цепи ДНК в комплексах в виде

связанных Cas9:нРНК действительно могут праймировать обратную транскрипцию с использованием связанной нРНК в качестве матрицы (фермент RT in trans). Далее будут определены различные линкеры нРНК , участки связывания праймера, и матрицы для синтеза для определения оптимальных правил конструирования *in vitro*. Затем будут оценены *in vitro* различные ферменты RT, действующие in trans или в виде продуктов слияния с Cas9. Наконец будут идентифицированы конструкции модифицированной нРНК, которые сохраняют эффективную активность в виде связывания и осуществления разрезов в клетках. Успешное достижение данной цели заложит основу для реализации синтеза мутагенных цепей в клетках.

Раздел 2

[1152] Реализация праймерного редактирования в клетках человека. На основе механизмов процессирования и репарации ДНК предполагается, что мутагенные цепи ДНК (одноцепочечные флэпы) могут применяться для обеспечения специфичного и эффективного редактирования нуклеотидов-мишеней. В рамках внушающих оптимизм предварительных исследования была определена реализуемость данной стратегии на основе демонстрации редактирования с использованием модели в виде плазмидных субстратов, содержащих мутагенные флэпы. Одновременно с Задачей 1 будут дополнительно проанализированы результаты репарации на основе систематически варьирующейся длины мутагенного флэпа, композиции последовательностей, идентичности целевых нуклеотидов, и 3′-конца. Также будут протестированы малые вставки и делеции длиной от 1 до 3 нуклеотидов. Параллельно и отдельно от Задачи 1 будет проанализирована архитектура Cas9-RT, включая слитые белки и стратегии нековалентного рекрутирования. Архитектура Cas9-RT и удлиненные нРНК будут проанализированы на предмет клеточного редактирования в нескольких участках-мишенях в геноме человека и затем будут оптимизированы для высокой эффективности. В случае успеха данная цель немедленно сделает возможным TPRT геномное редактирование (то есть праймерное редактирование) для использования в рамках фундаментальных наук.

Раздел 3

[1153] Реализация специфического в отношении участка редактирования патогенных мутаций в культивируемых клетках человека. Потенциальная универсальность данной технологии может обеспечивать редактирования трансверсивных мутаций и инделей, которые в настоящее время не могут быть скорректированы с помощью BE. На основе

результатов, полученных в рамках Задачи 1 и Задачи 2, патогенные трансверсивные мутации будут являться мишенями в рамках культивируемых клеток человека, включая вызывающую серповидноклеточную анемию мутацию в бета-глобине (для корректирования требуется трансверсия A•T на T•A) и наиболее частую мутацию при болезни Уилсона в ATP7B (для корректирования требуется трансверсия G•C на T•A). Также будет проанализирована коррекция мутаций в виде малых вставок и делеций , включая 3-нуклеотидную делецию ΔF508 в *CFTR*, которая вызывает муковисцидоз. В случае успеха это заложит основу для разработки мощных терапевтических подходов, которые направлены на лечение данных важных заболеваний человека.

Подход

[1154] Целью является разработать стратегию геномного редактирования, которая обеспечивает прямое внесение точечных мутаций в геномных участках-мишенях. В рамках фазы разработки технологии усилия сосредоточатся на конструировании белка и РНК для встраивания функциональности TPRT в систему CRISPR/Cas. Будут применяться *in vitro* анализы для осторожной оценки функции каждого этапа TPRT начиная с самого начала (Задача 1). На втором этапе будут оценены результаты редактирования в клетках млекопитающих с применением комбинации модели в виде субстратов и модифицированных CRISPR/Cas систем (Задача 2). Наконец, на этапе использования технология будет применяться для корректирования мутаций, которые являлись недостижимыми при геномном редактировании с использованием других способов (Задача 3).

[1155] Общая схема редактирования показана на Фиг.1A-1B. Никазы Cas9 содержат деактивирующие мутации в домене нуклеазы HNH (Spy Cas9 H840A или N863A), рестриктирующие расщепление ДНК до PAM-содержащей цепи (не являющейся мишенью цепи). Направляющие РНК (нРНК) сконструированы для включения матрицы для обратной транскрипции (конструкции подробно показаны на слайде 5). Показано 5' удлинение нРНК, но также могут быть предусмотрены и 3' удлинения. Никаза Cas9 слита с ферментом в виде обратной транскриптазы (RT) через C-конец или N-конец. Комплекс нРНК:Cas9-RT нацелен на представляющий интерес участок ДНК и образует R-петлю после вытеснения не являющейся мишенью цепи. Cas9 разрывает не являющуюся мишенью цепь ДНК. Высвобождение никированной цепи освобождает свободный 3′-OH конец, который может праймировать обратную транскрипцию с использованием удлиненной нРНК в качестве матрицы. Данная

реакция синтеза ДНК осуществляется с помощью фермента в виде слитой RT. нРНК матрица кодирует последовательность ДНК, которая гомологична исходному ДНК дуплексу за исключением нуклеотида, который является мишенью для редактирования. Продукт обратной транскрипции представляет собой одноцепочечный ДНК-флэп, который кодирует требуемое изменение. Данный флэп, который содержит свободный 3′-конец, может уравновешиваться со смежной цепью ДНК, что приводит к образованию 5′ флэпа. Предполагается, что данный вид флэпов служит в качестве эффективного субстрата для FEN1 (флэп-эндонуклеазы 1), представляющей собой фермент, который естественным образом обеспечивает эксцизию 5′ флэпов из фрагментов Оказаки при синтезе запаздывающей цепи ДНК и удаление 5′ флэпов после синтеза с замещением цепи, который осуществляется при репарации эксцизии длинных последовательностей оснований. Лигирование никированной ДНК приводит к образованию несовпадающей пары оснований. Данный промежуточный продукт может либо подвергнуться возврату к исходной паре оснований или конверсии в требуемую подвергнутую редактированию пару за счет процессов репарации несовпадений (MMR). В качестве альтернативы, полуконсервативная репликация ДНК может обеспечивать получение по одной копии каждого из продукта с возвратом и изменением.

1. Реализация обратной транскрипции мутагенных цепей ДНК с направляющей РНК в качестве матрицы.

Предпосылки и обоснование

[1156] В рамках предложенной стратегии геномного редактирования никированная Cas9 не являющаяся мишенью цепь ДНК (PAM-содержащая цепь, которая образует R-петлю) действует в качестве праймера для синтеза ДНК. Предполагается, что это возможно на основе нескольких фрагментов биохимических и структурных данных. Эксперименты по защите на основе нуклеазы[32], кристаллографические исследовани[33], и окна редактирования оснований[4,24] продемонстрировали высокую степень гибкости и разнообразия нуклеотидов не являющейся мишенью цепи от -20 до -10 в рамках так называемой R-петли Cas9-связанного комплекса (нумерация обозначает расстояние 5′ от первого нуклеотида PAM). Более того, удаленная от PAM часть расщепленной не являющейся мишенью цепи может быть вытеснена из тесно связанных третичных комплексов при добавлении комплементарной оцДНК in trans[20]. Данные исследования поддерживают вывод о том, что не являющаяся мишенью цепь имеет высокую

гибкость, доступна для ферментов, и что после никирования 3′-конец удаленного от PAM фрагмента высвобождается до диссоциации Cas9. Более того, предполагается, что нРНК могут быть удлинены для функционирования в качестве матриц для синтеза ДНК. Предшествующие исследования показали, что нРНК для SpCas9, SaCas9, и LbCas12a (ранее Cpf1) переносят удлинение нРНК РНК с использованием аптамеров[34], лиганд-индуцируемых саморасщепляемых рибозимов[35], и длинных некодирующих РНК[36]. Данное описание образует прецедент в плане двух основных свойств, которые будут использоваться. При оценке данной стратегии будут проанализированы некоторые системы CRISPR-Cas наряду с конструкциями 5′ и 3′ удлиненной нРНК с использованием комбинации *in vitro* и клеточных анализов (Фиг.2A-2C).

[1157] Конструкции модифицированных нРНК для TRT редактирования показаны на Фиг.3A-3B. Синтез ДНК осуществляется от 5′ к 3′ с копированием, таким образом, РНК матрицы в направлении от 3′ к 5′. Конструкция 5′ удлинения содержит линкерный участок, участок связывания праймера, где отжигается никированная цепь ДНК, и матрицу для синтеза ДНК с помощью обратной транскрипции. 3′ удлиненная нРНК содержит участок связывания праймера и матрицу обратной транскрипции. В некоторых случаях 3′ РНК шпилька ядра нРНК модифицирована для соответствия последовательности-мишени ДНК, поскольку *in vitro* эксперименты показали, что обратная транскрипция приводит к ~3-нуклеотидному удлинению в ядре нРНК для конструктов 3′ удлиненной нРНК (похоже, что модификация последовательности шпильки хорошо переносится при условии реализации компенсаторных изменений, которые сохраняют структуру РНК шпильки). Синтез ДНК осуществляется от 5′ к 3′ с добавлением нуклеотидов 3′ OH растущей цепи ДНК.

Предварительные результаты

[1158] Никированная Cas9 ДНК праймирует обратную транскрипцию нРНК матриц. Для оценки доступности никированной не являющейся мишенью цепи ДНК могут быть проведены *in vitro* биохимические анализы с использованием нуклеазы Cas9 из *S. pyogenes* (SpCas9) и флуоресцентно меченных Cy5 субстратов в виде ДНК дуплексов (51 пара оснований). Сначала была получена группа нРНК, содержащих 5′ удлинения с различной длиной матрицы для синтеза с помощью *in vitro* транскрипции (общая конструкция показана на Фиг.2B). Анализы сдвига электрофоретической подвижности (EMSA) с не имеющим нуклеазной активности Cas9 (dCas9) показали, что 5′ удлиненные нРНК сохраняют аффинность связывания мишени (данные не показаны).

Далее была протестирована активность TPRT на подвергнутых предварительному никированию меченных Cy5 субстратах в виде ДНК дуплексов с использованием dCas9, 5'-удлиненных нРНК, и обратной транскриптазы вируса лейкоза мышей Молони (M-MLV) (Superscript III). После 1 часа инкубирования при 37 ˚C продукты анализировали с помощью денатурирующего электрофореза на основе полиакриламидного геля (PAGE) и подвергали визуализации на основе флуоресценции Cy5 (Фиг.4A). Каждый из вариантов 5′-удлиненной нРНК приводил к значительному образованию продукта, при этом наблюдаемые размеры продукта в виде ДНК соответствовали длине удлиненных матриц (Фиг.4B). Является важным то, что при отсутствии dCas9 подвергнутые предварительному никированию субстраты удлинялись до полной 51 bp длины ДНК субстрата, что является сильным свидетельством того, что комплементарная цепь ДНК, а не нРНК, была использована в качестве матрицы для синтеза ДНК, когда dCas9 отсутствовал (Фиг.4C). Следует отметить, что система была сконструирована таким образом, что новая синтезированная цепь ДНК отражает продукт, который будет необходим для редактирования участка-мишени (гомологичной цепи с изменением одного нуклеотида). Данный результат приводит к тому, что связывание Cas9:нРНК высвобождает 3′-конец никированной не являющейся мишенью цепи, а также тому, что не являющаяся мишенью цепь является доступной для обратной транскрипции.

[1159] Далее оценивали не подвергнутые никированию дцДНК субстраты с использованием мутанта Cas9(H840A), который осуществляет разрыв не являющейся мишенью цепи ДНК. Сначала для тестирования никазной активности Cas9(H840A) с 5′-удлиненными нРНК проводили *in vitro* анализы расщепления, как описано ранее[37]. Несмотря на то, что никирование было нарушено по сравнению со стандартной нРНК, были образованы значительные продукты расщепления (Фиг.4D). Является важным то, что продукты RT также наблюдались при осуществлении TPRT реакций с 5′-удлиненными нРНК и Cas9(H840A), хотя и с более низким выходом, который, вероятно, может быть объяснен сниженной никирующей активностью (Фиг.4D). Данный результат показывает, что комплексы в виде 5′-удлиненная нРНК:Cas9(H840A) могут никировать ДНК и служить матрицей для обратной транскрипции.

[1160] Наконец 3′ удлинения нРНК оценивали на предмет никирования Cas9(H840A) и TPRT. В отличие от 5′-удлиненных нРНК, расщепление ДНК за счет 3′-удлиненных нРНК не ухудшалось до какой-либо детектируемой степени по сравнению со

стандартной нРНК. Является значимым то, что матрицы в виде 3′-удлиненных нРНК также поддерживают эффективную обратную транскрипцию как с подвергнутыми предварительному никированию, так и интактными субстратами в виде ДНК дуплексов при поступлении RT M-MLV in trans (Фиг.4E). Удивительным является то, что только один продукт наблюдался для 3′-удлиненных матриц, что свидетельствует о том, что обратная транскрипция терминируется в определенном участке в рамках каркаса нРНК. Добавление гомополимерного хвоста продукта с использованием концевой трансферазы с последующим удлинением фрагментом Klenow и секвенированием по Сэнгеру показало, что копировалась полная нРНК матрица для синтеза наряду с 3 концевыми нуклеотидами ядра нРНК. В будущем флэп-конец будет перепрограммирован для модифицирования концевой последовательности нРНК[38,39]. Этот результат показывает, что 3′-удлиненные нРНК могут служить в качестве эффективных направляющих для нацеливания нуклеаз и могут служить матрицами для обратной транскрипции.

[1161] Cas9-TPRT использует никированную ДНК и нРНК in cis. Эксперименты на основе двойного цвета использовали для определения того, осуществляется ли предпочтительным образом RT реакция с нРНК in cis (связывание с тем же комплексом) (см. Фиг.8). Два отдельных эксперимента проводили для 5′-удлиненных и 3′-удлиненных нРНК. Для указанного эксперименте в отдельных пробирках получали третичные комплексы dCas9, нРНК, и ДНК субстрата. В одной пробирке нРНК кодирует длинный RT продукт, а ДНК субстрат помечен Cy3 (красный); в другой пробирке нРНК кодирует короткий RT продукт, а ДНК субстрат помечен Cy5 (синий). После короткого инкубирования комплексы смешивали и затем обрабатывали ферментом RT и dNTP. Продукты разделяли с помощью денатурирующего PAGE с применением мочевины и визуализировали с помощью флуоресценции в Cy3 и Cy5 каналах. Определяли, что продукты реакции предпочтительно образуются с использованием нРНК матрицы, которая предварительно образовывала комплексы с ДНК субстратом, что свидетельствует о том, что RT реакция вероятно происходит in cis. Эти результаты поддерживают вывод о том. что один комплекс Cas9:нРНК может нацеливаться на участок ДНК и служить матрицей для обратной транскрипции мутагенной цепи ДНК.

Тестирование TPRT с другими Cas системами

[1162] Эксперименты, аналогичные описанным в предыдущих разделах, будут проведены с использованием других Cas систем, включая Cas9 из *S. aureus* и Cas12a из *L. bacterium* (см. Фиг.2A-2C). Если TPRT также будет продемонстрирована для данных вариантов Cas, то потенциальный объем редактирования и вероятность общего успеха в клетках будут повышены.

Тестирование TPRT со слитыми белками RT-Cas9

[1163] Сначала будет проанализирована группа коммерчески доступных или способных быть подвергнутыми очистке ферментов RT in trans на предмет TPRT активности. Помимо уже протестированной RT из M-MLV, будет протестирована RT из вируса миобластоза птиц (AMV), интрон Группы II *Geobacillus stearothermophilus* (GsI-IIC)[41,42], и интрон Группы II *Eubacterium rectale* (Eu.re.I2)[43,44]. Является значимым то, что последние две RT реализуют TPRT в рамках их природных биологических контекстов. Где это возможно будут протестированы деактивирующие РНКазу мутации и другие потенциально полезные модификации ферментов RT. После идентификации функциональных RT при их поступлении in trans каждая из них будет проанализирована в виде слитого белка с вариантами Cas9. Будут протестированы как N-концевые, так и N-концевые ориентация продукта слияния наряду с различной длиной и архитектурой линкера. Будут использованы эксперименты на основе кинетического времени для определения того, может ли TPRT осуществляться при использовании ферментов RT in cis. Если может быть сконструирована архитектура продукта слияния RT-Cas9, обеспечивающая эффективные химические процессы в рамках TPRT, то это значительно повысит вероятность функционального редактирования в контексте клетки.

Нацеливание Cas9 с использованием модифицированных нРНК в клетках

[1164] Модифицированные нРНК-кандидаты, разработанные в рамках предыдущих подзадач, будут проанализированы в рамках эксперимента на основе культуры клеток человека (HEK293) для подтверждения эффективности нацеливания Cas9. Путем использования стандартных анализов образования инделей на основе SpCas9 дикого типа[45], модифицированные нРНК будут подвергнуты всестороннему сравнению со стандартными нРНК в рамках 5 или более участков в геноме человека. Эффективность геномного редактирования будет описана на основе мультиплексного секвенирования ампликонов с помощью расположенной в лаборатории платформы Illumina MiSeq. Предполагается, что результаты, полученные в рамках этого и предыдущих разделов,

обеспечат возможность получения информации для последующих итераций цикла разработки-получения-тестирования, в рамках которых нРНК могут быть оптимизированы как в плане обратной транскрипции на основе матриц, так и эффективного нацеливания Cas9 в клетках.

[1165] Результаты *in vitro* анализов показаны на Фиг.5-7. *In vitro* эксперименты продемонстрировали, что никированная не являющаяся мишенью цепь ДНК является гибкой и доступной для праймирования синтеза ДНК и что удлинение нРНК может служить матрицей для обратной транскрипции (см. Фиг.5). В рамках данного набора экспериментов использовались 5′-удлиненные нРНК (сконструированные как показано на Фиг.3A-3B) с матрицами для синтеза различной длины (перечислены слева). В данном наборе экспериментов флуоресцентно меченные (Cy5) ДНК-мишени использовали в качестве субстратов и подвергали предварительному никированию. Использованный в этих экспериментах Cas9 представляет собой каталитически мертвый Cas9 (dCas9), поэтому он не может осуществлять разрывы ДНК, но все же может эффективно связываться. Superscript III, являющийся коммерческой RT, получаемой на основе вируса лейкоза мышей Молони (M-MLV), поступал in trans. Сначала из очищенных компонентов были образованы комплексы dCas9:нРНК. Затем добавляли флуоресцентно меченный субстрат ДНК совместно с dNTP и ферментом RT. Спустя 1 час инкубирования при 37 ºC продукты реакции анализировали с помощью денатурирующего электрофореза на основе полиакриламидного геля с применением мочевины (PAGE). Изображение на геле показывает удлинение исходной цепи ДНК до показателей длины, которые соответствуют длине матрицы обратной транскрипции. Следует отметить, что реакции осуществляли при отсутствии полученных с помощью dCas9 продуктов ДНК с длиной в виде 51 нуклеотида независимо от использованной нРНК. Данный продукт соответствует использованию комплементарной цепи ДНК в качестве матрицы для синтеза ДНК, а не РНК (данные не показаны). Таким образом, для обеспечения синтеза ДНК на РНК матрице необходимо связывание с Cas9. Данный набор *in vitro* экспериментов похож на эксперименты, показанные на Фиг.5, за исключением того, что ДНК субстрат не подвергнут предварительному никированию и используется никаза Cas9 (мутант SpyCas9 H840A). Как показано на геле, никаза эффективно расщепляет цепь ДНК при использовании стандартной нРНК (нРНК_0, дорожка 3). Наблюдается множество продуктов расщепления, что соответствует предыдущим биохимическим исследованиям SpyCas9. 5′ удлинение нарушает

никирующую активность (дорожки 4-8), однако некоторый продукт RT все же наблюдается. На Фиг.7 показано, что 3′ удлинения поддерживают синтез ДНК и не оказывают значительного влияния на активность никазы Cas9. Подвергнутые предварительному никированию субстраты (черная стрелка) практически полностью конвертируются в продукты RT при использовании dCas9 или Cas9 никазы (дорожки 4 и 5). Более чем 50% конверсия в продукт RT (красная стрелка) наблюдается с полными субстратами (дорожка 3). Для определения длины и последовательности продукта RT полосу продукта выделяли из геля, отделяли и секвенировали. Это показывало, что RT обеспечивало 3-нуклеотидное удлинение в 3′-концевой шпильке ядра нРНК. Последующие эксперименты (не показаны) демонстрировали, что данные три нуклеотида могут быть изменены для соответствия последовательностям-мишеням ДНК при условии, что были реализованы комплементарные изменения для сохранения структуры шпильки РНК.

Потенциальные сложности и альтернативы

[1166] (1) RT не функционирует в качестве продукта слияния: молекулярная скученность и/или нежелательные геометрические свойства могут препятствовать полимеразному удлинению за счет слитых с Cas9 ферментов RT. Сначала может быть протестирована оптимизация линкера. Будут проанализированы циркулярно пермутированные варианты Cas9, которые могут обеспечивать переориентацию пространственной связи между ДНК праймером, нРНК и ферментом RT. Могут быть протестированы стратегии нековалентного рекрутирования RT, как подробно описано в Задаче 2. (2) Сниженная эффективность нацеливания Cas за счет вариантов удлиненной нРНК: это скорее всего является проблемой для 5′-удлиненных нРНК. На основе структурных данных[24] могут быть сконструированы и подвергнуты скринингу мутанты Cas9 для идентификации вариантов с большей устойчивостью к удлинению нРНК. Кроме того, библиотеки нРНК могут быть подвергнуты скринингу в клетках на предмет линкеров, которые повышают связанную с нацеливанием активность.

Значимость

[1167] Данные предварительные результаты показывают, что никазы Cas9 и удлиненные нРНК могут инициировать мишень-инициированную обратную транскрипцию на связанных ДНК-мишенях с использованием обратной транскриптазы, поступающей in trans. Важно, что было определено, что связывание Cas9 является критически важным для образования продукта. Хотя это вероятно не

является абсолютным требованием для геномного редактирования в клетках, дальнейшее развитие системы, включающей функцию фермента RT in cis, будет значительно повышать вероятность успеха в рамках вариантов использования на основе клеток. Реализация остальных аспектов в рамках этой Задачи заложит молекулярную основу для осуществления точного геномного редактирования в контексте генома человека.

2. Реализация праймерного редактирования в клетках человека.

Предпосылки и обоснование

[1168] В рамках предлагаемой стратегии комплекс в виде модифицированной RT-Cas9:нРНК обеспечит внесение мутагенных 3′ ДНК-флэпов в геномные участки-мишени. Предполагается, что мутагенные 3′ флэпы, содержащие одно несовпадение, будут встроены с помощью механизма репарации ДНК за счет энергетически доступного уравнивания со смежными 5′ флэпами, которые будут предпочтительно удалены (Фиг.1C-1D). Механизмы репликации и репарации ДНК встречают 5′ оцДНК-флэпы при процессировании фрагментов Оказаки[46] и при репарации эксцизии длинных последовательностей оснований (LP-BER)[47]. 5′ флэпы являются предпочтительными субстратами для широко экспрессируемой флэп-эндонуклеазы FEN1, которая рекрутируется в участки репарации ДНК за счет гомотримерного скользящего зажимного комплекса PCNA[48]. PCNA также служит каркасом для одновременного рекрутирования других факторов репарации, включая ДНК-лигазу Lig1[49]. Действуя в качестве "пояса с инструментами", PCNA ускоряет последовательное расщепление и лигирование флэпов, что является критически важным для процессирования миллионов фрагментов Оказаки, образованных при каждом делении клеток[50,51]. На основе аналогии с данными природными промежуточными продуктами ДНК, предполагается, что мутагенные цепи будут встроены с помощью уравнивания с 5′ флэпами с последующей скоординированной эксцизией и лигированием 5′ флэпов. Затем должна произойти репарация несовпадений (MMR) на любой цепи с равной вероятностью, что ведет к редактированию или возврату (Фиг.1C-1D). В качестве альтернативы сначала может произойти репликация ДНК, которая ведет напрямую к встраиванию изменения в новую синтезированную дочернюю цепь. В то время как наиболее высокий ожидаемый выход данного процесса составляет 50%, множественные попытки редактирования субстратов могут сдвигать реакцию к завершению вследствие необратимости репарации изменения.

Предварительные результаты

[1169] ДНК-флэпы индуцируют специфический в отношении участка мутагенез в модели в виде плазмидных субстратов в дрожжах и клетках HEK. Для тестирования предложенной стратегии редактирования были инициированы исследования с использованием модели в виде плазмидных субстратов, содержащих мутагенные 3′ флэпы, которые похожи на продукт TPRT. Был создан двойной флуоресцентный репортерный белок, который кодирует стоп-кодон между GFP и mCherry. Мутагенные флэпы кодируют корректирование стоп-кодона (Фиг.9A), что обеспечивает синтез mCherry. Таким образом, эффективность мутагенеза может быть количественно проанализирована на основе соотношений GFP:mCherry. Плазмидные субстраты были получены *in vitro* и внесены в дрожжи (*S. cerevisiae*) или клетки человека (HEK293). В обеих системах наблюдался высокочастотный мутагенез (Фиг.9B), при этом выделенные колонии дрожжей содержали либо подвергнутое возврату основание, мутированное основание, или смесь обоих продуктов (Фиг.9C). Детектирование последнего свидетельствует о том, что плазмидная репликации произошла до MMR в данных случаях, и дополнительно свидетельствует о том, что эксцизия и лигирование флэпа предшествует MMR. Данный результат свидетельствует об осуществимости редактирования ДНК с использованием 3′ мутагенных цепей.

- ***Системные исследования с использованием модели в виде флэп-субстратов***

[1170] На основе описанных выше предварительных результатов будет проанализирован более широкий спектр флэп-субстратов в клетках HEK для определения принципов эффективного редактирования. 3′ оцДНК-флэпы будут систематически варьироваться для определения влияния образований пар с несовпадениями, расположения мутагенного нуклеотида в рамках флэпа, и идентичности концевого нуклеотида (Фиг.9D). Также будут протестированы вставки и делеции отдельных нуклеотидов. Для анализа точности редактирования будет использоваться ампликонное секвенирование. Данные результаты помогут определить конструкцию нРНК матриц для обратной транскрипции.

[1171] *In vitro* TPRT на плазмидных субстратах ведет к эффективным результатам редактирования. TPRT реакции, разработанные в рамках Задачи 1, использовали для индуцирования мутагенеза в плазмидном субстрате. Реакцию осуществляли на плазмидных субстратах циркулярной ДНК (см. Фиг.10). Это исключает вероятность диссоциации цепи ДНК в качестве механизма удлинения RT в рамках предыдущих *in*

vitro экспериментов. Это также обеспечивало возможность тестирования репарации ДНК флэп-субстратов в клетках. Репортерную плазмиду с двойной флуоресценцией конструировали для экспрессии в дрожжах (S. cerevisiae). Эта плазмида кодирует GFP (зеленый флуоресцентный белок) и mCherry (красный флуоресцентный белок) с промежуточным стоп-кодоном (TGA). Экспрессия данного конструкта в дрожжах продуцирует только GFP. Плазмида была использована в качестве субстрата для *in vitro* TRT [Cas9(H840A) никаза, модифицированная нРНК, фермент RT MLV, dNTP]. Удлинение нРНК кодирует мутацию в стоп-кодоне. Цепь флэпа используется для репарации стоп-кодона и, как ожидается, будет продуцировать плазмиду, которая экспрессирует как GFP, так и mCherry в виде слитого белка. На Фиг.10 показаны дрожжевые плазмидные трансформанты с двойным флуоресцентным белком. Трансформирование исходной плазмидой или *in vitro* Cas9(H840A) никированной плазмидой приводит к получению колоний, экспрессирующих только зеленый GFP. Реакция TRT с 5′-удлиненными или 3′-удлиненными нРНК продуцирует смесь зеленых и желтых колоний. Последние экспрессируют как GFP, так и mCherry. Еще больше желтых колоний наблюдается при использовании 3′-удлиненной нРНК. Также показан положительный контроль, который не содержит стоп-кодон.

[1172] Этот результат показывает, что длинные двуцепочечные субстраты могут подвергаться TPRT, и что продукты TPRT индуцируют редактирование в эукариотических клетках.

[1173] Был осуществлен другой эксперимент, аналогичных указанному выше эксперименту на основе праймерного редактирования, но вместо внесения точечной мутации в стоп-кодон TRT редактирование обеспечивает внесение вставки (слева) или делеции (справа) в виде одного нуклеотида, которая корректирует мутацию сдвига рамки считывания и обеспечивает возможность синтеза mCherry ниже (см. Фиг.11). В обоих экспериментах использовались 3′-удлиненные нРНК. Отдельные колонии из TRT трансформантов отбирали и анализировали с помощью секвенирования по Сэнгеру (см.Фиг.12). Зеленые колонии содержали плазмиды с исходной последовательность ДНК, в то время как желтые колонии содержали точную мутацию, осуществленную с помощью нРНК для TRT редактирования. Другие точечные мутации или индели не наблюдались.

Реализация праймерного редактирования в клетках HEK с использованием архитектур RT-Cas9

[1174] Оптимизированные конструкты из предшествующих задач будут адаптированы для экспрессии в млекопитающих и редактирования в участках-мишенях в геноме человека. Будет протестировано множество ферментов RT и архитектур слияния, помимо смежного нацеливания с использованием вторичных нРНК (являющихся усеченными для предотвращения никирования). Также будет проанализировано нековалентное рекрутирование RT с использованием системы Sun-Tag[52] и системы на основе аптамера MS2[53]. Будут использоваться анализы образования инделей для оценки эффективности нацеливания с использованием стандартных нРНК и продуктов слияния RT-Cas9 (как описано выше). Затем для каждого участка генома будут проанализированы удлиненные нРНК и пары RT-Cas9 на предмет изменения в виде одного нуклеотида. Результаты редактирования будут проанализированы с помощью MiSeq.

[1175] Исходные эксперименты на основе клеток HEK были проведены с использованием продуктов слияния Cas9-RT. Редактирование с помощью компонентов, экспрессируемых в клетках, требует никазы Cas9(H840A), обратной транскриптазы (экспрессируемой в виде продукта слияния или поступающей in trans), и модифицированной нРНК с 3′ удлинение (см. Фиг.14). Предварительные исследования показали, что длина участка связывания праймера в удлинении нРНК являлась важной в плане повышения эффективности редактирования в клетках человека (см. Фиг.15).

Оптимизация параметров праймерного редактирования в клетках HEK

[1176] После идентификации архитектур Cas9-RT, которые могут обеспечивать реализацию праймерного редактирования в клетках, компоненты и конструкция будут оптимизированы для достижения высокоэффективного редактирования. Расположение и идентичность нуклеотидов кодируемой точечной мутации, а также общая длина новой синтезированной цепи ДНК будут варьироваться для оценки объема редактирования и потенциальных ограничений. Также будут проанализированы мутации в виде коротких вставок и делеций. Конструкты для экспрессии белков будут являться кодон-оптимизированными. В случае успеха будет обеспечено эффективное праймерное редактирование в клетках млекопитающих.

[1177] Предварительные результаты. Были сконструированы дополнительные нРНК для доведения фермента RT до более высокой местной концентрации в локусе редактирования для случая, если внутримолекулярная обратная транскрипция за счет

слитого фермента RT была бы невозможной. Эти вспомогательные направляющие подвергались усечению по 5′-концу (14-15 nt спейсер), что, как было показано ранее, предотвращает образование разрезов за счет Cas9, но сохраняет связывание (см. Фиг.16). Для анализа данной стратегии был выбран локус HEK3.

Потенциальные сложности и альтернативы

[1178] 1) Расщепление нРНК в клетках: если концы удлиненной нРНК являются усеченными в клетке, то могут быть внесены стабилизирующие вторичные структуры, или могут быть протестированы синтетические нРНК со стабилизирующими модификациями. (2) Отсутствие наблюдаемого редактирования в клетках человека: будут проанализированы дополнительные стратегии, включая вторичное нацеливание продуктов слияния RT-Cas9 на смежные участки генома[54]. Кроме того, могут быть проанализированы потенциальные стратегии на основе направленной эволюции в E. coli или S. cerevisiae.

Значимость

[1179] Если праймерное редактирование может быть осуществлено в экспериментальных линиях клеток, то это окажет немедленное влияние на основные биомедицинские исследования за счет обеспечения быстрого получения и описания большого количества точечных мутаций в генах человека. Универсальность данного способа и его ортогональное окно редактирования относительно редакторов оснований обеспечит подход для внесения многих мутаций, которые недоступны в настоящее время. Более того, если праймерное редактирование может быть оптимизировано для достижения высокой эффективности и чистоты продукта, то его потенциальная применимость для корректирования связанных с заболеваниями мутаций в других типах клеток человека будет значительной.

[1180]

3. Реализация специфического в отношении участка редактирования патогенных мутаций в культивируемых клетках человека.

Предпосылки и обоснование.

[1181] Большое количество патогенных мутаций не может быть скорректировано с помощью доступных в настоящее время редакторов оснований вследствие связанных с PAM ограничений или потребности в корректировании трансверсивных или индельных мутаций. При праймерном редактировании все транзиции и трансверсии являются теоретически возможными, как и малые вставки и делеции. Более того, в

плане PAM окно праймерного редактирования (предполагается от -3 до +4) отличается от окна редакторов оснований (от -18 до -12) (Фиг.13). Менделевские условия, которые в настоящее время не могут быть скорректированы с помощью редакторов оснований, включают: (1) фаундер-мутация Glu6Val в гемоглобине бета при серповидноклеточной анемии (для корректирования требуется трансверсия A•T на T•A); (2) наиболее частый вариант His1069Gln в ATP7B при болезни Уилсона (для корректирования требуется трансверсия G•C на T•A); и (3) мутация ΔPhe508 в *CFTR*, которая вызывает муковисцидоз (требуется 3-нуклеотидная вставка). Каждая из данных мишеней содержит подходящим образом расположенный PAM для нацеливания SpCas9 и праймерного редактирования.

Предварительные результаты.

- ***Редактирование T на A в клетках HEK3 не является реализуемым с помощью доступного в настоящее время редактирования оснований, но реализуемо с помощью TPRT редактирования (см. Фиг.17A-17C).***

[1182] На Фиг.17A показан график, демонстрирующий % конверсии T в A в нуклеотиде-мишени после трансфекции компонентов в эмбриональные клетки почки человека (HEK). Эти данные показывают результаты с использованием N-концевого слияния обратной транскриптазы MLV дикого типа и никазы Cas9(H840A) (32-аминокислотный линкер). Эффективность редактирования была значительно повышена, когда длина участка связывания праймера удлиняется от 7 нуклеотидов до 11 или 12 нуклеотидов. Кроме того, вспомогательная направляющая A, которая располагается непосредственно выше локуса редактирования (см. Фиг.16), значительно повышает активность в виде редактирования, в частности для участков связывания праймера более короткой длины. Эффективность редактирования подвергали количественному анализу с помощью ампликонного секвенирования с использованием платформы Illumina MiSeq. На Фиг.17B также показан % конверсии T на A в нуклеотиде-мишени после трансфекции компонентов в эмбриональные клетки почки человека (HEK), однако эти данные показывают результаты с применением C-концевого слияния фермента RT. Здесь вспомогательная направляющая A не имеет такого эффекта, при этом эффективность редактирования в целом является более высокой. На Фиг.17C показаны данные, которые демонстрируют результаты с использованием N-концевого слияния обратной транскриптазы MLV дикого типа и

никазы Cas9(H840A), аналогичных использованным на Фиг.17A; однако линкер между RT MLV и Cas9 имеет длину 60 аминокислот вместо 32 аминокислот.

- ***Результаты редактирования T на A в участке HEK3 с помощью TPRT редактирования демонстрируют высокую чистоту.***

[1183] На Фиг.18 показан результат анализа на основе секвенирования с использованием высокопроизводительного ампликонного секвенирования. Результаты демонстрируют наиболее часто встречающиеся генотипы подвергнутых редактированию клеток. Следует отметить, что не было получено заметных продуктов в виде инделей, при этом требуемая точечная мутация (T на A) была точно внесена без побочных изменений. Первая последовательность показывает исходный генотип. Два верхних продукта представляют собой начальный генотип, содержащий эндогенный полиморфизм (G или A). Два нижних продукта соответствуют правильным образом отредактированным генотипам.

- ***Мутанты RT MLV обеспечивают улучшение редактирования.***

[1184] Мутантные обратные транскриптазы, описанные в Baranauskas, et al (doi:10.1093/ protein/gzs034), тестировали в виде C-концевых продуктов слияния с никазой Cas9(H840A) для нацеленного нуклеотидного редактирования в эмбриональных клетках почки человека (HEK). Плазмида редактора Cas9-RT была котрансфицирована с плазмидой, кодирующей 3′ направляющую РНК для праймерного редактирования, которая служит матрицей для обратной транскрипции. На Фиг.19 показа эффективность редактирования в нуклеотиде-мишени (синие столбцы) наряду с показателями инделей (оранжевые столбцы). WT обозначает фермент RT MLV дикого типа. Мутантные ферменты (от M1 до M4) содержат перечисленные справа мутации. Показатели редактирования количественно анализировали с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

Никирование комплементарной цепи с использованием второй нРНК обеспечивает улучшение редактирования.

[1185] В рамках данного эксперимента оценивается эффективность редактирования нуклеотида-мишени при внесении одноцепочечного разрыва в комплементарную цепь ДНК рядом с нуклеотидом-мишенью, при этом предполагается, что это направит репарацию несовпадений к предпочтительному удалению исходного нуклеотида и конвертации пары оснований с реализацией требуемого изменения. Конструкт

Cas9(H840A)-RT для редактирования был котрансфицирован с двумя плазмидами, кодирующими направляющие РНК, одна из которых служит матрицей для реакции обратной транскрипции, а другая нацелена на комплементарную цепь ДНК для никирования. Было протестировано никирование на различном расстоянии от нуклеотида-мишени (оранжевые треугольники) (см. Фиг.20). Показана эффективность редактирования в паре азотистых оснований-мишеней (синие столбцы) наряду с показателями образования инделей (оранжевые столбцы). Пример "нет" не содержит никирующей комплементарную цепь направляющей РНК. Показатели редактирования количественно анализировали с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[1186] На Фиг.21 демонстрируются обработанные данные высокопроизводительного секвенирования, показывающие требуемую трансверсивнаю мутацию в виде конверсии T в A и общее отсутствие других значительных побочных продуктов редактирования генома.

[1187] Объем. На Фиг.13 показан потенциальный объем новой технологии редактирования и сравнен с технологиями редактирования оснований на основе дезаминазы. Ранее разработанные редакторы основания нацелены на участок ~15 ± 2 bp выше PAM. Путем конвертации нуклеотидов-мишеней C или A в T или G, соответственно, ранее разработанные редакторы оснований обеспечивают все переходные мутации (конверсии A:T в G:C). Однако ранее разработанные редакторы оснований не способны вносить трансверсивные мутации (A на T, A на C, G на T, G на C, T на A , T на G, C на A, C на G). Более того, при наличии нескольких нуклеотидов-мишеней в окне редактирования могут возникнуть дополнительные нежелательные изменения.

[1188] Новая технология праймерного редактирования теоретически может обеспечивать внесение конверсии любого нуклеотида и пары оснований, а также потенциально изменений в виде малой вставки и делеции. В плане PAM окна праймерного редактирования начинаются в участке никирования ДНК (на 3 основания выше от PAM) и заканчиваются в пока ещё не определенном положении ниже PAM. Следует отметить, что данное окно редактирования отличается от окна редактирования редакторов оснований на основе дезаминазы. поскольку TPRT системы осуществляют редактирование с помощью ферментов в виде ДНК-полимеразы, они потенциально имеют все их полезные свойства, включая универсальность и точность.

Корректирование патогенных мутаций в полученных от пациента линиях клеток.

[1189] Линии клеток, несущих соответствующие мутации (серповидноклеточная анемия: CD34 гемопоэтические стволовые клетки; болезнь Уилсона: культивированные фибробласты; муковисцидоз: культивированные клетки бронхиального эпителия) будут получены от ATCC, Coriell Biobank, или сотрудничающих аффилированных лабораторий Harvard/Broad. Эффективность редактирования будет оценена с помощью высокопроизводительного секвенирования, при этом эффективность подвергнутого корректированию генотипа будет протестирована с помощью фенотипических анализов (ВЭЖХ на основе гемоглобина, иммуноокрашивания ATP7B, и анализа мембранного потенциала CFTR).

Описание активности в виде нецелевого редактирования.

[1190] Потенциальное нецелевое редактирование будет подвергнуто скринингу с использованием стандартных способов, таких как GUIDE-seq[55] и CIRCLE-seq[56] с применением нРНК-мишеней, соединенных с Cas9 дикого типа. При идентификации побочных мишеней данные локусы будут проанализированы в рамках подвергнутых TPRT редактированию клетках для идентификации случаев нецелевого редактирования.

Потенциальные сложности и альтернативы.

[1191] (1) Низкая эффективность редактирования: праймерный редактор (PE) может требовать оптимизации для каждой мишени. В этом случае могут быть протестированы библиотеки нРНК для идентификации вариантов для конкретных видов использования с наивысшими показателями функционирования. Экспрессия продукта слияния RT-Cas и ядерная локализация могут быть оптимизированы. Для ограничения нецелевого редактирования может использоваться липосомная доставка на основе RNP.

Последующие эксперименты.

[1192] Оптимизация конструкций нРНК может быть достигнута путем дополнительного анализа длины участка связывания праймера и удлинения матрицы для синтеза. Объем тестирования и универсальность будут включать различные нуклеотидные конверсии, малые вставки и делеции, а также различные положения редактирования относительно PAM и несколько участков в геноме человека. Оптимизация компонента в виде RT будет включать анализ мутаций в RT MLV для

повышения активности (деактивация РНКазы H, повышение аффинности связывания праймера-матрицы, модифицирование процессирования) и новые ферменты RT (RT интронов группы II, другие ретровирусные RT).

Значимость.

[1193] Изменения в виде одного нуклеотида в отдельных генах вызывают множество генетических заболеваний. Разработка описанной здесь технологи геномного редактирования и ее применение в рамках связанных с заболеванием типов клеток заложит основу для перехода к клиническому использованию. Для некоторых заболеваний, таких как серповидноклеточная анемия, одна точечная мутация соответствует доминантному генотипу во всей популяции. Однако для многих других генетических заболеваний наблюдается значительная гетерогенность различных точечных мутаций в одном гене у популяции пациентов, каждая из которых приводит к аналогичному фенотипу заболевания. Таким образом, в качестве универсального способа геномного редактирования, который теоретически может быть нацелен на большое количество подобных мутаций, данная технология может обеспечивать огромную потенциальную пользу для многих из таких пациентов, а также их семей. Если доказательство принципа для данных вариантов использования может быть достигнуто в клетках, то это заложит основу для исследований в рамках моделей заболеваний с использованием животных.

Преимущества

[1194] Точность: требуемое изменение напрямую кодируется в последовательности нуклеиновой кислоты. Универсальность: в теории может являться возможным осуществлять любую конверсию пар оснований, включая изменения в виде трансверсий, а также малые вставки или делеции. Существует окно редактирование, отличное от окна редактирования редакторов оснований в плане последовательности смежного с протоспейсером мотива (PAM) Cas9. Этот способ обеспечивает достижение многих возможностей редактирования на основе направленной гомологичной репарации (HDR), но без значительных ограничений HDR (которая является неэффективной в большинстве типов клеток и обычно сопровождающейся избытком нежелательных побочных продуктов, таких как индели). Кроме того, при нем не осуществляются двуцепочечные разрывы ДНК (DSB, то есть малое количество инделей, транслокаций, крупных делеций, активации p53 и так далее).

[1195]

ПРИМЕР 2 – ПОДВЕРЖЕННОЕ ОШИБКАМ ПРАЙМЕРНОЕ РЕДАКТИРОВАНИЕ (PE)

[1196] Описанные здесь системы праймерного редактирования также могут применяться совместно с ферментом в виде подверженной ошибкам обратной транскриптазы для внесения мутаций в геном.

[1197] На Фиг.22 показан вариант осуществления, который представляет собой схему примера процесса осуществления направленного мутагенеза с использованием подверженной ошибкам обратной транскриптазы в локусе-мишени с помощью программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) в виде комплекса с удлиненной направляющей РНК. Данный процесс может быть обозначен как вариант осуществления праймерного редактирования для направленного мутагенеза. Удлиненная направляющая РНК включает удлинение на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК. На этапе (a) комплекс napDNAbp/нРНК контактирует с молекулой ДНК, при этом нРНК направляет napDNAbp для связывания с подвергаемым мутагенезу локусом-мишенью. На этапе (b) осуществляется разрыв в одной из цепей ДНК локуса-мишени (например с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3′-конца в одной из цепей локуса-мишени. В определенных вариантах осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, то есть цепи, которая не гибридизируется с последовательностью направляющей РНК. На этапе (c) 3′-концевая цепь ДНК взаимодействует с удлиненной частью направляющей РНК для праймирования обратной транскрипции. В определенных вариантах осуществления 3′-концевая цепь ДНК гибридизируется с определенной прайминговой последовательностью RT на удлиненной части направляющей РНК. На этапе (d) вносится подверженная ошибкам обратная транскриптаза, которая синтезирует подвергнутую мутагенезу одиночную цепь ДНК от 3′-конца праймированного участка к 3′-концу направляющей РНК. Примеры мутаций указаны с помощью звездочки "*". Это обеспечивает образование одноцепочечного ДНК-флэпа, включающего требуемый подвергнутый мутагенезу участку. На этапе (e) napDNAbp и направляющая РНК высвобождаются. Этапы (f) и (g) связаны с коррекцией одноцепочечного ДНК-флэпа (включающего подвергнутый мутагенезу участок), так что требуемый подвергнутый мутагенезу участок становится встроенным в локус-мишень. Данный процесс может быть смещен в сторону образования

требуемого продукта путем удаления соответствующего 5′ эндогенного ДНК-флэпа, который образуется когда 3′ одноцепочечный ДНК-флэп вносится и гибридизируется с комплементарной последовательностью на другой цепи. Процесс также может быть смещен в сторону образования продукта с никированием второй цепи, как показано в качестве примера на Фиг.1F. После процессов репарации и/или репликации эндогенной ДНК мутагенизированный участок становится встроенным в обе цепи ДНК локуса ДНК.

ПРИМЕР 3 – СОКРАЩЕНИЕ ТРИНКЛЕОТИДНЫХ ПОВТОРОВ С ПОМОЩЬЮ PE

[1198] Описанная здесь система праймерного редактирования (PE) может применяться для сокращения мутаций в виде тринуклеотидных повторов (или "заболеваний, связанных с экспансией тринуклеотидных повторов") для лечения состояний, таких как болезнь Хантингтона и другие заболевания, связанные с тринуклеотидными повторами). Без связи с теорией, экспансия триплетов вызывается проскальзыванием при репликации ДНК или при репаративном синтезе ДНК. Поскольку тандемные повторы имеют идентичную последовательность относительно друг друга, спаривание оснований между двумя цепями ДНК может происходить в нескольких точках вдоль последовательности. Это может приводить к образованию "выступающих в виде петель" структур при репликации ДНК или репаративном синтезе ДНК. Это может приводить к повторяющемуся копированию повторенной последовательности, что увеличивает количество повторов. Были предложены дополнительные механизмы, связанные с промежуточными продуктами в виде РНК:ДНК гибридов. Праймерное редактирование может применяться для снижения или устранения данных участков экспансии триплетов за счет делеции одного или нескольких триплетов в виде нежелательных повторяющихся кодонов. В рамках варианта осуществления, соответствующего данному варианту использования, на Фиг.23 приведена схема конструкции PEgRNA для сокращения или уменьшения последовательностей тринуклеотидных повторов с помощью праймерного редактирования.

[1199] Таким образом, праймерное редактирование может быть использовано для корректирования любого заболевания, связанного с тринуклеотидными повторами, включая болезнь Хантингтона, синдром ломкой Х-хромосомы и атаксию Фридрейха.

[1200] Наиболее часто встречающийся тринуклеотидный повтор содержит триплеты CAG, хотя также встречаются триплеты GAA (атаксия Фридрейха), и триплеты CGG

(синдром ломкой Х-хромосомы). Триплеты CAG кодируют глутамин (Q), таким образом повторы CAG приводят к полиглутаминовым трактам в кодирующих участках связанных с заболеванием белков. Данный конкретный класс заболеваний, связанных с тринуклеотидными повторами, также называется "полиглутаминовыми (PolyQ) заболеваниями". Другие тринуклеотидные повторы могут вызывать изменения в регулировании генов и обозначаются как "неполиглутаминовые заболевания". Унаследование предрасположенности к экспансии или приобретение уже экспансированного родительского аллеля повышает вероятность возникновения заболевания. Патогенные экспансии тринуклеотидных повторов могут быть скорректированы с помощью праймерного редактирования.

[1201] Праймерное редактирование может быть реализовано для сокращения участков экспансии триплетов за счет никирования участка выше участка в виде триплетного повтора с помощью праймерного редактора, включающего PEgRNA, которая подходящим образом нацелена на участок разрыва. Затем праймерный редактор обеспечивает синтез новой цепи ДНК (оцДНК-флэпа) на основе PEgRNA в качестве матрицы (то есть матрицы для ее редактирования), которая кодирует здоровое количество триплетных повторов (которое зависит от конкретного гена и заболевания). Новая синтезированная цепь оцДНК, включающая здоровую последовательность триплетных повторов, также синтезируется таким образом, чтобы включать короткий участок гомологии (то есть плечо гомологии), который соответствует последовательности, являющейся смежной относительного другого конца повтора (красная цепь). Внесение новой синтезированной цепи и последующее замещение эндогенной ДНК новым синтезированным оцДНК-флэпом ведет к сокращению повторяющегося аллеля.

ПРИМЕР 4 – ПЕПТИДНОЕ МЕЧЕНИЕ С ПОМОЩЬЮ PE

[1202] Описанные здесь системы праймерного редактирования (например системы PE) также могут применяться для внесения различных пептидных меток в кодирующие белки гены. Такие метки могут включать гистидиновые метки *HEXA*, FLAG-метку, V5-метку, GCN4-метку, HA-метку, Myc-метку и другие. Данный подход может применяться для вариантов использования, которые включают флуоресцентное мечение белка, иммуноосаждение, иммуноблоттинг, иммуногистохимию, рекрутирование белка, индуцируемые белковые дегроны, и скрининг в рамках всего генома. Варианты осуществления изображены на Фиг.25 и 26.

[1203] Фиг.25 представляет собой схему, показывающую конструкцию нРНК для генов пептидного мечения в эндогенных геномных локусах и пептидное мечение с помощью геномного редактирования на основе TPRT (то есть праймерного редактирования). Системы мечения FlAsH и ReAsH включают две части: (1) флуорофор-бимышьяковый зонд и (2) генетически кодируемый пептид, содержащий мотив тетрацистеина, например последовательность FLNCCPGCCMEP (SEQ ID NO: 1). При экспрессии в клетках белки, содержащие мотив тетрацистеина, могут подвергаться флуоресцентному окрашиванию флуорофор-мышьяковыми зондами (см. J. Am. Chem. Soc., 2002, 124 (21), стр. 6063–6076. DOI: 10.1021/ja017687n). В рамках системы "sortagging" (на основе опосредованного сортазой лигирования) используются бактериальные ферменты сортазы, которые ковалентно конъюгируют меченые пептидные зонды с белками, содержащими подходящие пептидные субстраты (см. Nat. Chem. Biol. 2007 Nov;3(11):707-8. DOI: 10.1038/nchembio.2007.31). FLAG-метка (DYKDDDDK (SEQ ID NO: 2)), V5-метка (GKPIPNPLLGLDST (SEQ ID NO: 3)), GCN4-метка (EELLSKNYHLENEVARLKK (SEQ ID NO: 4)), HA-метка (YPYDVPDYA (SEQ ID NO: 5)), и Myc-метка (EQKLISEEDL (SEQ ID NO: 6)) обычно используются в качестве эпитопных меток для иммуноанализов. π-клэмп кодирует пептидную последовательность (FCPF) (SEQ ID NO: 622), которая может быть помечена пентафтор-ароматическими субстратами (ref: Nat. Chem. 2016 Feb;8(2):120-8. doi: 10.1038/nchem.2413).

[1204] На Фиг.26 показано точное внесение His6-метки и FLAG-метки в геномную ДНК. С помощью матрицы обратной транскрипции была сконструирована направляющая РНК, нацеленная на локус HEK3, которая кодирует либо вставку 18-nt His-метки, или вставку 24-nt FLAG-метки. Эффективность редактирования в трансфицированных HEK клетках анализировали с помощью ампликонного секвенирования. Следует обратить внимание, что полная 24-nt последовательность FLAG-метки располагается вне наблюдаемого изображения (секвенирование подтвердило осуществление полной и точной вставки).

[1205]

ПРИМЕР 5 – ПРЕДОТВРАЩЕНИЕ ИЛИ ЛЕЧЕНИЕ ПРИОННОГО ЗАБОЛЕВАНИЯ С ПОМОЩЬЮ РЕ

[1206] Данное изобретение может помочь решить проблему прионного заболевания у человека, скота, или диких животных. Ни одна из ранее описанных стратегий

редактирования не является достаточно эффективной и чистой для внесения защитных мутаций или надежного нокдауна *PRNP*. Могут применяться нуклеазы Cas9 и HDR, однако они будут продуцировать по большей части смесь индельных вариантов *PRNP*, некоторые из которых, как считается, являются патогенными. Более того, HDR не работает в большинстве типов клеток. Праймерное редактирование является надежным и эффективным в плане внесения обоих типов мутаций без продуцирования избытка двуцепочечных разрывов ДНК или возникающих инделей.

[1207] В рамках данного изобретения описывается внесение защитной мутации в *PRNP*, которая предотвращает или останавливает прогрессирования прионного заболевания. Данный участок является консервативным у млекопитающих, поэтому помимо лечения заболевания у человека он также может использоваться для получения коров и овец с иммунитетом против прионного заболевания или даже осуществления лечения диких популяций животных, которые страдают прионным заболеванием. Праймерное редактирование уже использовалось для достижения ~25% внесения встречающегося в природе защитного аллеля в клетки человека, при этом предшествующие эксперименты на мышах показывают, что данный уровень внесения является достаточным для возникновения иммунитета против большинства прионных заболеваний. Данный способ является первым и потенциально единственным доступным в настоящее время способом внесения данного аллеля с настолько высокой эффективностью в большинство типов клеток. Другая возможная стратегия лечения заключается в применении праймерного редактирования для снижения или устранения экспрессии *PRNP* путем внесения в ген раннего стоп-кодона. Многие исследователи предсказывают, что это обеспечит лечение заболевания.

[1208] Три потенциальные терапевтические стратегии включают праймерное редактирование для снижения экспрессии PrP. Данная задача может быть решена за счет внесения мутаций, которые вызывают появление преждевременного стоп-кодона в *PRNP*, устраняют начальный кодон, обеспечивают мутацию или делецию основных аминокислотных кодонов, вносят или удаляют участки сплайсинга с продуцированием неправильного транспорта, или изменяют регуляторные элементы, которые снижают уровни транскрипта. Праймерное редактирование для устранения мутаций, связанных с заболеванием. Были описаны многие варианты *PRNP* (ncbi.nlm.nih.gov/pmc/articles/PMC6097508/#b154-ndt-14-2067), которые приводят к повышению вероятности заразиться заболеванием. Каждый известный вариант может

быть реверсирован с помощью праймерного редактирования, поскольку праймерное редактирование может обеспечивать реализацию всех возможных типов точечных мутаций, локальных вставок, и локальных делеций. Праймерное редактирование для внесения одной или нескольких защитных мутаций в *PRNP*, которые нарушают образование и/или трансмиссию прионов. Например, было продемонстрировано, что G127V в гене *PRNP* человека обеспечивает защиту против многих форм прионного заболевания (ncbi.nlm.nih.gov/pmc/articles/PMC4486072/). Позже было описано, что данная мутация препятствует образованию прионов за счет предотвращения образования стабильных бета-листов и димеров (ncbi.nlm.nih.gov/pubmed/30181558, ncbi.nlm.nih.gov/pubmed/26906032). Помимо внесения полиморфизма на основе отдельных нуклеотидов, вставка или делеция последовательностей в *PRNP*, которая будет нарушать образование прионов, также может быть использована для защиты от прионного заболевания или лечения прионного заболевания.

[1209] Третья терапевтическая стратегия является особенно полезной, поскольку внесение защитных вариантов может обеспечивать пользу даже когда относительно малое количество клеток подвергается редактированию. Более того, внесение защитных вариантов, особенно тех, которые естественным образом встречаются у человека, таких как G127V, не будет, как ожидается, иметь каких-либо вредных последствий, в то время как снижение экспрессии прионного белка в качестве стратегии 1 может быть связано с некоторыми вредоносными фенотипами, как было описано у нокаутных по *PRNP* мышей (ncbi.nlm.nih.gov/pmc/articles/PMC4601510/, ncbi.nlm.nih.gov/pmc/articles/PMC2634447/).

[1210] Ранее было продемонстрировано, что мыши, экспрессирующие прионный белок человека дикого типа и защитный вариант G127V прионного белка человека при соотношении приблизительно 2:1 (приблизительно 33% экспрессия защитного варианта), имели полный иммунитет против большинства протестированных форм прионного заболевания и также являлись устойчивыми к варианту болезни Крейтцфельдта-Якоба (vCJD), представляющей собой заболевание человека, передающееся от губчатой энцефалопатии коров (BSE или болезни коровьего бешенства)[91]. Мыши, которые экспрессировали только защитный вариант G127V, имели полный иммунитет против всех протестированных прионных заболеваний, включая vCJD.

[1211] Здесь показано, что защитная мутация G127V может быть эффективным образом внесена в клетки человека в рамках культуры ткани с помощью праймерного редактирования (см. Фиг.27).

[1212] На основании данных результатов описаны три стратегии, в рамках которых может применяться редактирование *PRNP*. Один тип редактирования *PRNP*, который может применяться, представляет собой праймерное редактирование у представляющих собой людей пациентов для предотвращения или лечения прионного заболевания. Второй тип редактирования *PRNP*, который может применяться, представляет собой праймерное редактирование у скота для предотвращения возникновения и распространения прионного заболевания. Как коровы, так и овцы подвергаются спорадическому возникновению прионного заболевания, вызываемого белком, продуцируемым геном *PRNP*. Помимо ослабляющего и смертельного заболевания, которым страдают животные, данные случаи также экономически разрушительны, частично вследствие усилий, которые должны быть предприняты для предотвращения распространения данного крайне заразного заболевания. У одной молочной коровы в штате Вашингтон была обнаружена BSE в декабре 2003 г., что привело к убыткам в плане продаж говядины на следующий год, оцениваемым в размере 2,8-4,2 миллиардов долларов (bookstore.ksre.ksu.edu/pubs/MF2678.pdf). Ген *PRNP* является очень консервативным у млекопитающих. Внесение мутации *PRNP*, такой как G127V, в зародышевую линию клеток скота может устранять возникновение BSE или скрапи, которая является проявлением прионного заболевания у овец. Третий тип редактирования *PRNP*, который может применяться, представляет собой праймерное редактирование у диких животных для предотвращения распространения прионного заболевания в дикой природе. В настоящее время популяции цервидов, включая оленей, настоящих оленей, и лосей в Северной Америке, страдают хроническим истощающим заболеванием (CWD), являющимся проявлением прионного заболевания, вызванного PNRP у данных видов. Сообщается, что частота возникновения составляет вплоть до 25% в некоторых популяциях (cdc.gov/prions/cwd/occurrence.html). О CWD также сообщалось в Норвегии, Финляндии, и Южной Корее. Пока не известно, является ли заболевание переносимым от данных видов человеку (cdc.gov/prions/cwd/transmission.html) или скоту. Внесение мутаций *PRNP*, таких как G127V, в зародышевую линию клеток данных видов может защищать их от CWD и снижать риск переноса другим видам, включая человека.

[1213] Данный способ может применяться для лечения болезни Крейтцфельдта-Якоба (CJD), куру, болезни Герстмана-Штрауслера-Шейнкера, фатальной семейной бессонницы (FFI), коровьей губчатой энцефалопатии (BSE; болезнь коровьего бешенства), скрапи (у овец), и хронического истощающего заболевания (CWD; у оленей, настоящих оленей, и лосей).

[1214] Способ должен быть объединен со способами доставки для зародышей или взрослых нейронов, например микроинъекцией, липидными наночастицами, или AAV векторами.

ПРИМЕР 6 – МЕЧЕНИЕ И МАНИПУЛИРОВАНИЕ РНК С ПОМОЩЬЮ PE

[1215] Здесь описан новый способ внесения мотивов в генетические последовательности, которые обеспечивают мечение или иное манипулирование РНК в относящихся к млекопитающим, эукариотических, или бактериальных клетках. В то время как считается, что только 1% генома человека кодирует белки, практически весь геном транскрибируется с определенным уровнем. Является открытым вопрос о том, какое количество полученной некодирующей РНК (нкРНК) играет функциональную роль, а тем более каковы роли большинства из данных предполагаемых РНК. "Мечение" данных молекул РНК путем вставки в представляющие интерес гены новой РНК-кодирующей последовательности с полезным свойством является полезным способом исследования биологических функций молекул РНК в клетках. Также может являться полезным внесение меток в кодирующие белок иРНК в качестве средства для изменения и, таким образом, лучшего понимания того, как модификации иРНК могут влиять на клеточную функцию. Например, повсеместная природная РНК метка — полиаденилирование — используется клетками для управления транспортом иРНК в цитоплазму. Различные типы сигналов в виде полиаденилирования приводят к различной скорости транспорта и различному времени жизни иРНК, и, таким образом, различиям в уровнях, на которых экспрессируется кодируемый белок.

[1216] Стандартный подход для экспрессии меченых РНК в клетках заключается в экзогенном введении синтетического конструкта с помощью (i) транзиентной плазмидной трансфекции, которая продуцирует кратковременное повышение генной экспрессии, зачастую со сверхфизиологическими уровнями; или (ii) перманентного встраивания гена меченой РНК в геном (в случайных участках) с использованием лентивирусного встраивания или транспозонов, что обеспечивает пролонгированную экспрессию. Оба данных подхода ограничены в плане продуцирования измененных

уровней экспрессии, а также отсутствия естественных механизмов, которые регулируют экспрессию или активность гена. Альтернативная стратегия заключается в прямом мечении представляющего интерес гена в его эндогенном локусе с помощью направленной гомологичной репарации (HDR) двуцепочечных разрывов ДНК, индуцированных Cas9 или другими нацеленными ДНК-нуклеазами. В то время как данный подход обеспечивает получение широкого диапазона эндогенно меченных генов, HDR является заметно неэффективной и, таким образом, требует значительного скрининга для идентификации требуемой клональной популяции клеток, которые были успешно помечены. Более того, HDR обычно является очень неэффективной или полностью неактивной у большого количества типов клеток, в особенности постмитотических клеток. Низкая эффективность HDR дополнительно усугубляется продуцированием нежелательных индельных продуктов, что может являться особенно проблематичным в случае генов РНК, поскольку они могут приводить к продуцированию РНК, активность которой нарушает функцию нормальных аллелей. Наконец, исследователям зачастую требуется провести скрининг различных положений мечения в молекуле РНК для достижения оптимальной эффективности. Данные недостатки совместным образом делают HDR менее желательным способом внесения меток в РНК.

[1217] Праймерное редактирование представляет собой новую технологию геномного редактирования, которая обеспечивает нацеленное редактирование геномных локусов за счет переноса генетической информации из РНК в ДНК. С помощью праймерного редактирования гены РНК могут быть помечены рядом компонентов, таких как РНК аптамеры, рибозимы, или другие мотивы РНК. Праймерное редактирование имеет потенциал быть более быстрым, дешевым, и эффективным в рамках большего ряда типов клеток по сравнению со стратегиями на основе HDR. Таким образом, описанное изобретение представляет собой новый, полезный, и неочевидный инструмент для исследования биологических свойств генов РНК в здоровом состоянии и при заболевании. Здесь описан новый способ внесения мотивов РНК в генетические последовательности, которые обеспечивают мечение или иное манипулирование РНК с помощью праймерных редакторов (PE). PE способны обеспечивать специфическую в отношении участка вставку, мутирование, и/или делецию нескольких нуклеотидов в требуемом геномном локусе, который может являться мишенью для системы CRISPR/Cas. PE состоят из продуктов слияния между доменами нуклеазы Cas9 и

доменами обратной транскриптазы. Они направляются к их геномной мишени с помощью сконструированных PEgRNA (направляющих РНК для праймерного редактирования), которые содержат часть в виде направляющего спейсера для нацеливания на ДНК, а также матрицу для обратной транскрипции, которая кодирует требуемое геномное изменение (см. Фиг.28A). Предполагается, что PE может применяться для вставки мотивов, которые являются функциональными на уровне РНК (далее — мотивов РНК) для мечения или иным образом манипулирования некодирующими РНК или иРНК. Данные мотивы могут служить для повышения генной экспрессии, снижения генной экспрессии, изменения сплайсинга, изменения посттрансляционной модификации, влияния на субклеточное расположение РНК, обеспечения выделения или определения внутри- или внешнеклеточного расположения РНК (с помощью, например, флуоресцентных РНК аптамеров, например Spinach, Spinach2, Baby Spinach, или Broccoli), рекрутирования эндогенного и экзогенного белка или связывающихся с РНК элементов, внесения sgРНК или индуцирования процессирования РНК за счет саморасщепления или РНКаз (см. Фиг.28B). Вследствие гибкости праймерного редактирования является невозможным привести исчерпывающий перечень мотивов РНК, которые могут быть внесены в геном. Здесь показана серия примеров, которые широким образом иллюстрируют предполагаемый объем вносимых с помощью PE мотивов РНК, которые могут применяться для мечения генов РНК. В настоящее время невозможно эффективно и достаточно точно произвести данные изменения в большинстве типов клеток (включая многие из тех, которые не поддерживают HDR) с помощью любого иного описанного способа геномного редактирования помимо PE.

[1218] Генная экспрессия может быть изменена за счет кодирования 3′ нетранслируемого участка (UTR), что приводит к изменениям в ядерном транспорте или удержании или времени жизни иРНК. Например, polyA-хвост из полиомавируса обезьяны 40 (SV40) имеет дополнительные хелперные последовательности, которые обеспечивают эффективную терминацию транскрипции и могут повышать генную экспрессию относительно других 3′ UTR[57, 58]. Пример последовательности polyA-хвоста SV40:

POLYA-XBOCT SV40	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT TGTGGTTTGTCCAAACTCATCAATGTATCTTA (SEQ ID NO: 331)

[1219] Сигналы посттрансляционной модификации, помимо сигналов полиаденилирования, также могут кодироваться РЕ. Они включают сигналы, включающие модификации на основе N6-метиладенозина, N1- метиладенозина, 5-метилцитозина, и псевдоуридина[59]. За счет использования РЕ для включения последовательностей, связанных ферментами, которые записывают или удаляют эти модификации в РНК транскрипте, будет являться возможным индуцировать их запись или стирание. Это может быть использовано в качестве инструмента для исследования эффектов данных маркеров, для индуцирования клеточной дифференциации, влияния на реакции на стресс, или, с учетом того, что функции данных маркеров еще не исследованы, влияния на клетки-мишени иным образом.

[1220] РЕ может обеспечивать кодирование мутаций, которые влияют на субклеточную локализацию. Например, встраивание тРНК-Lys в иРНК теоретически может приводить к транспорту в митохондрию[60], в то время как различные 3′ UTR могут приводить к удержанию в ядре или транспорту[61].

[1221] Примеры:

[1222] Сигнал polyA SV40 приводит к транспорту.

POLYA-ХВОСТ SV40	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT TGTGGTTTGTCCAAACTCATCAATGTATCTTA (SEQ ID NO: 331)

[1223] 3′ бокс мяРНК U1 приводит к удержанию.

3′ БОКС МЯРНК U1	TTCATTCAGCAAGTTCAGAGAAATCTGAACTTGCTGGATTTTTG GAGCAGGGAGATGGAATAGGAGCTTGCTCCGTCCACTCCACGC ATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGCACCCACTT TCTGGAGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTT TTTGTTT GGTTTGTGTCTTGGTTGGCGTCTTAA (SEQ ID NO: 625)

[1224] Определение субклеточной локализации эндогенной РНК может быть трудным и требовать добавления экзогенных флуоресцентно меченных нуклеотидных зондов, как в случае FISH, или медленного и потенциально неточного фракционирования клеток с последующим детектированием РНК. Кодирование зонда в эндогенной РНК обеспечит устранение многих из данных проблем. Один пример заключается в кодировании флуоресцентного РНК аптамера, такого как Spinach[62] или Broccoli, в эндогенной РНК, тем самым обеспечивая визуализацию присутствия данной РНК за счет добавления низкомолекулярного протофлуорофора.

[1225] Аптамер Broccoli:

АПТАМЕР BROCCOLI	GAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTG GGCTC (SEQ ID NO: 357)

[1226] РЕ может обеспечивать вставку или удаление последовательностей, которые кодируют РНК, которая распознается связывающимися с РНК белками, что влияет на стабильность, экспрессию, локализацию, или модификацию РНК (например, см. перечисленные белки[63]).

[1227] РЕ может обеспечивать вставку последовательностей, которые кодируют мяРНК в геноме, в качестве защитного механизма против вируса или рака. Аналогичным образом, оно может использоваться для вставки микроРНК (например премикроРНК) для обеспечения сайлесинга генов-мишеней.

[1228] РЕ может обеспечивать вставку последовательностей, приводящих к процессингу РНК, либо самостоятельно, или за счет внешних факторов, либо в качестве терапии, или инструмента для исследования функции различных частей РНК. Например, рибозим HDV[64] при включении в последовательность РНК приводит к процессингу РНК непосредственно 5′ от рибозима, в то время как рибозим типа "головка молотка" (hammerhead) расщепляется до третьей шпильки в рамках рибозима[65]. Другие саморасщепляющиеся рибозимы включают "пистолет" (pistol)[166], "топор" (hatchet)[66], "шпильку" (hairpin)[67], сателлит Neuropora Varkud[68], glmS[69], "твистер"[70], и "твистер-систер" (twister sister)[66]. Другие последовательности могут включать относящиеся к дикому типу или сконструированные или подвергнутые эволюции версии рибозимов. Большинство данных рибозимов могут иметь различные последовательности в зависимости от участка РНК, с которым они ассоциированы, в зависимости от расположения участка разрыва рибозима. Также могут быть получены последовательности, которые будут обеспечивать процессинг РНК за счет внешних факторов, таких как специфические в плане последовательности РНКазы[71], то есть РНКазы, которые распознают специфические структуры[72], такие как Dicer[73] и Drosha[74].

[1229] Рибозим HDV:

РИБОЗИМ HDV	GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAA CATGCTTCGGCATGGCGAATGGGAC (SEQ ID NO: 365)

[1230] Ссылки для Примера 6

[1231] Следующие ссылки полностью приведены здесь в качестве ссылки.

1. Schek N, Cooke C, Alwine JC. *Molecular and Cellular Biology*. 1992.

2. Gil A, Proudfoot NJ. *Cell.* 1987.

3. Zhao, B.S., Roundtree, I.A., He, C. Nat Rev Mol Cell Biol. 2017.

4. Rubio, M.A.T., Hopper, A.K. Wiley Interdiscip Rev RNA 2011.

5. Shechner, D.M., Hacisuleyman E., Younger, S.T., Rinn, J.L. Nat Methods. 2015.

6. Paige, J.S., Wu, K.Y., Jaffrey, S.R. *Science* 2011.
7. Ray D., … Hughes TR. Nature 2013.
8. Chadalavada, D. M., Cerrone-Szakal, A. L., Bevilacqua, P.C. RNA 2007.
9. Forster AC, Symons RH. Cell. 1987.
10. Weinberg Z, Kim PB, Chen TH, Li S, Harris KA, Lünse CE, Breaker RR. Nat. Chem. Biol. 2015.
11. Feldstein PA, Buzayan JM, Bruening G. Gene 1989.
12. Saville BJ, Collins RA. Cell. 1990.
13. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. Nature 2004.
14. Roth A, Weinberg Z, Chen AG, Kim PG, Ames TD, Breaker RR. Nat Chem Biol. 2013.
15. Choudhury R, Tsai YS, Dominguez D, Wang Y, Wang Z. Nat Commun. 2012.
16. MacRae IJ, Doudna JA. Curr Opin Struct Biol. 2007.
17. Bernstein E, Caudy AA, Hammond SM, Hannon GJ Nature 2001.
18. Filippov V, Solovyev V, Filippova M, Gill SS. Gene 2000.

ПРИМЕР 7 – ПОЛУЧЕНИЕ БИБЛИОТЕК ГЕНОВ С ПОМОЩЬЮ РЕ

[1232] Здесь описан новый способ клеточного получения сложных библиотек белок- или РНК-кодирующих генов с определенными или варьирующимися вставками, делециями, или определенными аминокислотными/нуклеотидными конверсиями, а также их применение в рамках высокопроизводительного скрининга и направленной эволюции. Ссылки, приведенные в рамках данного Примера, основаны на перечне ссылок, включенном в конце данного Примера.

[1233] Получение вариабельных генетических библиотек наиболее часто достигалось с помощью мутагенной ПЦР[1]. Данный способ основан либо на использовании реакционных условий, которые снижают точность ДНК-полимеразы, или использовании модифицированных ДНК-полимераз с более высокими показателями мутации. Таким образом, отличия в рамках данных полимераз отражаются на продукте в виде библиотеки (например предпочтение мутаций в виде транзиций против трансверсий). Врожденным ограничением данного подхода для конструирования библиотеки является относительная неспособность влиять на размер подвергаемого варьированию гена. Большинство ДНК-полимераз имеет крайне низкие показатели индельных мутаций[2] (вставок и делеций), при этом большинство из них будут приводить к мутациям рамки считывания в кодирующих белок участках, делая элементы библиотеки неспособными пройти какой-либо последующий отбор. Кроме

того, отклонения при ПЦР и клонировании могут усложнить получение отдельных библиотек, состоящих из генов различных размеров. Данные ограничения могут существенно ограничить эффективность направленной эволюции для стимулирования существующих или конструирования новых функций белков. При естественной эволюции крупные изменения функции или эффективности белка обычно связаны со вставкой и делецией мутаций, которые имеют малую вероятность возникновения при стандартном получении библиотеки для мутагенеза. Более того, данные мутации наиболее часто возникают в участках представляющего интерес белка, которые, как предполагается, будут образовывать петли, в отличие от гидрофобного ядра. Таким образом, большинство инделей, полученных с помощью традиционного подхода без отклонений, вероятно будут являться либо вредоносными, или неэффективными.

[1234] Библиотеки, которые могут обеспечивать образование данных мутаций в участках в белке, где они с наибольшей вероятностью будут являться полезными, например в участках в виде петель, будут иметь значительное преимущество по сравнению с традиционными библиотеками при условии, что все библиотеки имеют доступ лишь к небольшой части возможного пространства мутаций. Наконец, хотя является возможным получать генетические библиотеки со специфичными в плане участков мутациями в виде инделей с помощью многоэтапной ПЦР и клональной сборки с помощью праймеров NNK или с помощью тасования ДНК, данные библиотеки не могут подвергаться дополнительным раундам "образования инделей" при непрерывной эволюции. Непрерывная эволюция представляет собой тип направленной эволюции с минимальным вмешательством пользователя. Одним подобным примером является PACE[3]. Поскольку непрерывная эволюция осуществляется при минимальном вмешательстве пользователя, любое повышение разнообразия библиотеки при эволюции должно осуществляться с помощью нативного механизма репликации. Таким образом, несмотря на то, что библиотеки генов со вставленными или удаленными кодонами в качестве специфических локусов могут быть получены и подвергнуты скринингу в рамках PACE, дополнительные раунды "образования инделей" являются невозможными.

[1235] Предполагается, что программируемость праймерного редактирования (PE) может быть использована для получения имеющих высокую сложность программируемых генетических библиотек для применения в рамках высокопроизводительного скрининга и направленной эволюции (см. Фиг.29A). PE

может обеспечивать вставку, изменение или удаление определенных количеств нуклеотидов из определенных генетических локусов с помощью информации, кодируемой в направляющей РНК для праймерного редактирования (см. Фиг.29B). Это обеспечивает получение нацеленных библиотек с одной или несколькими аминокислотами, вставленными или удаленными из участков в виде петель, где мутации с наибольшей вероятностью вызовут изменения в плане функции без побочного внесения нефункциональных мутаций рамки считывания (см. Фиг.29C). PE может использоваться для внесения определенных наборов мутаций безотносительно отклонений, присущих ДНК-полимеразе или подвергаемой мутированию последовательности.

[1236] Например, в то время как конвертация кодона CCC в стоп-кодон будет иметь малую вероятность осуществления при стандартном получении библиотеки, поскольку она потребует трех последовательных мутаций, включая две трансверсии, PE может быть использован для конвертации любого необходимого кодона в стоп-кодон TGA за один этап. Они также могут применяться для внесения программируемого разнообразия в необходимых положениях, например путем внесения кодонов, кодирующих любую гидрофобную аминокислоту в необходимом участке, в то же время не кодирую какие-либо другие. Более того, вследствие программируемости PE несколько PEgRNA могут применяться для получения нескольких различных изменений в нескольких участках одновременно, обеспечивая получение высокопрограммируемых библиотек (см. Фиг.29D). Кроме того, является возможным использовать обратные транскриптазы с меньшей точностью для получения участков мутагенеза в рамках иным образом невариабельной библиотеки (например обратную транскриптазу ВИЧ-1[4] или обратную транскриптазу фага Bordetella[5]).

[1237] Также предполагается возможность итеративных раундов PE в рамках одного участка, что обеспечивает, например, возможность повторной вставки кодонов в одном участке. Наконец, предполагается, что все из описанных выше подходов могут быть встроены в непрерывную эволюцию, обеспечивая получение новых подвергаемых эволюции in situ библиотек (см. Фиг.30). Они также могут применяться для конструирования этих библиотек в рамках других типов клеток, где иным образом будет сложно собрать крупные библиотеки, например в рамках клеток млекопитающих. Получение кодируемых PE бактериальных штаммов, которые были оптимизированы для направленной эволюции, будет являться полезным

дополнительным инструментом для идентификации белков и РНК с улучшенной или новой функциональностью. Все из данных вариантов применения РЕ не являются очевидными вследствие новой природы РЕ. В заключение, получение библиотек с помощью РЕ будет являться крайне полезным инструментом в рамках синтетической биологии и направленной эволюции, а также для высокопроизводительного скрининга белка и комбинаторных мутантов РНК.

Альтернативные подходы

[1238] Основной способ, с помощью которого получают разнообразные библиотеки в настоящее время, представляет собой описанную выше мутагенную ПЦР[1]. Вставки или делеции могут быть внесены с помощью расщепленных праймеров NNK в определенных участках во время ПЦР, однако внесение подобных мутаций в нескольких участках требует нескольких раундов итеративной ПЦР и клонирования перед конструированием более разнообразной библиотеки с помощью мутагенной ПЦР, что делает данный способ медленным. Альтернативный комплементарный способ представляет собой тасование ДНК, при котором фрагменты библиотеки генов, полученной с помощью обработки ДНКазой, вносят в ПЦР реакцию без праймеров, что приводит к отжигу различных фрагментов друг на друге и быстрому получению более разнообразных библиотек, чем при использовании мутагенной ПЦР самостоятельно[6]. Несмотря на то, что этот подход может теоретически обеспечивать получение индельных мутаций, он чаще приводит к мутациям рамки считывания, которые уничтожают функцию гена. Более того, тасование ДНК требует высокой степени гомологии между фрагментами генов. Оба данных способа должны быть реализованы *in vitro* с трансформацией полученной библиотеки в клетки, в то время как библиотеки, полученные с помощью РЕ, могут быть сконструированы *in situ*, обеспечивая возможность их применения в рамках непрерывной эволюции. В то время как библиотеки могут быть сконструированы *in situ* с помощью *in vivo* мутагенеза, данные библиотеки основаны на клеточном механизме хозяина и демонстрируют отклонения в плане инделей. Аналогичным образом, несмотря на то, что традиционные способы клонирования могут применяться для получения специфичных в плане участка мутационных профилей, они не могут применяться *in situ* и в целом собираются по одному за раз *in vitro* перед трансформацией в клетки. Эффективность и широкая функциональность РЕ как в прокариотических, так и эукариотических типах клеток дополнительно свидетельствует о том, что данные библиотеки могут быть

сконструированы напрямую в представляющем интерес типе клеток, в отличие от клонирования в организм-модель, такой как E. coli, с последующим трансфицированием в представляющую интерес клетку или организм. Другой альтернативный подход для направленной диверсификации представляет собой автоматизированную мультиплексную геномную инженерию или MAGE, в рамках которой несколько одноцепочечных ДНК-олигонуклеотидов могут быть встроены в репликационные вилки и приводят к программируемым мутациям[7]. Однако MAGE требует значительного модифицирования штамма хозяина и может приводить к 100-кратному повышению нецелевых или фоновых мутаций[8], в то время как PE является более высокопрограммируемым и, как ожидается, будет приводить к меньшему количеству нецелевых эффектов. Кроме того, MAGE не была продемонстрирована в рамках широкого ряда типов клеток, включая клетки млекопитающих. Праймерное редактирование представляет собой новый и неочевидный дополнительный способ получения библиотек.

Примеры PE в рамках направленной эволюции для получения библиотек генов

[1239] В одном примере PE может применяться в рамках эксперимента на основе направленной эволюции для внесения вариантов белка в библиотеки генов в рамках эксперимента в виде непрерывной эволюции с помощью PACE, обеспечивая итеративное накопление как точечных мутаций, так и инделей таким образом, который не является возможным с использованием традиционных подходов. Уже было показано, что PE может обеспечивать специфическую в плане участка и программируемую вставку нуклеотидов в генетическую последовательность *E. coli*. В рамках указанной направленной эволюции предлагается идентифицировать монотела с повышенным связыванием в отношении определенного эпитопа с использованием модифицированного отбора в рамках PACE на основе двухгибридного связывания белок:белок. Специфичные и высоковариабельные петли в рамках данных монотел вносят значительный вклад в аффинность и специфичность. Повышенное связывание монотел может быть быстро достигнуто в рамках PACE путем варьирования длины и композиции данных петель определенным образом. Однако варьирование длины последовательности не является стандартной функциональностью PACE. В то время как библиотека с петлями варьирующегося размера может использоваться в качестве начальной точки для PACE, в рамках отбора на основе PACE не будут осуществляться последующие улучшения в плане длины, что ограничивает доступ к полезным

синергетическим комбинациям точечных мутаций и индельных мутаций. Встраивание PE в отбор в рамках PACE обеспечит *in situ* получение и эволюцию монотел с петлями различной длины. Для этого предполагается внести дополнительную PE плазмиду в штамм-хозяин *E. coli*, кодирующую фермент PE и одну или несколько PEgRNA. Экспрессия фермента PE и PEgRNA будет контролироваться малой молекулой, доставляемой в "лагуну" PACE (lagoon, сосуд фиксированного объема) со скоростью, выбранной экспериментатором.

[1240] В различных вариантах осуществления компоненты PEgRNA будут содержать спейсер, направляющий PE к представляющему интерес участку на фаге для отбора, и будет сконструирована таким образом, что несколько из трех нуклеотидов могут быть вставлены в участок-мишень, так что будет осуществляться внесение нового участка связывания PEgRNA, обеспечивая итеративную вставку одного или нескольких кодонов в участке-мишени.

[1241] Параллельно с этим другой штамм-хозяин *E. coli* может включать PEgRNA, которые будут служить матрицей для удаления одного или нескольких кодонов, обеспечивая уменьшение размера петли при эволюции. В рамках эксперимента PACE может использоваться смесь обоих штаммов или обеспечиваться их изменение для реализации медленного и контролируемого добавления или удаления последовательностей петель.

[1242] Следует отметить, что данных способ также может применяться для эволюции антител. Принципы связывания, присущие антителам, являются очень похожими на принципы, присущие монотелам: длина петель, соответствующих участку определения комплементарности антитела, является критически важной для их функции связывания. Кроме того, было обнаружено, что более длинные петли являются критически важными в плане развития редких антител с широкой защитной активностью против ВИЧ-1 и других вирусных инфекций[9]. Применение PE как описано выше в отношении антитела или полученной на основе антитела молекулы будет обеспечивать получение антитела с различной длиной петель и варьирующейся последовательностью петель. В комбинации с PACE подобный подход будет обеспечивать улучшенное связывание за счет геометрии петель, которая является недостижимой при стандартном PACE, и, таким образом, обеспечивать эволюцию высокофункциональных антител.

[1243] Эксперименты покажут возможность использования PE для корректирования вредоносной мутации в бактериофаге M3 в рамках прерывистой эволюции с помощью фагов (PANCE), что является первым необходимым этапом для использования PE в рамках непрерывной эволюции (см. Фиг.69).

[1244] Ссылки для Примера 7

[1245] Следующие ссылки полностью приведены здесь в качестве ссылки.

1. Cadwell RC and Joyce GF. PCR Methods Appl. 1992.
2. McInerney P, Adams P, and Hadi MZ. Mol Biol Int. 2014.
3. Esvelt KM, Carlson JC, and Liu DR. Nature. 2011.
4. Naorem SS, Hin J, Wang S, Lee WR, Heng X, Miller JF, Guo H. Proc Natl Acad Sci USA 2017.
5. Martinez MA, Vartanian JP, Wain-Hobson S. Proc Natl Acad Sci USA 1994.
6. Meyer AJ, Ellefson JW, Ellington AD. Curr Protoc Mol Biol. 2014.
7. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Nature. 2009.
8. Nyerges Á et al. Proc Natl Acad Sci USA. 2016.
9. Mascola JR, Haynes BF. Immunol Rev. 2013.

ПРИМЕР 8 – ВСТАВКА ИММУНОЭПИТОПА С ПОМОЩЬЮ PE

[1246] Технологии точного геномного нацеливания на основе использования системы CRISPR/Cas недавно были протестированы в рамках ряда вариантов использования, включая вставку сконструированных последовательностей ДНК в геномные локусы-мишени. Ранее для данного варианта использования применяли направляемую на основе гомологии репарацию (HDR), которая требовала донорской оцДНК матрицы и инициирования репарации с помощью двуцепочечного разрыва ДНК (DSB). Данная стратегия обеспечивает наиболее широкий диапазон возможных реализуемых изменений в клетках и является единственным способом, доступным для вставки крупных последовательностей ДНК в клетки млекопитающих. Однако HDR связана с затруднениями в виде нежелательных клеточных побочных эффектов, возникающих вследствие ее инициирующего DSB, таких как высокие уровни образования инделей, транслокация ДНК, крупные делеции и активация P53. Помимо данных недостатков, HDR ограничивается низкой эффективностью в рамках многих типов клеток (T-клетки являются заметным исключением из данного наблюдения). Недавние попытки преодолеть данные недостатки включают слияния мутантов Rad51 человека с никазой D10A Cas9 (RDN), приводящие к получению HDR системы без DSB, которая включает

улучшенные соотношения продукт HDR:индели и более низкое нецелевое редактирования, но все же ограничена зависимостями в плане типов клеток и лишь умеренной эффективностью HDR редактирования.

[1247] Недавно разработанные продукты слияния Cas9 с обратными транскриптазами ("праймерные редакторы"), соединенные с PEgRNA, представляют собой новую технологию геномного редактирования, которая обеспечивает ряд преимуществ по сравнению с существующими способами геномного редактирования, включая способность вносить любое нуклеотидное замещение и обеспечивать вставку или делецию любой короткой последовательности нуклеотидов (до, по меньшей мере, нескольких дюжин оснований) специфичным в плане участка образом. Следует отметить, что изменения на основе PE осуществляются со в целом низкими показателями нежелательных инделей. Таким образом, PE обеспечивает возможность нацеленного редактирования на основе вставок, которое ранее являлось невозможным или непрактичным.

[1248] Данное изобретение предусматривает способ применения праймерного редактирования в качестве средства для вставки известных иммуногенетических эпитопов в эндогенную или чужеродную геномную ДНК, что приводит к модифицированию соответствующих белков для терапевтических или биотехнологических вариантов использования (см. Фиг.31 и 32). До изобретения праймерного редактирования такие вставки могли быть реализованы лишь неэффективным образом с высокими показателями образования инделей из DSB. Праймерное редактирование устраняет проблему в виде высокого образования инделей при изменениях в виде вставок, в то же время в целом обеспечивая более высокую эффективность, чем HDR. Данный низкий показатель образования инделей является значительным преимуществом праймерного редактирования над HDR в качестве способа нацеленных вставок ДНК, особенно в рамках описанного использования в виде вставки иммуногенных эпитопов. Длина эпитопов соответствует диапазону от нескольких оснований до сотен оснований. Праймерный редактор представляет собой наиболее эффективную и точную технологию реализации таких нацеленных вставок в клетках млекопитающих.

[1249] Ключевая концепция изобретения заключается в применении праймерных редакторов для вставки нуклеотидной последовательности, содержащей ранее описанные иммуногенные эпитопы, в эндогенную или чужеродную геномную ДНК

для снижения и/или устранения их продуктов в виде белков и/или экспрессирующих типов клеток. Нуклеотидные последовательности для вставки иммуногенных эпитопов будут нацелены на гены для получения слитых белков с белком, кодируемым геном, на который осуществляется нацеливание, и трансляции соответствующего белка вставленного иммуногенного эпитопа. Иммунная система пациента была ранее подвергнута тренировке для распознавания данных эпитопов в результате стандартной предшествующей иммунизации на основе обычной вакцинации против, например, столбняка, дифтерии, или кори. В результате иммуногенной природы слитых эпитопов ожидается, что иммунная система пациента будет распознавать и деактивировать подвергнутый праймерному редактированию белок (а не только вставленный эпитоп) и потенциально клетки, из которых он был экспрессирован.

[1250] Продукты слияния с генами-мишенями будут сконструированы соответствующим образом для обеспечения видимости трансляции вставленного эпитопа белка для распознавания иммунной системой. Это может включать нацеленные нуклеотидные вставки, приводящие к трансляции белка с образованием С-концевых продуктов слияния иммуногенных эпитопов с генами-мишенями, N-концевых продуктов слияния иммуногенных эпитопов с генами-мишенями, или вставке нуклеотидов в гены, так что иммуногенные эпитопы кодируются внутри расположенных на поверхности участков структуры белка.

[1251] В качестве части данного изобретения может потребоваться сконструировать белковые линкеры, кодируемые в виде нуклеотидов, вставленных между последовательностью гена-мишени и нуклеотидной последовательностью вставленного иммуногенного эпитопа, для обеспечения распознавания иммунной системой, клеточного трафика, функции белка, или фолдинга белка, связанного с геном-мишенью. Данные вставленные кодируемые нуклеотидами белковые линкеры могут включать (но без ограничения) различную длину и последовательности линкера XTEN или различную длину и последовательности линкеров на основе глицина-серина. Данные сконструированные линкеры ранее применялись для успешной реализации слияния белков.

[1252] Отличительные признаки данного изобретения включают способность использовать ранее приобретенные иммунные ответы на определенные аминокислотные последовательности в качестве средства индуцирования иммунного ответа против иным образом неиммуногенных белков. Другой отличительный признак

представляет собой способность обеспечивать вставку нуклеотидных последовательностей данных иммуногенных эпитопов в мишени таким образом, при котором не осуществляется индуцирование высоких уровней нежелательных инделей в качестве сопутствующего продукта редактирования и который является эффективным при их вставке. Описанное здесь изобретение также имеет способность обеспечивать комбинирование специфичных в плане типа клеток способов доставки (например серотипов AAV) для вставки эпитопов в типы клеток, которые представляют интерес в плане запуска иммунного ответа на них.

[1253] Праймерное редактирование в качестве средства вставки иммуногенных эпитопов в патогенные гены может применяться для программирования иммунной системы пациента для борьбы с широким диапазоном заболеваний (не ограничиваясь раком, как в случае иммуно-онкологических стратегий). Очевидным вариантом применения данной технологии будет использование в качестве лекарства от рака, поскольку она может нарушать механизм опухоли, связанный с ускользанием от иммунитета, за счет индуцирования иммунного ответа на соответствующий онкоген, подобный HER2, или факторы роста, подобные EGFR. Подобный подход может выглядеть аналогичным модифицированию Т-клеток, однако одно новое преимущество данного подхода заключается в том, что он может применяться в рамках многих типов клеток и для заболеваний помимо рака без необходимости в получении и введении модифицированных Т-клеток пациентам.

[1254] PE используется для вставки иммуногенного эпитопа, против которого большинство пациентов уже являются вакцинированными (столбняк, коклюш, дифтерия, корь, эпидемический паротит, краснуха и так далее), в чужеродный или эндогенный ген, который вызывает заболевание, так что иммунная система пациента учится деактивировать данный белок.

[1255] Заболевания, которые могут быть связаны с потенциальной терапевтической пользой от вышеупомянутой стратегии, включают заболевания, вызываемые накоплением токсичных белков, например фатальную семейную бессонницу. Другие заболевания, которые могут быть связаны с пользой, включают заболевания, вызываемые патогенной сверхэкспрессией иным образом нетоксичного эндогенного белка, а также вызываемые чужеродными патогенами.

[1256] Первичные терапевтические показания включают те, которые были указаны выше, такие как лекарства для лечения рака, прионных и иных нейродегенеративных

заболеваний, инфекционных заболеваний, и профилактическую медицину. Вторичные терапевтические показания могут включать профилактический уход для пациентов с поздно развившимися генетическими заболеваниями. Ожидается, что текущий стандарт в плане поддерживающих лекарственных средств может быть использован совместно с праймерным редактированием для некоторых заболеваний, таких как особо агрессивные виды рака, или в случаях, когда лекарства помогают смягчать симптомы заболевания до полного излечения заболевания.

Ниже приведены примеры иммуногенных эпитопов, которые могут быть вставлены с помощью праймерного редактирования:

Вакцина	Заболевание	Аминокислотная последовательность эпитопа	Пример последовательности нуклеиновой кислоты (8)
1	Столбнячный токсоид	QYIKANSKFIGITEL (SEQ ID NO: 396)	CATGATATAAAAGCAAATTCTAAATTTATAGGTATAACTGAACTA (SEQ ID NO: 397)
2	Мутант дифтерийного токсина CRM197	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVG	GGCGCCGACGACGTGGTGGACAGCAGCAAGAGCTTCGTGATGGAGAACTTCAGCAGCTACCACGGCACCAAGCCCGGCTACGTGGACAGCATCCAGAAGGGCATCCAGAAGCCCAAGAGCGGCACCCAGGGCAACTACGACGACGACTGGAAGGAGTTCTACAGCACCGACAACAAGTACGACGCCGCCGGCTACAGCGTGGACAACGAGAACCCCCTGAGCGGCAAGGCCGGCGGCGTGGTGAAGGTGACCTACCCCGGCCTGACCAAGGTGCTGGCCCTGAAGGTGGACAACGCCGAGACCATCAAGAAGGAGCTGGGCCTGAGCCTGACCGAGCCCCTGATGGAGCAGGTGGGCACCGAGGAGTTCATCAAGAGGTTCGGCGACGGCGCCAGCAGGGTGGTGCTGAGCCTGCCCTTCGCCGAGGGCAGCAGCAGCGTGGAGTACATCAACAACTGGGAGCAGGCCAAGGCCCTGAGCGTGGAGCTGGAGATCAACTTCGAGACCAGGGGCAAGAGGGGCCAGGACGCCATGTACGAGTACATGGCCCAGGCCTGCGCCGGCAACAGGGTGAGGAGGAGCGTGGGCAGCAGCCTGAGCTGCATCAACCTGGACTGGGACGTGATCAGGGACAAGACCAAGACCAAGATCGAGAGCCTGAAGGAGCACGGCCCCATCAAGAACAAGATGAGCGAGAGCCCCAACAAGACCGTGAGCGAGGAGAAGGCCAAGCAGTACCTGGAGGAGTTCCACCAGACCGCCCTGGAGCACCCCGAGCTGAGCGAGCTGAAGACCGTGACCGGCACCAACCCCGTGTTCGCCGGCGCCAACTACGCCGCCTGGGCCGTGAACGTGGCCCAGGTGATCGACAGCGAGACCGCCGACAACCTGGAGAAGACCACCGCCGCCCTGAGCATCCTGCCCGGCATCGGCAGCGTGATGGGCA

		NGVHANLHVAFH RSSSEKIHSNEISS DSIGVLGYQKTV DHTKVNSKLSLFF EIKS (SEQ ID NO: 630)	TCGCCGACGGCGCCGTGCACCACAA CACCGAGGAGATCGTGGCCCAGAGC ATCGCCCTGAGCAGCCTGATGGTGG CCCAGGCCATCCCCCTGGTGGGCGA GCTGGTGGACATCGGCTTCGCCGCC TACAACTTCGTGGAGAGCATCATCA ACCTGTTCCAGGTGGTGCACAACAG CTACAACAGGCCCGCCTACAGCCCC GGCCACAAGACCCAGCCCTTCCTGC ACGACGGCTACGCCGTGAGCTGGAA CACCGTGGAGGACAGCATCATCAGG ACCGGCTTCCAGGGCGAGAGCGGCC ACGACATCAAGATCACCGCCGAGAA CACCCCCCTGCCCATCGCCGGCGTG CTGCTGCCCACCATCCCCGGCAAGC TGGACGTGAACAAGAGCAAGACCCA CATCAGCGTGAACGGCAGGAAGATC AGGATGAGGTGCAGGGCCATCGACG GCGACGTGACCTTCTGCAGGCCCAA GAGCCCCGTGTACGTGGGCAACGGC GTGCACGCCAACCTGCACGTGGCCT TCCACAGGAGCAGCAGCGAGAAGAT CCACAGCAACGAGATCAGCAGCGAC AGCATCGGCGTGCTGGGCTACCAGA AGACCGTGGACCACACCAAGGTGAA CAGCAAGCTGAGCCTGTTCTTCGAG ATCAAGAGC (SEQ ID NO: 399)
3	Паротит	GTYRLIPNARANL TA (SEQ ID NO: 400)	GGCACCTACAGGCTGATCCCCAACG CCAGGGCCAACCTGACCGCC (SEQ ID NO: 401)
4	Паротит	PSKFFTISDSATFA PGPVSNA (SEQ ID NO: 402) PSKLFIMLDNATF APGPVVNA (SEQ ID NO: 404)	Ccgagcaaatttttaccattagcgatagcgcgacctttgcg ccgggcccggtgagcaacgcg (SEQ ID NO: 403) Ccgagcaaactgtttattatgctggataacgcgacctttgcg ccgggcccggtggtgaacgcg (SEQ ID NO: 405) Избранные примеры на основе разнообразия гемагглутинин-нейраминидазы (HN) среди связанных с заболеванием штаммов (таблица 1), расхождение между штаммом в рамках вакцины JL5 и связанными с заболеванием штаммами (таблица 2)
5	Вирус рубеллы	TPPPYQVSCGGES DRASARVIDPAA	ACCCCCCCCCCCTACCAGGTGAGCT GCGGCGGCGAGAGCGACAGGGCCA

	(RV)	QS (SEQ ID NO: 406)	GCGCCAGGGTGATCGACCCCGCCGC CCAGAGC (SEQ ID NO: 407)
6	Гемагглутинин	PEYAYKIVKNKK MEDGFLQGMVD GWYGHHSNEQGS GLMENERTLDKA NPNNDLCSWSDH EASSNNTNQEDLL QRESRRKKRIGTS TLNQRGNCNTKC QTEEARLKREEVS LVKSDQCSNGSL QCRANNSTEQVD (SEQ ID NO: 408)	CCCGAGTACGCCTACAAGATCGTGA AGAACAAGAAGATGGAGGACGGCT TCCTGCAGGGCATGGTGGACGGCTG GTACGGCCACCACAGCAACGAGCAG GGCAGCGGCCTGATGGAGAACGAG AGGACCCTGGACAAGGCCAACCCCA ACAACGACCTGTGCAGCTGGAGCGA CCACGAGGCCAGCAGCAACAACACC AACCAGGAGGACCTGCTGCAGAGGG AGAGCAGGAGGAAGAAGAGGATCG GCACCAGCACCCTGAACCAGAGGGG CAACTGCAACACCAAGTGCCAGACC GAGGAGGCCAGGCTGAAGAGGGAG GAGGTGAGCCTGGTGAAGAGCGACC AGTGCAGCAACGGCAGCCTGCAGTG CAGGGCCAACAACAGCACCGAGCA GGTGGAC (SEQ ID NO: 409)
7	Нейраминидаза	TKSTNSRSGGISG PDNEAPVGEAPSP YGDNPRPNDGNN IRIGSKGYNGIITD TIEESCSCYPDAK VVKSVELDSTIWT SGSSPNQKIITIGW DPNGWTGTPMSP NGAYGTDGPSNG QANQHQAESISA GNSSLCPIRDNWH GSNRSWSWPDGA E (SEQ ID NO: 410)	ACCAAGAGCACCAACAGCAGGAGC GGCGGCATCAGCGGCCCCGACAACG AGGCCCCCGTGGGCGAGGCCCCCAG CCCCTACGGCGACAACCCCAGGCCC AACGACGGCAACAACATCAGGATCG GCAGCAAGGGCTACAACGGCATCAT CACCGACACCATCGAGGAGAGCTGC AGCTGCTACCCCGACGCCAAGGTGG TGAAGAGCGTGGAGCTGGACAGCAC CATCTGGACCAGCGGCAGCAGCCCC AACCAGAAGATCATCACCATCGGCT GGGACCCCAACGGCTGGACCGGCAC CCCCATGAGCCCCAACGGCGCCTAC GGCACCGACGGCCCCAGCAACGGCC AGGCCAACCAGCACCAGGCCGAGA GCATCAGCGCCGGCAACAGCAGCCT GTGCCCCATCAGGGACAACTGGCAC GGCAGCAACAGGAGCTGGAGCTGGC CCGACGGCGCCGAG (SEQ ID NO: 411)
8	TAP (презентация транспортного антигена) на гемагглутинине вируса	EKIVLLLAMMEKI VLLLAKCQTPMG AIKAVDGVTNKC PYLGSPSF (SEQ ID NO: 412)	GAGAAGATCGTGCTGCTGCTGGCCA TGATGGAGAAGATCGTGCTGCTGCT GGCCAAGTGCCAGACCCCCATGGGC GCCATCAAGGCCGTGGACGGCGTGA CCAACAAGTGCCCCTACCTGGGCAG CCCCAGCTTC (SEQ ID NO: 413)

	H5N1		
9	TAP (презентация транспортного антигена) на нейраминидазе вируса H5N1	IRPCFWVELNPNQKIITIRPCFWVELICYPDAGEIT (SEQ ID NO: 414)	ATCAGGCCCTGCTTCTGGGTGGAGCTGAACCCCAACCAGAAGATCATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGATCTGCTACCCCGACGCCGGCGAGATCACC (SEQ ID NO: 415)
10	Эпитопы гемагглутинина к HLA класса I	MEKIVLLLAEKIVLLLAMCPYLGSPSFKCQTPMGAIKAVDGVTNK (SEQ ID NO: 416)	ATGGAGAAGATCGTGCTGCTGCTGGCCGAGAAGATCGTGCTGCTGCTGGCCATGTGCCCCTACCTGGGCAGCCCCAGCTTCAAGTGCCAGACCCCCATGGGCGCCATCAAGGCCGTGGACGGCGTGACCAACAAG (SEQ ID NO: 417)
11	Эпитопы нейраминидазы к HLA класса I	NPNQKIITICYPDAGEITIRPCFWVELRPCFWVELI (SEQ ID NO: 418)	AACCCCAACCAGAAGATCATCACCATCTGCTACCCCGACGCCGGCGAGATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGAGGCCCTGCTTCTGGGTGGAGCTGATC (SEQ ID NO: 419)
12	Эпитопы гемагглутинина к HLA класса II	MVSLVKSDQIGTSTLNQR (SEQ ID NO: 420)	ATGGTGAGCCTGGTGAAGAGCGACCAGATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 421)
13	Эпитопы нейраминидазы к HLA класса II	YNGIITDTI (SEQ ID NO: 422)	TACAACGGCATCATCACCGACACCATC (SEQ ID NO: 423)
14	HLA класса I и класса II, связанный с эпитопом гемагглутинина H5N1	MEKIVLLLAEKIVLLAMMVSLVKSDQCPYLGSPSFIGTSTLNQRKCQTPMGAIKAVDGVTNK (SEQ ID NO: 424)	ATGGAGAAGATCGTGCTGCTGCTGGCCGAGAAGATCGTGCTGCTGCTGGCCATGATGGTGAGCCTGGTGAAGAGCGACCAGTGCCCCTACCTGGGCAGCCCCAGCTTCATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 425)
15	HLA класса I и класса II, связанный с эпитопом нейраминидазы H5N1	NPNQKIITIYNGIITDTICYPDAGEITIRPCFWVELRPCFWVELI (SEQ ID NO: 426)	AACCCCAACCAGAAGATCATCACCATCTACAACGGCATCATCACCGACACCATCTGCTACCCCGACGCCGGCGAGATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGAGGCCCTGCTTCTGGGTGGAGCTGATC (SEQ ID NO: 427)

[1257] Ниже приведены дополнительные примеры эпитопов, которые могут быть встроены в ген-мишень для иммуноэпитопного мечения.

Ссылки, приведенные в Примере 8

[1258] Следующие ссылки полностью приведены здесь в качестве ссылки.

1. X. Wen, K. Wen, D. Cao, G. Li, R. W. Jones, J. Li, S. Szu, Y. Hoshino, L. Yuan, Inclusion of a universal tetanus toxoid CD4(+) T cell epitope P2 significantly enhanced the immunogenicity of recombinant rotavirus ΔVP8* subunit parenteral vaccines. *Vaccine* 32, 4420-4427 (2014).

2. G. Ada, D. Isaacs, Carbohydrate-protein conjugate vaccines. *Clin Microbiol Infect* 9, 79-85 (2003).

3. E. Malito, B. Bursulaya, C. Chen, P. L. Surdo, M. Picchianti, E. Balducci, M. Biancucci, A. Brock, F. Berti, M. J. Bottomley, M. Nissum, P. Costantino, R. Rappuoli, G. Spraggon, Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. Proceedings of the National Academy of Sciences 109, 5229 (2012).

4. J. de Wit, M. E. Emmelot, M. C. M. Poelen, J. Lanfermeijer, W. G. H. Han, C. van Els, P. Kaaijk, The Human CD4(+) T Cell Response against Mumps Virus Targets a Broadly Recognized Nucleoprotein Epitope. *J Virol* 93, (2019).

5. M. May, C. A. Rieder, R. J. Rowe, Emergent lineages of mumps virus suggest the need for a polyvalent vaccine. *Int J Infect Dis* 66, 1-4 (2018).

6. M. Ramamurthy, P. Rajendiran, N. Saravanan, S. Sankar, S. Gopalan, B. Nandagopal, Identification of immunogenic B-cell epitope peptides of rubella virus E1 glycoprotein towards development of highly specific immunoassays and/or vaccine. *Conference Abstract*, (2019).

7. U. S. F. Tambunan, F. R. P. Sipahutar, A. A. Parikesit, D. Kerami, Vaccine Design for H5N1 Based on B- and T-cell Epitope Predictions. Bioinform Biol Insights 10, 27-35 (2016).

ПРИМЕР 9 – *IN VIVO* ДОСТАВКА АГЕНТОВ РЕ

[1259] Технологии точного геномного нацеливания с использованием системы CRISPR/Cas9 недавно были рассмотрены в рамках широкого ряда вариантов

использования, включая генную терапию. Значительным ограничением использования Cas9 и основанных на Cas9 агентов для геномного редактирования в рамках генной терапии является размер Cas9 (>4 kb), что ограничивает его эффективную доставку с помощью рекомбинантного адено-ассоциированного вируса (rAAV). Недавно разработанные продукты слияния Cas9 с обратными транскриптазами ("праймерные редакторы"), представляют собой новую технологию геномного редактирования, которая обеспечивает ряд преимуществ по сравнению с существующими способами геномного редактирования, включая способность вносить любое нуклеотидное замещение и обеспечивать вставку или делецию любой произвольно определенной короткой (<~20) последовательности нуклеотидов специфичным в плане участка образом. Таким образом, данный способ обеспечивает возможность редактирования патогенных вариантов человека, которые ранее являлись недостижимыми для корректирования. Доставка реагентов для праймерного редактирования может обеспечивать корректирование генетических последовательностей, которые вызывают заболевание человека, или обеспечивать внесение предотвращающих заболевание вариантов генов.

[1260] В рамках данного изобретения описаны способы доставки праймерных редакторов в клетки *in vitro* и *in vivo*. Праймерные редакторы были разработаны и описаны исключительно в культивированных клетках. Не один известный способ не может обеспечить доставку праймерных редакторов *in vivo*. Ранее описанные способы доставки праймерных редакторов с помощью rAAV или предварительно собранных рибонуклеопротеиновых (RNP) комплексов обеспечат преодоление некоторых барьеров для *in vivo* доставки. Например, ДНК, кодирующая праймерные редакторы, является более крупной, чем верхний предел упаковывания rAAV, и поэтому требует специальных решений. Одно подобное решение заключается в конструировании редактора, слитого с парами разделенного интеина, которые упакованы в две отдельные rAAV частицы, которые при совместной доставке в клетку обеспечивают восстановление функционального белка редактора. Описаны некоторые другие специальные факторы, которые следует учитывать в плане уникальных свойств праймерного редактирования, включая оптимизацию мишеней никирования второго участка и правильное упаковывание праймерных редакторов в вирусные векторы, включая лентивирусы и rAAV.

[1261] Отличительные признаки включают использование составов для рибонуклеопротеиновой (RNP) доставки, при этом праймерные редакторы и соседние мишени никирования могут предварительно образовывать комплексы с их соответствующей sgРНК/PEgRNA. Это увеличит диапазон возможных участков нацеливания и обеспечит лучшую оптимизацию эффективности редактирования относительно текущих данных, связанных с доставкой в ДНК. При использовании составов для доставки на основе RNP или иРНК могут применяться варианты белков Cas, каждый из которых образует комплекс со своим собственным вариантом направляющей РНК. Это также обеспечит большее разнообразие потенциальных локусов никирования, поэтому предполагается, что оптимизация может быть реализована для большей эффективности в рамках любого варианта использования. На основе предыдущих описаний RNP предполагается, что использование RNP повысит специфичность редактирования (Rees et al., 2017). Это снизит нецелевое праймерное редактирование. Описаны потенциальные архитектуры для разделения праймерных редакторов на два AAV вектора для доставки *in vivo* или *ex vivo*. Упаковывание праймерного редактора в двойную AAV систему требует учета оптимизации конструкции, включая участки разделения, способы восстановления (например интеины), и архитектуру направляющей экспрессии. При использовании смеси вируса и RNP для доставки праймерного редактора ожидается, что редактирование будет контролироваться с течением времени, поскольку RNP в конечном итоге расщепляется *in vivo*, что остановит праймерное редактирование после прекращения поступления RNP.

[1262] Рибонуклеобелок (RNP) праймерного редактора, иРНК с направляющей РНК праймерного редактора, или ДНК могут быть упакованы в липидные наночастицы, rAAV, или лентивирус, и инъецированы, проглочены или ингалированы для изменения геномной ДНК *in vivo* и *ex vivo*, включая цели, связанные с реализацией моделей заболеваний человека с использованием животных, тестирование и научные исследования в рамках моделей заболевания человека с использованием животных, и лечение заболевания у людей.

[1263] Праймерные редакторы могут быть использованы для корректирования значительной доли всех генетических заболеваний (~89% патогенных генетических вариантов человека в рамках Clinvar), если разработаны подходящие средства доставки в соответствующие типы клеток *in vivo*. Заболевания крови, заболевания сетчатки, и

заболевания печения являются наиболее вероятными первыми вариантами использования вследствие наличия систем доставки для других реагентов. В комбинации с данным изобретением должны будут использоваться капсиды AAV, другие подвергнутые эволюции или модифицированные вирусные векторы, а также составы на основе липидных наночастиц.

[1264] В определенных вариантах осуществления один или несколько доменов праймерного редактора (например домен napDNAbp или домен RT) могут быть модифицированы с использованием последовательности интеина.

ПРИМЕР 10 – ПРИМЕНЕНИЕ PE ДЛЯ ИДЕНТИФИКАЦИИ НЕЦЕЛЕВОГО РЕДАКТИРОВАНИЯ

[1265] В настоящее время отсутствуют описанные способы детектирования нецелевого редактирования при использовании праймерных редакторов (описание самого праймерного редактирования еще не было опубликовано). Эти способы позволили бы исследователю идентифицировать потенциальные участки нецелевого редактирования с помощью праймерных редакторов, что являлось бы важным фактором, если данная технология использовалась бы для лечения генетического заболевания у пациентов.

[1266] Описанные здесь способы также могут применяться для идентификации нецелевых мишеней нуклеаз Cas. Данные нецелевые мишени ранее были идентифицированы с помощью BLESS, Guide-Seq, CIRCLE-Seq, и Digenome-Seq. Однако данный способ является полезным вследствие чувствительности и простоты процесса.

[1267] Ключевая концепция данного аспекта заключается в идее применения праймерного редактирования для вставки адаптерной последовательности или участка связывания праймера, матрицей для которых служит PEgRNA, для обеспечения быстрой идентификации геномных участков нецелевой модификации нуклеаз Cas или праймерных редакторов.

[1268] Неизвестен ни один объективный способ идентификации нецелевых участков праймерного редактирования. Данный способ является отличным от других способов, которые обеспечивают идентификацию нуклеазных нецелевых участков, поскольку адаптерная последовательность вставляется одновременно со связыванием и никированием ДНК, что упрощает дальнейшее процессирование.

[1269] Настоящее изобретение включает идентификацию нецелевых участков редактирования при редактировании внутри живой клетки, в культуре ткани, или моделях с использованием животных (см. Фиг.33). Для реализации данного способа получают PEgRNA, которая имеет спейсер, являющийся идентичным конечному требуемому редактору (и, если смотреть на нецелевые участки праймерного редактирования, последовательность участка связывания праймера, которая идентична конечному требуемому редактору), но включает необходимые последовательности для встраивания адаптера или участка связывания праймера после обратной транскрипции в рамках праймерного редактирования. *In vivo* редактирование осуществляется с помощью праймерного редактора или RT-слитой нуклеазы и выделения геномной ДНК. Геномная ДНК фрагментируется ферментативным или механическим способом и обеспечивает присоединение другого адаптера к участкам фрагментирования ДНК. ПЦР используется для амплификации от одного адаптера к адаптеру, внесенному с помощью PEgRNA. Полученный продукт подвергается глубокому секвенированию для идентификации всех модифицированных участков.

[1270] Изобретение также включает идентификацию нецелевых участков редактирования с помощью *in vitro* модификации геномной ДНК (см. Фиг.33). Для реализации данного способа собирают RNP из очищенного белка праймерного редактора и PEgRNA, который обеспечит внесение адаптера или участка связывания праймера, но в другой стороны является таким же, что и представляющая интерес PEgRNA. Данная RNP инкубируется с выделенной геномной ДНК до или после фрагментации ДНК и присоединения различных адаптеров к участкам разрывов ДНК. Используют ПЦР для амплификации из фрагментированного участка в адаптер, который был внесен с помощью PE. Используют глубокое секвенирование для идентификации участков модификации. Данный *in vitro* способ редактирования должен повышать чувствительность детектирования, поскольку репарация клеточной ДНК никогда не устранит подвергнутый обратной транскрипции ДНК-адаптер, добавленный с помощью праймерного редактора.

[1271] Данные способы могут применяться для идентификации нецелевого редактирования для любого праймерного редактора или любого геномного редактора, в рамках которого используется направляющая РНК для распознавания участка-мишени разрыва (как в случае большинства Cas нуклеаз).

[1272] Данные способы могут применяться для любых генетических заболеваний, для лечения которых могут использоваться геномные редакторы.

ПРИМЕР 11 – ПРИМЕНЕНИЕ РЕ ДЛЯ ОБЕСПЕЧЕНИЯ ХИМИЧЕСКИ ИНДУЦИРУЕМОЙ ДИМЕРИЗАЦИИ БЕЛКОВ-МИШЕНЕЙ *IN VIVO*

[1273] Описанные здесь праймерные редакторы также могут применяться для реализации индуцируемых димеризацией биологических процессов, например передачи сигнала рецептора, под контролем удобного низкомолекулярного лекарства за счет геномного встраивания генов, кодирующих связывающие малую молекулу белки, с помощью праймерного редактирования, как описано здесь. С помощью описанных здесь праймерных редакторов последовательность гена, кодирующего белок, связывающий малую молекулу, может быть вставлена в ген, кодирующий представляющий интерес белок-мишень в живой клетке или пациенте. Данное изменение самостоятельно не должно иметь физиологического эффекта. При введении низкомолекулярного лекарства, которое обычно представляет собой димерную малую молекулу, которая может одновременно связываться с двумя доменами белка, связывающими лекарство, каждый из которых слит с копией белка-мишени, малая молекула вызывает димеризацию белка-мишени. Данная димеризация белка-мишени затем индуцирует передачу биологического сигнала, например эритропоэз или передачу сигнала инсулина.

ПРИМЕР 12 - ПРАЙМЕРНОЕ РЕДАКТИРОВАНИЕ: ВЫСОКОУНИВЕРСАЛЬНОЕ И ТОЧННОЕ ГЕНОМНОЕ РЕДАКТИРОВАНИЕ ТИПА ПОИСК-И-ЗАМЕНА В КЛЕТКАХ ЧЕЛОВЕКА БЕЗ ДВУЦЕПОЧЕЧНЫХ РАЗРЫВОВ ДНК

[1274] Современные способы геномного редактирования могут обеспечивать изменение, делецию, или вставку генов-мишеней с соответствующими сопутствующими продуктами двуцепочечных разрывов ДНК с использованием программируемых нуклеаз и внесение четырех и внесение четырех переходных точечных мутаций в локусах-мишенях с помощью редакторов оснований. Малые вставки, малые делеции, и восемь трансверсивных точечных мутаций, однако, совместно представляют собой наиболее патогенные генетические варианты, но не могут быть скорректированы эффективным образом и без избытка побочных

продуктов в большинстве типов клеток. Здесь описано праймерное редактирование, являющееся высокоуниверсальным и точным способом геномного редактирования, который обеспечивает прямую запись новой генетической информации в указанный участок ДНК с использованием каталитически ослабленного Cas9, слитого со сконструированной обратной транскриптазой, запрограммированной с использованием сконструированной направляющей РНК для праймерного редактирования (PEgRNA), которая как указывает участок-мишень, так и кодирует требуемое изменение. В клетках человека было реализовано более чем 175 отдельных изменений для определения того, что праймерное редактирование может обеспечивать реализацию нацеленных вставок, делеций, всех 12 возможных типов точечных мутаций, и их комбинаций эффективным образом (обычно 20-60%, вплоть до 77% в неотсортированных клетках) и с низким показателем побочных продуктов (обычно 1-10%) без необходимости двуцепочечных разрывов и донорских матриц ДНК. Праймерное редактирование использовалось в клетках человека для корректирования первичных генетических причин серповидноклеточной анемии (требующей трансверсии A•T на T•A в *HBB*) и болезни Тея-Сакса (требующей делеции 4 оснований в *HEXA*), что в обоих случаях эффективным образом приводило к возврату патогенных геномных аллелей к дикому типу с минимальными побочными продуктами. Праймерное редактирование также использовалось для создания линий клеток человека с данными патогенными мутациями в виде трансверсии *HBB* и инсерции в *HEXA* для внесения мутации G127V в *PRNP*, которая придает устойчивость к прионному заболеванию (требующему трансверсии G•C на T•A), а также эффективной вставки His6 метки, эпитопной метки FLAG, и удлиненного участка LoxP в локусы-мишени в клетках человека. Праймерное редактирование обеспечивает преимущества в виде эффективности и чистоты продукта по сравнению с HDR, а также дополнительные сильные и слабые стороны по сравнению с редактированием оснований. В соответствии с его механизмом на основе "поиска-и-замены", который требует трех отдельных событий в виде спаривания оснований, праймерное редактирование намного менее предрасположено к нецелевой модификации ДНК в известных нецелевых участках Cas9, чем Cas9. Праймерное редактирование по существу расширяет объем и возможности геномного редактирования и в принципе может обеспечивать корректирование ~89% известных патогенных генетических вариантов человека.

[1275] Способность осуществлять практически любое нацеленное изменение в геноме любой живой клетки или организме является давним стремлением в рамках наук о жизни. Несмотря на быстрое развитие технологий геномного редактирования, большинство из >75000 известных генетических вариантов человека, ассоциированных с заболеваниями[111], не могут быть скорректированы или внесены в большинство терапевтически значимых клеток (Фиг.38A). Программируемые нуклеазы, такие как CRISPR-Cas9, могут осуществлять двуцепочечные разрывы ДНК (DSB), которые могут изменять гены за счет индуцирования смесей вставок или делеций (инделей) в участках-мишенях[112-114]. Нуклеазы также могут использоваться для делеции генов-мишеней[115,116] или вставки экзогенных генов[117-119] за счет независимого от гомологии процесса. Однако двуцепочечные разрывы ДНК также ассоциированы с нежелательными результатами, включая комплексные смеси продуктов, транслокации[120], и активацию p53[121,122]. Более того, подавляющее большинство патогенных аллелей отличается от их непатогенных аналогов наличием малых вставок, делеций, или замещений оснований, которые требуют для корректирования намного более точных технологий редактирования (Фиг.38A). Направленная гомологичная репарация (HDR), стимулируемая индуцируемыми нуклеазой DSB[123], широко использовалась для внесения ряда точных изменений ДНК. Однако HDR основана на экзогенных донорских матрицах для репарации ДНК, обычно приводит к получению избытка побочных продуктов в виде инделей после репарации DSB со связыванием концов, и является неэффективной в рамках большинства терапевтически релевантных типов клеток (при этом Т-клетки и некоторые стволовые клетки являются важными исключениями)[124,125]. В то время как повышение эффективности и точности DSB-опосредованного геномного редактирования остается целью заметных усилий[126-130], данные проблемы требуют анализа альтернативных стратегий точного геномного редактирования.

[1276] Редактирование оснований может эффективным образом обеспечивать внесение или корректирование четырех типов переходных мутаций (C на T, G на A, A на G, и T на C) без необходимости DSB в рамках широкого ряда типов клеток и организмов, включая млекопитающих[128-131], но в настоящее время не может обеспечивать реализацию какой-либо из восьми трансверсивных мутаций (C на A, C на G, G на C, G на T, A на C, A на T, T на A, и T на G), например мутации T•A на A•T, которая необходима для прямого корректирования наиболее частой причины

серповидноклеточной анемии (*HBB* E6V)[132]. Кроме того, не сообщается ни об одном способе без DSB для реализации нацеленных делеций, таких как удаление дупликации в виде 4 оснований, которая вызывает болезнь Тея-Сакса (*HEXA* 1278+TATC)[133], или нацеленных вставок, например точной вставки в виде 3 оснований, требуемой для прямого корректирования наиболее частой причины муковисцидоза (*CFTR* ΔF508)[134]. Таким образом, нацеленный трансверсивные точечные мутации, вставки, и делеции являются сложными в плане внесения или эффективного корректирования и при отсутствии избытка побочных продуктов в большинстве типов клеток, даже несмотря на то, что они совместно связаны с большинством известных патогенных аллелей (Фиг.38A).

[1277] Здесь описана разработка праймерного редактирования, новой технологии геномного редактирования типа "поиск-и-замена", которая обеспечивает возможность реализации нацеленных вставок, делеций, и всех 12 возможных конверсий типа основание-на-основание в локусах-мишенях в клетках человека без необходимости двуцепочечных разрывов ДНК или донорских матриц ДНК. В рамках праймерных редакторов, примером которого первоначально был PE1, используется обратная транскриптаза, слитая с программируемой никазой и удлиненной направляющей РНК для праймерного редактирования (PEgRNA), для прямого копирования генетической информации из удлинения на PEgRNA в геномный локус-мишень. В рамках праймерного редактора второго поколения (PE2) используется сконструированная обратная транскриптаза для значительного повышения эффективности редактирования с минимальным (обычно <2%) образованием инделей, в то время как в PE3 системе третьего поколения добавлена вторая направляющая РНК для никирования не подвергаемой редактированию цепи, тем самым обеспечивая предпочтительное замещение не подвергаемой редактированию цепи и дополнительно повышая эффективность редактирования, обычно на 20-50% в клетках человека с приблизительно 1-10% образованием инделей. PE3 обеспечивает намного меньшее количество побочных продуктов и более высокую или аналогичную эффективность относительно оптимизированной инициируемой нуклеазой Cas9 HDR, а также дополнительные сильные и слабые стороны по сравнению с редакторами оснований текущего поколения.

[1278] PE3 использовался в геномных локусах в клетках HEK293T человека для реализации эффективной конверсии *HBB* E6V в *HBB* дикого типа, делеции

вставленных TATC для восстановления *HEXA 1278+TATC* до *HEXA* дикого типа, внесения мутации G127V в *PRNP*, которая придает устойчивость к прионному заболеванию135 (требующему трансверсии G•C на T•A), и нацеленной вставки His_6-метки (18 bp), эпитопной метки FLAG (24 bp), и удлиненного участка *LoxP* для Cre-опосредованной рекомбинации (44 bp). Праймерное редактирование также было успешным в рамках трех других линий клеток человека, а также в постмитотических первичных нейронах коры головного мозга мыши с различными показателями эффективности. Вследствие высокой степени гибкости на расстоянии между исходным разрывом и месторасположением изменения, праймерное редактирование по существу не ограничено необходимостью PAM для Cas9 и в принципе может обеспечивать нацеливание на подавляющее большинство геномных локусов. Нецелевое праймерное редактирование является намного более редким, чем нецелевое редактирование при использовании Cas9 в известных нецелевых локусах Cas9, вероятно вследствие необходимости в трех отдельных событиях в виде спаривания оснований ДНК для осуществления продуктивного праймерного редактирования. За счет обеспечения возможности реализации точных нацеленных вставок, делеций, и всех 12 возможных классов точечных мутаций в широком диапазоне геномных локусов без необходимости DSB или донорских матриц ДНК праймерное редактирование обладает потенциалом в плане продвижения изучения и корректирования многих вариантов генов.

Результаты

Стратегия для переноса информации из удлиненной направляющей РНК в локус-мишень ДНК

[1279] Cas9 нацеливается на ДКН с помощью направляющей РНК, содержащей спейсерную последовательность, которая гибридизируется с участком-мишенью ДНК[112-114,136,137]. Целью являлось сконструировать направляющие РНК как для указания ДНК мишени (как в природных системах CRISPR)[138,139], так и для включения новой генетической информации, которая замещает соответствующие нуклеотиды ДНК в локусе-мишени. Прямой перенос генетической информации из удлиненной направляющей РНК в указанный участок ДНК с последующим замещением оригинальной не подвергаемой редактированию ДНК в принципе может обеспечить универсальное средство внесения нацеленных изменений последовательности ДНК в живых клетках без зависимости от DSB или донорских матриц ДНК. Для реализации

данного прямого переноса информации задачей являлось использовать геномную ДНК, никированную в участке-мишени, для высвобождения 3′-гидроксильной группы для праймирования обратной транскрипции генетической информации из удлинения на сконструированной направляющей РНК (здесь и далее обозначаемой как направляющая РНК для праймерного редактирования или PEgRNA) напрямую в участок-мишень (Фиг.38A).

[1280] Данные начальные этапы никирования и обратной транскрипции, которые похожи на механизмы, используемые некоторыми природными подвижными генетическими элементами[140], приводят к получению разветвленного промежуточного продукта с двумя лишними одноцепочечными ДНК-флэпами на одной цепи: 5′ флэпом, который содержит не подвергаемую редактированию последовательность ДНК, и 3′ флэпом, который содержит подвергаемую редактированию последовательность, скопированную из PEgRNA (Фиг.38B). Для реализации успешного изменения данный разветвленный промежуточный продукт должен быть подвергнут коррекции, так что отредактированный 3′ замещает не подвернутый редактированию 5′ флэп. В то время как гибридизация 5′ флэпа с не подвергаемой редактированию цепью вероятно будет термодинамически более благоприятной, поскольку подвергаемый редактированию 3′ флэп может образовывать меньшее количество пар оснований с не подвергаемой редактированию цепью, 5′ флэпы являются предпочтительным субстратом для специфичных в плане структуры эндонуклеаз, таких как FEN1[141], которые обеспечивают эксцизию 5′ флэпов, образованных при синтезе запаздывающей цепи ДНК и репарации эксцизии длинной последовательности оснований. Было предположено, что предпочтительная эксцизия 5′ флэпа и лигирование 3′ флэпа могут обеспечивать встраивание подвергаемой редактированию цепи ДНК, создавая гетеродуплексную ДНК, содержащую одну подвергаемую редактированию цепь и одну не подвергаемую редактированию цепь (Фиг.38B).

[1281] Перманентное внесение изменения может возникать из последующей репарации ДНК, которая обеспечивает корректирование несовпадения между данными двумя цепями ДНК таким образом, при котором обеспечивается копирование информации из подвергаемой редактированию цепи в комплементарную цепь ДНК (Фиг.38C). На основе аналогичной стратегии, разработанной для максимизации эффективности редактирования оснований ДНК[131-133], было предположено, что никирование не подвергаемой редактированию цепи ДНК на достаточно далеком расстоянии от

участка начального разрыва для минимизации образования двуцепочечных разрывов может направлять репарацию ДНК к предпочтительному замещению не подвергаемой редактированию цепи.

Анализ этапов праймерного редактирования in vitro и в клетках дрожжей

[1282] После расщепления PAM-содержащей цепи ДНК с помощью нуклеазного домена RuvC Cas9 удаленный от PAM фрагмент данной цепи может диссоциироваться от иным образом стабильных комплексов Cas9:sgРНК:ДНК[143]. Было предположено, что 3′-конец данной высвобожденной цепи может являться достаточно доступным для праймирования полимеризации ДНК. Усилия по конструированию направляющих РНК[144-146] и кристаллических структур комплексов Cas9:sgРНК:ДНК[147-149] свидетельствуют о том, что 5′ и 3′-концы sgРНК могут быть удлинены без устранения активности Cas9:sgРНК. PEgRNA были сконструированы путем удлинения sgРНК для включения двух критических компонентов: участка связывания праймера (PBS), который обеспечивает гибридизацию 3′-конца никированной цепи ДНК с PEgRNA, и матрицы в виде обратной транскриптазы (RT), содержащей требуемое изменение, которое будет напрямую скопировано в участок геномной ДНК, поскольку 3′-конец никированной цепи ДНК удлиняется вдоль РНК матрицы за счет полимеразы (Фиг.38C).

[1283] Данные предположения были протестированы *in vitro* с помощью очищенного белка Cas9 *S. pyogenes*. Конструировали группу PEgRNA-кандидатов путем удлинения sgРНК на любом конце с использованием последовательности PBS (от 5 до 6 нуклеотидов, nt) и RT матрицы (от 7 до 22 nt). Было подтверждено, что 5′-удлиненные PEgRNA обеспечивают связывание Cas9 с ДНК-мишенью, а также то, что как 5′-удлиненные PEgRNA, так и 3′-удлиненные PEgRNA поддерживают Cas9-опосредованное никирование мишени *in vitro* и связанную с расщеплением ДНК активность в клетках млекопитающих (Фиг.44A-44C). Данные конструкции-кандидаты PEgRNA были протестированы с использованием подвергнутых предварительному никированию 5′ Cy5-меченных субстратов дцДНК, каталитически мертвого Cas9 (dCas9), и коммерческого варианта обратной транскриптазы вируса лейкоза мышей Молони (M-MLV) (Фиг.44D). После внесения всех компонентов наблюдали эффективную конверсию флуоресцентно меченной цепи ДНК в более длинные продукты ДНК с подвижностью на основе геля, что соответствует обратной

транскрипции вдоль RT матрицы (Фиг.38D, Фиг.44D-44E). Продукты требуемой длины образовывались при использовании либо 5′-удлиненных, или 3′-удлиненных PEgRNA (Фиг.38D-38E). Пропуск dCas9 приводит к продуктам трансляции разрыва, полученным на основе опосредованной обратной транскриптазой полимеризации ДНК на матрице ДНК без переноса информации PEgRNA (Фиг.38D). Продукты полимеризации ДНК не наблюдались при замещении PEgRNA стандартной sgРНК, что подтверждает необходимость компонентов PEgRNA в виде PBS и RT матрицы (Фиг.38D). Данные результаты демонстрируют, что Cas9-опосредованное плавление ДНК высвобождает одноцепочечную R-петлю, которая в случае никирования может праймировать обратную транскрипцию из 5′-удлиненной или 3′-удлиненной PEgRNA.

[1284] Далее были протестированы субстраты дцДНК без разрывов с использованием никазы Cas9 (мутант H840A), которая никирует исключительно PAM-содержащую цепь[112]. В рамках данных реакций 5′-удлиненные PEgRNA образовывали продукты обратной транскрипции неэффективным образом, вероятно вследствие нарушенной активности никазы Cas9 (Фиг.44F). Однако 3′-удлиненные PEgRNA обеспечивали устойчивое никирование за счет Cas9 и эффективную обратную транскрипцию (Фиг.38E). Использование 3′-удлиненных PEgRNA приводило к получению только одного заметного продукта, несмотря на возможную терминацию обратной транскрипции в любом месте в рамках остатка PEgRNA. Секвенирование ДНК продуктов реакций с использованием никазы Cas9, RT, и 3′-удлиненных PEgRNA показало, что полная последовательность RT матрицы подвергалась обратной транскрипции в субстрат ДНК (Фиг.44G). Данные эксперименты подтвердили, что 3′-удлиненные PEgRNA могут служить матрицей для обратной транскрипции новых цепей ДНК в то же время сохраняя способность обеспечивать активность никазы Cas9.

[1285] Для оценки результатов репарации ДНК 3′ флэпов эукариотических клеток, образованных в рамках PEgRNA-программируемой обратной транскрипции *in vitro*, осуществляли никирование ДНК и обратную транскрипцию с использованием PEgRNA, никазы Cas9, и RT *in vitro* на субстратах в виде репортерной плазмиды, при этом продукты реакции затем трансформировали в клетки дрожжей (*S. cerevisiae*) (Фиг.45A). Когда плазмиды подвергались редактированию in vitro с использованием 3′-удлиненных PEgRNA, кодирующих трансверсию T•A на A•T, которая обеспечивает корректирование преждевременного стоп-кодона, 37% дрожжевых трансформантов обнадеживающим образом экспрессировали как GFP, так и mCherry белки (Фиг.38F,

Фиг.45C). В полном соответствии с результатами, показанными на Фиг.38E и Фиг.44F, реакции редактирования, осуществленные in vitro с использованием 5′-удлиненных PEgRNA, приводили к меньшему количеству двойных положительных в плане GFP и mCherry колоний (9%), чем при использовании 3′-удлиненных PEgRNA (Фиг.38F и Фиг.45D). Продуктивное редактирование также наблюдалось при использовании 3′-удлиненных PEgRNA, которые обеспечивают вставку одного нуклеотида (15% двойных положительных трансформантов) или делецию одного нуклеотида (29% двойных положительных трансформантов) для корректирования мутаций сдвига рамки считывания (Фиг.38F и Фиг.45E-45F). Секвенирование ДНК подвергнутых редактированию плазмид, выделенных из двойных положительных колоний дрожжей, подтвердило, что кодируемое трансверсивное изменение было осуществлено в необходимом положении последовательности (Фиг.45G). Данные результаты демонстрируют, что репарация ДНК в эукариотических клетках может обеспечивать корректирование 3′ ДНК-флэпов, образованных при праймерном редактировании, для включения точных изменений ДНК, включая трансверсии, вставки, и делеции.

Конструирование праймерного редактора 1 (PE1)

[1286] На основе обнадеживающих результатов *in vitro* и в дрожжах были предприняты попытки разработать систему праймерного редактирования с минимальным количеством компонентов, способных обеспечивать редактирование геномной ДНК в клетках млекопитающих. Было предположено, что 3′-удлиненные PEgRNA (здесь и далее обозначаемые просто как PEgRNA, Фиг.39A) и продукты прямого слияния Cas9 H840A с обратной транскриптазой через гибкий линкер могут представлять собой функциональную двухкомпонентную систему праймерного редактирования. HEK293T (иммортализованные эмбриональные клетки почки человека) трансфицировали с использованием одной плазмиды, кодирующей продукт слияния обратной транскриптазы M-MLV дикого типа с любым концом никазы Cas9 H840A, а также второй плазмиды, кодирующей PEgRNA. Первоначальные попытки приводили к отсутствию детектируемой конверсии T•A на A•T в локусе-мишени HEK3.

[1287] Удлинение PBS в PEgRNA до 8-15 оснований (Фиг.39A), однако, приводило к детектируемому редактированию T•A на A•T в участке-мишени HEK3 (Фиг.39B) с более высокой эффективностью для конструктов праймерного редактора, в которых

RT была слита с С-концом никазы Cas9 (максимальная конверсия T•A на A•T 3,7% с длиной PBS в диапазоне 8-15 nt), по сравнению с N-концевыми продуктами слияния RT-никазы Cas9 (максимальная конверсия T•A на A•T 1,3%) (Фиг.39B; все приведенные здесь данные для клеток млекопитающих включают значения для всей подвергнутой обработке популяции клеток без отбора или сортировки, если не указано иное). Данные результаты свидетельствуют о том, что RT M-MLV дикого типа, слитая с Cas9, требует более длинных последовательностей PBS для геномного редактирования в клетках человека по сравнению с тем, что требуется *in vitro* при использовании коммерческого варианта RT M-MLV, поступающего in trans. Данная обратная транскриптаза M-MLV дикого типа первого поколения, слитая с С-концом никазы Cas9 H840A, была обозначена как PE1.

[1288] Тестировалась способность PE1 обеспечивать точное внесение трансверсивных точечных мутаций в четырех дополнительных геномных-участках мишенях, указанных с помощью PEgRNA (Фиг.39C). Аналогично редактированию в локусе HEK3, эффективность в данных геномных участках зависела от длины PBS, при этом максимальная эффективность редактирования была в диапазоне 0,7-5,5% (Фиг.39C). Уровень инделей в рамках PE1 был низким, в среднем составляя 0,2±0,1% для пяти участков при условиях, которые максимизировали эффективность редактирования каждого участка (Фиг.46A). PE1 также был способен обеспечивать внесение нацеленных вставок и делеций, примером которых служит делеция одного нуклеотида (эффективность 4,0%), вставка одного нуклеотида (9,7%), и вставка трех нуклеотидов (17%) в локусе HEK3 (Фиг.39C). Данные результаты подтверждают способность PE1 напрямую обеспечивать внесение нацеленных трансверсий, вставок, и делеций без необходимости двуцепочечных разрывов ДНК или матриц ДНК.

[1289] Конструирование праймерного редактора 2 (PE2)

[1290] В то время как PE1 может обеспечивать внесение множества изменений в нескольких локусах в клетках HEK293T, эффективность редактирования в целом являлась низкой (обычно ≤5%) (Фиг.39C). Было предположено, что модифицирование обратной транскриптазы в PE1 может повышать эффективность синтеза ДНК в рамках уникальных конформационных факторов комплекса для праймерного редактирования, приводящих к более высоким показателям геномного редактирования. Ранее сообщалось о мутациях RT M-MLV, которые повышают термостабильность

ферментов[150,151], процессирование[150], и аффинность гетеродуплексного ДНК:РНК субстрата[152], и которые деактивируют активность РНКазы H[153]. Было сконструировано 19 вариантов PE1, содержащих ряд мутаций обратной транскриптазы для оценки эффективности их праймерного редактирования в клетках человека.

[1291] Сначала анализировали группу вариантов RT M-MLV, которые были получены ранее с помощью лабораторной эволюции на основе их способности поддерживать обратную транскрипцию при повышенных температурах[150]. Последующее внесение трех из данных аминокислотных замещений (D200N, L603W, и T330P) в RT M-MLV, далее обозначаемых как M3, приводило в среднем к 6,8-кратному повышению эффективности редактирования в виде трансверсии и вставки среди пяти геномных локусов в клетках HEK293T по сравнению с эффективностью PE1 (Фиг.47A-47S).

[1292] Далее в комбинации с M3 тестировали дополнительные мутации обратной транскриптазы, которые, как было ранее показано, повышают связывание с комплексом матрица:PBS, процессирование ферментов, и термостабильность[152]. Среди данных дополнительных проанализированных 14 мутантов вариант с замещениями T306K и W313F, наряду с мутациями M3, дополнительно обеспечивал от 1,3-кратного до 3,0-кратного повышения эффективности редактирования по сравнению с M3 для шести изменений в виде трансверсии или вставки среди пяти геномных участков в клетках человека (Фиг.47A-47S). Данный пентамутант обратной транскриптазы M-MLV, встроенный в архитектуру PE1 (Cas9 H840A–RT M-MLV (D200N L603W T330P T306K W313F)), далее обозначается как PE2.

[1293] PE2 обеспечивает внесение мутаций в виде трансверсии, вставки, и делеции одного нуклеотида со значительно более высокой эффективностью, чем PE1 (Фиг.39C), и является совместимым с более короткими последовательностями PBS PEgRNA (Фиг.39C), что соответствует повышенной способности продуктивно связывать транзиентные комплексы геномная ДНК:PBS. В среднем PE2 приводил к от 1,6 до 5,1-кратному повышению эффективности точечной мутации в рамках праймерного редактирования по сравнению с PE1 (Фиг.39C), и при этом в некоторых случаях — к крайне значительному повышению выхода редактирования вплоть до 46-кратного (Фиг.47F и Фиг.47I). PE2 также обеспечивал осуществление нацеленных вставок и делеций более эффективным образом, чем PE1, позволяя достигать реализацию нацеленной вставки 24 bp эпитопной метки FLAG в локус HEK3 с 4,5% эффективностью, то есть 15-кратным повышением по сравнению с эффективностью

внесения данной вставки с использованием PE1 (Фиг.47D), и опосредовал 1 bp делецию в HEK3 с 8,3% эффективностью, то есть в 2,1 раза большей по сравнению с PE1 (Фиг.39C). Данные результаты подтверждают, что PE2 является более эффективным праймерным редактором, чем PE1.

[1294] Оптимизация свойств PEgRNA

[1295] Связь между архитектурой PEgRNA и эффективностью праймерного редактирования была систематически проанализирована в пяти геномных локусах в клетках HEK293T в рамках PE2 (Фиг.39C). В целом, участки праймирования с более низким содержанием GC требовали более длинных последовательностей PBS (EMX1 и RNF2, содержащие 40% и 30% GC, соответственно, в первых 10 nt выше разрыва), в то время как имеющие более высокое содержание GC поддерживали праймерное редактирование с использованием более коротких последовательностей PBS (HEK4 и FANCF, содержащие 80% и 60% GC, соответственно, в первых 10 nt выше разрыва) (Фиг.39C), что соответствует энергетическим требованиям в плане гибридизации никированной цепи ДНК с PBS PEgRNA. Длина PBS или уровень содержания GC не являлись строго определяющими эффективность праймерного редактирования, при этом другие факторы, такие как вторичная структура в праймере ДНК или удлинение PEgRNA, также могут влиять на активность редактирования. Рекомендуется начинать с длины PBS ~13 nt для стандартной последовательности-мишени и использовать различную длину PBS, если последовательность отклоняется от уровня содержания GC ~40-60%. При необходимости оптимальные последовательности PBS должны быть определены эмпирически.

[1296] Далее исследовали детерминанты эффективности части PEgRNA в виде RT матрицы. PEgRNA с RT матрицами длиной 10-20 nt систематически анализировали в пяти геномных участках-мишенях при использовании PE2 (Фиг.39D) и более длинных RT матриц длиной 31 nt в трех геномных участках (Фиг.48A-48C). Как и в случае с длиной PBS, длина RT матрицы также могла варьироваться для максимизации эффективности праймерного редактирования, хотя в целом многие RT матрицы с длиной ≥10 nt поддерживают более эффективное праймерное редактирование (Фиг.39D). Поскольку для некоторых участков-мишеней были предпочтительны более длинные RT матрицы (>15 nt) для достижения более высокой эффективности редактирования (FANCF, EMX1), в то время как для других локусов

предпочтительными являлись короткие RT матрицы (HEK3, HEK4) (Фиг.39D), рекомендуется тестировать как короткие, так и длинные RT матрицы при оптимизации PEgRNA, начиная с ~10-16 nt.

[1297] Является важным то, что RT матрицы, которые обеспечивали размещение C в качестве нуклеотида, являющегося смежным относительно концевой шпильки каркаса sgРНК, в целом приводили к более низкой эффективности редактирования по сравнению с другими PEgRNA при использовании RT матриц такой же длины (Фиг.48A-48C). На основе структуры sgРНК, связанных с Cas9[148,149], было предположено, что присутствие C в качестве первого нуклеотида 3′ удлинения стандартной sgРНК может нарушать складку каркаса sgРНК за счет связывания с G81, представляющим собой нуклеотид, который нативным образом образует π-стэк с Tyr 1356 в Cas9 и нестандартную пару оснований с A68 sgРНК. Поскольку различные показатели длины RT матрицы обеспечивают возможность праймерного редактирования, рекомендуется выбирать PEgRNA, в которых первое основание 3′ удлинения (последнее подвергаемое обратной транскрипции основание RT матрицы) не является C.

[1298] Конструирование систем праймерного редактора 3 (PE3 и PE3b)

[1299] В то время как PE2 может обеспечивать перенос генетической информации из PEgRNA в локус-мишень более эффективно, чем PE1, способ, с помощью которого клетка корректирует образованную гетеродуплексную ДНК, созданную на основе одной подвергаемой редактированию цепи и одной не подвергаемой редактированию цепи, определяет устойчивость изменения. Предыдущая разработка редактирования оснований столкнулась с аналогичной проблемой, поскольку исходный продукт дезаминирования цитозина или аденина представляет собой гетеродуплексную ДНК, содержащую одну подвергаемую редактированию и одну не подвергаемую редактированию цепь. Для повышения эффективности редактирования оснований использовали никазу Cas9 D10A для внесения разрыва в не подвергаемой редактированию цепи и прямой репарации ДНК в данной цепи с использованием подвергнутой редактированию цепи в качестве матрицы[129,130,142]. Для использования данного принципа для повышения эффективности праймерного редактирования тестировали аналогичную стратегию в виде никирования не подвергаемой редактированию цепи с использованием никазы Cas9 H840A, которая уже

присутствует в PE2, и простой sgРНК для индуцирования предпочтительного замещения не подвергаемой редактированию цепи клеткой (Фиг.40А). Поскольку подвергаемая редактированию цепь ДНК также подвергалась никированию для инициирования праймерного редактирования, тестировали несколько sgРНК-программируемых месторасположений разрыва на не подвергаемой редактированию цепи для минимизации образования двуцепочечных разрывов ДНК, которые приводят к инделям.

[1300] Данную стратегию PE3 сначала тестировали в пяти геномных участках клеток HEK293T путем скрининга sgРНК, которые индуцируют разрывы, расположенные на расстоянии от 14 до 116 оснований от участка индуцированного PEgRNA разрыва, либо 5′ или 3′ относительно PAM. В четырех из пяти протестированных участках никирование не подвергаемой редактированию цепи повысило количество продуктов праймерного редактирования без инделей по сравнению с PE2 системой от 1,5 до 4,2-кратно, до 55% (Фиг.40В). В то время как оптимальное положение никирования варьировалось в зависимости от геномного участка, разрывы, расположенные 3′ относительно PAM (положительные расстояния на Фиг.40В) на расстоянии приблизительно 40-90 bp от индуцированного PEgRNA разрыва, в целом продуцировали благоприятное повышение эффективности праймерного редактирования (в среднем 41%) без избытка образования инделей (в среднем 6,8% инделей для sgРНК, что приводило к наивысшей эффективности редактирования для каждого из пяти протестированных участков) (Фиг.40В). Как и ожидалось, в некоторых участках размещение разрыва не подвергаемой редактированию цепи на расстоянии до 40 bp от индуцированного PEgRNA разрыва приводило к значительному повышению образования инделей до 22% (Фиг.40В), предположительно вследствие образования двуцепочечного разрыва из-за никирования обеих цепей рядом друг с другом. Однако в других участках никирование на расстоянии лишь 14 bp от индуцированного PEgRNA разрыва продуцировало только 5% инделей (Фиг.40В), что свидетельствует о том, что локус-зависимые факторы контролируют конверсию проксимальных двойных разрывов в двуцепочечные разрывы ДНК. В одном протестированном участке (HEK4) разрывы комплементарной цепи не обеспечивали полезного результата или приводили к уровням инделей, которые превышали эффективность редактирования (до 26%), даже при размещении на расстояниях >70 bp от индуцированного PEgRNA разрыва, что согласуется с необычной

предрасположенностью подвергаемой редактированию цепи в данном участке к никированию клеткой или лигированию неэффективным образом. Рекомендуется начинать с разрывов не подвергаемой редактированию цепи на расстоянии приблизительно 50 bp от опосредованного PEgRNA разрыва и тестировать альтернативные места разрывов, если частота инделей превышает приемлемые уровни.

[1301] Данная модель, основанная на том, что никирование комплементарной цепи повышало эффективность праймерного редактирования (Фиг.40A), обеспечила предсказание того, что никирование не подвергаемой редактированию цепи только *после* корректирования флэпа подвергаемой редактированию цепи может минимизировать присутствие одновременных разрывов, снижая частоту двуцепочечных разрывов, которые приводят к образованию инделей. Для достижения временно́го контроля над никированием не подвергаемой редактированию цепи были сконструированы sgРНК со спейсерными последовательностями, которые совпадают с подвергаемой редактированию цепью, но не исходным аллелем. За счет использования данной стратегии, обозначаемой далее как PE3b, несовпадения между спейсером и не подвергаемым редактированию аллелем должны способствовать отсутствию никирования за счет sgРНК до реализации редактирования на включающей PAM цепи. Данный подход PE3b был протестирован с использованием пяти различных изменений в трех геномных участках в клетках HEK293T, а результаты были сравнены с результатами, полученными при использовании систем PE2 и PE3. Во всех случаях PE3b был ассоциирован с существенно более низкими уровнями инделей по сравнению с PE3 (в 3,5-12 раз более низкими инделями, или 0,85%) без какого-либо явного снижения общей эффективности редактирования по сравнению с PE3 (Фиг.40C). Таким образом, при расположении изменения в рамках второго протоспейсера, система PE3b может обеспечивать снижение инделей, в то же время повышая эффективность редактирования по сравнению с PE2, зачастую до уровней, аналогичных соответствующим PE3 (Фиг.40C).

[1302] Данные результаты совместно позволили определить, что системы PE3 (никаза Cas9–оптимизированная обратная транскриптаза + PEgRNA + sgРНК) обеспечивали ~3-кратное повышение эффективности редактирования по сравнению с PE2 (Фиг.40B-40C). PE3 сопровождалось более широкими диапазонами инделей, чем PE2, как и ожидалось с учетом дополнительной связанной с никированием активности PE3. Применение PE3 рекомендуется при приоритизации эффективности праймерного

редактирования. Если критической является минимизация инделей, PE2 обеспечивает ~10-кратно более низкую частоту инделей. При возможность использовать sgРНК, которая распознает внесенное изменение, для никирования не подвергаемой редактированию цепи, система PE3b может обеспечивать соответствующие PE3 уровни редактирования, в то же время значительно снижая образование инделей.

[1303] Для демонстрации объема нацеливания и универсальности праймерного редактирования при использовании PE3 было проанализировано внесение всех возможных замещений одного нуклеотида среди положений от +1 до +8 (считая первое основание 3′ индуцированного PEgRNA разрыва как положение +1) участка-мишени HEK3 с использованием PE3 и PEgRNA с 10-нуклеотидными RT матрицами (Фиг.41A). Данные 24 отдельных изменения совместно покрывают все четыре переходные мутации и все восемь трансверсивных мутаций и осуществляются с эффективностью редактирования (без инделей) в среднем 33±7,9% (в диапазоне от 14% до 48%) со средним показателем инделей 7,5±1,8%.

[1304] Является важным то, что длинные RT матрицы также могут обеспечивать эффективное редактирование в рамках PE3. Например, при использовании PE3 с RT матрицей длиной 34 nt были внесены точечные мутации в положениях +12, +14, +17, +20, +23, +24, +26, +30, и +33 (от 12 до 33 оснований от индуцированного PEgRNA разрыва) в локусе HEK3 со средней эффективностью 36±8,7% и 8,6±2,0% инделями (Фиг.41B). Несмотря на отсутствие попыток реализации изменений на расстоянии +10 в других локусах, другие RT матрицы длиной ≥30 nt в трех альтернативных участках также обеспечивают эффективное редактирование (Фиг.48A-C). Эффективность длинных RT матриц обеспечила эффективное праймерное редактирование для дюжин нуклеотидов из исходного участка разрыва. Поскольку PAM NGG на любой цепи ДНК встречается в среднем через каждые ~8 bp, то есть намного меньше, чем максимальные расстояния между изменением и PAM, которые поддерживают эффективное праймерное редактирование, праймерное редактирование не является существенным образом ограниченным в плане доступности близлежащей последовательности PAM, в отличие от других способов точного геномного редактирования[125,142,154]. С учетом предполагаемой связи между вторичной структурой РНК и эффективностью праймерного редактирования, при конструировании PEgRNA для изменений на больших расстояниях является разумным протестировать RT матрицы с различной

длиной и, если необходимо, различной композицией последовательности (например синонимичными кодонами) для оптимизации эффективности редактирования.

[1305] Для дополнительного тестирования объема и ограничений системы PE3 для внесения переходных и трансверсивных точечных мутаций были протестированы 72 дополнительных изменения, покрывающие все 12 возможных типов точечных мутаций среди шести дополнительных геномных участков-мишеней (Фиг.41C-41H). В целом эффективность редактирования без инделей в среднем составляла 25±14%, в то время как образование инделей в среднем составляло 8,3±7,5%. Поскольку RT матрица PEgRNA включала последовательность PAM, праймерное редактирование могло индуцировать изменения последовательности PAM. В этих случаях наблюдалась более высокая эффективность редактирования (в среднем 39±9,7%) и более низкое образование инделей (в среднем 5,0±2,9%) (Фиг.41A-41K, точечные мутации в положениях +5 или +6). Данное повышение эффективности и снижение образования инделей для изменений PAM может быть связано с неспособностью никазы Cas9 повторно связываться и никировать подвергаемую редактированию цепь до репарации комплементарной цепи. Поскольку праймерное редактирование поддерживает комбинированные изменения без заметной потери эффективности редактирования, редактирование PAM, помимо других требуемых изменений, является рекомендуемым, когда это возможно.

[1306] Далее были осуществлены 14 нацеленных малых вставок и 14 нацеленных малых делеций в семи геномных участках с использованием PE3 (Фиг.41I). Нацеленные вставки длиной 1 bp были осуществлены со средней эффективностью 32±9,8%, в то время как вставки длиной 3 bp были внесены со средней эффективностью 39±16%. Нацеленные делеции длиной 1 bp и 3 bp также являлись эффективными и осуществлялись со средним выходом 29±14% и 32±11%, соответственно. Образование инделей (помимо нацеленной вставки или делеции) в среднем составляло 6,8±5,4%. Поскольку вставки и делеции, внесенные между положениями +1 и +6, изменяют положение или структуру PAM, было предположено, что изменения в виде вставки и делеции в данном диапазоне обычно являются более эффективными вследствие неспособности никазы Cas9 повторно связывать и никировать подвергаемую редактированию цепь ДНК до репарации комплементарной цепи, аналогично точечным мутациям, которые обеспечивают изменение PAM.

[1307] PE3 также тестировали на его способность опосредовать более крупные точные делеции длиной от 5 bp до 80 bp в участке HEK3 (Фиг.41J). Наблюдали очень высокую эффективность редактирования (от 52 до 78%) для делеций длиной 5, 10, и 15 bp при использовании PBS длиной 13 nt и RT матрицы, которая содержала 29, 24, или 19 bp гомологии относительно локуса-мишени, соответственно. Использование RT матрицы длиной 26 nt обеспечивало поддержку более крупных делеций длиной 25 bp с 72±4,2% эффективностью, в то время как RT матрица длиной 20 nt обеспечивала делецию длиной 80 bp с эффективностью 52±3,8%. Данные нацеленные делеции сопровождались в среднем 11±4,8% частотой образования инделей (Фиг.41J).

[1308] Наконец, была протестирована способность PE3 опосредовать 12 комбинаций множественных изменений в одном локусе-мишени, состоящих из вставок и делеций, вставок и точечных мутаций, делеций и точечных мутаций, или двух точечных мутаций в рамках трех геномных участков. Данные комбинированные изменения были очень эффективными, соответствуя в среднем 55% изменения-мишени с 6,4% инделями (Фиг.41K) и демонстрируя способность праймерного редактирования обеспечивать реализацию комбинаций точных вставок, делеций, и точечных мутаций в отдельных участках-мишенях с высокой эффективностью и низкой частотой инделей.

[1309] Примеры в рамках Фиг.41A-41K совместно отражают 156 отдельных переходных, трансверсивных, связанных со вставкой, связанных с делецией, и комбинированных изменений среди семи геномных локусов человека. Данные результаты подтверждают универсальность, точность, и гибкость нацеливания праймерного редактирования.

[1310] Сравнение праймерного редактирования и редактирования оснований

[1311] Редакторы оснований на основе цитидина (CBE) и редакторы оснований на основе аденина (ABE) текущего поколения могут обеспечивать внесение переходных мутаций C•G на T•A и переходных мутаций A•T на G•C с высокой эффективностью и низкими показателями инделей[129,130,142]. Использование редактирования оснований может быть ограничено присутствием нескольких оснований в виде цитидина или аденина в окне активности редактирования оснований (обычно шириной ~5 bp), что приводит к нежелательных побочным изменениям[129,130,142,155], или отсутствием PAM, расположенного на расстоянии приблизительно 15±2 от нуклеотида-мишени[142,156]. Предполагалось, что праймерное редактирование может являться особенно полезным

для точного внесения переходных мутаций без побочных изменений или когда недостаток подходящим образом расположенных PAM делает невозможным требуемое расположение нуклеотида-мишени в рамках окна активности CBE или ABE.

[1312] Праймерное редактирование и редактирование оснований на основе цитозина сравнивали на основе редактирования трех геномных локусов, которые содержат несколько цитидинов-мишеней в стандартном окне редактирования оснований (положения протоспейсера 4-8, если считать PAM в качестве положений 21-23), с использованием оптимизированных CBE[157] без никазной активности (BE2max) или с никазной активностью (BE4max), или с использованием аналогичных систем праймерного редактирования PE2 и PE3. Среди всех девяти цитозинов-мишеней в окнах редактирования оснований трех участков BE4max приводил к в среднем в 2,2 раза более высокой общей конверсии C•G на T•A, чем PE3, для оснований в центре окна редактирования оснований (положения протоспейсера 5-7, Фиг.42A). Аналогичным образом, не осуществляющий никирование BE2max превосходил PE2 в среднем в 1,4 раза в рамках данных удобно расположенных оснований (Фиг.42A). Однако PE3 превосходил BE4max в 2,7 раза, а PE2 превосходил BE2max в 2,0 раза для цитозинов за пределами центра окна редактирования оснований (среднее редактирование 40±17% для PE3 против 15±18% для BE4max, и 22±11% для PE2 против 11±13% для BE2max). В целом, показатели частоты инделей для PE2 были очень низкими (в среднем 0,86±0,47%), а для PE3 были аналогичными или умеренно более высокими, чем для BE4max (диапазон BE4max: от 2,5% до 14%; диапазон PE3: от 2,5% до 21%) (Фиг.42B).

[1313] При сравнении эффективности редактирования оснований с праймерным редактированием в плане внесения *точных* изменений в виде C•G на T•A (без какого-либо побочного редактирования) эффективность праймерного редактирования значительно превосходила эффективность редактирования оснований в рамках указанных выше участков, которые, подобно большинству участков геномной ДНК, содержат множество цитозинов в рамках окна редактирования оснований длиной ~5 bp (Фиг.42C). В данных участках, таких как EMX1, которые содержат цитозины в положениях протоспейсера C5, C6, и C7, BE4max приводил к образованию небольшого количества продуктов, содержащих только одну конверсию-мишень пар оснований без побочных изменений. С другой стороны, праймерное редактирование в данном участке может быть использовано для избирательного внесения изменения C•G на T•A в

любом положении или комбинации положений (C5, C6, C7, C5+C6, C6+C7, C5+C7, или C5+C6+C7) (Фиг.42C). Все точные изменения в виде одного основания или двух оснований (то есть изменения, которые не модифицируют какие-либо другие соседние основания) являлись более эффективными при использовании PE3 или PE2, чем при использовании BE4max или BE2, соответственно, в то время как изменение трех оснований C•G на T•A было более эффективным при использовании BE4max (Фиг.42C), что отражает склонность редакторов оснований обеспечивать редактирование всех оснований-мишеней в рамках окна активности. Данные результаты совместно демонстрируют, что редакторы оснований на основе цитозина могут приводить к более высоким уровням редактирования в оптимально расположенных основаниях-мишенях, чем PE2 или PE3, однако праймерное редактирование может превосходить редактирование оснований в неоптимально расположенных основаниях-мишенях и может обеспечивать изменения с намного более высокой точностью в рамках нескольких изменяемых оснований.

[1314] Изменение A•T на G•C сравнивали в двух геномных локусах с использованием оптимизированного не осуществляющего никирование ABE (ABEmax[152] с dCas9 вместо никазы Cas9, далее обозначаемый как ABEdmax) против PE2, а также с использованием оптимизированного никирующего редактора оснований ABEmax на основе аденина против PE3. В участке, который содержит два аденина-мишени в окне редактирования оснований (HEK3), ABE являлись более эффективными, чем PE2 или PE3 в плане конверсии A5, однако PE3 являлся более эффективным в плане конверсии A8, который расположен на краю окна редактирования ABEmax (Фиг.42D). При сравнении эффективности точных изменений, в рамках которых конверсии подвергается только один аденин, PE3 превосходил ABEmax как в плане A5, так и A8 (Фиг.42E). В целом, ABE продуцировали намного меньше инделей в HEK3, чем праймерные редакторы (0,19±0,02% для ABEdmax против 1,5±0,46% для PE2, и 0,53±0,16% для ABEmax против 11±2,3% для PE3, Фиг.42F). В FANCF, в котором присутствует только один A в рамках окна редактирования оснований, ABE2 и ABEmax превосходили свои связанные с праймерным редактированием аналоги в плане общей конверсии пар-мишеней оснований в 1,8-2,9 раза, при этом почти все продукты, подвергнутые редактированию как в рамках редактирования оснований, так и праймерного редактирования содержали только точное изменение (Фиг.42D-42E).

Как и в случае с участком HEK3, ABE продуцировали намного меньше инделей в участке FANCF (Фиг.42F).

[1315] Данные результаты совместно показывают, что редактирование оснований и праймерное редактирование связаны с дополнительными сильными и слабыми сторонами в плане поучения нацеленных переходных мутаций. Для случаев, в рамках которых один нуклеотид-мишень присутствует в рамках окна редактирования оснований или когда побочные изменения являются приемлемыми, текущие редакторы оснований обычно являются более эффективными и приводят к меньшему количеству инделей, чем праймерные редакторы. При присутствии множества цитозинов или аденинов и нежелательности побочных изменений, или когда основания-мишени являются неудобным образом расположенными для редактирования оснований относительно доступных PAM, праймерные редакторы обеспечивают существенные преимущества.

[1316] Нецелевое праймерное редактирование

[1317] Для реализации продуктивного редактирования праймерное редактирование требует связывания локуса-мишени:спейсера PEgRNA, комплементарного домену Cas9; комплементарности локуса-мишени:PBS PEgRNA для инициирования праймируемой PEgRNA обратной транскрипции, и комплементарности продукта локуса-мишени:обратной транскриптазы для корректирования флэпа. Было предположено, что данные три отдельных требований в плане гибридизации ДНК могут минимизировать нецелевое праймерное редактирование по сравнению с нецелевым редактированием в рамках других способов геномного редактирования. Для тестирования данной возможности клетки HEK293T обрабатывали PE3 или PE2 и конструировали всего 16 PEgRNA для нацеливания на четыре целевых геномных локуса с использованием Cas9 и четырех соответствующих sgРНК, нацеленных на одни и те же протоспейсеры, или с использованием Cas9 и тех же 16 PEgRNA. Данные четыре локуса-мишени были выбраны, поскольку каждый из них имеет, по меньшей мере, четыре хорошо описанных нецелевых участках, в отношении которых Cas9 и соответствующая целевая sgРНК в клетках HEK293T, как известно, вызывают существенные нецелевые модификации ДНК[118,159]. После обработки четыре целевых локуса и четыре главных известных нецелевых участка Cas9 для каждого целевого спейсера были секвенированы в виде всего 16 нецелевых участков (Таблица 1).

[1318] В полном соответствии с предыдущими исследованиями[118], Cas9 и четыре sgРНК-мишени обеспечивали модифицирование всех 16 из ранее описанных нецелевых локусов (Фиг.42G). Эффективность нецелевой модификации Cas9 среди четырех нецелевых участков для локуса-мишени HEK3 в среднем составляла 16%. Cas9 и sgРНК, нацеленная на HEK4, приводили в среднем к 60% модификации четырех протестированных известных нецелевых участков. Аналогичным образом, нецелевые участки для EMX1 и FANCF были модифицированы Cas9:sgРНК со средней частотой 48% и 4,3%, соответственно (Фиг.42G). Было отмечено, что PEgRNA с нуклеазой Cas9 обеспечивала модифицирование целевых участков в среднем с аналогичной (в 1,0-1,5 раза более низкой) эффективностью по сравнению с sgРНК, в то время как PEgRNA с нуклеазой Cas9 обеспечивала модифицирование нецелевых участков в среднем с в ~4 раза более низкой эффективностью, чем sgРНК.

[1319] Удивительным является то, что PE3 или PE2 с теми же 16 протестированными PEgRNA, содержащими данные четыре спейсера-мишени, приводили к намного более низкому нецелевому редактированию (Фиг.42H). Из 16 участков, которые, как известно, подвергаются нецелевому редактированию Cas9+sgРНК, PE3+PEgRNA или PE2+PEgRNA приводили к детектируемому нецелевому праймерному редактированию только в 3 из 16 нецелевых участках, при этом только 1 из 16 демонстрировал эффективность нецелевого редактирования ≥1% (Фиг.42H). Среднее нецелевое праймерное редактирование для PEgRNA, нацеленных на HEK3, HEK4, EMX1, и FANCF, в данных 16 известных нецелевых участках Cas9 составляло <0,1%, <2,2±5,2%, <0,1%, и <0,13±0,11%, соответственно (Фиг.42H). Следует отметить, что в нецелевом участке 3 HEK4, редактирование которого Cas9+PEgRNA1 обеспечивает с 97% эффективностью, PE2+PEgRNA1 приводит лишь к 0,7% нецелевому редактированию, несмотря на ту же спейсерную последовательность, что демонстрирует то, как два дополнительных события в виде гибридизации ДНК, требуемые для праймерного редактирования, в отличие от редактирования Cas9, могут значительно снижать нецелевое редактирование. Данные результаты совместно свидетельствуют о том, что PE3 и PEgRNA индуцируют намного более низкое нецелевое редактирование ДНК в клетках человека, чем Cas9 и sgРНК, которые нацелены на те же протоспейсеры.

[1320] Обратная транскрипция 3′-удлиненных PEgRNA в принципе может продолжаться в каркас направляющей РНК. Если образованный 3′ флэп, несмотря на

недостаток комплементарности на его 3′-конце относительно не подвергаемой редактированию цепи ДНК, является встроенным в локус-мишень, то результатом является вставка нуклеотидов каркаса PEgRNA, которая вносит вклад в частоту инделей. Авторы проанализировали данные, полученные с помощью секвенирования на основе 66 экспериментов в виде редактирования с использованием PE3 в четырех локусах в клетках HEK293T и отметили низкую частоту вставки каркаса PEgRNA, в среднем составляющую 1,7±1,5% любого количества нуклеотидов каркаса PEgRNA (Фиг.56A-56D). Предполагается, что недоступность каркаса направляющей РНК для обратной транскриптазы вследствие связывания домена Cas9, а также клеточная эксцизия при корректировании флэпа несовпадающего 3′-конца 3′ флэпа, которая возникает вследствие обратной транскрипции каркаса PEgRNA, обеспечивает минимизацию продуктов, которые включают нуклеотиды каркаса PEgRNA. В то время как такие события являются редкими, дополнительные усилия по конструированию PEgRNA или белков праймерного редактора, которые минимизируют встраивание каркаса PEgRNA, могут обеспечивать дополнительное снижение частоты инделей.

[1321] Дезаминазы в некоторых редакторах оснований могут действовать независимо от Cas9, что приводит к имеющему низкий уровень, но повсеместному нецелевому редактированию ДНК среди CBE первого поколения (то не ABE)[160-162] и нецелевому редактированию РНК среди CBE и ABE первого поколения[163-165], хотя более новые варианты CBE и ABE с модифицированными дезаминазами значительно снижают Cas9-независимое нецелевое редактирование ДНК и РНК[163-165]. Праймерные редакторы не имеют ферментов для модификации оснований, таких как дезаминазы, и, таким образом, не имеют врожденной способности модифицировать основания ДНК или РНК Cas9-независимым образом.

[1322] В то время как домен обратной транскриптазы в праймерных редакторах в принципе может процессировать правильно праймированные РНК или ДНК матрицы в клетках, было отмечено, что все из ретротранспозонов, таких как в семействе LINE-1[166], эндогенных ретровирусов[167,168], и теломеразы человека обеспечивали наличие активных эндогенных обратных транскриптаз человека. Их естественное присутствие в клетках человека свидетельствует о том, что активность обратной транскриптазы сама по себе не является по существу токсичной. Действительно, PE3-зависимые различия в плане жизнеспособности клеток HEK293T по сравнению с контролями, экспрессирующими dCas9, никазу Cas9 H840A, или PE2 с мутациями R110S+K103L

(PE2-dRT), которые деактивируют обратную транскриптазу[169], отсутствовали (Фиг.49A-49B).

[1323] Несмотря на приведенные выше данные и результаты анализов, необходимы дополнительные исследования для анализа нецелевого праймерного редактирования объективным образом в рамках всего генома, а также для описания степени влияния на клетки вариантов обратной транскриптазы в праймерных редакторах или промежуточных продуктах для праймерного редактирования.

[1324] Патогенные мутации в виде трансверсии, вставки, и делеции в клетках человека на основе праймерного редактирования

[1325] Была протестирована способность PE3 напрямую вносить или корректировать мутации в виде трансверсии, малой вставки, и малой делеции в клетках человека, которые вызывают генетические заболевания. Серповидноклеточная анемия наиболее часто вызывается трансверсивной мутацией A•T на T•A в HBB, что приводит к мутации Glu6→Val в бета-глобине. Обработка гемопоэтических стволовых клеток *ex vivo* нуклеазой Cas9 и донорской ДНК матрицей для HDR с последующим повышением содержания отредактированных клеток, трансплантацией, и приживлением представляет собой многообещающую потенциальную стратегию лечения серповидноклеточной анемии[170]. Однако данный подход все еще приводит к образованию многих содержащих индели побочных продуктов помимо правильным образом отредактированного аллеля HBB[170-171]. В то время как редакторы оснований в целом обеспечивают продуцирование намного меньшего количества инделей, они в настоящее время не могут обеспечивать реализацию трансверсивной мутации T•A на A•T, которая необходима для прямого восстановления нормальной последовательности в HBB.

[1326] PE3 использовали для внесения мутации HBB E6V в клетки HEK293T с 44% эффективностью и 4,8% инделей (Фиг.43A). Из смеси обработанных PE3 клеток выделяли шесть линий клеток HEK293T, которые являлись гомозиготными (триплоидными) по аллелю HBB E6V (Фиг.53A-53D), что демонстрирует способность праймерного редактирования обеспечивать получение линий клеток человека с патогенными мутациями. Для корректирования аллеля HBB E6V к HBB дикого типа проводили обработку гомозиготных HBB E6V клеток HEK293T с использованием PE3 и PEgRNA, запрограммированной для прямого возврата мутации HBB E6V к HBB

дикого типа. Всего было протестировано 14 конструкций PEgRNA. Спустя три дня секвенирование ДНК показало, что все 14 PEgRNA при комбинировании с PE3 обеспечивали эффективное корректирование HBB E6V с возвратом к HBB дикого типа (≥26% HBB дикого типа без инделей) и средние уровни инделей 2,8±0,70% (Фиг.50A). Лучшая PEgRNA приводила к 52% коррекции HBB E6V в дикий тип с 2,4% инделей (Фиг.43A). Внесение молчащей мутации, которая модифицирует PAM, распознаваемый PEgRNA, умеренно повышало эффективность редактирования и чистоту продукта до 58% коррекции с 1,4% инделей (Фиг.43A). Данные результаты подтверждают, что праймерное редактирование может обеспечивать внесение и корректирование патогенной трансверсивной точечной мутации в линии клеток человека с высокой эффективностью и минимальными побочными продуктами.

[1327] Болезнь Тея-Сакса наиболее часто вызывается 4 bp вставкой в ген HEXA (HEXA 1278+TATC)[136]. PE3 использовали для внесения данной 4 bp вставки в клетки HEK293T с 31% эффективностью и 0,8% инделей (Фиг.43B) и выделяли две линии клеток HEK293T, которые являлись гомозиготными в плане аллеля HEXA 1278+TATC (Фиг.53A-53D). Данные клетки использовали для тестирования 43 PEgRNA и трех никирующих sgРНК с системами PE3 или PE3b на предмет корректирования данной патогенной вставки в HEXA (Фиг.50B) с помощью полного возврата к аллелю дикого типа или 4 bp делеции со сдвигом, которая нарушает PAM и вносит молчащую мутацию. Девятнадцать из 43 протестированных PEgRNA приводили к ≥20% редактированию. Полный возврат к HEXA дикого типа с помощью PE3 или PE3 и лучшая PEgRNA обеспечивали аналогичную среднюю эффективность (30% для PE3 против 33% для PE3b), однако система PE3b сопровождалась в 5,3 раза более низким количеством продуктов в виде инделей (1,7% для PE3 против 0,32% для PE3b) (Фиг.43B и Фиг.50B). Эти результаты демонстрируют способность праймерного редактирования обеспечивать реализацию точных малых вставок и делеций, которые вносят или корректируют патогенный аллель в клетках млекопитающих эффективным образом и с минимальным количеством побочных продуктов.

[1328] Наконец, было протестировано внесение защитного SNP в PRNP, который представляет собой ген, кодирующий прионный белок человека (PrP). Неправильный фолдинг PrP вызывает прогрессирующее и смертельное нейродегенеративное прионное заболевание, которое может возникать спонтанным образом за счет унаследованных доминантных мутаций в гене PRNP или за счет подвергания

воздействию подвергнутого неправильному фолдингу PrP[172]. Встречающийся в природе мутантный аллель G127V PRNP придает устойчивость к прионному заболеванию у людей[138] и мышей[173]. PR3 использовали для внесения мутации G127V в аллель PRNP человека в клетках HEK293T, которая требует трансверсии G•C на T•A. С системой PE3 анализировали четыре PEgRNA и три никирующие sgРНК. Спустя три дня подвергания воздействую наиболее эффективного PE3 и PEgRNA секвенирование ДНК выявило 53±11% эффективность внесения мутации G127V и уровни инделей 1,7±0,7% (Фиг.43C). Данные результаты совместно подтверждают способность праймерного редактирования обеспечивать внесение или корректирование мутаций в виде трансверсии, вставки, или делеции в клетках человека, которые вызывают или эффективным образом придают устойчивость к заболеванию с минимальным уровнем побочных продуктов.

[1329] Праймерное редактирование в различных линиях клеток человека и первичных нейронах мыши

[1330] Далее праймерное редактирование тестировали на его способность обеспечивать редактирование эндогенных участков в трех дополнительных линиях клеток человека. PE использовали в клетках K562 (относящихся к костному мозгу при лейкозе) для реализации трансверсивных изменений в участках HEK3, EMX1, и FANCF, а также 18 bp вставки метки 6xHis в HEK3. Наблюдали среднюю 15-30% эффективность редактирования для каждого из данных четырех опосредованных PE3 изменений со средним уровнем инделей 0,85-2,2% (Фиг.43A). В клетки U2OS (остесаркома) вносили трансверсивные мутации в HEK3 и FANCF, а также 3 bp вставку и вставку в виде 6xHis метки в HEK3 с 7,9-22% эффективностью редактирования, которая превосходила уровень образования инделей в 10-76 раз (Фиг.43A). Наконец в клетки HeLa (рак шейки матки) осуществляли 3 bp вставку в HEK3 с 12% средней эффективностью и 1,3% уровнем инделей (Фиг.43A). Эти данные совместно указывают на то, что многие другие линии клеток, помимо клеток HEK293T, поддерживают праймерное редактирование, однако эффективность редактирования варьируется в зависимости от типа клеток и в целом является более низкой, чем в клетках HEK293T. Соотношения редактирование:индели оставалось высоким во всех протестированных линиях клеток человека. Наконец в клетки HeLa (рак шейки матки) осуществляли 3 bp вставку в HEK3 с 12% средней эффективностью

и 1,3% уровнем инделей (Фиг.43A). Эти данные совместно указывают на то, что многие другие линии клеток, помимо клеток HEK293T, поддерживают праймерное редактирование, однако эффективность редактирования варьируется в зависимости от типа клеток и в целом является более низкой, чем в клетках HEK293T. Соотношения редактирование:индели оставалось высоким во всех протестированных линиях клеток человека.

[1331] Для определения возможности праймерного редактирования в постмитотических необратимо дифференцированных первичных клетках первичные нейроны коры головного мозга, собранные у мышей E18.5, трансдуцировали с использованием двойной лентивирусной системы доставки разделенного PE3, в рамках которой сплайсинг разделенного интеина[203] обеспечивает восстановление белка PE2 из N-концевой и C-концевой половин, каждая из которых доставляется с помощью отдельного вируса. Для ограничения редактирования постмитотическими нейронами промотор синапсина человека, который является высокоспецифичным в отношении зрелых нейронов[204], использовали для обеспечения экспрессии обоих компонентов белка PE2. GFP был слит с помощью саморасщепляемого пептида P2A[205] с N-концевой половиной PE2. Ядра из нейронов были выделены спустя две недели после двойной вирусной трансдукции и напрямую подвергались секвенированию или подвергались сортировке на основе экспрессии GFP перед секвенированием. Наблюдали 7,1±1,2% уровень среднего праймерного редактирования для внесения трансверсии в локусе DNMT1 с 0,58±0,14% средним уровнем инделей в отсортированных ядрах (Фиг.43D). Нуклеаза Cas9 в рамках той же двойной лентивирусной системы на основе разделенного интеина приводила к 31±5,5% инделей среди отсортированных ядер нейронов коры головного мозга (Фиг.43D). Эти данные свидетельствуют о том, что постмитотические необратимо дифференцированные первичные клетки могут поддерживать праймерное редактирование и, таким образом, подтверждают, что праймерное редактирование не требует клеточной репликации.

[1332] Сравнение праймерного редактирования и Cas9-инициируемой HDR

[1333] Эффективность PE3 сравнивали с эффективностью оптимизированной Cas9-инициируемой HDR128,125 в линиях митотических клеток, которые поддерживают HDR[128]. Клетки HEK293T, HeLa, K562, и U2OS обрабатывали нуклеазой Cas9, sgРНК, и донорской олигонуклеотидной оцДНК матрицей, сконструированной для внесения

ряда изменений в виде трансверсий и вставок (Фиг.43E-43G, и Фиг.51A-51F). Инициированная Cas9 HDR во всех случаях успешно обеспечивала внесение требуемого изменения, но с намного более высокими уровнями побочных продуктов (в основном инделей), как и ожидается для обработки, которая вызывает образование двуцепочечных разрывов. С помощью PE3 в клетках HEK293T внесение и корректирование *HBB* E6V имело 42% и 58% среднюю эффективность редактирования с 2,6% и 1,4% средним уровнем инделей, соответственно (Фиг.43E и Фиг.43G). С другой стороны, осуществление тех же изменений с использованием нуклеазы Cas9 и матрицы HDR приводило к 5,2% и 6,7% средней эффективности редактирования с 79% и 51% средней частотой инделей (Фиг.43E и Фиг.43G). Аналогичным образом, PE3 обеспечивал внесение *PRNP* G127V с 53% эффективностью и 1,7% уровнем инделей, в то время как инициированная Cas9 HDR обеспечивала внесение данной мутации с 6,9% эффективностью и 53% уровнем инделей (Фиг.43E и Фиг.43G). Таким образом, соотношение редактирование:индели для внесения *HBB* E6V, корректирования *HBB* E6V, и внесения *PRNP* G127V в среднем было в 270 раз более высоким для PE3, чем для инициируемой Cas9 HDR.

[1334] Сравнения между PE3 и HDR в линиях клеток человека, отличных от HEK293T, показали аналогичные результаты, хотя и с более низкой эффективностью редактирования PE3. Например, в клетках K562 опосредованная PE3 3 bp вставка в HEK3 осуществлялась с 25% эффективностью и 2,8% уровнем инделей по сравнению с 17% эффективностью и 72% уровнем инделей для инициируемой Cas9 HDR, что соответствует 40-кратному превосходству соотношения редактирование:индели в пользу PE3 (Фиг.43F-43G). В клетках U2OS PE3 обеспечивал реализацию данной 3 bp вставки с 22% эффективностью и 2,2% уровнем инделей, в то время как инициируемая Cas9 HDR приводила к 15% редактированию с 74% уровнем инделей, что соответствует в 49 раз более низкому соотношению редактирование:индели (Фиг.43F-43G). В клетках HeLa PE3 обеспечивал осуществления данной вставки с 12% эффективностью и 1,3% уровнем инделей против 3,0% эффективности и 69% уровня инделей для инициируемой Cas9 HDR, что соответствует 210-кратной разнице в соотношении редактирование:индели (Фиг.43F-43G). Эти данные совместно показали, что HDR обычно приводит к аналогичной или более низкой эффективности редактирования и намного более высокому уровню инделей, чем PE3, в четырех протестированных линиях клеток (Фиг.51A-51F).

[1335] Обсуждение и дальнейшее направление

[1336] Способность обеспечивать вставку последовательностей ДНК с точностью на уровне отдельных нуклеотидов является особенно полезной возможностью праймерного редактирования. Например, PE3 использовали для точной вставки His_6 метки в локус HEK3 в клетках HEK293T (18 bp, 65% средняя эффективность), эпитопной метки FLAG (24 bp, 18% средняя эффективность), и удлиненного участка *LoxP* (44 bp, 23% средняя эффективность), который представляет собой нативный субстрат для рекомбиназы Cre. Средний уровень инделей был в диапазоне от 3,0% до 5,9% для данных примеров (Фиг.43H). Предполагаются многие биотехнологические, связанные с синтетической биологией, и терапевтические виды использования на основе способности эффективно и точно вносить новые последовательности ДНК в представляющие интерес участки-мишени в живых клетках.

[1337] Описанные здесь эксперименты на основе праймерного редактирования совместно обеспечили реализацию 18 вставок длиной до 44 bp, 22 делеций длиной до 80 bp, 113 точечных мутаций, включая 77 трансверсий, и 18 комбинированных изменений среди 12 эндогенных локусов в геноме человека и мыши в положениях от 3 bp выше до 29 bp ниже начала PAM без фактического осуществления двуцепочечных разрывов ДНК. Данные результаты подтверждают, что праймерное редактирование является высокоуниверсальным способом геномного редактирования. Поскольку подавляющее большинство (85-99%) вставок, делеций, инделей, и дупликаций в ClinVar имеют длину ≤30 bp (Фиг.52A-52D), праймерное редактирование в принципе может обеспечивать корректирование до ~89% из 75122 известных в настоящее время патогенных генетических вариантов человека в ClinVar (транзиций, трансверсий, вставок, делеций, инделей, и дупликаций на Фиг.38A) с дополнительным потенциалом облегчать заболевания, вызванные увеличением или уменьшением количества копий.

[1338] Является важным то, что для любого требуемого изменения гибкость праймерного редактирования обеспечивает множество возможных выборов расположений PEgRNA-индуцируемого разрыва, расположений sgРНК-индуцируемого второго разрыва, длины PBS, длины RT матрицы, и цепи, которая первой будет подвергнута редактированию, как подробно описано здесь. Данная гибкость, которая отличается от более ограниченных возможностей, которые обычно доступны в рамках других способов точного геномного редактирования[125,142,154], обеспечивает

возможность оптимизации эффективности редактирования, чистоты продуктов, специфичности ДНК, или других параметров для соответствия потребностям в рамках конкретного варианта использования, как показано на Фиг.50A-50B, где тестирование 14 и 43 PEgRNA, соответствующих диапазону стратегий праймерного редактирования, обеспечивало оптимизацию корректирования патогенных аллелей HBB и HEXA, соответственно.

[1339] Требуется множество дополнительных исследований для дальнейшего понимания и улучшения праймерного редактирования. Могут потребоваться дополнительные модификации систем праймерных редакторов для расширения их совместимости для включения других типов клеток, таких как постмитотические клетки. Объединение праймерного редактирования с вирусными и невирусными *in vitro* и *in vivo* стратегиями доставки необходимо для полного анализа потенциала праймерного редактирования для обеспечения широкого диапазона вариантов использования, включая исследование и лечение генетических заболеваний. За счет обеспечения высокоточных нацеленных транзиций, трансверсий, малых вставок, и малых делеций в геномах клеток млекопитающих без необходимости двуцепочечных разрывов или HDR праймерное редактирование обеспечивает новую способность типа "поиск-и-замена", которая существенно расширяет объем геномного редактирования.

Способы

[1340] Общие способы

[1341] Амплификацию ДНК осуществляли с помощью ПЦР с использованием мастер-микса для ПЦР Phusion U Green Multiplex (ThermoFisher Scientific) или мастер-микса Q5 Hot Start High-Fidelity 2x (New England BioLabs), если не указано иное. Олигонуклеотиды ДНК, включая Cy5-меченные нуклеотиды ДНК, белок dCas9, и белок Cas9 H840A получали от Integrated DNA Technologies. Репортерные плазмиды дрожжей получали из ранее описанных плазмид 64 и клонировали с помощью способа на основе метода Гибсона. Все использованные здесь плазмиды млекопитающих для редактирования были собраны с помощью способа клонирования USER, как описано ранее[175]. Плазмиды, экспрессирующие sgРНК, конструировали путем лигирования подвергнутых отжигу нуклеотидов в BsmBI-расщепляемый акцепторный вектор. Плазмиды, экспрессирующие PEgRNA, конструировали на основе сборки согласно

методу Гибсона или сборки Golden Gate с использованием модифицированной акцепторной плазмиды (см. дополнительное описание "сборки Golden Gate"). Последовательности используемых здесь конструктов sgРНК и PEgRNA перечислены в Таблицах 2A-2C и Таблицах 3A-3R. Все векторы для экспериментов с использованием клеток млекопитающих были очищены с помощью наборов Plasmid Plus Midiprep (Qiagen) или наборов плазмидных минипрепаратов PureYield (Promega), которые включают этапы удаления эндотоксинов. Все эксперименты с использованием живых животных были одобрены Институционными комитетами по уходу и использованию животных Института Броуда. Мыши C57BL/6 дикого типа были получены от Charles River (#027).

[1342] In vitro биохимические анализы

[1343] PEgRNA и sgРНК транскрибировали *in vitro* с использованием набора для *in vitro* транскрипции HiScribe (New England Biolabs) из амплифицированных с помощью ПЦР матриц, содержащих последовательность промотора T7. РНК очищали с помощью денатурирующего электрофореза PAGE с применением мочевины и проверяли качество с помощью аналитического геля перед использованием. 5′-Cy5-меченные дуплексные ДНК-субстраты отжигались с использованием двух олигонуклеотидов (Cy5-AVA024 и AVA025; соотношение 1:1,1) для не подвергнутого никированию субстрата и трех нуклеотидов (Cy5-AVA023, AVA025 и AVA026; 1:1,1:1,1) для предварительно подвергнутого никированию субстрата путем нагревания до 95 °C в течение 3 минут с последующим медленным охлаждением до комнатной температуры (Таблицы 2A-2C). Реакции расщепления Cas9 и обратной транскрипции проводили в 1x буфере для расщепления[205] с добавлением dNTP (20 мМ HEPES-K, pH 7,5; 100 мМ KCl; 5% глицерин; 0,2 мМ ЭДТА, pH 8,0; 3 мМ MgCl2; 0,5 мМ dNTP микс; 5 мМ ДТТ). dCas9 или Cas9 H840A (конечная концентрация 5 мкМ) и sgРНК или PEgRNA (конечная концентрация 5 мкМ) предварительно инкубировали при комнатной температуре в 5 мкл реакционной смеси в течение 10 минут перед добавлением дуплексного ДНК-субстрата (конечная концентрация 400 нМ) и далее с добавлением обратной транскриптазы Superscript III (ThermoFisher Scientific), неописанного варианта RT M-MLV, где это необходимо. Реакции проводили при 37 °C в течение 1 часа, затем разбавляли до объема 10 мкл водой, обрабатывали 0,2 мкл раствора протеиназы K (20 мг/мл, ThermoFisher Scientific), и инкубировали при

комнатной температуре в течение 30 минут. После тепловой деактивации при 95 °C в течение 10 минут продукты реакции соединяли с 2x буфером для загрузки формамидного геля (90% формамид; 10% глицерин; 0,01% бромфеноловый синий), денатурировали при 95 °C в течение 5 минут, и отделяли с помощью геля для денатурирующего PAGE с применением мочевины (15% TBE-мочевина, 55 °C, 200В). ДНК-продукты визуализировали на основе сигнала флуоресценции Cy5 с использованием биомолекулярного сканера Typhoon FLA 7000.

[1344] Проводили анализы сдвига электрофоретической подвижности в 1x связывающем буфере (1x буфер расщепления + 10 мкг/мл гепарин) с использованием предварительно инкубированных комплексов dCas9:sgРНК или dCas9:PEgRNA (диапазон конечной концентрации от 5 нМ до 1 мкМ) и Cy5-меченной дуплексной ДНК (Cy5- AVA024 и AVA025; конечная концентрация 20 нМ) Спустя 15 минут инкубирования при 37 °C образцы анализировали с помощью нативного геля PAGE (10% TBE) для визуализировали на основе флуоресценции Cy5.

[1345] Для секвенирования ДНК продуктов обратной транскрипции флуоресцентные полосы выделяли и очищали из гелей PAGE с применением мочевины, затем соединяли 3′ с концевой трансферазой (TdT; New England Biolabs) в присутствии dGTP или dATP в соответствии с протоколом производителя. Присоединенные ДНК-продукты 10-кратно разбавляли связывающим буфером (40% насыщенный водный хлорид гуанидина + 60% изопропанол) и очищали с помощью спин-колонки QIAquick (Qiagen), затем использовали в качестве матрицы для праймерного удлинения с использованием фрагмента Klenow (New England Biolabs) с применением праймера AVA134 (A-присоединенные продукты) или AVA135 (G-присоединенные прдукты) (Таблицы 2A-2C). Удлинение амплифицировали с помощью ПЦР в рамках 10 циклов с использованием праймеров AVA110 и AVA122, затем секвенировали с AVA037 с использованием способа Сангера (Таблицы 2A-2C).

[1346] Анализы на основе флуоресцентного репортера дрожжей

[1347] Двойные флуоресцентные репортерные плазмиды, содержащие находящийся в рамке считывания стоп-кодон, +1 сдвиг рамки считывания, или -1 сдвиг рамки считывания, подвергали *in vitro* реакциям праймерного редактирования с использованием 5′-удлиненной PEgRNA или 3′-удлиненной PEgRNA, как описано выше. После инкубирования при 37 °C в течение 1 часа реакции разбавляли водой и

осаждали плазмидную ДНК с использованием 0,3 М ацетата натрия и 70% этанола. Ресуспендированную ДНК трансформировали в S. cerevisiae с помощью электропорации, как описано ранее[67], и высевали на синтетическую полную среду без лейцина (SC(глюкоза), L-). Сигналы флуоресценции GFP и mCherry визуализировали из колоний с использованием биомолекулярного сканера Typhoon FLA 7000.

[1348] Общие условия культивирования клеток млекопитающих

[1349] Клетки HEK293T (ATCC CRL-3216), U2OS (ATTC HTB-96), K562 (CCL-243), и HeLa (CCL-2) приобретали у ATCC и культивировали и пассировали в модифицированной по способу Дульбекко среде Игла (DMEM) плюс GlutaMAX (ThermoFisher Scientific), среде МакКоя 5A (Gibco), среде RPMI 1640 плюс GlutaMAX (Gibco), или минимальной основной среде Игла (EMEM, ATCC), соответственно, с добавлением к каждой 10% (об./об.) фетальной бычьей сыворотки (Gibco, квалифицированная) и 1x пенициллин-стрептомицина (Corning). Все типы клеток инкубировали, поддерживали, и культивировали при 37 °C с 5% CO_2. Линии клеток были проверены их соответствующими поставщиками и протестированы на отсутствие микоплазмы.

[1350] Протокол трансфекции культуры ткани HEK293T и подготовка геномной ДНК

[1351] Выращенные клетки HEK293T высевали на 48-луночные покрытые поли-D-лизином планшеты (Corning). Через 16-24 часа после высевания клетки трансфицировали приблизительно с 60% конфлюентностью 1 мкл Lipofectamine 2000 (Thermo Fisher Scientific) в соответствии с протоколами производителя и 750 нг плазмиды PE, 250 нг плазмиды PEgRNA, и 83 нг плазмиды sgРНК (для PE3 и PE3b). Если не указано иное, клетки культивировали в течение 3 дней после трансфекции, после чего среду удаляли, клетки промывали 1x PBS раствором (Thermo Fisher Scientific), и выделяли геномную ДНК путем добавления 150 мкл свежеприготовленного лизисного буфера (10 мМ Tris-HCl, pH 7,5; 0,05% SDS; 25 мкг/мкл Протеиназы K (ThermoFisher Scientific)) напрямую в каждую лунку планшета для культивирования ткани. Смесь геномной ДНК инкубировали при 37 °C в течение от 1 до 2 часов с последующим этапом в виде деактивации ферментов при 80 °C в течение 30 минут. Праймеры, использованные для амплификации геномной ДНК клеток млекопитающих, перечислены в Таблице 4. Для экспериментов на основе HDR

с использованием клеток HEK293T 231 нг плазмиды экспрессии нуклеазы, 69 нг плазмиды экспрессии sgРНК, 50 нг (1,51 пмоль) 100 nt донорской оцДНК матрицы (очищенной с помощью PAGE; Integrated DNA Technologies) были липофектированы с использованием 1,4 мкл Lipofectamine 2000 (ThermoFisher) в расчете на одну лунку. Геномную ДНК из всех экспериментов на основе HDR очищали с помощью набора Agencourt DNAdvance Kit (Beckman Coulter) в соответствии с протоколом производителя.

[1352] Высокопроизводительное секвенирование ДНК образцов геномной ДНК

[1353] Представляющие интерес геномные участки амплифицировали из образцов геномной ДНК и секвенировали на Illumina MiSeq, как описано ранее, со следующими модификациями[129,130]. Кратко, праймеры амплификации, содержащие прямые и обратные адаптеры Illumina (Таблица 4), использовали для первого этапа ПЦР-амплификации (ПЦР 1) представляющего интерес геномного участка. Осуществляли 25 мкл ПЦР 1 реакции с 0,5 мкМ каждого из прямого и обратного праймера, 1 мкл экстракта геномной ДНК, и 12,5 мкл ПЦР мастер-микса Phusion U Green Multiplex. ПЦР реакции осуществляли следующим образом: 98 °C в течение 2 минут, затем 30 циклов из [98 °C в течение 10 секунд, 61 °C в течение 20 секунд, и 72 °C в течение 30 секунд], и далее конечная стадия при 72 °C в течение 2 минут. В рамках вторичной ПЦР реакции (ПЦР 2) в каждый образец добавляли уникальные обеспечивающие мечение пары праймеров Illumina. ПЦР 2 реакции мечения осуществляли следующим образом: 98 °C в течение 2 минут, затем 12 циклов из [98 °C в течение 10 секунд, 61 °C в течение 20 секунд, и 72 °C в течение 30 секунд], и далее конечная стадия при 72 °C в течение 2 минут. Продукты ПЦР анализировали аналитическим образом с помощью электрофореза в 1,5% агарозном геле. Продукты ПЦР 2 (объединенные на основе общих ампликонов) очищали с помощью электрофореза с 1,5% агарозным гелем с использованием набора QIAquick Gel Extraction Kit (Qiagen) и элюировали 40 мкл воды. Концентрацию ДНК измеряли с помощью флуорометрического количественного анализа (Qubit, ThermoFisher Scientific) или кПЦР (набор KAPA Library Quantification Kit- Illumina, KAPA Biosystems) и секвенировали на инструменте Illumina MiSeq в соответствии с протоколами производителя.

[1354] Считывания в рамках секвенирования демультиплексировали с помощью MiSeq Reporter (Illumina). Выравнивание последовательностей ампликонов относительно

контрольной последовательности осуществляли с помощью CRISPResso2[178]. Для количественного анализа редактирования в виде точечных мутаций запускали CRISPResso2 в стандартном режиме с включенной настройкой "discard_indel_reads" ("игнорировать считывания с инделями"). Эффективность редактирования вычисляли следующим образом: (частота указанной точечной мутации в неигнорируемых считываниях) x (# неигнорируемых считываний) ÷ всего считываний. Для изменений в виде вставки или делеции CRISPResso2 запускали в режиме HDR с использованием требуемого аллеля в качестве ожидаемого аллеля (флаг e) и с включенной настройкой "discard_indel_reads" ("игнорировать считывания с инделями"). Выход редактирования вычисляли в виде количества считываний HDR, разделенного на общее количество считываний. Для всех изменений выходы инделей вычисляли в виде количества игнорируемых считываний, разделенного на общее количество считываний.

[1355] Нуклеофекция клеток U2OS, K562, и HeLa

[1356] Нуклеофекцию осуществляли в рамках всех экспериментов с использованием клеток K562, HeLa, и U2OS. Для условий PE в рамках данных типов клеток 800 нг плазмиды экспрессии праймерного редактора, 200 нг плазмиды экспрессии PEgRNA, и 83 нг никирующей плазмиды подвергали нуклеофекции в рамках конечного объема 20 мкл в 16-луночном стрипе Nucleocuvette (Lonza). Для условий HDR в рамках данных трех типов клеток 350 нг плазмиды экспрессии нуклеазы, 150 нг плазмиды экспрессии sgРНК, и 200 пмоль (6,6 мкг) 100 nt донорской оцДНК-матрицы (очищенной с помощью PAGE; Integrated DNA Technologies) подвергали нуклеофекции в рамках конечного объема 20 мкл в 16-луночном стрипе Nucleocuvette (Lonza). Клетки K562 подвергали нуклеофекции с использованием набора SF Cell Line 4D-Nucleofector X Kit (Lonza) с 5×10^5 клетками в расчете на один образец (программа FF-120) в соответствии с протоколом производителя. Клетки U2OS подвергали нуклеофекции с использованием набора SF Cell Line 4D-Nucleofector X Kit (Lonza) с 3–4 $\times 10^5$ клетками в расчете на один образец (программа DN-100) в соответствии с протоколом производителя. Клетки HeLa подвергали нуклеофекции с использованием набора SF Cell Line 4D-Nucleofector X Kit (Lonza) с 2×10^5 клетками в расчете на один образец (программа CN-114) в соответствии с протоколом производителя. Клетки собирали спустя 72 часа после нуклеофекции для выделения геномной ДНК.

[1357] Выделение геномной ДНК для экспериментов на основе HDR

[1358] Геномную ДНК из всех экспериментов на основе сравнения HDR в клетках HEK293T, HEK293T *HBB* E6V, K562, U2OS, и HeLa очищали с помощью набора Agencourt DNAdvance Kit (Beckman Coulter) в соответствии с протоколом производителя.

[1359] Сравнение между PE2, PE3, BE2, BE4max, ABEdmax, и ABEmax

[1360] Клетки HEK293T высевали на 48-луночные покрытые поли-D-лизином планшеты (Corning). Спустя 16-24 часа клетки трансфицировали с приблизительно 60% конфлюентностью. Для редактирования оснований с использованием конструктов CBE или ABE клетки трансфицировали 750 нг плазмиды редактора оснований, 250 нг плазмиды экспрессии sgРНК, и 1 мкл Lipofectamine 2000 (Thermo Fisher Scientific). Трансфекции PE осуществляли как описано выше. Выделение геномной ДНК для PE и BE осуществляли как описано выше.

[1361] Определение активности PE3 в известных нецелевых участках Cas9

[1362] Для оценки активности нецелевого редактирования PE3 в известных нецелевых участках Cas9 геномную ДНК, выделенную из клеток HEK293T спустя 3 дня после трансфекции PE3, использовали в качестве матрицы для ПЦР амплификации 16 ранее описанных нецелевых геномных участков Cas9[118, 159] (главных четырех нецелевых участков для каждого из HEK3, EMX1, FANCF, и HEK4 спейсеров; праймерные последовательности перечислены в Таблице 4). Данные образцы геномной ДНК были идентичны тем, которые применялись для количественного анализа активности в виде целевого редактирования PE3, показанного на Фиг.41A-41K; последовательности PEgRNA и никирующей sgРНК перечислены в Таблицах 3A-3R. После ПЦР амплификации нецелевых участков ампликоны секвенировали на платформе Illumina MiSeq, как описано выше (HTS анализ). Для определения нуклеазы Cas9, никазы Cas9 H840A, dCas9, и активности в виде целевого и нецелевого редактирования PE2-dRT клетки HEK293T трансфицировали 750 нг плазмиды для редактирования (нуклеаза Cas9, никаза Cas9 H840A, dCas9, или PE2-dRT), 250 нг плазмиды PEgRNA или sgРНК, и 1 мкл Lipofectamine 2000. Геномную ДНК выделяли из клеток спустя 3 дня после трансфекции, как описано выше. Целевые и нецелевые геномные локусы

амплифицировали с помощью ПЦР с использованием праймерных последовательностей из Таблицы 4 и секвенировали на Illumina MiSeq.

[1363] Анализ данных HTS проводили с использованием CRISPResso2[178]. Эффективность редактирования нуклеазы Cas9, никазы Cas9 H840A, и dCas9 подвергали количественному анализу в виде процентной доли всех считываний в рамках секвенирования, содержащих индели. Для количественного анализа нецелевых мишеней PE3 и PE3-dRT подвергнутые выравниванию результаты считывания в рамках секвенирования анализировали на наличие точечных мутаций, вставок, или делеций, которые соответствовали ожидаемому продукту обратной транскрипции PEgRNA, инициированной в участке разрыва Cas9. Вариации отдельных нуклеотидов с общей частотой <0,1% среди всех результатов считывания в образце исключали из анализа. Для считываний, содержащих вариации отдельных нуклеотидов, которые как встречались с частотой ≥ 0,1%, так и были частично совпадающими с кодируемым PEgRNA изменением, использовали t-тест (неспаренный, односторонний, $\alpha = 0.5$) для определения того, встречались ли данные варианты со значительно более высокими уровнями по сравнению с образцами, обработанными PEgRNA, которые содержали тот же спейсер, но кодировали другие изменения. Для устранения различий в ошибках секвенирования сравнения проводили между образцами, которые были секвенированы одновременно в рамках одного прогона MiSeq. Варианты, которые не соответствовали критерию p-значения > 0,05, исключали. Затем вычисляли активность в виде нецелевого редактирования PE3 в виде процентной доли всех считываний в рамках секвенирования, которые соответствовали приведенным выше критериям.

[1364] Получение линии клеток HEK293T, содержащих мутацию HBB E6V, с помощью Cas9-инициируемой HDR

[1365] Клетки HEK293T высевали в 48-луночный планшет и трансфицировали приблизительно с 60% конфлюентностью 1,5 мкл Lipofectamine 2000, 300 нг плазмиды никазы Cas9 D10A, 100 нг плазмиды sgРНК, и 200 нг донорской оцДНК-матрицы со 100 повторяющимися единицами (Таблица 5). Спустя три дня после трансфекции среду заменяли на свежую среду. Спустя четыре дня после трансфекции клетки подвергались диссоциации с использованием 30 мкл раствора TrypLE и суспендировались в 1,5 мл среды. Отдельные клетки выделяли в отдельные лунки двух 96-луночных планшетов с помощью сортировки клеток с активированной флуоресценцией (FACS) (Beckman-

Coulter Astrios). См. соответствующие примеры FACS сортировки на Фиг.53A-53B. Клетки подвергали экспансии в течение 14 дней перед проведением секвенирования геномной ДНК, как показано выше. Было обнаружено, что из выделенных клональных популяций ни одна не являлась гомозиготной в плане мутации *HBB* E6V, поэтому был проведен второй раунд редактирования на основе липофекции, сортировки, и роста в параллельно редактируемой линии клеток для получения линии клеток, являющихся гомозиготными в плане аллеля E6V.

[1366] Получение линии клеток HEK293T, содержащих мутацию HBB E6V, с помощью PE3

[1367] 2,5 x 10^4 клеток HEK293T, выращенных при отсутствии антибиотиков, высевали в 48-луночные покрытые поли-D-лизином планшеты (Corning). Через 16-24 часа после высевания клетки трансфицировали приблизительно с 70% конфлюентностью 1 мкл Lipofectamine 2000 (Thermo Fisher Scientific) в соответствии с протоколами производителя и 750 нг плазмиды PE2-P2A-GFP, 250 нг плазмиды PEgRNA, и 83 нг плазмиды sgРНК. Спустя 3 дня после трансфекции клетки промывали фосфатно-буферным солевым раствором (Gibco) и подвергали диссоциации с использованием TrypLE Express (Gibco). Затем клетки разбавляли DMEM плюс GlutaMax (Thermo Fisher Scientific) с добавлением 10% (об./об.) FBS (Gibco) и пассировали через 35 мкм клеточный фильтр (Corning) перед проведением сортировки. Проточную цитометрию осуществляли на клеточном сортере LE-MA900 (Sony). Клетки обрабатывали 3 нМ DAPI (BioLegend) за 15 минут перед сортировкой. После гейтинга для исключения дублетов отдельные DAPI-отрицательные клетки с GFP флуоресценцией, превышающей флуоресценцию GFP-отрицательной контрольной популяции клеток, сортировали в 96-луночные планшеты для клеточных культур с плоским дном (Corning), наполненные предварительно охлажденной DMEM с GlutaMax с добавлением 10% FBS. См. соответствующие примеры FACS сортировки на Фиг.53A-53B. Клетки культивировали в течение 10 дней перед выделением геномной ДНК и анализом с помощью HTS, как описано выше. Было идентифицировано всего шесть линий клональных клеток, которые являлись гомозиготными в плане мутации E6V в *HBB*.

[1368] Получение линии клеток HEK293T, содержащих HEXA 1278 + вставку TATC, с помощью PE3

[1369] Клетки HEK293T, содержащие HEXA 1278+TATC аллель, получали в соответствии с протоколом, описанным выше для создания линии клеток HBB E6V; последовательности PEgRNA и sgРНК перечислены в Таблицах 2A-2C в соответствии с подзаголовками на Фиг.43A-43H. После трансфекции и сортировки клетки культивировали в течение 10 дней перед выделением геномной ДНК и анализом с помощью HTS, как описано выше. Были выделены две линии гетерозиготных клеток, которые содержали 50% *HEXA* 1278+TATC аллелей, при этом были выделены две линии гомозиготных клеток, содержащих 100% *HEXA* 1278+TATC аллелей.

[1370] Анализы жизнеспособности клеток

[1371] Клетки HEK293T высевали в 48-луночные планшеты и трансфицировали с приблизительно 70% конфлюентностью 750 нг плазмиды для редактирования (PE3, PE3 R110S K103L, никаза Cas9 H840A, или dCas9), 250 нг нацеленной на HEK3 плазмиды PEgRNA, и 1 мкл Lipofectamine 2000, как описано выше. Жизнеспособность клеток измеряли каждые 24 часа после трансфекции в течение 3 дней с помощью анализа CellTiter-Glo 2.0 (Promega) в соответствии с протоколом производителя. Люминесценцию измеряли в 96-луночных полистироловых микропланшетах с плоским дном (Corning) с использованием сканера для микропланшетов M1000 Pro (Tecan) с временем интеграции, соответствующим 1 секунде.

[1372] Получение лентивируса

[1373] Лентивирус получали как описано ранее[206]. Флаконы T-75 с быстро делящимися клетками HEK293T (ATCC; Manassas, VA, USA) трансфицировали хелперными плазмидами pVSV-G и psPAX2 для получения лентивируса в комбинации с модифицированными геномами lentiCRISPR_v2, несущими редактор PE2 на основе разделенного интеина, с применением FuGENE HD (Promega, Madison, WI, USA) в соответствии с инструкциями производителя. Были сконструированы четыре конструкта редактора на основе разделенного интеина: 1) вирусный геном, кодирующий кассету экспрессии U6-PEgRNA и N-концевую часть (1-573) никазы Cas9 H840A, слитые с N-интеином Npu, саморасщепляемым пептидом P2A, и GFP-KASH; 2) вирусный геном, кодирующий C-интеин Npu, слитый с C-концевым остатком PE2; 3)

вирусный геном, кодирующий C-интеин Npu, слитый с C-концевым остатком Cas9 для контроля Cas9; и 4) никирующая sgРНК для DNMT1. Разделенный интеин опосредует транс-сплайсинг для соединения двух половин PE2 или Cas9, в то время как P2A GFP-KASH обеспечивает котрансляционное продуцирование GFP, локализованного в ядерной мембране. Спустя 48 часов собирали супернатант, центрифугировали при 500 g в течение 5 минут для удаления клеточного дебриса, и фильтровали с помощью 0,45 мкм фильтра. Отфильтрованный супернатант концентрировали с помощью PEG-it Virus Precipitation Solution (System Biosciences, Palo Alto, CA, USA) в соответствии с инструкциями производителя. Полученный сгусток ресуспендировали в Opti-MEM (Thermo Fisher Scientific, Waltham, MA, USA) с использованием 1% исходного объема среды. Ресуспендированный сгусток подвергали шоковой заморозке и хранили при -80°C до использования.

[1374] Иссечение и культивирование первичных кортикальных нейронов мышей

[1375] Диссоциированные кортикальные культуры E18.5 собирали от мышей C57BL/6 с известным сроком беременности (Charles River). Зародышей собирали из беременных мышей после умерщвления с помощью CO_2 с последующим обезглавливанием. Кортикальные лоскуты иссекали в ледяном Hibernate-E с добавлением пенициллина/стрептомицина (Life Technologies). После промывания ледяным Hibernate-E ткани расщепляли при 37 °C в течение 8 минут в папаине/ДНКазе (Worthington/Sigma). Ткани растирали в NBActiv4 (BrainBits) с добавлением ДНКазы. Клетки подсчитывали и высевали в 24-луночные планшеты в количестве 100000 на одну лунку. Половину среды заменяли два раза в неделю.

[1376] Праймерное редактирование в первичных нейронах и выделение ядер

[1377] В DIV 1 добавляли 15 мкл лентивируса при 10:10:1 соотношении N-концевой:C-концевой:никирующая sgРНК. В DIV 14 ядра нейронов выделяли с помощью буфера EZ-PREP (Sigma D8938) в соответствии с протоколом производителя. Все этапы осуществляли на льду или при 4 °C. Из подвергнутых диссоциации культур удаляли среду, при этом культуры промывали ледяным PBS. PBS аспирировали и замещали 200 мкл раствора EZ-PREP. После 5-минутного инкубирования на льду EZ-PREP вносили с помощью пипетки вдоль площади лунки для удаления оставшихся клеток. Образец центрифугировали при 500 g в течение 5 минут и удаляли супернатант. Образцы

промывали 200 мкл EZ-PREP и центрифигурували снова при 500 g в течение 5 минут. Образцы осторожно ресуспендировали с помощью пипетки в 200 мкл ледяного буфера Nuclei Suspension Buffer (NSB), состоящего из 100 мкг/мл BSA и 3,33 мкМ Vybrant DyeCycle Ruby (Thermo Fisher) в 1xPBS, затем центрифугировали при 500 g в течение 5 минут. Супернатант удаляли и ресуспендировали ядра в 100 мкл NSB и сортировали в 100 мкл лизисного буфера Agencourt DNAdvance с использованием MoFlo Astrios (Beckman Coulter) в помещении для проточной цитометрии Института Броуда. Геномную ДНК очищали в соответствии с инструкциями Agencourt DNAdvance производителя.

[1378] РНК-секвенирование и анализ данных

[1379] Клетки HEK293T котрансфицировали нацеленными на PRNP или нацеленными на HEXA PEgRNA и PE2, PE2-dRT, или никазой Cas9 H840A. Спустя 72 часа после трансфекции всю РНК собирали из клеток с помощью реагента TRIzol (Thermo Fisher) и очищали с помощью набора RNeasy Mini (Qiagen), включая обработку ДНКазой I на колонке. Содержание рибосом в общей РНК уменьшали с использованием протокола удаления рРНК набора TruSeq Stranded Total для получения библиотек РНК (Illumina) и далее промывали частицами RNAClean XP (Beckman Coulter). Библиотеки секвенирования получали с помощью РНК с пониженным содержанием рибосом на системе SMARTer PrepX Apollo для получения NGS-библиотек (Takara) в соответствии с протоколом производителя. Полученные библиотеки визуализировали на 2200 TapeStation (Agilent Technologies), подвергали нормализации с использованием набора Qubit dsDNA HS Assay Kit (Thermo Fisher), и секвенировали на NextSeq 550 с использованием высокопроизводительной проточной ячейки v2 (Illumina) в виде 75 bp парных считываний. Fastq файлы были получены с помощью bcl2fastq2 версии 2.20 и обработаны с помощью TrimGalore версии 0.6.2 (https://github.com/FelixKrueger/TrimGalore) для удаления низкокачественных оснований, неспаренных последовательностей, и адаптерных последовательностей. Обработанные результаты считывания были оптимизированы на основе геномной сборки GRCh[148] Homo sapiens с использованием модифицированной записи, соответствующей гену H840A Cas9, с помощью RSEM версии 1.3.1[207]. Использовали пакет limma-voom[208] для нормализации уровней генной экспрессии и реализации дифференциального анализа экспрессии с серийной коррекцией эффектов.

Дифференциально экспрессируемые гены были определены на основе пороговых значений в виде FDR-скорректированного p-значения < 0,05 и кратности изменения >2 , при этом результаты были визуализированы с помощью R.

[1380] Анализ ClinVar

[1381] Описание вариантов ClinVar скачивали с NCBI (по состоянию на 15 июля 2019), при этом содержащуюся в нем информацию использовали для всех последующих анализов. Перечень всех указанных вариантов фильтровали на основе ID аллеля для удаления дубликатов и на основе клинической значимости для ограничения анализа патогенными вариантами. Перечень патогенных вариантов последовательно фильтровали на основе типа вариантов для вычисления доли патогенных вариантов, которые представляют собой вставки, делеции, и так далее. Варианты на основе одного нуклеотида (SNV) разделяли на две категории (транзиции и трансверсии) на основе указанного обозначения и измененных аллелей. SNV без указания обозначения или измененных аллелей исключали из анализа.

[1382] Показатели длины указанных вставок, делеций, и дубликатов вычисляли с помощью исходных/измененных аллелей, положений начала/конца вариантов, или подходящей идентификаторной информации в названии варианта. Варианты без указания перечисленной выше информации исключали из анализа. Показатели длины указанных инделей (отдельные варианты, которые включают как вставки, так и делеции, по сравнению с исходным геномом) вычисляли путем определения количества несовпадений или промежутков при наилучшем попарном выравнивании между исходными и измененными аллелями. Распределение частоты длины вариантов вычисляли с помощью GraphPad Prism 8.

[1383] Доступность данных

[1384] Данные высокопроизводительного секвенирования внесены в базу данных архива считываний секвенирования NCBI. Плазмиды, кодирующие векторы экспрессии PE1, PE2/PE3, и PEgRNA, будут доступны в Addgene.

[1385] Доступность кода

[1386] Скрипт, использованный для количественного анализа вставки каркаса PEgRNA, представлен на Фиг.60A-60B.

[1387] Дополнительная информация: Таблицы и последовательности

[1388] Таблица 1: Активность праймерных редакторов, нуклеазы Cas9, никазы Cas9 H840A, и PE2-dRT в целевых и нецелевых участках HEK3, HEK4, EMX1, и FANCF. Редактирование PE2/PE3 показано в виде % праймерного редактирования наряду с % уровня инделей (в скобках). % уровня инделей показан для Cas9, никазы Cas9 H840A (nCas9), и PE2-dRT в главных четырех ранее описанных нецелевых участках[179,180]. Последовательности sgРНК и PEgRNA приведены в Таблицах 3A-3R под заголовком, соответствующим Фиг.42A-42H. Все значения представляют собой среднее значение трех независимых биологических повторов.

PE РЕнРНК Участок	HEK3 (PE3)					HEK4 (PE2)					EMX1 (PE3)					FANCF (PE3)				
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
Целевой		44.2 (11.9)	61.2 (6.8)	40.4 (16.5)	48.4 (3.3)		16.2 (0.9)	14.4 (1.8)	9.8 (2.0)	7.9 (2.2)		28.6 (3.5)	14.1 (2.4)	35.7 (3.3)	15.4 (2.9)		56.8 (9.3)	32.4 (16.7)	42.8 (13.6)	47.6 (12.0)
Нецелевой 1		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	0.4 (<0.1)	<0.1 (<0.1)	0.4 (<0.1)		<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)		<0.1 (0.1)	0.6 (0.1	<0.1 (0.1)	<0.1 (0.1)
Нецелевой 2		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)		<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (<0.1)
Нецелевой 3		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		0.2 (0.5)	6.8 (1.9)	19.2 (0.5)	7.9 (3.5)		<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Нецелевой 4		<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1		<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (<0.1		<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)

Cas9 РЕнРНК Участок	HEK3					HEK4					EMX1					FANCF				
	sgPHK	1	2	3	4	sgPHK	1	2	3	4	sgPHK	1	2	3	4	sgPHK	1	2	3	4
Целевой	91.8	87.5	89.2	89.1	86.8	71.8	68.6	72.9	72.8	70.9	85.6	79.7	70.6	76.6	76.0	78.7	55.9	58.3	51.8	52.0
Нецелевой 1	17.2	1.9	5.5	5.2	1.8	54.2	39.5	48.4	49.7	49.2	81.1	63.5	48.1	53.0	59.6	12.6	1.9	1.9	1.7	1.7
Нецелевой 2	38.0	6.5	12.6	11.8	4.7	42.5	19.5	29.4	27.3	30.3	58.3	12.0	6.0	8.2	12.9	1.1	0.2	0.2	0.2	0.1
Нецелевой 3	8.8	0.6	1.7	1.5	0.5	98.1	96.9	97.3	97.6	97.5	14.8	4.2	3.1	3.6	4.8	2.4	0.2	<0.1	0.2	0.2
Нецелевой 4	0.3	<0.1	<0.1	0.1	<0.1	45.3	16.9	28.0	27.5	29.7	39.5	1.3	0.9	0.6	1.3	1.0	0.2	0.2	0.2	0.2

nCas9 РЕнРНК Участок	HEK3					HEK4					EMX1					FANCF				
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
Нецелевой 1		<0.1	<0.1	<0.1	<0.1		<0.1	<0.1	<0.1	<0.1		0.1	0.1	<0.1	0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 2		<0.1	<0.1	<0.1	<0.1		<0.1	<0.1	< 0.1	<0.1		0.3	0.1	<0.1	<0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 3		<0.1	<0.1	<0.1	<0.1		0.3	0.5	0.7	0.7		<0.1	<0.1	<0.1	<0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 4		<0.1	<0.1	<0.1	<0.1		<0.1	0.1	<0.1	<0.1		1.5	0.1	0.1	0.4		<0.1	<0.1	<0.1	<0.1

PE2-dRT РЕнРНК Участок	HEK3					HEK4					EMX1					FANCF				
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
Нецелевой 1		<0.1	<0.1	<0.1	<0.1		0.1	0.2	<0.1	<0.1		0.1	0.1	0.1	0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 2		<0.1	<0.1	<0.1	<0.1		<0.1	0.1	<0.1	<0.1		<0.1	0.3	<0.1	<0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 3		<0.1	<0.1	<0.1	<0.1		<0.1	0.1	1.4	0.9		<0.1	<0.1	<0.1	<0.1		<0.1	<0.1	<0.1	<0.1
Нецелевой 4		<0.1	<0.1	<0.1	<0.1		0.1	0.1	0.1	0.2		0.9	0.1	0.1	0.2		<0.1	<0.1	<0.1	<0.1

[1389]

[1390] Таблицы 2A-2C: Последовательности олигонуклеотидов ДНК, PEgRNA, и sgРНК, использованных для *in vitro* экспериментов.

[1391] Таблица 2A: Олигонуклеотиды ДНК

ОЛИГОНУКЛЕОТИД	ПОСЛЕДОВАТЕЛЬНОСТЬ
AVA023	5CY5-CCTGGGTCAATCCTTGGGGCCCAGACTGAGCACG (SEQ ID NO: 374)
AVA024	5CY5-CCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGG (SEQ ID NO: 375)
AVA025	5PHOS-

	CCTTTCCTCTGCCATCACGTGCTCAGTCTGGGCCCCAAGGATTGACCCAGG (SEQ ID NO: 376)
AVA026	5PHOS-TGATGGCAGAGGAAAGG (SEQ ID NO: 377)
AVA037	GCAGGCTTTAAAGGAACCAATTC (SEQ ID NO: 378)
AVA110	GCAGGCTTTAAAGGAACCAATTCCCTGGGTCAATCCTTGGGGC (SEQ ID NO: 379)
AVA122	CTCTGGAGGATCTAGCGGAG (SEQ ID NO: 380)
AVA134	CTCTGGAGGATCTAGCGGAGTTTTTTTTTTTTTTTTTT (SEQ ID NO: 381)
AVA135	CTCTGGAGGATCTAGCGGAGCCCCCCCCCCCCCC (SEQ ID NO: 382)

[1392] Таблица 2B: 5′-удлиненные PEgRNA

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	ПОСЛЕДОВАТЕЛЬНОСТЬ 5′ УДЛИНЕНИЯ	ДЛИНА ЛИНКЕРА (NT)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
PEGRNA 1	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 383)	GGCTAACCGTGCCATTTGATCAGGTCA (SEQ ID NO: 429)	15	5	7
PEGRNA 2	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 384)	GGCTAACCGTGCAAATTAACAAACTAA(SEQ ID NO: 430)	15	5	7
PEGRNA 3	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 385)	GGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 431)	15	5	8
PEGRNA 4	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 386)	GGTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 432)	15	5	15
PEGRNA 5	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 387)	GGCTTCCTTTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 433)	15	5	22
5′-PEGRNA_RT_7_A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 388)	GGCTAACCGTGCCATTTGATCAGGTCA(SEQ ID NO: 434)	15	5	7
5′-PEGRNA_RT_7_B	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 389)	GGCTAACCGTGCAAATTAACAAACTAA(SEQ ID NO: 435)	15	5	7
5′-PEGRNA_RT_8	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 390)	GGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 436)	15	5	8
5′-PEGRNA_RT_15	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 391)	GGTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 437)	15	5	15
5′-	GGCCCAGACTG	GGCTTCCTTTCCTCTG	15	5	22

PEGRNA_RT_22	AGCACGTGA (SEQ ID NO: 392)	CCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 438)			

[1393] Таблица 2C: 3′-удлиненные PEgRNA

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′-УДЛИНЕНИЯ	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
3′-PEGRNA_10	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 506)	TCTGCCATCTCGTGCTC (SEQ ID NO: 439)	7	10
3′-PEGRNA_YEAST_TTOA	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 507)	TCTGCCATCTCGTGCTC(SEQ ID NO: 440)	7	10
3′-PEGRNA_YEAST_+1AINS	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 508)	TCTGCCATCATCGTGCTC(SEQ ID NO: 441)	7	11
3′-PEGRNA_YEAST_+1TDEL	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 509)	TCTGCCATCCGTGCTC(SEQ ID NO: 442)	7	9

[1394] Таблицы 3A-3R: Последовательность PEgRNA и sgРНК, использованные в экспериментах на основе клеток млекопитающих. Все последовательности показаны в ориентации от 5′ к 3′. Для конструирования PEgRNA перечисленные ниже спейсерные последовательности добавляли к 5′-концу каркаса sgРНК, а перечисленные ниже 3′ удлинения, содержащие участок связывания праймера и RT матрицу, добавляли к 3′-концу каркаса sgРНК. Последовательность каркаса sgРНК представляет собой: GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 131).

[1395] Таблица 3A: PEgRNA на Фиг.39A-39D

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 2890-2996)	3′ УДЛИНЕНИЕ (SEQ ID NO:2997-3103)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_2B-C_8	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCA	8	10
HEK3_2B-C_9	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAG	9	10
HEK3_2B-C_10	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGT	10	10
HEK3_2B-	GGCCCAGACTGAG	TCTGCCATCTCGTGCTCA	11	10

C_11	CACGTGA	GTC		
HEK3_2B-C_12	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCT	12	10
HEK3_2B-C_13	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_2B-C_14	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTGG	14	10
HEK3_2B-C_15	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTGGG	15	10
HEK3_2C_16	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTGGGC	16	10
HEK3_2C_17	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTGGGCC	17	10
EMX1_2C_9	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGC	9	13
EMX1_2C_10	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCT	10	13
EMX1_2C_11	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTC	11	13
EMX1_2C_12	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCG	12	13
EMX1_2C_13	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGG	13	13
EMX1_2C_14	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGGA	14	13
EMX1_2C_15	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGGAC	15	13
EMX1_2C_16	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGGACT	16	13
EMX1_2C_17	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGGACTC	17	13
FANCF_2C_8	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGA	8	17
FANCF_2C_9	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAA	9	17
FANCF_2C_10	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAG	10	17
FANCF_2C_11	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGG	11	17
FANCF_2C_12	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGG	12	17
FANCF_2C_13	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	17
FANCF_2C_14	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGAT	14	17
FANCF_2C_15	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGATT	15	17
FANCF_2C	GGAATCCCTTCTGC	GGAAAAGCGATCAAGGT	16	17

_16	AGCACC	GCTGCAGAAGGGATTC		
FANCF_2C_17	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGATTCC	17	17
RNF2_2C_9	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACT	9	11
RNF2_2C_10	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTA	10	11
RNF2_2C_11	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAA	11	11
RNF2_2C_12	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAG	12	11
RNF2_2C_13	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGA	13	11
RNF2_2C_14	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGAT	14	11
RNF2_2C_15	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGATG	15	11
RNF2_2C_16	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGATGA	16	11
RNF2_2C_17	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGATGAC	17	11
HEK4_2C_7	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAG	7	13
HEK4_2C_8	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGC	8	13
HEK4_2C_9	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCC	9	13
HEK4_2C_10	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCG	10	13
HEK4_2C_11	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGC	11	13
HEK4_2C_12	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGCA	12	13
HEK4_2C_13	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGCAG	13	13
HEK4_2C_14	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGCAGT	14	13
HEK4_2C_15	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGCAGTG	15	13
HEK3_2C_1TDEL	GGCCCAGACTGAGCACGTGA	TCTGCCATCCGTGCTCAGTCTG	13	10
HEK3_2C_1AINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCATCGTGCTCAGTCTG	13	10
HEK3_2C_1CTTINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	10
HEK3_2D_10	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_2D_1	GGCCCAGACTGAG	CTCTGCCATCTCGTGCTC	13	11

1	CACGTGA	AGTCTG		
HEK3_2D_12	GGCCCAGACTGAGCACGTGA	CCTCTGCCATCTCGTGCTCAGTCTG	13	12
HEK3_2D_13	GGCCCAGACTGAGCACGTGA	TCCTCTGCCATCTCGTGCTCAGTCTG	13	13
HEK3_2D_14	GGCCCAGACTGAGCACGTGA	TTCCTCTGCCATCTCGTGCTCAGTCTG	13	14
HEK3_2D_15	GGCCCAGACTGAGCACGTGA	TTTCCTCTGCCATCTCGTGCTCAGTCTG	13	15
HEK3_2D_16	GGCCCAGACTGAGCACGTGA	CTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	16
HEK3_2D_17	GGCCCAGACTGAGCACGTGA	CCTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	17
HEK3_2D_18	GGCCCAGACTGAGCACGTGA	TCCTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	18
HEK3_2D_19	GGCCCAGACTGAGCACGTGA	TTCCTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	19
HEK3_2D_20	GGCCCAGACTGAGCACGTGA	CTTCCTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	20
EMX1_2D_10	GAGTCCGAGCAGAAGAAGAA	GGAGCCCTTGTTCTTCTGCTCGG	13	10
EMX1_2D_11	GAGTCCGAGCAGAAGAAGAA	GGGAGCCCTTGTTCTTCTGCTCGG	13	11
EMX1_2D_12	GAGTCCGAGCAGAAGAAGAA	TGGGAGCCCTTGTTCTTCTGCTCGG	13	12
EMX1_2D_13	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCCCTTGTTCTTCTGCTCGG	13	13
EMX1_2D_14	GAGTCCGAGCAGAAGAAGAA	GATGGGAGCCCTTGTTCTTCTGCTCGG	13	14
EMX1_2D_15	GAGTCCGAGCAGAAGAAGAA	TGATGGGAGCCCTTGTTCTTCTGCTCGG	13	15
EMX1_2D_16	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTTGTTCTTCTGCTCGG	13	16
EMX1_2D_17	GAGTCCGAGCAGAAGAAGAA	TGTGATGGGAGCCCTTGTTCTTCTGCTCGG	13	17
EMX1_2D_18	GAGTCCGAGCAGAAGAAGAA	ATGTGATGGGAGCCCTTGTTCTTCTGCTCGG	13	18
EMX1_2D_19	GAGTCCGAGCAGAAGAAGAA	GATGTGATGGGAGCCCTTGTTCTTCTGCTCGG	13	19
EMX1_2D_20	GAGTCCGAGCAGAAGAAGAA	TGATGTGATGGGAGCCCTTGTTCTTCTGCTCGG	13	20
FANCF_2D_10	GGAATCCCTTCTGCAGCACC	CGATCAAGGTGCTGCAGAAGGGA	13	10
FANCF_2D_11	GGAATCCCTTCTGCAGCACC	GCGATCAAGGTGCTGCAGAAGGGA	13	11
FANCF_2D_12	GGAATCCCTTCTGCAGCACC	AGCGATCAAGGTGCTGCAGAAGGGA	13	12
FANCF_2D	GGAATCCCTTCTGC	AAGCGATCAAGGTGCTG	13	13

_13	AGCACC	CAGAAGGGA		
FANCF_2D_14	GGAATCCCTTCTGCAGCACC	AAAGCGATCAAGGTGCTGCAGAAGGGA	13	14
FANCF_2D_15	GGAATCCCTTCTGCAGCACC	AAAAGCGATCAAGGTGCTGCAGAAGGGA	13	15
FANCF_2D_16	GGAATCCCTTCTGCAGCACC	GAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	16
FANCF_2D_17	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	17
FANCF_2D_18	GGAATCCCTTCTGCAGCACC	CGGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	18
FANCF_2D_19	GGAATCCCTTCTGCAGCACC	TCGGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	19
FANCF_2D_20	GGAATCCCTTCTGCAGCACC	CTCGGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	20
RNF2_2D_10	GTCATCTTAGTCATTACCTG	AACACCTCATGTAATGACTAAGATG	15	10
RNF2_2D_11	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTAAGATG	15	11
RNF2_2D_12	GTCATCTTAGTCATTACCTG	CGAACACCTCATGTAATGACTAAGATG	15	12
RNF2_2D_13	GTCATCTTAGTCATTACCTG	ACGAACACCTCATGTAATGACTAAGATG	15	13
RNF2_2D_14	GTCATCTTAGTCATTACCTG	AACGAACACCTCATGTAATGACTAAGATG	15	14
RNF2_2D_15	GTCATCTTAGTCATTACCTG	CAACGAACACCTCATGTAATGACTAAGATG	15	15
RNF2_2D_16	GTCATCTTAGTCATTACCTG	ACAACGAACACCTCATGTAATGACTAAGATG	15	16
RNF2_2D_17	GTCATCTTAGTCATTACCTG	TACAACGAACACCTCATGTAATGACTAAGATG	15	17
RNF2_2D_18	GTCATCTTAGTCATTACCTG	TTACAACGAACACCTCATGTAATGACTAAGATG	15	18
RNF2_2D_19	GTCATCTTAGTCATTACCTG	GTTACAACGAACACCTCATGTAATGACTAAGATG	15	19
RNF2_2D_20	GTCATCTTAGTCATTACCTG	AGTTACAACGAACACCTCATGTAATGACTAAGATG	15	20
HEK4_2D_7	GGCACTGCGGCTGGAGGTGG	ACCCCAACCTCCAGCCGC	11	7
HEK4_2D_8	GGCACTGCGGCTGGAGGTGG	AACCCCAACCTCCAGCCGC	11	8
HEK4_2D_9	GGCACTGCGGCTGGAGGTGG	TAACCCCAACCTCCAGCCGC	11	9
HEK4_2D_10	GGCACTGCGGCTGGAGGTGG	TTAACCCCAACCTCCAGCCGC	11	10
HEK4_2D_11	GGCACTGCGGCTGGAGGTGG	TTTAACCCCAACCTCCAGCCGC	11	11
HEK4_2D_1	GGCACTGCGGCTG	CTTTAACCCCAACCTCC	11	12

2	GAGGTGG	AGCCGC		
HEK4_2D_13	GGCACTGCGGCTGGAGGTGG	GCTTTAACCCCAACCTCCAGCCGC	11	13
HEK4_2D_14	GGCACTGCGGCTGGAGGTGG	CGCTTTAACCCCAACCTCCAGCCGC	11	14
HEK4_2D_15	GGCACTGCGGCTGGAGGTGG	CCGCTTTAACCCCAACCTCCAGCCGC	11	15
HEK4_2D_16	GGCACTGCGGCTGGAGGTGG	TCCGCTTTAACCCCAACCTCCAGCCGC	11	16
HEK4_2D_17	GGCACTGCGGCTGGAGGTGG	CTCCGCTTTAACCCCAACCTCCAGCCGC	11	17
HEK4_2D_18	GGCACTGCGGCTGGAGGTGG	CTCCGCTTTAACCCCAACCTCCAGCCGC	11	18
HEK4_2D_19	GGCACTGCGGCTGGAGGTGG	CTCCGCTTTAACCCCAACCTCCAGCCGC	11	19

[1396] Таблица 3B: PEgRNA на Фиг.40A-40C

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3104-3112)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3113-3121)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
RNF2_3B	GTCATCTTAGTCATTACCTG	AACGAACACCTCATGTAATGACTAAGATG	15	14
EMX1_3B	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCACTTCTTCTTCTGCTCGGAC	15	13
FANCF_3B	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGATT	15	17
HE3_3B	GGCCCAGACTGAGCACGTGA	TCTGCCATGACGTGCTCAGTCTG	13	10
HEK4_3B	GGCACTGCGGCTGGAGGTGG	TTAACCCCAACCTCCAGCC	9	10
RNF2_3C_4ATOC	GTCATCTTAGTCATTACCTG	AACGAACACCGCAGGTAATGACTAAGATG	15	14
RNF2_3C_4ATOG	GTCATCTTAGTCATTACCTG	AACGAACACCCCAGGTAATGACTAAGATG	15	14
FANCF_3C_5GTOT	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA	13	17
FANCF_3C_7ATOC	GGAATCCCTTCTGCAGCACC	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT	14	17

[1397] Таблица 3C: Последовательности никирующей sgРНК на Фиг.40A-40C

НИКИРУЮЩАЯ SGPHK	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	SEQ ID NO:
RNF2_2B_+41	GTCAACCATTAAGCAAAACAT	**3122**

RNF2_2B_+67	GTCTCAGGCTGTGCAGACAAA	**3123**
EMX1_2B_-116	GGGGCACAGATGAGAAACTC	**3124**
EMX1_2B_-57	GCCGTTTGTACTTTGTCCTC	**3125**
EMX1_2B_+14	GCGCCACCGGTTGATGTGAT	**3126**
EMX1_2B_+27	GCTTCGTGGCAATGCGCCAC	**3127**
EMX1_2B_+53	GACATCGATGTCCTCCCCAT	**3128**
EMX1_2B_+80	GTGGTTGCCCACCCTAGTCAT	**3129**
FANCF_2B_-78	GCGACTCTCTGCGTACTGAT	**3130**
FANCF_2B_-50	GCCCTACTTCCGCTTTCACCT	**3131**
FANCF_2B_-27	GGATTCCATGAGGTGCGCGA	**3132**
FANCF_2B_-17	GCTGCAGAAGGGATTCCATG	**3133**
FANCF_2B_+21	GCTTGAGACCGCCAGAAGCT	**3134**
FANCF_2B_+48	GGGGTCCCAGGTGCTGACGT	**3135**
HEK3_2B_-108	GCAGAAATAGACTAATTGCA	**3136**
HEK3_2B_-38	GGATTGACCCAGGCCAGGGC	**3137**
HEK3_2B_+26	GACGCCCTCTGGAGGAAGCA	**3138**
HEK3_2B_+37	GCTGTCCTGCGACGCCCTC	**3139**
HEK3_2B_+63	GCACATACTAGCCCCTGTCT	**3140**
HEK3_2B_+90	GTCAACCAGTATCCCGGTGC	**3141**
HEK4_2B_-95	TCCCTTCCTTCCACCCAGCC	**3142**
HEK4_2B_-51	CCCTGCCTGTCATCCTGCTT	**3143**
HEK4_2B_-26	GCAGTGCCACCGGGGCGCCG	**3144**
HEK4_2B_+52	GCGGGGGCTCAGAGAGGGCA	**3145**
HEK4_2B_+74	GAGACACACACACAGGCCTGG	**3146**
RNF2_2C_+41	GTCAACCATTAAGCAAAACAT	**3147**
RNF2_2C_4ATOC_+5	GTGAGTTACAACGAACACCGC	**3148**
RNF2_2C_4ATOG_+5	GTGAGTTACAACGAACACCCC	**3149**
FANCF_2C_+48	GGGGTCCCAGGTGCTGACGT	**3150**
FANCF_2C_5GTOT_+7	GAAGCTCGGAAAAGCGATCA	**3151**
FANCF_2C_7ATOC_+7	GAAGCTCGGAAAAGCGAGCC	**3152**
HEK3_2C_+90	GTCAACCAGTATCCCGGTGC	**3153**

[1398] Таблица 3D: PEgRNA на Фиг.41A-41K

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3154-3304)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3305-3455)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_4A_1TTOA	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_4A_1TTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATCGCGTGCTCAGTCTG	13	10
HEK3_4A_1TTOG	GGCCCAGACTGAGCACGTGA	TCTGCCATCCCGTGCTCAGTCTG	13	10
HEK3_4A_2GTOA	GGCCCAGACTGAGCACGTGA	TCTGCCATTACGTGCTCAGTCTG	13	10
HEK3_4A_2GTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATGACGTGCTCAGTCTG	13	10

HEK3_4A_2GTOT	GGCCCAGACTGAGCACGTGA	TCTGCCATAACGTGCTCAGTCTG	13	10
HEK3_4A_3ATOC	GGCCCAGACTGAGCACGTGA	TCTGCCAGCACGTGCTCAGTCTG	13	10
HEK3_4A_3ATOG	GGCCCAGACTGAGCACGTGA	TCTGCCACCACGTGCTCAGTCTG	13	10
HEK3_4A_3ATOT	GGCCCAGACTGAGCACGTGA	TCTGCCAACACGTGCTCAGTCTG	13	10
HEK3_4A_4TTOA	GGCCCAGACTGAGCACGTGA	TCTGCCTTCACGTGCTCAGTCTG	13	10
HEK3_4A_4TTOC	GGCCCAGACTGAGCACGTGA	TCTGCCGTCACGTGCTCAGTCTG	13	10
HEK3_4A_4TTOG	GGCCCAGACTGAGCACGTGA	TCTGCCCTCACGTGCTCAGTCTG	13	10
HEK3_4A_5GTOA	GGCCCAGACTGAGCACGTGA	TCTGCTATCACGTGCTCAGTCTG	13	10
HEK3_4A_5GTOC	GGCCCAGACTGAGCACGTGA	TCTGCGATCACGTGCTCAGTCTG	13	10
HEK3_4A_5GTOT	GGCCCAGACTGAGCACGTGA	TCTGCAATCACGTGCTCAGTCTG	13	10
HEK3_4A_6GTOA	GGCCCAGACTGAGCACGTGA	TCTGTCATCACGTGCTCAGTCTG	13	10
HEK3_4A_6GTOC	GGCCCAGACTGAGCACGTGA	TCTGGCATCACGTGCTCAGTCTG	13	10
HEK3_4A_6GTOT	GGCCCAGACTGAGCACGTGA	TCTGACATCACGTGCTCAGTCTG	13	10
HEK3_4A_7CTOA	GGCCCAGACTGAGCACGTGA	TCTTCCATCACGTGCTCAGTCTG	13	10
HEK3_4A_7CTOG	GGCCCAGACTGAGCACGTGA	TCTCCCATCACGTGCTCAGTCTG	13	10
HEK3_4A_7CTOT	GGCCCAGACTGAGCACGTGA	TCTACCATCACGTGCTCAGTCTG	13	10
HEK3_4A_8ATOC	GGCCCAGACTGAGCACGTGA	TCGGCCATCACGTGCTCAGTCTG	13	10
HEK3_4A_8ATOG	GGCCCAGACTGAGCACGTGA	TCCGCCATCACGTGCTCAGTCTG	13	10
HEK3_4A_8ATOT	GGCCCAGACTGAGCACGTGA	TCAGCCATCACGTGCTCAGTCTG	13	10
HEK3_4B_1TTOA	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCTCGTGCTCAGTCTG	13	34
HEK3_4B_12GTOC	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTGCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_14ATOT	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTATCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_	GGCCCAGACTGAG	TGGAGGAAGCAGGGCTT	13	34

17GTOC	CACGTGA	GCTTTCCTCTGCCATCAC GTGCTCAGTCTG		
HEK3_4B_ 20GTOC	GGCCCAGACTGAG CACGTGA	TGGAGGAAGCAGGGGTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
HEK3_4B_ 23CTOG	GGCCCAGACTGAG CACGTGA	TGGAGGAAGCACGGCTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
HEK3_4B_ 24TTOA	GGCCCAGACTGAG CACGTGA	TGGAGGAAGCTGGGCTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
HEK3_4B_ 26CTOG	GGCCCAGACTGAG CACGTGA	TGGAGGAACCAGGGCTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
HEK3_4B_ 30CTOG	GGCCCAGACTGAG CACGTGA	TGGACGAAGCAGGGCTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
HEK3_4B_ 33CTOG	GGCCCAGACTGAG CACGTGA	TCGAGGAAGCAGGGCTT CCTTTCCTCTGCCATCAC GTGCTCAGTCTG	13	34
RNF2_4C_ 1CTOA	GTCATCTTAGTCATT ACCTG	AACGAACACCTCATGTAA TGACTAAGATG	15	14
RNF2_4C_ 1CTOG	GTCATCTTAGTCATT ACCTG	AACGAACACCTCACGTA ATGACTAAGATG	15	14
RNF2_4C_ 1CTOT	GTCATCTTAGTCATT ACCTG	AACGAACACCTCAAGTA ATGACTAAGATG	15	14
RNF2_4C_ 2TTOA	GTCATCTTAGTCATT ACCTG	AACGAACACCTCTGGTAA TGACTAAGATG	15	14
RNF2_4C_ 2TTOG	GTCATCTTAGTCATT ACCTG	AACGAACACCTCCGGTA ATGACTAAGATG	15	14
RNF2_4C_ 3GTOC	GTCATCTTAGTCATT ACCTG	AACGAACACCTGAGGTA ATGACTAAGATG	15	14
RNF2_4C_ 4ATOC	GTCATCTTAGTCATT ACCTG	AACGAACACCGCAGGTA ATGACTAAGATG	15	14
RNF2_4C_ 4ATOT	GTCATCTTAGTCATT ACCTG	AACGAACACCACAGGTA ATGACTAAGATG	15	14
RNF2_4C_ 4ATOG	GTCATCTTAGTCATT ACCTG	AACGAACACCCCAGGTA ATGACTAAGATG	15	14
RNF2_4C_ 5GTOT	GTCATCTTAGTCATT ACCTG	AACGAACACATCAGGTA ATGACTAAGATG	15	14
RNF2_4C_ 6GTOA	GTCATCTTAGTCATT ACCTG	AACGAACATCTCAGGTAA TGACTAAGATG	15	14
RNF2_4C_ 7TTOC	GTCATCTTAGTCATT ACCTG	AACGAACGCCTCAGGTA ATGACTAAGATG	15	14
FANCF_4D _1ATOG	GGAATCCCTTCTGC AGCACC	GGAAAAGCGATCCAGGC GCTGCAGAAGGGAT	14	17
FANCF_4D _1ATOT	GGAATCCCTTCTGC AGCACC	GGAAAAGCGATCCAGGA GCTGCAGAAGGGAT	14	17

FANCF_4D_2CTOA	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGTTGCTGCAGAAGGGAT	14	17
FANCF_4D_3CTOG	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT	14	17
FANCF_4D_3CTOT	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT	14	17
FANCF_4D_4TTOA	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_4TTOG	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_5GTOA	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_6GTOC	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_7ATOC	GGAATCCCTTCTGCAGCACC	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_8TTOC	GGAATCCCTTCTGCAGCACC	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT	14	17
FANCF_4D_10GTOT	GGAATCCCTTCTGCAGCACC	GGAAAAGAGATCCAGGTGCTGCAGAAGGGAT	14	17
EMX1_4E_2ATOC	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA	14	16
EMX1_4E_2ATOT	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTACTTCTTCTGCTCGGA	14	16
EMX1_4E_3ATOG	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCCTCTTCTTCTGCTCGGA	14	16
EMX1_4E_4GTOC	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCGTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_5GTOA	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_5GTOT	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCACTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_7CTOA	GAGTCCGAGCAGAAGAAGAA	GTGATGGGATCCCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_8TTOA	GAGTCCGAGCAGAAGAAGAA	GTGATGGGTGCCCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_8TTOC	GAGTCCGAGCAGAAGAAGAA	GTGATGGGGGCCCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_8TTOG	GAGTCCGAGCAGAAGAAGAA	GTGATGGGCGCCCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_9CTOG	GAGTCCGAGCAGAAGAAGAA	GTGATGGCAGCCCTTCTTCTTCTGCTCGGA	14	16
EMX1_4E_9CTOT	GAGTCCGAGCAGAAGAAGAA	GTGATGGAAGCCCTTCTTCTTCTGCTCGGA	14	16
RUNX1_4F_1CTOA	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATCTCTTCCTCCTGAAAAT	15	15
RUNX1_4F_1CTOG	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATCCCTTCCTCCTGAAAAT	15	15

RUNX1_4F_1CTOT	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATCACTTCCTCCTGAAAAT	15	15
RUNX1_4F_2GTOA	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATTGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_3ATOC	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_3ATOG	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCACCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_3ATOT	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCAACGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_4TTOA	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCTTCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_4TTOC	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCGTCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_4TTOG	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCCTCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_5GTOT	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	15
RUNX1_4F_6GTOC	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGGCATCGCTTCCTCCTGAAAAT	15	15
VEGFA_4G_1TTOA	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCTCTTCTGGCCTGCAGA	13	22
VEGFA_4G_1TTOC	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCTCGTCTGGCCTGCAGA	13	22
VEGFA_4G_1TTOG	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCTCCTCTGGCCTGCAGA	13	22
VEGFA_4G_2GTOA	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCTTATCTGGCCTGCAGA	13	22
VEGFA_4G_3ATOC	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCGCATCTGGCCTGCAGA	13	22
VEGFA_4G_3ATOG	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCCCATCTGGCCTGCAGA	13	22
VEGFA_4G_3ATOT	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA	13	22
VEGFA_4G_5GTOT	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	22
VEGFA_4G_6GTOC	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGGCCTCATCTGGCCTGCAGA	13	22
VEGFA_4G_7CTOA	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA	13	22
VEGFA_4G_7CTOT	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAACCCTCATCTGGCCTGCAGA	13	22
VEGFA_4G_9CTOG	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGCAGCCCTCATCTGGCCTGCAGA	13	22
DNMT1_4H_1ATOC	GATTCCTGGTGCCAGAAACA	GTCACCCCTGGTTCTGGCACCAGG	13	11
DNMT1_4H_1ATOG	GATTCCTGGTGCCAGAAACA	GTCACCCCTGCTTCTGGCACCAGG	13	11

DNMT1_4H_2CTOA	GATTCCTGGTGCCAGAAACA	GTCACCCCTTTTTCTGGCACCAGG	13	11
DNMT1_4H_2CTOG	GATTCCTGGTGCCAGAAACA	GTCACCCCTCTTTCTGGCACCAGG	13	11
DNMT1_4H_2CTOT	GATTCCTGGTGCCAGAAACA	GTCACCCCTATTTCTGGCACCAGG	13	11
DNMT1_4H_3ATOT	GATTCCTGGTGCCAGAAACA	GTCACCCCAGTTTCTGGCACCAGG	13	11
DNMT1_4H_4GTOA	GATTCCTGGTGCCAGAAACA	GTCACCCTTGTTTCTGGCACCAGG	13	11
DNMT1_4H_5GTOT	GATTCCTGGTGCCAGAAACA	GTCACCACTGTTTCTGGCACCAGG	13	11
DNMT1_4H_6GTOC	GATTCCTGGTGCCAGAAACA	GTCACGCCTGTTTCTGGCACCAGG	13	11
DNMT1_4H_8TTOA	GATTCCTGGTGCCAGAAACA	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG	13	19
DNMT1_4H_8TTOC	GATTCCTGGTGCCAGAAACA	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG	13	19
DNMT1_4H_8TTOG	GATTCCTGGTGCCAGAAACA	GCCCTCCCGTCCCCCCTGTTTCTGGCACCAGG	13	19
HEK3_4J_DEL1-5	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCGTGCTCAGTCTG	13	29
HEK3_4J_DEL1-10	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCCGTGCTCAGTCTG	13	24
HEK3_4J_DEL1-15	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCCGTGCTCAGTCTG	13	19
HEK3_4J_DEL1-25	GGCCCAGACTGAGCACGTGA	TGTCCTGCGACGCCCTCTGGAGGAAGCGTGCTCAGTCTG	13	26
HEK3_4J_DEL1-30	GGCCCAGACTGAGCACGTGA	TGTCCTGCGACGCCCTCTGGACGTGCTCAGTCTG	13	21
HEK3_4J_DEL1-80	GGCCCAGACTGAGCACGTGA	AGTATCCCGGTGCAGGAGCTCGTGCTCAGTCTG	13	20
HEK3_4I_1AINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCATCGTGCTCAGTCTG	13	11
HEK3_4I_1CTTINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	13
HEK3_4I_1TDEL	GGCCCAGACTGAGCACGTGA	TCTGCCATCCGTGCTCAGTCTG	13	9
HEK3_4I_1-3TGADEL	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCACGTGCTCAGTCTG	13	31
RNF2_4I_1TINS	GTCATCTTAGTCATTACCTG	AACGAACACCTCAGAGTAATGACTAAGATG	15	15
RNF2_4I_1GTAINS	GTCATCTTAGTCATTACCTG	AACGAACACCTCAGTACGTAATGACTAAGATG	15	17

RNF2_4I_4ADEL	GTCATCTTAGTCATTACCTG	AACGAACACCCAGGTAATGACTAAGATG	15	13
RNF2_4I_3-5GAGDEL	GTCATCTTAGTCATTACCTG	AACGAACACAGGTAATGACTAAGATG	15	11
FANCF_4I_3CINS	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT	14	18
FANCF_4I_4GATINS	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT	14	20
FANCF_4I_6GDEL	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAGGTGCTGCAGAAGGGAT	14	16
FANCF_4I_5-7GGADEL	GGAATCCCTTCTGCAGCACC	GGAAAAGCGAAGGTGCTGCAGAAGGGAT	14	14
EMX1_4I_6TINS	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCACCTTCTTCTTCTGCTCGGA	14	17
EMX1_4I_1TGCINS	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA	14	19
EMX1_4I_5GDEL	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCTTCTTCTTCTGCTCGGA	14	15
EMX1_4I_4-6GGGDEL	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTTCTTCTTCTGCTCGGA	14	13
RUNX1_4I_1CINS	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATCGGCTTCCTCCTGAAAAT	15	16
RUNX1_4I_1ATGINS	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATCCATGCTTCCTCCTGAAAAT	15	18
RUNX1_4I_2GDEL	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCATGCTTCCTCCTGAAAAT	15	14
RUNX1_4I_2-4GATDEL	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCCGCTTCCTCCTGAAAAT	15	12
VEGFA_4I_4CINS	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCGCTCATCTGGCCTGCAGA	13	23
VEGFA_4I_2ACAINS	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCTTGTCATCTGGCCTGCAGA	13	25
VEGFA_4I_3ADEL	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCCCATCTGGCCTGCAGA	13	21
VEGFA_4I_2-4GAGDEL	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCCATCTGGCCTGCAGA	13	19
DNMT1_4I_4CINS	GATTCCTGGTGCCAGAAACA	TCCCGTCACCCGCTGTTTCTGGCACCAGG	13	16
DNMT1_4I_1TCAINS	GATTCCTGGTGCCAGAAACA	TCCCGTCACCCCTGTGATTTCTGGCACCAGG	13	18
DNMT1_4I_3ADEL	GATTCCTGGTGCCAGAAACA	TCCCGTCACCCCGTTTCTGGCACCAGG	13	14
DNMT1_4I_3-	GATTCCTGGTGCCAGAAACA	TCCCGTCACCGTTTCTGGCACCAGG	13	12

5AGGDEL				
HEK3_4K_1CTTINS_5GDEL	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCATCAAAGCGTGCTCAGTCTG	13	36
HEK3_4K_1CTTINS_2GTOC	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATGAAAGCGTGCTCAGTCTG	13	37
HEK3_4K_3TDEL_5GTOC	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCGATCCGTGCTCAGTCTG	13	33
HEK3_4K_3GTOC_6GTOT	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGACATGACGTGCTCAGTCTG	13	34
RNF2_4K_2AAINS_3-4GADEL	GTCATCTTAGTCATTACCTG	AACGAACACCATTGGTAATGACTAAGATG	15	14
RNF2_4K_1AINS_5GTOC	GTCATCTTAGTCATTACCTG	AACGAACACGTCAGTGTAATGACTAAGATG	15	15
RNF2_4K_1-2CTDEL_6GTOT	GTCATCTTAGTCATTACCTG	AACGAACAACTCGTAATGACTAAGATG	15	12
RNF2_4K_1CTOA_5GTOT	GTCATCTTAGTCATTACCTG	AACGAACACATCATGTAATGACTAAGATG	15	14
FANCF_4K_1TINS_3-4TGDEL	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCGGTAGCTGCAGAAGGGAT	14	16
FANCF_4K_1TINS_6GTOA	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATTCAGGTAGCTGCAGAAGGGAT	14	18
FANCF_4K_2CDEL_5GTOT	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGTGCTGCAGAAGGGAT	14	16

[1399]

[1400] Таблица 3E: Никирующая sgРНК на Фиг.41A-41K

НИКИРУЮЩАЯ SGPHK	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3456-3463)
HEK3_4A_+90	GTCAACCAGTATCCCGGTGC
HEK3_4B_+90	GTCAACCAGTATCCCGGTGC
RNF2_4C_+41	GTCAACCATTAAGCAAAACAT
FANCF_4D_+48	GGGGTCCCAGGTGCTGACGT
EMX1_4E_+53	GACATCGATGTCCTCCCCAT
RUNX1_4F_+38	GATGAAGCACTGTGGGTACGA
VEGFA_4G_+57	GATGTACAGAGAGCCCAGGGC

DNMT1_4H_+49	GCCCTTCAGCTAAAATAAAGG

[1401]

[1402] Таблица 3F: PEgRNA на Фиг.42A-42H

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3464-3478)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3479-3493)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_5A_C3	GGCCCAGACTGAGCACGTGA	TCTGTCATCACGTGCTCAGTCTG	13	10
HEK3_5A_C4	GGCCCAGACTGAGCACGTGA	TCTGCTATCACGTGCTCAGTCTG	13	10
HEK3_5A_C7	GGCCCAGACTGAGCACGTGA	TCTGCCATTACGTGCTCAGTCTG	13	10
FANCF_5A_C3	GGAATCCCTTCTGCAGCACC	GGAAAAGTGATCCAGGTGCTGCAGAAGGGAT	14	17
FANCF_5A_C7	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATTCAGGTGCTGCAGAAGGGAT	14	17
FANCF_5A_C8	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT	14	17
EMX1_5A_C5	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTCCTTCTTCTTCTGCTCGGA	14	16
EMX1_5A_C6	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA	14	16
EMX1_5A_C7	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCTTTCTTCTTCTGCTCGGA	14	16
EMX1_5C_C5_6	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTTCTTCTTCTTCTGCTCGGA	14	16
EMX1_5C_C5_7	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTCTTTCTTCTTCTGCTCGGA	14	16
EMX1_5C_C6_7	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCTTTTCTTCTTCTGCTCGGA	14	16
EMX1_5C_C5_6_7	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTTTTTCTTCTTCTGCTCGGA	14	16
HEK3_5D_A5	GGCCCAGACTGAGCACGTGA	TCTGCCGTCACGTGCTCAGTCTG	13	10
HEK3_5D_A8	GGCCCAGACTGAGCACGTGA	TCTGCCATCGCGTGCTCAGTCTG	13	10

[1403] Таблица 3G: Никирующая sgРНК на Фиг.42A-42H

НИКИРУЮЩАЯ SGРНК	ВОЗМОЖНАЯ СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 616-618)	ВОЗМОЖНАЯ СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 619-621)
HEK3_5A-F_+90	GTCAACCAGTATCCCGGTGC	GTCAACCAGTATCCCGGTGC
FANCF_5A-F_+48	GGGGTCCCAGGTGCTGACGT	GATGTACAGAGAGCCCAGGGC
EMX1_5A-	GATGTACAGAGAGCCCAGGG	GGGGTCCCAGGTGCTGACGT

F_+57	C	

[1404] Таблица 3H: sgРНК для редактирования оснований на Фиг.42A-42H

SGPHK ДЛЯ РЕДАКТИРОВАНИЯ ОСНОВАНИЙ	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
HEK3_5A-F_BE	GTGCCATCACGTGCTCAGTCT (SEQ ID NO: 455)
FANCF_5A-F_BE	GAGCGATCCAGGTGCTGCAGA (SEQ ID NO: 456)
EMX1_5A-F_BE	GGAGCCCTTCTTCTTCTGCT (SEQ ID NO: 455)

[1405] Таблица 3I: Целевая sgРНК на Фиг.42A-42H

ЦЕЛЕВАЯ SGPHK	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
HEK3_5G	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 510)
HEK4_5G	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 511)
EMX1_5G	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 512)
FANCF_5G	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 513)

[1406] Таблица 3J: Целевая PEgRNA на Фиг.42A-42H

ЦЕЛЕВАЯ PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 663-677)	3′ УДЛИНЕНИЕ (SEQ ID NO: 678-692)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_5G-H_PEGRNA_1	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_5G-H_PEGRNA_2	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	13
HEK3_5G-H_PEGRNA_3	GGCCCAGACTGAGCACGTGA	TCTGCCATCCGTGCTCAGTCTG	13	9
HEK3_5G-H_PEGRNA_4	GGCCCAGACTGAGCACGTGA	TCTGCGATCACGTGCTCAGTCTG	13	10
HEK4_5G-H_PEGRNA_1	GGCACTGCGGCTGGAGGTGG	TTAACGCCCACCTCCAGCC	9	10
HEK4_5G-H_PEGRNA_2	GGCACTGCGGCTGGAGGTGG	TTAACCCCCCCCTCCAGCC	9	10
HEK4_5G-	GGCACTGCGGCTGG	TTAACCCCTTACACCTCC	9	13

H_PEGRNA_3	AGGTGG	AGCC		
HEK4_5G-H_PEGRNA_4	GGCACTGCGGCTGGAGGTGG	TTAACCCCCCCTCCAGCC	9	9
EMX1_5G-H_PEGRNA_1	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCACTTCTTCTTCTGCTCGGA	14	16
EMX1_5G-H_PEGRNA_2	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA	14	16
EMX1_5G-H_PEGRNA_3	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA	14	19
EMX1_5G-H_PEGRNA_4	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTTCTTCTTCTGCTCGGA	14	13
FANCF_5G-H_PEGRNA_1	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT	14	17
FANCF_5G-H_PEGRNA_2	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT	14	17
FANCF_5G-H_PEGRNA_3	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT	14	20

[1407] Таблица 3K: PEgRNA на Фиг.49A-49B

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3494-3521)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3522-3540)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_6A_2GTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATGACGTGCTCAGTCTG	13	10
HEK3_6A_2GTOC	GGCCCAGACTGAGCACGTGA			
EMX1_6A_3GTOC	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCCCTTGTTCTTCTGCTCGG	13	13
EMX1_6A_3GTOC	GAGTCCGAGCAGAAGAAGAA			
FANCF_6A_5GTOT	GGAATCCCTTCTGCAGCACC	AAAAGCGATCAAGGTGCTGCAGAAGGGA	13	15
FANCF_6A_5GTOT	GGAATCCCTTCTGCAGCACC			
HEK3_6A_1HIS6INS	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATGATGGTGATGATGGTGCGTGCTCAGTCTG	13	52
HEK3_6A_1HIS6INS	GGCCCAGACTGAGCACGTGA			
HEK3_6A_5GTOT	GGCCCAGACTGAGCACGTGA	TCTGCAATCACGTGCTCAGTCTG	13	10
HEK3_6A_5GTOT	GGCCCAGACTGAGCACGTGA			
HEK3_6A_1C	GGCCCAGACTG	TCTGCCATCAAAGCGTG	13	10

TTINS	AGCACGTGA	CTCAGTCTG		
HEK3_6A_1C TTINS	GGCCCAGACTG AGCACGTGA			
*HBB*_6B_INS ALL	GCATGGTGCAC CTGACTCCTG	AGACTTCTCCACAGGAG TCAGGTGCAC	13	14
*HBB*_6B_INS ALL	GCATGGTGCAC CTGACTCCTG			
*HBB*_6B_COR RECT	GCATGGTGCAC CTGACTCCTG	AGACTTCTCCTCAGGAG TCAGGTGCAC	13	14
*HBB*_6B_COR RECT	GCATGGTGCAC CTGACTCCTG			
*HBB*_6B_COR RECT_W_SIL ENT	GCATGGTGCAC CTGACTCCTG	AGACTTCTCTTCAGGAG TCAGGTGCAC	13	14
*HBB*_6B_COR RECT_W_SIL ENT	GCATGGTGCAC CTGACTCCTG			
*HEXA*_6B_IN STALL	GTACCTGAACC GTATATCCTA	AGTCAGGGCCATAGGAT AGATATACGGTTC	12	14
*HEXA*_6B_CO RRECT	GATCCTTCCAGT CAGGGCCAT	ACCTGAACCGTATATCCT ATGGCCCTGACTG	10	21
*HEXA*_6B_CO RRECT_W_SI LENT	GATCCTTCCAGT CAGGGCCAT	GTACCTGAACCGTATATC TTATGGCCCTGACT	9	27
*PRNP*_6C	GCAGTGGTGGG GGGCCTTGG	ATGTAGACGCCAAGGCC CCCCACC	12	12
HEK3_6E-G_1TTOG	GGCCCAGACTG AGCACGTGA	TCTGCCATCCCGTGCTC AGTCTG	13	10
HEK3_6E-G_1CTTINS	GGCCCAGACTG AGCACGTGA	TCTGCCATCAAAGCGTG CTCAGTCTG	13	10
RNF2_6E-G_1CTOG	GTCATCTTAGTC ATTACCTG	AACGAACACCTCACGTA ATGACTAAGATG	15	14
*HBB*_6E-G_4ATOT	GCATGGTGCAC CTGACTCCTG	AGACTTCTCCACAGGAG TCAGGTGCAC	13	14
HEK3_6H_1H IS6INS	GGCCCAGACTG AGCACGTGA	TGGAGGAAGCAGGGCTT CCTTTCCTCTGCCATCAA TGATGGTGATGATGGTG CGTGCTCAGTCTG	13	52
HEK3_6H_1F LAGINS	GGCCCAGACTG AGCACGTGA	TGGAGGAAGCAGGGCTT CCTTTCCTCTGCCATCAC TTATCGTCGTCATCCTTG TAATCCGTGCTCAGTCT G	13	58

[1408] Таблица 3L: PEgRNA на Фиг. 47A-74D

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНО	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′ УДЛИНЕНИЯ (SEQ ID	ДЛИНА	ДЛИНА RT

	СТЬ (SEQ ID NO: 3541-3547)	NO: 3549-3556)	PBS (NT)	МАТР ИЦЫ (NT)
HEK3_ED4B_1 TDEL	GGCCCAGACTGAGC ACGTGA	TCTGCCATCCGTGCTCAG TCTG	13	9
HEK3_ED4B_1 AINS	GGCCCAGACTGAGC ACGTGA	TCTGCCATCATCGTGCTC AGTCTG	13	11
HEK3_ED4B_1 CTTINS	GGCCCAGACTGAGC ACGTGA	TCTGCCATCAAAGCGTGC TCAGTCTG	13	13
HEK3_ED4C_2 GTOC	GGCCCAGACTGAGC ACGTGA	TCTGCCATGACGTGCTCA GTCTG	13	10
HEK3_ED4D_1 FLAGINS	GGCCCAGACTGAGC ACGTGA	TGGAGGAAGCAGGGCTT CCTTTCCTCTGCCATCAC TTATCGTCGTCATCCTTGT AATCCGTGCTCAGTCTG	13	58
RNF2_ED4E_1 CTOA	GTCATCTTAGTCATT ACCTG	AACGAACACCTCATGTAA TGACTAAGATG	15	14
EMX1_ED4F_1 GTOC	GAGTCCGAGCAGAA GAAGAA	ATGGGAGCCCTTGTTCTT CTGCTCGG	13	13
*HBB*_ED4G_2T TOA	GTAACGGCAGACTT CTCCTC	ATCTGACTCCTGTGGAGA AGTCTGCC	12	14

[1409] Таблица 3М: PEgRNA на Фиг.48A-48C

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3557-3627)	ПОСЛЕДОВАТЕЛЬНОСТЬ 3' УДЛИНЕНИЯ (SEQ ID NO: 3628-3698)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
VEGFA_ED 5A_31	GATGTCTGCAGG CCAGATGA	CCCTCTGACAATGTGCCAT CTGGAGCACTCATCTGGCC TGCAGA	13	31
VEGFA_ED 5A_30	GATGTCTGCAGG CCAGATGA	CCTCTGACAATGTGCCATCT GGAGCACTCATCTGGCCTG CAGA	13	30
VEGFA_ED 5A_29	GATGTCTGCAGG CCAGATGA	CTCTGACAATGTGCCATCTG GAGCACTCATCTGGCCTGC AGA	13	29
VEGFA_ED 5A_28	GATGTCTGCAGG CCAGATGA	TCTGACAATGTGCCATCTG GAGCACTCATCTGGCCTGC AGA	13	28
VEGFA_ED 5A_27	GATGTCTGCAGG CCAGATGA	CTGACAATGTGCCATCTGG AGCACTCATCTGGCCTGCA GA	13	27
VEGFA_ED	GATGTCTGCAGG	TGACAATGTGCCATCTGGA	13	26

5A_26	CCAGATGA	GCACTCATCTGGCCTGCAGA		
VEGFA_ED5A_25	GATGTCTGCAGGCCAGATGA	GACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	25
VEGFA_ED5A_24	GATGTCTGCAGGCCAGATGA	ACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	24
VEGFA_ED5A_23	GATGTCTGCAGGCCAGATGA	CAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	23
VEGFA_ED5A_22	GATGTCTGCAGGCCAGATGA	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	22
VEGFA_ED5A_21	GATGTCTGCAGGCCAGATGA	ATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	21
VEGFA_ED5A_20	GATGTCTGCAGGCCAGATGA	TGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	20
VEGFA_ED5A_19	GATGTCTGCAGGCCAGATGA	GTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	19
VEGFA_ED5A_18	GATGTCTGCAGGCCAGATGA	TGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	18
VEGFA_ED5A_17	GATGTCTGCAGGCCAGATGA	GCCATCTGGAGCACTCATCTGGCCTGCAGA	13	17
VEGFA_ED5A_16	GATGTCTGCAGGCCAGATGA	CCATCTGGAGCACTCATCTGGCCTGCAGA	13	16
VEGFA_ED5A_15	GATGTCTGCAGGCCAGATGA	CATCTGGAGCACTCATCTGGCCTGCAGA	13	15
VEGFA_ED5A_14	GATGTCTGCAGGCCAGATGA	ATCTGGAGCACTCATCTGGCCTGCAGA	13	14
VEGFA_ED5A_13	GATGTCTGCAGGCCAGATGA	TCTGGAGCACTCATCTGGCCTGCAGA	13	13
VEGFA_ED5A_12	GATGTCTGCAGGCCAGATGA	CTGGAGCACTCATCTGGCCTGCAGA	13	12
VEGFA_ED5A_11	GATGTCTGCAGGCCAGATGA	TGGAGCACTCATCTGGCCTGCAGA	13	11
VEGFA_ED5A_10	GATGTCTGCAGGCCAGATGA	GGAGCACTCATCTGGCCTGCAGA	13	10
VEGFA_ED5A_9	GATGTCTGCAGGCCAGATGA	GAGCACTCATCTGGCCTGCAGA	13	9
VEGFA_ED5A_8	GATGTCTGCAGGCCAGATGA	AGCACTCATCTGGCCTGCAGA	13	8

DNMT1_ED 5B_31	GATTCCTGGTGC CAGAAACA	AGGACTAGTTCTGCCCTCC CGTCACCACTGTTTCTGGC ACCAGG	13	31
DNMT1_ED 5B_30	GATTCCTGGTGC CAGAAACA	GGACTAGTTCTGCCCTCCC GTCACCACTGTTTCTGGCA CCAGG	13	30
DNMT1_ED 5B_29	GATTCCTGGTGC CAGAAACA	GACTAGTTCTGCCCTCCCG TCACCACTGTTTCTGGCAC CAGG	13	29
DNMT1_ED 5B_28	GATTCCTGGTGC CAGAAACA	ACTAGTTCTGCCCTCCCGT CACCACTGTTTCTGGCACC AGG	13	28
DNMT1_ED 5B_27	GATTCCTGGTGC CAGAAACA	CTAGTTCTGCCCTCCCGTC ACCACTGTTTCTGGCACCA GG	13	27
DNMT1_ED 5B_26	GATTCCTGGTGC CAGAAACA	TAGTTCTGCCCTCCCGTCA CCACTGTTTCTGGCACCAG G	13	26
DNMT1_ED 5B_25	GATTCCTGGTGC CAGAAACA	AGTTCTGCCCTCCCGTCAC CACTGTTTCTGGCACCAGG	13	25
DNMT1_ED 5B_24	GATTCCTGGTGC CAGAAACA	GTTCTGCCCTCCCGTCACC ACTGTTTCTGGCACCAGG	13	24
DNMT1_ED 5B_23	GATTCCTGGTGC CAGAAACA	TTCTGCCCTCCCGTCACCA CTGTTTCTGGCACCAGG	13	23
DNMT1_ED 5B_22	GATTCCTGGTGC CAGAAACA	TCTGCCCTCCCGTCACCAC TGTTTCTGGCACCAGG	13	22
DNMT1_ED 5B_21	GATTCCTGGTGC CAGAAACA	CTGCCCTCCCGTCACCACT GTTTCTGGCACCAGG	13	21
DNMT1_ED 5B_20	GATTCCTGGTGC CAGAAACA	TGCCCTCCCGTCACCACTG TTTCTGGCACCAGG	13	20
DNMT1_ED 5B_19	GATTCCTGGTGC CAGAAACA	GCCCTCCCGTCACCACTGT TTCTGGCACCAGG	13	19
DNMT1_ED 5B_18	GATTCCTGGTGC CAGAAACA	CCCTCCCGTCACCACTGTT TCTGGCACCAGG	13	18
DNMT1_ED 5B_17	GATTCCTGGTGC CAGAAACA	CCTCCCGTCACCACTGTTT CTGGCACCAGG	13	17
DNMT1_ED 5B_16	GATTCCTGGTGC CAGAAACA	CTCCCGTCACCACTGTTTCT GGCACCAGG	13	16
DNMT1_ED 5B_15	GATTCCTGGTGC CAGAAACA	TCCCGTCACCACTGTTTCT GGCACCAGG	13	15

DNMT1_ED 5B_14	GATTCCTGGTGC CAGAAACA	CCCGTCACCACTGTTTCTG GCACCAGG	13	14
DNMT1_ED 5B_13	GATTCCTGGTGC CAGAAACA	CCGTCACCACTGTTTCTGG CACCAGG	13	13
DNMT1_ED 5B_12	GATTCCTGGTGC CAGAAACA	CGTCACCACTGTTTCTGGC ACCAGG	13	12
DNMT1_ED 5B_11	GATTCCTGGTGC CAGAAACA	GTCACCACTGTTTCTGGCA CCAGG	13	11
DNMT1_ED 5B_10	GATTCCTGGTGC CAGAAACA	TCACCACTGTTTCTGGCAC CAGG	13	10
DNMT1_ED 5B_9	GATTCCTGGTGC CAGAAACA	CACCACTGTTTCTGGCACC AGG	13	9
DNMT1_ED 5B_8	GATTCCTGGTGC CAGAAACA	ACCACTGTTTCTGGCACCA GG	13	8
RUNX1_ED 5C_31	GCATTTTCAGGA GGAAGCGA	AATGACTCAAATATGCTGTC TGAAGCAATCGCTTCCTCCT GAAAAT	15	31
RUNX1_ED 5C_30	GCATTTTCAGGA GGAAGCGA	ATGACTCAAATATGCTGTCT GAAGCAATCGCTTCCTCCT GAAAAT	15	30
RUNX1_ED 5C_29	GCATTTTCAGGA GGAAGCGA	TGACTCAAATATGCTGTCTG AAGCAATCGCTTCCTCCTG AAAAT	15	29
RUNX1_ED 5C_28	GCATTTTCAGGA GGAAGCGA	GACTCAAATATGCTGTCTGA AGCAATCGCTTCCTCCTGA AAAT	15	28
RUNX1_ED 5C_27	GCATTTTCAGGA GGAAGCGA	ACTCAAATATGCTGTCTGAA GCAATCGCTTCCTCCTGAA AAT	15	27
RUNX1_ED 5C_26	GCATTTTCAGGA GGAAGCGA	CTCAAATATGCTGTCTGAAG CAATCGCTTCCTCCTGAAAA T	15	26
RUNX1_ED 5C_25	GCATTTTCAGGA GGAAGCGA	TCAAATATGCTGTCTGAAGC AATCGCTTCCTCCTGAAAAT	15	25
RUNX1_ED 5C_24	GCATTTTCAGGA GGAAGCGA	CAAATATGCTGTCTGAAGCA ATCGCTTCCTCCTGAAAAT	15	24
RUNX1_ED 5C_23	GCATTTTCAGGA GGAAGCGA	AAATATGCTGTCTGAAGCAA TCGCTTCCTCCTGAAAAT	15	23
RUNX1_ED 5C_22	GCATTTTCAGGA GGAAGCGA	AATATGCTGTCTGAAGCAAT CGCTTCCTCCTGAAAAT	15	22

RUNX1_ED5C_21	GCATTTTCAGGAGGAAGCGA	ATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	21
RUNX1_ED5C_20	GCATTTTCAGGAGGAAGCGA	TATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	20
RUNX1_ED5C_19	GCATTTTCAGGAGGAAGCGA	ATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	19
RUNX1_ED5C_18	GCATTTTCAGGAGGAAGCGA	TGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	18
RUNX1_ED5C_17	GCATTTTCAGGAGGAAGCGA	GCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	17
RUNX1_ED5C_16	GCATTTTCAGGAGGAAGCGA	CTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	16
RUNX1_ED5C_15	GCATTTTCAGGAGGAAGCGA	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	15
RUNX1_ED5C_14	GCATTTTCAGGAGGAAGCGA	GTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	14
RUNX1_ED5C_13	GCATTTTCAGGAGGAAGCGA	TCTGAAGCAATCGCTTCCTCCTGAAAAT	15	13
RUNX1_ED5C_12	GCATTTTCAGGAGGAAGCGA	CTGAAGCAATCGCTTCCTCCTGAAAAT	15	12
RUNX1_ED5C_11	GCATTTTCAGGAGGAAGCGA	TGAAGCAATCGCTTCCTCCTGAAAAT	15	11
RUNX1_ED5C_10	GCATTTTCAGGAGGAAGCGA	GAAGCAATCGCTTCCTCCTGAAAAT	15	10
RUNX1_ED5C_9	GCATTTTCAGGAGGAAGCGA	AAGCAATCGCTTCCTCCTGAAAAT	15	9

[1410] Таблица 3N: PEgRNA на Фиг.48A-48C

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′ УДЛИНЕНИЯ	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_ED6_5GTOA	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 393)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 394)	13	10

[1411] Таблица 3O: Никирующая sgРНК на Фиг.48A-48C

НИКИРУЮЩАЯ SGPHK	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ

HEK3_ED6_+63	GCACATACTAGCCCCTGTCT (SEQ ID NO: 395)

[1412] Таблица 3P: PEgRNA на Фиг.50A-50B

PEGRNA	СПЕЙСЕР (SEQ ID NO: 3699-3754)	3′ УДЛИНЕНИЕ (ОТ 5′ К 3′)(SEQ ID NO: 3755-3810)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HBB 3.5	GTAACGGCAGACTTCTCCAC	AGACTTCTCCTCAGGAGTCAGGTGCAC	12	14
HBB 3.7	GCATGGTGCACCTGACTCCTG	AGACTTCTCTTCAGGAGTCAGGTGCAC	13	14
HBB 5.2	GCATGGTGCACCTGACTCCTG	TAACGGCAGACTTCTCCTCAGGAGTCAGGTGCAC	13	19
HBB 5.3	GCATGGTGCACCTGACTCCTG	ACGGCAGACTTCTCCTCAGGAGTCAGGTGCAC	13	17
HBB 5.4	GCATGGTGCACCTGACTCCTG	GGCAGACTTCTCCTCAGGAGTCAGGTGCAC	13	16
HBB 5.5	GCATGGTGCACCTGACTCCTG	GCAGACTTCTCCTCAGGAGTCAGGTGCAC	13	13
HBB 5.6	GCATGGTGCACCTGACTCCTG	GACTTCTCCTCAGGAGTCAGGTGCAC	13	12
HBB 5.7	GCATGGTGCACCTGACTCCTG	ACTTCTCCTCAGGAGTCAGGTGCAC	13	21
HBB 5.8	GCATGGTGCACCTGACTCCTG	TAACGGCAGACTTCTCCTCAGGAGTCAGGTGCA	12	19
HBB 5.9	GCATGGTGCACCTGACTCCTG	ACGGCAGACTTCTCCTCAGGAGTCAGGTGCA	12	17
HBB 5.10	GCATGGTGCACCTGACTCCTG	GGCAGACTTCTCCTCAGGAGTCAGGTGCA	12	16
HBB 5.11	GCATGGTGCACCTGACTCCTG	GCAGACTTCTCCTCAGGAGTCAGGTGCA	12	13
HBB 5.12	GCATGGTGCACCTGACTCCTG	GACTTCTCCTCAGGAGTCAGGTGCA	12	12
HBB 5.13	GCATGGTGCACCTGACTCCTG	ACTTCTCCTCAGGAGTCAGGTGCA	12	14
HEXA S 1	ATCCTTCCAGTCAGGGCCAT	ATATCTTATGGCCCTGACTGGAA	13	14
HEXA S 2	ATCCTTCCAGTCAGGGCCAT	TATATCTTATGGCCCTGACTGGAA	13	15
HEXA S 3	ATCCTTCCAGTCAGGGCCAT	GTATATCTTATGGCCCTGACTGGAA	13	16
HEXA S 4	ATCCTTCCAGTCAGGGCCAT	ACCGTATATCTTATGGCCCTGACTGGAA	13	19
HEXA S 5	ATCCTTCCAGTCAGGGCCAT	AACCGTATATCTTATGGCCCTGACTGGAA	13	20

HEXA S 6	ATCCTTCCAGTCA GGGCCAT	GAACCGTATATCTTATGGCCCTGAC TGGAA	13	21
HEXA S 7	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGGAA	13	22
HEXA S 8	ATCCTTCCAGTCA GGGCCAT	ATATCTTATGGCCCTGACT	9	14
HEXA S 9	ATCCTTCCAGTCA GGGCCAT	TATATCTTATGGCCCTGACT	9	15
HEXA S 10	ATCCTTCCAGTCA GGGCCAT	GTATATCTTATGGCCCTGACT	9	16
HEXA S 11	ATCCTTCCAGTCA GGGCCAT	ACCGTATATCTTATGGCCCTGACT	9	19
HEXA S 12	ATCCTTCCAGTCA GGGCCAT	AACCGTATATCTTATGGCCCTGACT	9	20
HEXA S 13	ATCCTTCCAGTCA GGGCCAT	GAACCGTATATCTTATGGCCCTGAC T	9	21
HEXA S 14	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CT	9	22
HEXA S 15	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA C	8	22
HEXA S 16	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTG	10	22
HEXA S 17	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGG	11	22
HEXA S 18	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGGA	12	22
HEXA S 19	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGGAA	13	22
HEXA S 20	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGGAAG	14	22
HEXA S 21	ATCCTTCCAGTCA GGGCCAT	TGAACCGTATATCTTATGGCCCTGA CTGGAAGG	15	22
HEXA S 22	ATCCTTCCAGTCA GGGCCAT	ACCTGAACCGTATATCTTATGGCCC TGACT	9	25
HEXA S 23	ATCCTTCCAGTCA GGGCCAT	TACCTGAACCGTATATCTTATGGCC CTGACT	9	26
HEXA S 24	ATCCTTCCAGTCA GGGCCAT	GTACCTGAACCGTATATCTTATGGC CCTGACT	9	27
HEXA S 25	ATCCTTCCAGTCA GGGCCAT	GGTACCTGAACCGTATATCTTATGG CCCTGACT	9	28
HEXA S 26	ATCCTTCCAGTCA GGGCCAT	TGGTACCTGAACCGTATATCTTATG GCCCTGACT	9	29
HEXA 5	ATCCTTCCAGTCA GGGCCAT	ACCTGAACCGTATATCCTATGGCCC TGACTGGAA	13	21
HEXA 6	ATCCTTCCAGTCA GGGCCAT	ACCGTATATCCTATGGCCCTGACTG GAA	13	15
HEXA 7	ATCCTTCCAGTCA GGGCCAT	ACCTGAACCGTATATCCTATGGCCC TGACTGGAAGG	15	21

HEXA 8	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTGGAAG	14	21
HEXA 9	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTGGA	12	21
HEXA 10	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTGG	11	21
HEXA 11	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTG	10	21
HEXA 12	ATCCTTCCAGTCAGGGCCAT	AACCGTATATCCTATGGCCCTGACTGGAA	13	16
HEXA 13	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCCTATGGCCCTGACTGGAA	13	18
HEXA 14	ATCCTTCCAGTCAGGGCCAT	TACCTGAACCGTATATCCTATGGCCCTGACTGGAA	13	22
HEXA 15	ATCCTTCCAGTCAGGGCCAT	TGGTACCTGAACCGTATATCCTATGGCCCTGACTGGAA	13	25
HEXA 16	ATCCTTCCAGTCAGGGCCAT	GTACCTGAACCGTATATCCTATGGCCCTGACTGGAA	13	23
HEXA 17	ATCCTTCCAGTCAGGGCCAT	AACCGTATATCCTATGGCCCTGACTG	10	16
HEXA 18	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCCTATGGCCCTGACTG	10	18
HEXA 19	ATCCTTCCAGTCAGGGCCAT	TACCTGAACCGTATATCCTATGGCCCTGACTG	10	22
HEXA 20	ATCCTTCCAGTCAGGGCCAT	TGGTACCTGAACCGTATATCCTATGGCCCTGACTG	10	25

[1413] Таблица 3Q: никирующая sgРНК на Фиг.50A-50B

НИКИРУЮЩАЯ SGРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
*HBB*_ED7A_+72	GCCTTGATACCAACCTGCCCA (SEQ ID NO: 626)
*HEXA*_ED7B_+60	GCTGGAACTGGTCACCAAGGC(SEQ ID NO: 627)
*HEXA*_ED7B_CORRECT_WT_PE3B	GTACCTGAACCGTATATCCTA(SEQ ID NO: 628)
*HEXA*_ED7B_CORRECT_SILENT_PE3B	GTACCTGAACCGTATATCTTA(SEQ ID NO: 629)

[1414] Таблица 3R: PEgRNA на Фиг.51A-51F

PEGRNA	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 632-640)	3′ УДЛИНЕНИЕ (SEQ ID NO: 641-649)	ДЛИНА PBS (NT)	ДЛИНА RT МАТРИЦЫ (NT)
HEK3_ED8_1TTOG	GGCCCAGACTGAGCACGTGA	TCTGCCATCCCGTGCTCAGTCTG	13	10

HEK3_ED8_3ATOC	GGCCCAGACTGAGCACGTGA	TCTGCCAGCACGTGCTCAGTCTG	13	10
HEK3_ED8_3ATOT	GGCCCAGACTGAGCACGTGA	TCTGCCAACACGTGCTCAGTCTG	13	10
HEK3_ED8_3ATOT_5-6GGTOTT	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGAAAACACGTGCTCAGTCTG	13	34
HEK3_ED8_1CTTINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	10
RNF2_ED8_1CTOA	GTCATCTTAGTCATTACCTG	AACGAACACCTCATGTAATGACTAAGATG	15	14
RNF2_ED8_1CTOG	GTCATCTTAGTCATTACCTG	AACGAACACCTCACGTAATGACTAAGATG	15	14
RNF2_ED8_1GTAINS	GTCATCTTAGTCATTACCTG	AACGAACACCTCAGTACGTAATGACTAAGATG	15	17
*HBB*_ED8_4ATOT	GCATGGTGCACCTGACTCCTG	AGACTTCTCCACAGGAGTCAGGTGCAC	13	14

[1415] Таблица 4: Последовательности праймеров, использованных для амплификации геномной ДНК клеток млекопитающих и HTS[181].

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3811-3863)
HEK3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATGTGGGCTGCCTAGAAAGG
HEK3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCAGCCAAACTTGTCAACC
RNF2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACGTCTCATATGCCCCTTGG
RNF2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGTAGGAATTTTGGTGGGACA
HEK4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAACCCAGGTAGCCAGAGAC
HEK4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCTTTCAACCCGAACGGAG
EMX1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGCTCAGCCTGAGTGTTGA
EMX1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCGTGGGTTTGTGGTTGC
FANCF FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCATTGCAGAGAGGCGTATCA
FANCF REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGTCCCAGGTGCTGAC
HBB FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGGGTTGGCCAATCTACTCCC
HBB REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTTCTCTGTCTCCACATGCC
PRNP FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTCAGTGGAACAAGCCGAGT

PRNP REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTTGGTTGGGGT AACGGTG
HEXA FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCATAC AGGTGTGGCGAGAGG
HEXA REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCAGCCTCCTTTG GTTAGCA
RUNX1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCACA AACAAGACAGGGAACTG
RUNX1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGATGTAGGGCTA GAGGGGTG
VEGFA FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACTTG GTGCCAAATTCTTCTCC
VEGFA REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTAAAGAGGGAATG GGCTTTGGA
DNMT FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCACAA CAGCTTCATGTCAGCC
DNMT REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGTTAATGTTTC CTGATGGTCC
HEK3 OFF-TARGET SITE 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCCC TGTTGACCTGGAGAA
HEK3 OFF-TARGET SITE 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTACTTGCC CTGACCA
HEK3 OFF-TARGET SITE 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTGGT GTTGACAGGGAGCAA
HEK3 OFF-TARGET SITE 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGATGTGGGC AGAAGGG
HEK3 OFF-TARGET SITE 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAGA GGGAACAGAAGGGCT
HEK3 OFF-TARGET SITE 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCCAAAGGCCC AAGAACCT
HEK3 OFF-TARGET SITE 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCTA GCACTTTGGAAGGTCG
HEK3 OFF-TARGET SITE 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCATCTTAATCT GCTCAGCC
HEK4 OFF-TARGET SITE 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGCAT GGCTTCTGAGACTCA
HEK4 OFF-TARGET SITE 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTCCCTTGCAC TCCCTGTCTTT
HEK4 OFF-TARGET SITE 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTGG CAATGGAGGCATTGG
HEK4 OFF-TARGET SITE 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAAGAGGCTGCC CATGAGAG
HEK4 OFF-TARGET SITE 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGTCT GAGGCTCGAATCCTG
HEK4 OFF-TARGET SITE 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGTGGCCTCCAT ATCCCTG
HEK4 OFF-TARGET SITE 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTCC ACCAGAACTCAGCCC

HEK4 OFF-TARGET SITE 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTCGGTTCCTCC ACAACAC
EMX1 OFF-TARGET SITE 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTGGG GAGATTTGCATCTGTGGAGG
EMX1 OFF-TARGET SITE 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTTTTATACCATC TTGGGGTTACAG
EMX1 OFF-TARGET SITE 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAATG TGCTTCAACCCATCACGGC
EMX1 OFF-TARGET SITE 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCATGAATTTGTG ATGGATGCAGTCTG
EMX1 OFF-TARGET SITE 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAGAA GGAGGTGCAGGAGCTAGAC
EMX1 OFF-TARGET SITE 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCATCCCGACCTTC ATCCCTCCTGG
EMX1 OFF-TARGET SITE 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTAGT TCTGACATTCCTCCTGAGGG
EMX1 OFF-TARGET SITE 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCAAACAAGGTG CAGATACAGCA
FANCF OFF-TARGET SITE 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCGGG CAGTGGCGTCTTAGTCG
FANCF OFF-TARGET SITE 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCTGGGTTTGGT TGGCTGCTC
FANCF OFF-TARGET SITE 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTCCT TGCCGCCCAGCCGGTC
FANCF OFF-TARGET SITE 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGGGGAAGA GGCGAGGACAC
FANCF OFF-TARGET SITE 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCAGT GTTTCCCATCCCCAACAC
FANCF OFF-TARGET SITE 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAATGGATCCCCC CCTAGAGCTC
FANCF OFF-TARGET SITE 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGGC CCACAGGTCCTTCTGGA

[1416] Таблица 5: Последовательности одноцепочечных донорских матриц в виде олигонуклеотидов ДНК со 100 повторяющимися единицами, использованные в рамках экспериментов на основе HDR и при создании линии клеток HEK293T *HBB* E6V. Олигонуклеотиды имеют длину 100-103 nt с плечами гомологии вокруг участка редактирования. Огинонуклеотиды были получены от Integrated DNA Technologies, очищены с помощью PAGE.

HEK3 +3 A HA T	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAG ACTGAGCACGTGTTGGCAGAGGAAAGGAAGCCCTGCTTC CTCCAGAGGGCGTCGCAGGAC (SEQ ID NO: 717)
HEK3 +3 A HA T, +5,6 GG HA TT	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAG ACTGAGCACGTGTTTTCAGAGGAAAGGAAGCCCTGCTTC CTCCAGAGGGCGTCGCAGGAC (SEQ ID NO: 718)
HEK3 +1 T HA G	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAG ACTGAGCACGGGATGGCAGAGGAAAGGAAGCCCTGCTTC CTCCAGAGGGCGTCGCAGGAC (SEQ ID NO: 719)
HEK3 +3 A HA C	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAG ACTGAGCACGTGCTGGCAGAGGAAAGGAAGCCCTGCTTC CTCCAGAGGGCGTCGCAGGAC (SEQ ID NO: 720)
HEK3 +1, ВСТАВКА CTT	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAG ACTGAGCACGCTTTGATGGCAGAGGAAAGGAAGCCCTGC TTCCTCCAGAGGGCGTCGCAGGAC (SEQ ID NO: 721)
RNF2 +1 C HA A	CCCAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTA GTCATTACATGAGGTGTTCGTTGTAACTCATATAAACTGAG TTCCCATGTTTTGCTTAA (SEQ ID NO: 722)
RNF2 +1 C HA G	CCCAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTA GTCATTACGTGAGGTGTTCGTTGTAACTCATATAAACTGAG TTCCCATGTTTTGCTTAA (SEQ ID NO: 723)
RNF2 +1, ВСТАВКА GTA	CAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTAGT CATTACGTACTGAGGTGTTCGTTGTAACTCATATAAACTGA GTTCCCATGTTTTGCTTA (SEQ ID NO: 724)
ВНЕСЕНИЕ *HBB* E6V (ТАКЖЕ ИСПОЛЬЗУЕТСЯ ДЛЯ СОЗДАНИИ ЛИНИИ КЛЕТОК HEK293T *HBB* E6V)	ACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGG CAGACTTCTCCACAGGAGTCAGATGCACCATGGTGTCTGT TTGAGGTTGCTAGTGAACAC (SEQ ID NO: 725)
КОРРЕКТИРУЮЩИЙ ПРОТОСПЕЙСЕР A *HBB* E6V	ACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGG CAGACTTCTCCTCAGGAGTCAGGTGCACCATGGTGTCTGT TTGAGGTTGCTAGTGAACAC (SEQ ID NO: 726)
КОРРЕКТИРУЮЩИЙ ПРОТОСПЕЙСЕР B *HBB* E6V	GTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCT GACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGC AAGGTGAACGTGGATGAAGT (SEQ ID NO: 727)
КОРРЕКТИРУЮЩИЙ ПРОТОСПЕЙСЕР B *HBB* E6V, МОЛЧАЩАЯ	GTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCT GACTCCTGATGAGAAGTCTGCCGTTACTGCCCTGTGGGGC AAGGTGAACGTGGATGAAGT (SEQ ID NO: 728)

МУТАЦИЯ PAM MUTATION	
G127V *PRNP*	CACATGGCTGGTGCTGCAGCAGCTGGGGCAGTGGTGGGG GGCCTTGGCGTCTACATGCTGGGAAGTGCCATGAGCAGGC CCATCATACATTTCGGCAGTG (SEQ ID NO: 729)

[1417] Дополнительные последовательности

[1418] Использованные здесь последовательности двойных флуоресцентных репортерных плазмид дрожжей

p425-GFP_stop_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGAT
GGTGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTAC
TTACGGTAAATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATG
GCCAACCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGA
TCATATGAAACAACATGACTTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAG
AAAGAACTATTTTTTTCAAAGATGACGGTAACTACAAGACCAGAGCTGAAGTCAA
GTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAAGGTATTGATTTTAAAGA
AGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTCACAATGTTTA
CATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATT
GGTGATGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTA
TCCAAAGATCCAAACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTG
CTGCTGGTATTACCCATGGTATGGATGAATTGTACAAAGCTAGCAACCTGGGTCAA
TCCTTGG**GGCCCAGACTGAGCACGTGATGG**CAGAGCACAGGAGACGTC*ATGGT*
TTCAAAAGGTGAAGAAGATAATATGGCTATTATTAAAGAATTTATGAGATTTAAAGTTCATA
TGGAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAAGGTGAAGGTAGACCA
TATGAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACCATTTGCTT
GGGATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTGA
TATTCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAATT
TTGAAGATGGTGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTAT
TTATAAAGTTAAATTGAGAGGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAAA
ACTATGGGTTGGGAAGCTTCATCAGAAAGAATGTATCCAGAAGATGGTGCTTTAAAAGG
TGAAATTAAACAAAGATTGAAATTAAAAGATGGTGGTCATTATGATGCTGAAGTTAAAACT
ACTTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTATAATGTTAATATTAAATTGGA
TATTACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAGCTGAAGGTAGAC
ATTCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATTGATT
*GA*GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTTG
CATTTAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTT
AATGACATTTTCGATTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAG
CTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTC
ATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGTAACTCACATT
AATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTG
CATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT
CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT
ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA

GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC
CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA
TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG
CTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTG
GTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC
CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC
GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGT
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGT
CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTT
CTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA
CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGG
GTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCA
TTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTG
AGCTGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAA
GAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTC
AAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCT
ATTTTACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCATCCCGAGA
GCGCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTAT
AATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGC
TACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTT
ACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATT
ATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGA
TGATTCTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACT
ACGTATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCT
TACTACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAACATAAAAAATGTAGAG
GTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGG
GATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGC
GGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAA

AGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTCCTATACTTT
CTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGC
TTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCT
ATATCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTT
ATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTAC
CTCCTGTGATATTATCCCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCC
TTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCTCATCCTTCAATGCTA
TCATTTCCTTTGATATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTAT
AAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTG
AAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGA
TGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCG
GGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCG
ACTACGTCGTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCAT
TATGGGAAATGCTTCAAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCAT
TGAGTGTTTTTTATTTGTTGTATTTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAA
TTAGGAATCGTAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTGCCCTC
CTCCTTGTCAATATTAATGTTAAAGTGCAATTCTTTTTCCTTATCACGTTGAGCCATT
AGTATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCTCCTTCTTGATAAA
TGTATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATT
TCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAA
AAGAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGA
CTTCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAA
ACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAAT
TTCAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTCAACCTTATTCTTTGG
CAAATCTGGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTG
TCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGAACCTGGGATA
ACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCA
TTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATG
TTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCA
TAATCTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGT
GGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGG
AACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTT
ACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAG
CAAATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACA
TGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTC
AGGTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAACAAAACG
GCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAG
CATCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAA
CGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACC
AACGTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGA
CATTAGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAG
AAAAGTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCC
TTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTCATGAACGCTT
CTCTATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTTTCTCCCAATTT
TTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCC
AGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGG
AAAAAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTA
TTCCCACAGTTAACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTC
GGCCAAACAACCAATTACTTGTTGAGAAATAGAGTATAATTATCCTATAAATATAAC

GTTTTTGAACACACATGAACAAGGAAGTACAGGACAATTGATTTTGAAGAGAATG
TGGATTTTGATGTAATTGTTGGGATTCCATTTTTAATAAGGCAATAATATTAGGTATG
TGGATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGTGTGAAATACCGCACAG
ATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTTAAA
ATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGG
CAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAG
TTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAA
AAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTT
TTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGAT
TTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAA
GCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTA
ACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGC
CATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCC
AGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATAC
GACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCG
ATAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATA
AAAAACACGCTTTTTCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAAT
ACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTAGCCTT
TTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATAT
ATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATG
GAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAAATA
TTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAG
AACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAA
CCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCA
TTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAA
AAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAG
ACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACT
TTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAAC
ACACATAAACAAACAAAGAATTC (SEQ ID NO: 730)

p425-GFP_+1fs_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGAT
GGTGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTAC
TTACGGTAAATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATG
GCCAACCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGA
TCATATGAAACAACATGACTTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAG
AAAGAACTATTTTTTTCAAAGATGACGGTAACTACAAGACCAGAGCTGAAGTCAA
GTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAAGGTATTGATTTTAAAGA
AGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTCACAATGTTTA
CATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATT
GGTGATGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTA
TCCAAAGATCCAAACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTG
CTGCTGGTATTACCCATGGTATGGATGAATTGTACAAAGCTAGCAAACCTGGGTCA
ATCCTTGG**<u>GGCCCAGACTGAGCACGTGA</u>TGG**CAGAGCACAGGAGACGTC*ATGG*
TTTCAAAAGGTGAAGAAGATAATATGGCTATTATTAAAGAATTTATGAGATTTAAAGTTCAT
ATGGAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAAGGTGAAGGTAGACC
ATATGAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACCATTTGCT

TGGGATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTG
ATATTCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAAT
TTTGAAGATGGTGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTA
TTTATAAAGTTAAATTGAGAGGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAA
AACTATGGGTTGGGAAGCTTCATCAGAAAGAATGTATCCAGAAGATGGTGCTTTAAAAG
GTGAAATTAAACAAAGATTGAAATTAAAAGATGGTGGTCATTATGATGCTGAAGTTAAAAC
TACTTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTATAATGTTAATATTAAATTGG
ATATTACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAGCTGAAGGTAGA
CATTCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATTGAT
*TGA*GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTT
GCATTTAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATT
TTAATGACATTTTCGATTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGG
AGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATG
GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGG
AGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGTAACTCACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT
GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGG
TATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC
CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA
TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG
CTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTG
GTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC
AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC
CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC
GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGT
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGT
CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTC
TTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTT
CTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA
CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGG
GTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC

GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCA
TTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTG
AGCTGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAA
GAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTC
AAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCT
ATTTTACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCATCCCGAGA
GCGCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTAT
AATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGC
TACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTT
ACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATT
ATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGA
TGATTCTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACT
ACGTATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCT
TACTACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAACATAAAAAATGTAGAG
GTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGG
GATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGC
GGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAA
AGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTCCTATACTTT
CTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGC
TTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCT
ATATCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTT
ATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTAC
CTCCTGTGATATTATCCCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCC
TTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCTCATCCTTCAATGCTA
TCATTTCCTTTGATATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTAT
AAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTG
AAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGA
TGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCG
GGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCG
ACTACGTCGTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCAT
TATGGGAAATGCTTCAAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCAT
TGAGTGTTTTTTATTTGTTGTATTTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAA
TTAGGAATCGTAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTGCCCTC
CTCCTTGTCAATATTAATGTTAAAGTGCAATTCTTTTTCCTTATCACGTTGAGCCATT
AGTATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCTCCTTCTTGATAAA
TGTATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATT
TCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAA
AAGAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGA
CTTCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAA
ACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAAT
TTCAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTCAACCTTATTCTTTGG
CAAATCTGGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTG
TCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGAACCTGGGATA
ACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCA
TTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATG
TTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCA
TAATCTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGT

GGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGG
AACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTT
ACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAG
CAAATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACA
TGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTC
AGGTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAACAAAACG
GCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAG
CATCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAA
CGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACC
AACGTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGA
CATTAGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAG
AAAAGTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCC
TTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTCATGAACGCTT
CTCTATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTTTCTCCCAATTT
TTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCC
AGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGG
AAAAAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTA
TTCCCACAGTTAACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTC
GGCCAAACAACCAATTACTTGTTGAGAAATAGAGTATAATTATCCTATAAATATAAC
GTTTTTGAACACACATGAACAAGGAAGTACAGGACAATTGATTTTGAAGAGAATG
TGGATTTTGATGTAATTGTTGGGATTCCATTTTTAATAAGGCAATAATATTAGGTATG
TGGATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGTGTGAAATACCGCACAG
ATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTTAAA
ATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGG
CAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAG
TTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAA
AAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTT
TTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGAT
TTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAA
GCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTA
ACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGC
CATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCC
AGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATAC
GACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCG
ATAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATA
AAAAACACGCTTTTTCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAAT
ACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTAGCCTT
TTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATAT
ATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATG
GAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAAATA
TTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAG
AACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAA
CCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCA
TTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAA
AAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAG
ACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACT
TTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAAC
ACACATAAACAAACAAAGAATTC (SEQ ID NO: 731)

p425-GFP_-1fs_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGAT
GGTGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTAC
TTACGGTAAATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATG
GCCAACCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGA
TCATATGAAACAACATGACTTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAG
AAAGAACTATTTTTTTCAAAGATGACGGTAACTACAAGACCAGAGCTGAAGTCAA
GTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAAGGTATTGATTTTAAAGA
AGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTCACAATGTTTA
CATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATT
GGTGATGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTA
TCCAAAGATCCAAACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTG
CTGCTGGTATTACCCATGGTATGGATGAATTGTACAAAGCTAGCAAACCTGGGTCA
ATCCTTGG**<u>GGCCCAGACTGAGCACGTGA</u>TGG**CAGAGCACAGGACGTC*ATGGTTT*
CAAAAGGTGAAGAAGATAATATGGCTATTATTAAAGAATTTATGAGATTTAAAGTTCATATG
GAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAAGGTGAAGGTAGACCATAT
GAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACCATTTGCTTGG
GATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTGATAT
TCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAATTTT
GAAGATGGTGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTATTT
ATAAAGTTAAATTGAGAGGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAAAAC
TATGGGTTGGGAAGCTTCATCAGAAAGAATGTATCCAGAAGATGGTGCTTTAAAAGGTG
AAATTAAACAAAGATTGAAATTAAAAGATGGTGGTCATTATGATGCTGAAGTTAAAACTAC
TTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTATAATGTTAATATTAAATTGGATAT
TACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAGCTGAAGGTAGACAT
TCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATTGATTGA
GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTTGCA
TTTAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAA
TGACATTTTCGATTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCT
CCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCAT
AGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGGAGCC
GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG
CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCA
GCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA
AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG
TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT
ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA
AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT

TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
TGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG
CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC
CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCA
GTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATA
AACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT
AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT
AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA
ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG
GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGA
GCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA
ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTAT
TGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCATTTTG
TAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTG
CATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATC
TGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACA
AAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTT
ACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCATCCCGAGAGCGC
TATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTATAATG
CAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGCTACT
TTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTTACTG
ATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATTATATT
CTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGAT
TCTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGT
ATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACT
ACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTC
GAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGGGATA
TAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGCGGTA
TTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAAAGTG
CGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTCCTATACTTTCTAG
AGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCC
GAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCTATATC
TGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGC
TTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCC
TGTGATATTATCCCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTA
GCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCTCATCCTTCAATGCTATCAT
TTCCTTTGATATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTATAAA
AATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAA
ACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGC
CGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGG

CTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACT
ACGTCGTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCATTAT
GGGAAATGCTTCAAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCATTGA
GTGTTTTTTATTTGTTGTATTTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAATTA
GGAATCGTAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTGCCCTCCTC
CTTGTCAATATTAATGTTAAAGTGCAATTCTTTTTCCTTATCACGTTGAGCCATTAGT
ATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCTCCTTCTTGATAAATGT
ATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATTTC
GTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAA
GAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACT
TCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAAC
CTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTT
CAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTCAACCTTATTCTTTGGCA
AATCTGGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTGTC
TGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGAACCTGGGATAAC
GGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCATT
TAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTT
GAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATA
ATCTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGG
CTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGAA
CGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTAC
CAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCA
AATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATG
GTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTCAG
GTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAACAAAACGGC
ATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGCAT
CGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGA
ACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAAC
GTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGACATT
AGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAA
AGTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCCTTA
AATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTCATGAACGCTTCTC
TATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTTTCTCCCAATTTTTC
AGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCCAG
TCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAA
AAAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTC
CCACAGTTAACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTCGGC
CAAACAACCAATTACTTGTTGAGAAATAGAGTATAATTATCCTATAAATATAACGTT
TTTGAACACACATGAACAAGGAAGTACAGGACAATTGATTTTGAAGAGAATGTGG
ATTTTGATGTAATTGTTGGGATTCCATTTTTAATAAGGCAATAATATTAGGTATGTGG
ATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGTGTGAAATACCGCACAGATG
CGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTTAAAATT
CGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAA
AATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTT
GGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAA
CCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTG
GGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTA
GAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG
AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACC

ACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCAT
TCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACG
CCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGG
GTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGAC
TCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAA
GCTTGATATCGAATTCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATAAAA
AACACGCTTTTTCAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACG
TAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTAGCCTTTTA
ATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATA
ACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAATGGA
GCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAAATATT
GTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGA
ACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACC
TGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATT
TTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAA
AGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGA
CGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTT
TTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACA
CACATAAACAAACAAAGAATTC (SEQ ID NO: 732)

ОБОЗНАЧЕНИЯ:
открытая рамка считывания GFP
Линкер, содержащий стоп-кодон +1 сдвиг рамки считывания или -1 сдвиг рамки считывания
Открытая рамка считывания mCherry
Каркас плазмиды (содержащий промотор GPD, маркер Leu2, и AmpR)
Протоспейсер (подчеркнут)
PAM (обозначен жирным шрифтом)

[1419] Последовательности ДНК плазмид млекопитающих для праймерного редактирования и пример плазмиды PEgRNA

pCMV-PE1:

ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAA
AGTCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGG
GCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGC
AACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC
AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATA
CACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATG
GCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA
GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTG
GCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACA
GCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAA
GTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGT
GGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAAC
CCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCA
AGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATG
GCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAG
CAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGA

CGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTT
CTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGA
ACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGA
GCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGA
GAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATT
GACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAA
AAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTG
CGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGA
GAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACA
ACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCC
TCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC
CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAG
CTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTG
CCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAG
TGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGC
AGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTC
CGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAA
ATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAG
ATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCT
GAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCG
GAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGA
CAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAAC
AGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCC
AGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCT
GGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGA
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAAT
GGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAA
TGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAAC
ACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGAC
TACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACA
ACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCC
TCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCC
AAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC
CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGG
CAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACG
ACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGC
TGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAA
CTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATC
AAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACG
ACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCA
AGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCC
AACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAG
ATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGC
CCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAG
AGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACT
GGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT
GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAG

AGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA
CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCT
GCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCT
GCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAAC
TTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATG
AGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCG
AGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAA
AGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGA
GAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGT
ACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG
ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCT
GTCTCAGCTGGGAGGTGAC*TCTGGAGGATCTAGCGGAGGATCCTCTGGCAGCGAGA*
CACCAGGAACAAGCGAGTCAGCAACACCAGAGAGCAGTGGCGGCAGCAGCGGCGGC
*AGCAGC***ACCCTAAATATAGAAGATGAGTATCGGCTACATGAGACCTCAAAAGA**
GCCAGATGTTTCTCTAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGG
GCGGAAACCGGGGGCATGGGACTGGCAGTTCGCCAAGCTCCTCTGATCATA
CCTCTGAAAGCAACCTCTACCCCCGTGTCCATAAAACAATACCCCATGTCACA
AGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGACTGTTGGACCAGGG
AATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCGTTAAG
AAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACA
AGCGGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAG
CGGGCTCCCACCGTCCCACCAGTGGTACACTGTGCTTGATTTAAAGGATGCC
TTTTTCTGCCTGAGACTCCACCCCACCAGTCAGCCTCTCTTCGCCTTTGAGT
GGAGAGATCCAGAGATGGGAATCTCAGGACAATTGACCTGGACCAGACTCC
CACAGGGTTTCAAAAACAGTCCCACCCTGTTTGATGAGGCACTGCACAGAGA
CCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTACAGTACGTG
GATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTC
GGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATCGGGCCTCGGCCAAGA
AAGCCCAAATTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCTAAAAGA
GGGTCAGAGATGGCTGACTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCC
TACTCCGAAGACCCCTCGACAACTAAGGGAGTTCCTAGGGACGGCAGGCTT
CTGTCGCCTCTGGATCCCTGGGTTTGCAGAAATGGCAGCCCCCCTGTACCCT
CTCACCAAAACGGGGACTCTGTTTAATTGGGGCCCAGACCAACAAAAGGCCT
ATCAAGAAATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAGA
TTTGACTAAGCCCTTTGAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAA
GGTGTCCTAACGCAAAAACTGGGACCTTGGCGTCGGCCGGTGGCCTACCTG
TCCAAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGG
TAGCAGCCATTGCCGTACTGACAAAGGATGCAGGCAAGCTAACCATGGGACA
GCCACTAGTCATTCTGGCCCCCCATGCAGTAGAGGCACTAGTCAAACAACCC
CCCGACCGCTGGCTTTCCAACGCCCGGATGACTCACTATCAGGCCTTGCTTT
TGGACACGGACCGGGTCCAGTTCGGACCGGTGGTAGCCCTGAACCCGGCTA
CGCTGCTCCCACTGCCTGAGGAAGGGCTGCAACACAACTGCCTTGATATCCT
GGCCGAAGCCCACGGAACCCGACCCGACCTAACGGACCAGCCGCTCCCAGA
CGCCGACCACACCTGGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACA
GCGTAAGGCGGGAGCTGCGGTGACCACCGAGACCGAGGTAATCTGGGCTAA
AGCCCTGCCAGCCGGGACATCCGCTCAGCGGGCTGAACTGATAGCACTCAC
CCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTTATACTGATAGC
CGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGCGTGG
GTTGCTCACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCC

CTACTAAAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGG
ACATCAAAAGGGACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCA
AGCGGCCCGAAAGGCAGCCATCACAGAGACTCCAGACACCTCTACCCTCCT
CATAGAAAATTCATCACCCTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGC
GAATTCGAGCCCAAGAAGAAGAGGAAAGTCTAACCGGTCATCATCACCATCAC
CATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATC
TGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTG
TCCTTTCCTAATAAAATGAGAAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA
GCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAG
CTGGGGCTCGATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGA
AGCATAAAGTGTAAAGCCTAGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC
GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTC
CTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGC
TGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT
CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG
ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA
GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTT
TTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCA
GCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA
TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG
TTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT
TGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAC
CGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG
TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAA
AAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTT
GAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCT
CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC

GCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCGATCTCC
CGATCCCCTAGGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCC
AGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGG
TTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGA
TTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAAC
GACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG
TACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAAT
GGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGT
ACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCA
ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC
GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTA
ACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA
TATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAGATCCGCGGCCGCTA
ATACGACTCACTATAGGGAGAGCCGCCACC (SEQ ID NO: 733)

pCMV-PE2:

ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAA
AGTCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGG
GCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGC
AACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC
AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATA
CACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATG
GCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA
GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTG
GCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACA
GCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAA
GTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGT
GGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAAC
CCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCA
AGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATG
GCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAG
CAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGA
CGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTT
CTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGA
ACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGA
GCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGA
GAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATT
GACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAA
AAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTG
CGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGA
GAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACA
ACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCC
TCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC
CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAG
CTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTG
CCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAG

TGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGC
AGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTC
CGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAA
ATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAG
ATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCT
GAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCG
GAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGA
CAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAAC
AGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCC
AGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCT
GGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGA
CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAAT
GGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAA
TGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAAC
ACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGAC
TACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACA
ACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCC
TCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCC
AAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC
CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGG
CAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACG
ACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGC
TGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAA
CTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATC
AAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACG
ACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCA
AGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCC
AACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAG
ATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGC
CCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAG
AGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACT
GGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT
GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAG
AGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA
CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCT
GCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCT
GCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAAC
TTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATG
AGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCG
AGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAA
AGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGA
GAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGT
ACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG
ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCT
GTCTCAGCTGGGAGGTGACTCTGGAGGATCTAGCGGAGGATCCTCTGGCAGCGAG
ACACCAGGAACAAGCGAGTCAGCAACACCAGAGAGCAGTGGCGGCAGCAGCGG
CGGCAGCAGCACCCTAAATATAGAAGATGAGTATCGGCTACATGAGACCTCAAAA

GAGCCAGATGTTTCTCTAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGC
GGAAACCGGGGGCATGGGACTGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTG
AAAGCAACCTCTACCCCCGTGTCCATAAAACAATACCCCATGTCACAAGAAGCCA
GACTGGGGATCAAGCCCCACATACAGAGACTGTTGGACCAGGGAATACTGGTACC
CTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCGTTAAGAAACCAGGGACTAAT
GATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACAAGCGGGTGGAAGACATCC
ACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCACCGTCCCACCA
GTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCCACCCCA
CCAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAGG
ACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCTGTTT
AATGAGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGA
TCCTGCTACAGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGCTAGACTGC
CAACAAGGTACTCGGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATCGGGCCT
CGGCCAAGAAAGCCCAAATTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCT
AAAAGAGGGTCAGAGATGGCTGACTGAGGCCAGAAAAGAGACTGTGATGGGGCA
GCCTACTCCGAAGACCCCTCGACAACTAAGGGAGTTCCTAGGGAAGGCAGGCTTC
TGTCGCCTCTTCATCCCTGGGTTTGCAGAAATGGCAGCCCCCCTGTACCCTCTCAC
CAAACCGGGGACTCTGTTTAATTGGGGCCCAGACCAACAAAAGGCCTATCAAGAA
ATCAAGCAAGCTCTTCTAACTGCCCCAGCCCTGGGGTTGCCAGATTTGACTAAGC
CCTTTGAACTCTTTGTCGACGAGAAGCAGGGCTACGCCAAAGGTGTCCTAACGCA
AAAACTGGGACCTTGGCGTCGGCCGGTGGCCTACCTGTCCAAAAAGCTAGACCC
AGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGTAGCAGCCATTGCCGTACTG
ACAAAGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTCTGGCCCCCC
ATGCAGTAGAGGCACTAGTCAAACAACCCCCCGACCGCTGGCTTTCCAACGCCCG
GATGACTCACTATCAGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCGGACCG
GTGGTAGCCCTGAACCCGGCTACGCTGCTCCCACTGCCTGAGGAAGGGCTGCAAC
ACAACTGCCTTGATATCCTGGCCGAAGCCCACGGAACCCGACCCGACCTAACGGA
CCAGCCGCTCCCAGACGCCGACCACACCTGGTACACGGATGGAAGCAGTCTCTTA
CAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGACCACCGAGACCGAGGTAATC
TGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGGCTGAACTGATAGCAC
TCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTTATACTGATAGC
CGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGCGTGGGTG
GCTCACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTACTA
AAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAA
GGGACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAA
GGCAGCCATCACAGAGACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCAC
CCTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGCCCAAGAAGA
AGAGGAAAGTCTAACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGAT
CAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTG
CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGAA
AATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGG
TGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATACCGTCGAC
CTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTA
TCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTAG
GGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTT
CCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGG
AGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG

TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAG
CAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG
CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTC
GCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG
CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCT
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACG
CGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACG
CTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG
GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTAT
ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCT
CAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAA
CTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTG
CCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCA
TTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG
TTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCA
CTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGC
TTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGA
ACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA
ATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT
CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAA
AGTGCCACCTGACGTCGACGGATCGGGAGATCGATCTCCCGATCCCCTAGGGTCG
ACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCT
TGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGC
AAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCT
GCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTA
ATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTAC
ATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTG
ACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGT
CAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA
TATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAG
TCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGC
GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG
TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATT
GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGT

TTAGTGAACCGTCAGATCCGCTAGAGATCCGCGGCCGCTAATACGACTCACTATAG
GGAGAGCCGCCACC (SEQ ID NO: 734)

N-концевая NLS + Cas9 H840A
Гибкий линкер
Обратная транскриптаза M-MLV + C-концевая NLS
Каркас плазмиды (содержащий промотор CMV и AmpR)

pU6-HEK3_PEgRNA_CTTin:

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAAT
GGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC
TTGTGGAAAGGACGAAACACC*GGCCCAGACTGAGCACGTGA*GTTTTAGAGCTAGA
AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAG
TCGGTCC**TCTGCCATCAAAGCGTGCTCAGTCTG**TTTTTTTAAGCTTGGGCCGCT
CGAGGTACCTCTCTACATATGACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCT
GCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT
AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT
GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGG
ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC
CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGG
CTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTC
CAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTC
CTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAA
GTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG
GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC
CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCA
CTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAG
TACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCC
GGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCA
CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAA
TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAA
GCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA
ATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGC

TAGCTGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGT
ACCAAGGTCGGGCAGGAA (SEQ ID NO: 735)

Последовательность промотора U6
Спейсерная последовательность
Каркас sgРНК
3′ удлинение (содержит PBS и RT матрицу)
Каркас (содержит AmpR)

pLenti-hSyn-N-PE2-NpuN-P2A-GFP-KASH_U6-DNMT1-PEgRNA:

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCT
GATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTA
GTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAG
AATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCG
TTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGC
CCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT
TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTG
ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTGCC
TGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGG
GAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCC
GTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAA
ATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGC
TCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGC
GACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGT
GCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTT
AAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAG
CTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAAT
ACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATA
ATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGA
AGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGC
GGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATT
ATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGA
GAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGT
TCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGG
CCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAG
GCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAA
TCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTC
TGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTC
TGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA
CACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAA
CAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAA
TTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAAT

AGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT
CAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAA
GGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCGGCACTGCGT
GCGCCAATTCTGCAGACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGG
ATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAA
CTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGAC
AGCAGAGATCCAGTTTGGTTAATTAAGGTACCAAGGTCGGGCAGGAAGAGGGCCTATT
TCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATT
AATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATT
TCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTA
ACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC**GC**
GGGCTGGAGCTGTTCGCGCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT
AGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAAGATGCAAGCGCG
AACAGCTCCAGTTTTTTTGAATTC*AGTGCAAGTGGGTTTTAGGACCAGGATGAGGCGG*
GGTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCCAACCCCCATT
CCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGATGCGGCGAGGCGCG
TGCGCACTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGCG
CCACCGCCGCCTCAGCACTGAAGGCGCGCTGACGTCACTCGCCGGTCCCCCGCAAACTCC
CCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGCCGCCGGCCCAGCCGGACCGCACCACG
CGAGGCGCGAGATAGGGGGGCACGGGCGCGACCATCTGCGCTGCGGCGCCGGCGACTCA
*GCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCGTGCCTGAGAGCGCAG****TC***
GAGAATCTAGAGCGCTGCCACCATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTC
ACCAAAGAAGAAGCGGAAAGTCGACAAGAAGTACAGCATCGGCCTGGACATCGGCAC
CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTC
AAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTG
CTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAG
AAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA
GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCT
ACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG
ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGG
CCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT
CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGC
GGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAA
AATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTG
CCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGC
CAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCA
GATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATC
CTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCT
CTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGT
GCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGG
CTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAG
CCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAG
GACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG
ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCC
TCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCAT
CACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATC
GAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCAC

AGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGA
CCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGG
ACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACT
TCAAGAAAATCGAG*TGCCTGTCCTACGAGACAGAGATCCTGACAGTGGAGTATGGCCTGC*
TGCCAATCGGCAAGATCGTGGAGAAGAGGATCGAGTGTACCGTGTACTCTGTGGATAACAAT
GGCAACATCTATACACAGCCCGTGGCACAGTGGCACGATAGGGGAGAGCAGGAGGTGTTC
GAGTATTGCCTGGAGGACGGCAGCCTGATCAGGGCAACCAAGGACCACAAGTTCATGACAG
TGGATGGCCAGATGCTGCCCATCGACGAGATTTTCGAGCGGGAGCTGGACCTGATGAGAGT
*GGATAACCTGCCTAAT**AGCGGAGGCAGTAAAAGAACAGCAGACGGGAGTGAGTTTGAGC***
CCAAGAAAAAGAGAAAGGTGGGATCCGGCGCAACAAACTTCTCTCTGCTGAAACAAGCC
GGAGATGTCGAAGAGAATCCTGGACCGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGG
TGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACC
ACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGC
AGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC
CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAG
CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGA
TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACAC
CCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCG
CCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGATCTCGAGA
GGAGGAGGAGGAGACAGACAGCAGGATGCCCCACCTCGACAGCCCCGGCAGCTCCCAG
CCGAGACGCTCCTTCCTCTCAAGGGTGATCAGGGCAGCGCTACCGTTGCAGCTGCTTCT
GCTGCTGCTGCTGCTCCTGGCCTGCCTGCTACCTGCCTCTGAAGATGACTACAGCTGCA
CCCAGGCCAACAACTTTGCCCGATCCTTCTACCCCATGCTGCGGTACACCAACGGGCCA
CCTCCCACCTAAACGCGTTAAGTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAG
ATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATG
CCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTG
GTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGC
ACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCC
TTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGC
CTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGT
CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGC
GGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGG
CCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGA
TCTCCCTTTGGGCCGCCTCCCCGCGTCGACTTTAAGACCAATGACTTACAAGGCAGCTG
TAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCA
ACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAG
CCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCT
TGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCT
CAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGGGCCCGTTTAAACCCGCTGATCA
GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC
TTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTT
CTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCG
GCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAG

CGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTT
TCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGC
ACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTG
ATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTT
GCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
AATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGC
AGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAG
GCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGT
CCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGC
TATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCC
GGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCATCGGCA
TAGTATATCGGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGAC
CAGTGCCGTTCCGGTGCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACC
GACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGG
ACGACGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCC
TGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCG
TGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGC
CGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGG
CCGAGGAGCAGGACTGACACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGA
AAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGG
GATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTAC
AAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT
TGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCT
AGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA
TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT
GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGT
CGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGA
GCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACG
CAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC
GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT
GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC
CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA
CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT
CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC
AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT
GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCAC
CTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC
TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATT
TCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT
ACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT
TTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTT

TATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGA
TCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA
ATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCT
TTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA
TGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC
ACCTGAC (SEQ ID NO: 736)

<u>Промотор U6</u>
PEgRNA
Промотор hSyn
N-концевой PE2
<u>N-концевой Npu</u>
<u>P2A-GFP-KASH</u>

pLenti-hSyn-C-PE2-NpuC:

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCT
GATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTA
GTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAG
AATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCG
TTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGC
CCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT
TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTG
ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTGCC
TGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGG
GAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCC
GTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAA
ATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGC
TCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGC
GACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGT
GCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTT
AAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAG
CTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAAT
ACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATA
ATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGA
AGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGC

GGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATT
ATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGA
GAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGT
TCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGG
CCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAG
GCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAA
TCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTC
TGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTC
TGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA
CACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAA
CAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAA
TTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAAT
AGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT
CAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAA
GGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCGGCACTGCGT
GCGCCAATTCTGCAGACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGG
ATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAA
CTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGAC
AGCAGAGATCCAGTTTGGTTAATTAAGGTACCAGTGCAAGTGGGTTTTAGGACCAGGA
TGAGGCGGGGTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACC
CAACCCCCATTCCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGAT
GCGGCGAGGCGCGTGCGCACTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCCCCC
GCCTGGCGGCGCGCGCCACCGCCGCCTCAGCACTGAAGGCGCGCTGACGTCACTCGC
CGGTCCCCCGCAAACTCCCCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGCCGCCGGC
CCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGGCACGGGCGCGACCATCTG
CGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTC
GTGTCGTGCCTGAGAGCGCAG*TCGAGAATCTAGAGCGCTGCCACCATGAAACGGACAGCC*
GACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGTCATCAAGATTGCTACACGGA
AATACCTGGGAAAGCAGAACGTGTACGACATCGGCGTGGAGCGGGATCACAACTTCGCCCT
*GAAGAATGGCTTTATCGCCAGCAAT***TGCTTCGACTCCGTGGAAATCTCCGGCGTGGAA**
GATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGG
ACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCT
GACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTA
TGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACAC
CGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGT
CCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACT
TCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGC
CCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGG
CAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCT
CGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAG
AGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGC
GGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCG
TGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATG
GGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACG
ATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA
GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCG
AAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGC
TGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGA
GCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGA

TCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACG
AGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGG
TGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTA
CCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAA
AAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGA
CGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAA
GTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCA
ACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAG
ATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATG
CCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAA
GAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGAC
TGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTG
CTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAA
AGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCAT
CGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAA
GCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGC
CTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGT
GAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGA
TAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGAT
CATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTG
GACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAG
GCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCT
TCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGG
TGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGA
TCGACCTGTCTCAGCTGGGAGGTGACTCTGGCGGCTCAAAAAGAACCGCCGACG
GCAGCGAATTCGAGCCCAAGAAGAAGAGGAAAGTCTAAACGCGTTAAGTCGACAA
TCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCC
TTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATG
GCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGG
CCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTG
GTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCT
ATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGC
TGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTG
CTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGC
CCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGC
GTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCGTCGA
CTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAG
GGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTA
CTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAA
CCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTC
TGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATC
TCTAGCAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGC
CATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACT
GTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATT
CTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
GCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGC
TCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG
TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTC
TTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCT

CCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG
GTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTG
GAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTAT
CTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAAT
GAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGG
TGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATT
AGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAA
GCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCC
CTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATG
CAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTT
TGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTG
ATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAA
GGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGC
GACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCG
TGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGT
CCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGA
CGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGG
GCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGA
CCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTGACACGTGCTACG
AGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGG
ACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCC
AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA
AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATA
AAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTC
ACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA
CGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACT
CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT
CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT
TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT
GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT
GAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC
GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGG
GCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTG
CCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG

CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCC
CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGG
CAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCG
GCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGG
AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCG
ATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCT
GGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACG
GAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTAT
TGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGAC (SEQ ID NO: 737)
Промотор hSyn
С-концевой Npu
С-концевой Cas9 дикого типа

pLenti-U6-DNMT1_micking_sgRNA:

TAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTT
TAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCAT
TATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCAGA*GAGGGCCTATTTCCCATG*
ATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGT
AAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCA
GTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTC
*TTGGCTTTATATATCTTGTGGAAAGGACGAAACACC**GCCGCGCGCGCGAAAAAGCCGGTTT***
TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA
CCGAGTCGGTGCTTTTTTAAGCTTGGCGTAACTAGATCTTGAGACAAATGGCAGTATTC
ATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATA
GTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAA
TTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTTGGCGCCGGCTCGAG
GGGGCCCGGGTGCAAAGATGGATAAAGTTTTAAACAGAGAGGAATCTTTGCAGCTAAT
GGACCTTCTAGGTCTTGAAAGGAGTGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGA
GCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGATCCGG
TGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGC
CTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCT
TTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGG
CCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCA
GTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGCCT
TGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGG
GCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCT
CTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTT
GTAAATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGGGGCCGCGGGCGGCG
ACGGGGCCCGTGCGTCCCAGCGCACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGC
CACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCT
CGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTT
GCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGG
ACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTT
TCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCAC
CTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGTTGGGGGGAGGGGTTTTA
TGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCAC

TTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCTCAAG
CCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTCAGGTGTCGTGACGTACGGCCAC
CATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGC
CGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGAT
CCGGACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCG
GGCTCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCCGTGGCGGTCTGGA
CCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGG
CCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGC
CGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGAGTCTCGCCCGACC
ACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGC
GCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGA
GCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTG
GTGCATGACCCGCAAGCCCGGTGCCTGAACGCGTTAAGTCGACAATCAACCTCTGGAT
TACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGT
GGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCT
CCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGC
AACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGC
CACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGG
AACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGA
CAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTG
CCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCG
GACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCG
CCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCGTCGACTTTAAGACCAAT
GACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAA
GGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTG
GTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAG
CCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTC
TGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTACGT
ATAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGAAATGAATATCAGA
GAGTGAGAGGAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCA
TCAATGTATCTTATCATGTCTGGCTCTAGCTATCCCGCCCCTAACTCCGCCCATCCCGCC
CCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTAT
GCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTT
TTGGAGGCCTAGGGACGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCAC
TGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCG
CCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGAT
CGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCG
CATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGC
CCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCC
CCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACC
TCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAG
ACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCA
AACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC
GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA
ACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCC
CTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTG
ATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGC
CCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGT

GAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGAT
CTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGA
GCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG
CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCAC
AGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCA
TGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCT
AACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGG
AGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGC
AACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAAT
TAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCC
GGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA
TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGG
AGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGA
TTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAA
TCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG
ATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCAC
CGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTA
ACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGG
CCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTAC
CAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATA
GTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAG
CTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGC
GCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGA
ACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTG
TCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG
AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGC
CTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT
GAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCC
GATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGC
AACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTT
CCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTA
TGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAG
CTGCAAGCTTAATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGA
GTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGAAGT
AAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGACGGGTCTGACATGGATTGGA
CGAACCACTGAATTGCCGCATTGCAGAGATATTGTATTTAAGTGCCTAGCTCGATACATA
AACGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGA
ACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGT
CTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAAT
CTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTC
TCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGA
CTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGC
GAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAA
GGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCT
AGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATA
CTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAA
TACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAA

GCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCG
GCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTA
TATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGA
GAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGT
TCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGG
CCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAG
GCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAA
TCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTC
TGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTC
TGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA
CACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAA
CAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACA (SEQ
ID NO: 738)
Промотор U6
sgРНК

[1420] Аминокислотные последовательности использованных здесь вариантов обратной транскриптазы вируса лейкоза мышей Молони (RT M-MLV).

PE1 RT M-MLV:

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTP
VSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLR
EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR
DPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAAT
SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET
VMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAY
QEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV
AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMT
HYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD
ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM
AEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLS
IIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 739)

M3 RT M-MLV (D200N, T330P, L603W) (см. Baranauskas et al.[182]):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTP
VSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLR
EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR
DPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAAT
SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET
VMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAY
QEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV
AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMT
HYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD
ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM
AEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRL

SIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 740)

PE2 RT M-MLV (D200N, T306K, W313F, T330P, L603W):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTP
VSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLR
EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR
DPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAAT
SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET
VMGQPTPKTPRQLREFLG**K**AGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAY
QEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV
AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMT
HYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD
ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM
AEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRL
SIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 741)

M3-мертвая RT M-MLV:

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTP
VSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVK**L**PGTNDY**S**PVQDLR
EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR
DPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAAT
SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKET
VMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAY
QEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV
AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMT
HYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD
ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM
AEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRL
SIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 742)

Ссылки для Примера 12

[1421] Каждая из следующих ссылок приведена в Примере 12, при этом каждая из них приведена здесь в качестве ссылки.

1. Landrum, M. J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862–D868 (2016).

2. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

3. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819– 823 (2013).

4. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339**, 823–826 (2013).

5. Yang, H. *et al.* One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. *Cell* **154**, 1370–1379 (2013).

6. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).

7. Orlando, S. J. *et al.* Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.* **38**, e152–e152 (2010).

8. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPRCas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).

9. Suzuki, K. *et al. In vivo* genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144–149 (2016).

10. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).

11. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).

12. Ihry, R. J. *et al.* p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).

13. Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci.* **91**, 6064–6068 (1994).

14. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* **47**, 497–510 (2012).

15. Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* **21**, 121–131 (2015).

16. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016).

17. Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).

18. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).

19. Rees, H. A., Yeh, W.-H. & Liu, D. R. Development of hRad51–Cas9 nickase fusions that mediate HDR without double-stranded breaks. *Nat. Commun.* **10**, 1–12 (2019).

20. Shen, M. W. *et al.* Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646–651 (2018).

21. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770 (2018).

22. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).

23. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

24. Gao, X. *et al.* Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* **553**, 217–221 (2018).

25. Ingram, V. M. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* **178**, 792–794 (1956).

26. Myerowitz, R. & Costigan, F. C. The major defect in Ashkenazi Jews with Tay-Sachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase. *J. Biol. Chem.* **263**, 18587–18589 (1988).

27. Zielenski, J. Genotype and Phenotype in Cystic Fibrosis. *Respiration* **67**, 117–133 (2000).

28. Mead, S. *et al.* A Novel Protective Prion Protein Variant that Colocalizes with Kuru Exposure. *N. Engl. J. Med.* **361**, 2056–2065 (2009).

29. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).

30. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

31. Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).

32. Hille, F. *et al.* The Biology of CRISPR-Cas: Backward and Forward. *Cell* **172**, 1239–1259 (2018).

33. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).

34. Liu, Y., Kao, H.-I. & Bambara, R. A. Flap endonuclease 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* **73**, 589–615 (2004).

35. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).

36. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

37. Shechner, D. M., Hacisuleyman, E., Younger, S. T. & Rinn, J. L. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664–670 (2015).

38. Tang, W., Hu, J. H. & Liu, D. R. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat. Commun.* **8**, 15939 (2017).

39. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* **343**, 1247997 (2014).

40. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935–949 (2014).

41. Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9–guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477–1481 (2015).

42. Baranauskas, A. *et al.* Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Eng. Des. Sel.* **25**, 657–668 (2012).

43. Gerard, G. F. *et al.* The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.* **30**, 3118–3129 (2002).

44. Arezi, B. & Hogrefe, H. Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res.* **37**, 473–481 (2009).

45. Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M. & Gerard, G. F. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16**, 265–277 (1988).

46. Thuronyi, B. W. *et al.* Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat. Biotechnol.* (2019). doi:10.1038/s41587-019-0193-0

47. Kim, Y. B. *et al.* Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* **35**, 371–376 (2017).

48. Koblan, L. W. *et al.* Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* (2018). doi:10.1038/nbt.4172

49. Kleinstiver, B. P. *et al.* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off target effects. *Nature* **529**, 490–495 (2016).

50. Zuo, E. *et al.* Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* **364**, 289–292 (2019).

51. Jin, S. *et al.* Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* **364**, 292–295 (2019).

52. Kim, D., Kim, D., Lee, G., Cho, S.-I. & Kim, J.-S. Genome-wide target specificity of CRISPR RNA guided adenine base editors. *Nat. Biotechnol.* **37**, 430–435 (2019).

53. Grunewald, J. *et al.* Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019).

54. Zhou, C. *et al.* Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* **571**, 275–278 (2019).

55. Rees, H. A., Wilson, C., Doman, J. L. & Liu, D. R. Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* **5**, eaax5717 (2019).

56. Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. *Annu. Rev. Genet.* **35**, 501–538 (2001).

57. Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biol.* **2**, REVIEWS1017 (2001).

58. Berkhout, B., Jebbink, M. & Zsiros, J. Identification of an Active Reverse Transcriptase Enzyme Encoded by a Human Endogenous HERV-K Retrovirus. *J. Virol.* **73**, 2365–2375 (1999).

59. Halvas, E. K., Svarovskaia, E. S. & Pathak, V. K. Role of Murine Leukemia Virus Reverse Transcriptase Deoxyribonucleoside Triphosphate-Binding Site in Retroviral Replication and In Vivo Fidelity. *J. Virol.* **74**, 10349–10358 (2000).

60. Dever, D. P. *et al.* CRISPR/Cas9 Beta-globin Gene Targeting in Human Hematopoietic Stem Cells. *Nature* **539**, 384–389 (2016).

61. Park, S. H. *et al.* Highly efficient editing of the β-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* doi:10.1093/nar/gkz475

62. Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550 (2001).

63. Asante, E. A. *et al.* A naturally occurring variant of the human prion protein completely prevents prion disease. *Nature* **522**, 478–481 (2015).

64. Zettler, J., Schutz, V. & Mootz, H. D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* **583**, 909–914 (2009).

65. Kugler, S., Kilic, E. & Bahr, M. Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther.* **10**, 337–347 (2003).

66. de Felipe, P., Hughes, L. E., Ryan, M. D. & Brown, J. D. Co-translational, intraribosomal cleavageof polypeptides by the foot-and-mouth disease virus 2A peptide. *J. Biol. Chem.* **278**, 11441–11448 (2003).

ПРИМЕР 13 - ЗАПИСЬ КЛЕТОЧНЫХ ДАННЫХ И ОТСЛЕЖИВАНИЕ ПРОИСХОЖДЕНИЯ С ПОМОЩЬЮ РЕ

Предпосылки

[1422] Модифицирование генома может быть использовано для исследования и записи клеточных процессов и развития. Соединение клеточных событий, таких как деление клеток или активация каскадов передачи сигналов, с модификацией последовательности ДНК обеспечивает хранение клеточной истории в виде интерпретируемых изменений последовательности ДНК, которые описывают наступление определенных клеточных событий. Редактирование ДНК является необходимым для данных вариантов использования, поскольку ДНК точным образом передается от одной клетки к другой, в отличие от РНК и белков. Информация, связанная с состоянием клетки и происхождением, в целом теряется при осуществлении изменений в белках и молекулах РНК с коротким временем жизни. Запись клеточных событий в отдельной клетке является мощным способом, обеспечивающим понимание того, как болезненные состояния инициируются, поддерживаются, и изменяются относительно здоровых контролей. Способность анализировать данные вопросы связана с трудностью понимания развития рака, неврологических заболеваний, и множества других важных проблем в области

здоровья человека. Праймерное редактирование (PE) обеспечивает систему создания нацеленных и специфичных в плане последовательности геномных вставок, делеций, или мутаций. Повторная модификация ДНК-мишени, которая может быть подвергнута секвенированию с помощью нацеленного ампликонного севенирования и/или секвенирования РНК (которое является особенно ценным для экспериментов в виде записи с использованием отдельных клеток), может применяться для записи множества важных биологических процессов, включая активацию каскадов передачи сигнала, метаболических состояний, и программ клеточной дифференциации. Соединение внутренних и внешних клеточных сигналов с модификацией последовательности в геноме теоретически является возможным для любого сигнала, для которого существует отвечающий на сигнал промотор. Считается, что PE обеспечивает значительно расширенный набор инструментов для анализа происхождения клеток и истории передачи сигнала эукариотических и прокариотических клеток как в условиях культивирования, так и *in vivo*.

[1423] Предыдущие стандарты

[1424] Нацеленная вставка, делеция, или мутация последовательности может использоваться для исследования ряда важных биологических вопросов, включая отслеживание происхождения и запись клеточных стимулов. Современный набор инструментов для продуцирования данных сигнатур в геноме является ограниченным. На сегодняшний день был разработан мутагенез локусов-мишеней с использованием как ДНК-нуклеаз, так и редакторов оснований.

[1425] Разрезание последовательностей-мишеней с использованием CRISPR/Cas9 продуцирует стохастические изменения последовательности для получения большого количества продуктов в виде вставок или делеций (инделей). Значительное количество результатов изменения последовательности вследствие разрезания Cas9 обеспечивает ясное определение последовательностей, которые были разрезаны нуклеазой. Способность отличать разрезанные и неразрезанные последовательности использовалась двумя основными способами.

[1426] Во-первых, экспрессия нуклеазы Cas9 и/или ее одиночной направляющей РНК (sgРНК) была связана с клеточными сигналами. Кроме того, записывалось то, являлся ли данный сигнал возникшим на основе модификаций последовательности нацеливаемого Cas9 геномного локуса. Однако данный подход является ограниченным, поскольку каждый сигнал требует уникального локуса-мишени, что

делает отслеживание относительного значения времени нескольких сигналов сложным для интерпретирования. Другое ограничение данного подхода заключается в том, что для конкретной sgРНК необходимо несколько локусов-мишеней, поскольку продуцирование инделя зачастую значительно затрудняет дополнительный мутагенез локуса-мишени; это зачастую означает, что предварительно сконструированные локусы-мишени встраиваются в клетки для редактирования вместо прямого мутагенеза эндогенных локусов.

[1427] Во-вторых, индели Cas9 использовались для отслеживания происхождения клеток. Как описано здесь, значительное количество возможных индельных состояний, продуцированных вследствие активности нуклеазы Cas9, обеспечивает получение деревьев клеточного развития, которые свидетельствуют о том, какие клетки возникли друг из друга с течением времени. Данный подход является мощным способом, обеспечивающим понимание того, как клетки возникают друг из друга, при этом он применялся для помощи в идентификации уникальных состояний и типов клеток с течением времени развития путем осуществления секвенирования РНК в отношении выбранных групп клеток. Данный подход не может независимым образом обеспечивать предоставление информации о событиях, связанных с передачей сигналов в клетках, их порядке, а также может иметь отклонения в плане предоставления информации о состоянии клеток-предшественников по сравнению с состоянием необратимо дифференцированных клеткок.

[1428] Опосредованное нуклеазой Cas9 отслеживание происхождения и запись сигналов являются мощными способами, которые связаны с некоторыми важными нюансами. Запись сигналов с использованием нуклеазы Cas9 зачастую является технически очень сложной. Осуществляемое Cas9 разрезание нарушает локус-мишень (повторное разрезание становится сложным после образования инделя), что затрудняет запись долговременных стимулов на уровне отдельной клетки. Кинетические свойства осуществляемого Cas9 разрезания могут быть модифицированы для обеспечения записи более долгосрочных событий, однако способность интегрировать порядок, интенсивность, и продолжительность нескольких стимулов остается очень трудной технической проблемой, которая может являться недостижимой при использовании данного инструмента. Эксперименты по отслеживанию происхождения с использованием Cas9 были невероятно успешными, однако они страдают от небольших технических проблем, связанных с разрушением последовательности

вследствие одновременных осуществляемых Cas9 разрезов в локусе-мишени. Данные эксперименты по отслеживанию происхождения требуют редактирования предварительно сконструированных локусов-мишеней, что ограничивает гибкость данного подхода.

[1429] Редактирование оснований ДНК также использовалось для отслеживания событий передачи сигнала в клетках. Редактирование оснований не является хорошо подходящим для отслеживания происхождения вследствие низкого количества результатов в виде состояния, продуцируемых связанным с редактированием событием по сравнению с количеством состояний, продуцируемых инделями Cas9; однако предопределенная природа модификаций последовательностей, осуществляемых с помощью редактора оснований, является особенно полезной для отслеживания внутренних и внешних клеточных стимулов. Как описано здесь, экспрессия редактора оснований или sgРНК может быть соединена с определенным биологически или химическим стимулом. Активность в виде редактирования оснований использовалась для отслеживания большого количества отдельных стимулов как в клетках млекопитающих, так и бактериальных клетках. Данный подход также использовался для отслеживания последующих стимулов, когда первое событие в виде редактирования было необходимым перед осуществления второго редактирования.

[1430] Запись сигналов на основе редактирования оснований является важным первым шагом в данной области, однако она имеет ряд ограничений. Одно подобное ограничение заключается в том, что редактирование оснований нарушает свою цель после редактирования, что ограничивает динамический диапазон данного способа. Это означает, что использование эндогенных мишеней для записи событий зачастую является сложным и ограниченным при записи нескольких видов активности вместо активности на уровне отдельной клетки. Альтернативой является внесение предварительно сконструированного повторяющегося локуса для записи, однако это не было осуществлено на сегодняшний день. Также существуют проблемы при записи двух сигналов. Данные эксперименты по записи двух сигналов обеспечивают получение информации лишь о наличии второго стимула после первого; они не обеспечивают получение информации о том, какой стимул произошел первым, или о длительности присутствия стимула. Это фундаментальным образом ограничивает биологическое понимание, обеспечиваемое в рамках данного эксперимента.

[1431] Было предположено, что отслеживание происхождения с помощью PE может обеспечивать как отслеживание происхождения, так и запись клеточных сигналов с помощью модифицирования геномных последовательностей-мишеней или встроенных предварительно сконструированных последовательностей. В рамках PE используется синтетический слитый белок, включающий фрагмент никазы Cas9 (зачастую вариант H840A SpCas9) и домен обратной транскриптазы (RT), наряду со сконструированной направляющей нРНК для праймерного редактирования (PEgRNA). Данные компоненты совместно нацеливаются на определенную геномную последовательность и обеспечивают внесение предварительного определенного изменения. Поскольку PEgRNA обеспечивает указание как геномной последовательности-мишени, так и результата редактирования, может быть достигнуто высокоспецифическое и контролируемое геномное модифицирование одновременно при использовании нескольких PEgRNA в рамках одной клетки. Вспомогательные геномные модификации включают все замещения одиночных нуклеотидов, вставки последовательностей с размером от малого до среднего, и делеции с размером от малого до среднего. Универсальность данной технологии геномного редактирования должна обеспечить возможность временно связанной сигнал-специфичной записи в клетках.

[1432] Использование записи происхождения и передачи сигнала в клетках на основе PE

[1433] Запись клеточных сигналов может быть осуществлена рядом способов. Одним важным первым вариантом использования данного подхода является соединение событий модификации ДНК с сигналами, ассоциированными с клеточным циклом, например экспрессией циклинов, CDK, или других белков, являющихся специфическими в плане фаз времени жизни клетки, для получения клеточных часов. Клеточные часы позволяют исследователям понимать порядок различных сигналов, которые принимаются и процессируются отдельными клетками. Молекулярные часы также могут обеспечить определение долговременной передачи сигнала против кратковременных вспышек в виде передачи сигнала. С помощью праймерного редактирования компоненты, которые могут обеспечивать редактирование только один раз в рамках клеточного цикла, также могут приводить к получению молекулярных часов. Если редактирование может осуществляться только один раз в рамках клеточного цикла без продолжительной модификации ДНК (вероятно за счет отсутствия никирования не подвергаемой редактированию цепи ДНК), то можно

представить систему, которая может быть отредактирована только второй нацеленной PEgRNA при последующем делении клетки. PE является особенно полезным в качестве клеточных часов, поскольку оно может обеспечивать повторную вставку, делецию, или мутирование локусов предварительно определенным образом, при этом вставки являются особенно ценными, поскольку повторные регулярные вставки могут быть реализованы в любом геномном локусе-мишени.

[1434] Другой важный вариант использования, связанный с записью клеточных сигналов, представляет собой параллельную запись большого количества входящих клеточных сигналов. Связывание событий в виде передачи сигнала в клетках с модификацией ДНК обеспечивает запись того, произошло ли такое событие в виде передачи сигнала. Аналогично системам записи на основе нуклеазы Cas9 или редактирования оснований, запись клеточных событий может быть связана с экспрессией нРНК или редактора. В отличие от данных подходов, праймерное редактирование для отслеживания происхождения также должно быть способно обеспечивать запись порядка, интенсивности, и продолжительности событий в виде передачи сигнала без необходимости строгих мотивов последовательности для упорядоченного редактирования. Действительно, праймерное редактирование для отслеживания происхождения должно быть способно обеспечивать встраивание описанного выше клеточного счетчика со специфическими в плане сигнала вставками, делециями, или мутациями для исследования порядка, интенсивности, и продолжительности биологических сигналов. Вследствие программируемой природы праймерного редактирования данный подход может быть реализован в геномных локусах, которые предварительно существуют в представляющей интерес клетке-мишени (независимо от того, относится ли она к бактериям, мышам, крысам, обезьянам, свиньям. человеку, полосатому данио, *C. elegans*, и так далее). Также важно отметить, что запись на основе праймерного редактирования обеспечивает внесение меток зависимым от направляющей РНК образом, при этом количество входных сигналов ограничено количеством сигналов, для которых присутствуют надежные сигнал-специфичные кассеты экспрессии направляющей РНК (которое должно быть очень высоким вследствие способности связывать данные виды экспрессии с активностью Pol II промоторов РНК). Количество записываемых сигналов возрастает линейным образом одновременно с количеством требуемых PEgRNA.

[1435] PE также может применяться для отслеживания происхождения клеток. Повторные модификации последовательности могут применяться для получения уникальных клеточных меток для отслеживания отдельных клеток. Группы меток, их порядок, и размер могут использоваться для определения происхождения клеток таким же образом, что и при большом количеством состояний в виде инделей, продуцированных нуклеазой Cas9.

[1436] Методология праймерного редактирования для повторной модификации последовательностей

[1437] Был определен ряд отдельных модальностей для повторной модификации последовательностей с помощью праймерного редактирования (PE): мутагенез ДНК; делеция последовательностей; и вставка последовательностей. Следует отметить, что эти варианты использования могут быть реализованы в рамках либо существующих мишеней в виде геномной ДНК, или предварительно сконструированных последовательностей ДНК, которые исследователи встраивают в клетки-мишени. Данные способы последовательной модификации последовательности имеют ценность в плане записи информации и предварительно определенной или стохастической модификации локусов-мишеней непрерывным образом. Последовательная модификация локусов-мишеней может являться особенно полезной в плане получения библиотек вариантов в различных хозяевах.

[1438] Повторная мутация последовательности может использоваться для изменения либо геномной ДНК, или предварительно сконструированных последовательностей ДНК итеративным образом для предоставления информации о событиях, связанных с передачей сигнала в клетках. В рамках данной парадигмы мутации, внесенные на основе активности нРНК PE, будут соответствовать наличию клеточного сигнала. Данные точечные мутации могут обеспечивать внесение мотивов PAM, необходимых для последовательных событий в виде редактирования, а также точечных мутаций, которые могут соответствовать наличию определенных сигналов. Данная система будет требовать конструирования нРНК перед использованием, поскольку каждая последующая направляющая РНК будет использовать новый протоспейсер, однако она может быть особенно успешной для анализа отдельных представляющих интерес стимулов или малых количеств стимулов. Внесение данных мутаций может зависеть от отдельных биологических стимулов или может быть связано с устойчивыми клеточными процессами, обеспечивающими отсчет клеточного времени. Указанные

ниже последовательности соответствуют SEQ ID NO: 743, 744, 744, и 745.

Последовательность-мишень геномной ДНК	CGTATCGGTAACTGATCCGAT**GG**AAAGCCAGTTCAGAACCG
Последовательность-мишень после изменения #1 (внесение G и T)	CGTATCGGTAACTGATCCGAT**GG**AATGCCAGGTCAGAACCG
Последовательность-мишень после изменения #1 (использование PAM AGG)	CGTATCGGTAACTGATCCGAT**GG**AATGCC**AGG**TCAGAACCG
Последовательность-мишень после изменения #2 (внесение G и A)	CGTATCGGTAACTGATCCGAT**GG**AATGCC**AGG**TCAAAACGG

[1439] Другой интенсивный в плане направляющей РНК способ PE представляет собой повторную делецию в последовательностях-мишенях. Удаление отдельных последовательностей из локуса-мишени может обеспечивать способность воссоздавать события в виде передачи сигнала за счет потери ДНК-мотивов. Конструирование PEgRNA, которые удаляют последующие последовательности, обеспечит отслеживание последовательных сигналов. Это позволит исследователям идентифицировать случаи, в рамках которых один сигнал следует за другим, что обеспечит исследователям возможность определения порядка осуществления событий в виде передачи сигнала. Такая система была протестирована с использованием CAMERA, однако это потребовало предварительного выбора определенных локусов с уникальными требованиями в плане последовательности. Последующая делеция последовательностей с использованием PE позволит обеспечит параллельную запись парных событий в отдельных клетках, поскольку при этом не будут требоваться специфические детерминанты последовательности. Это обеспечит исследователям возможность анализировать парные события в виде передачи сигнала мультиплексным образом внутри любой представляющей интерес клетки-мишени. Указанные ниже последовательности соответствуют SEQ ID NO: 746, 747, и 748.

Пример последовательности-мишени ДНК	CGTATCGGTAACTGATCCGAT**GG**AAAGCCAGTTCAGAACCG
Последовательность-мишень после изменения #1 (удаление AAA)	CGTATCGGTAACTGATCCGAT**GG**GCCAGTTCAGAACCG
Последовательность-мишень после изменения #2 (удаление GCC)	CGTATCGGTAACTGATCCGAT**GG**AGTTCAGAACCG

[1440]

[1441] Вставка последовательности является третьим подходом для отслеживания событий в виде передачи сигнала в клетках. Некоторые варианты данной стратегии зависят от PEgRNA меньшим образом, чем мутагенез или делеции. Существует ряд различных стратегий на основе вставки — вставка коротких последовательностей,

вставка протоспейсеров, вставка протоспейсера и метки, вставка новых последовательностей гомологии, и вставка последовательностей гомологии с меткой.

[1442] Вставка коротких повторяющихся последовательностей является способом инкрементального повышения размера последовательности-мишени для измерения хода времени в клетке. В рамках данной системы вставка повторяющейся последовательности длиной 5 или более нуклеотидов может вызывать экспансию повторов совместно с ходом времени или непрерывным присутствием предварительно определенного стимула. Не зависящая от локуса природа PE для определения происхождения также обеспечивает возможность параллельного отслеживания нескольких уникальных экспансий последовательности совместно с отдельными биологическими сигналами. Это должно обеспечивать возможность измерения интенсивности нескольких биологических сигналов с течением клеточного времени в рамках отдельных клеток. Указанные ниже последовательности соответствуют SEQ ID NO: 749, 750, и 751.

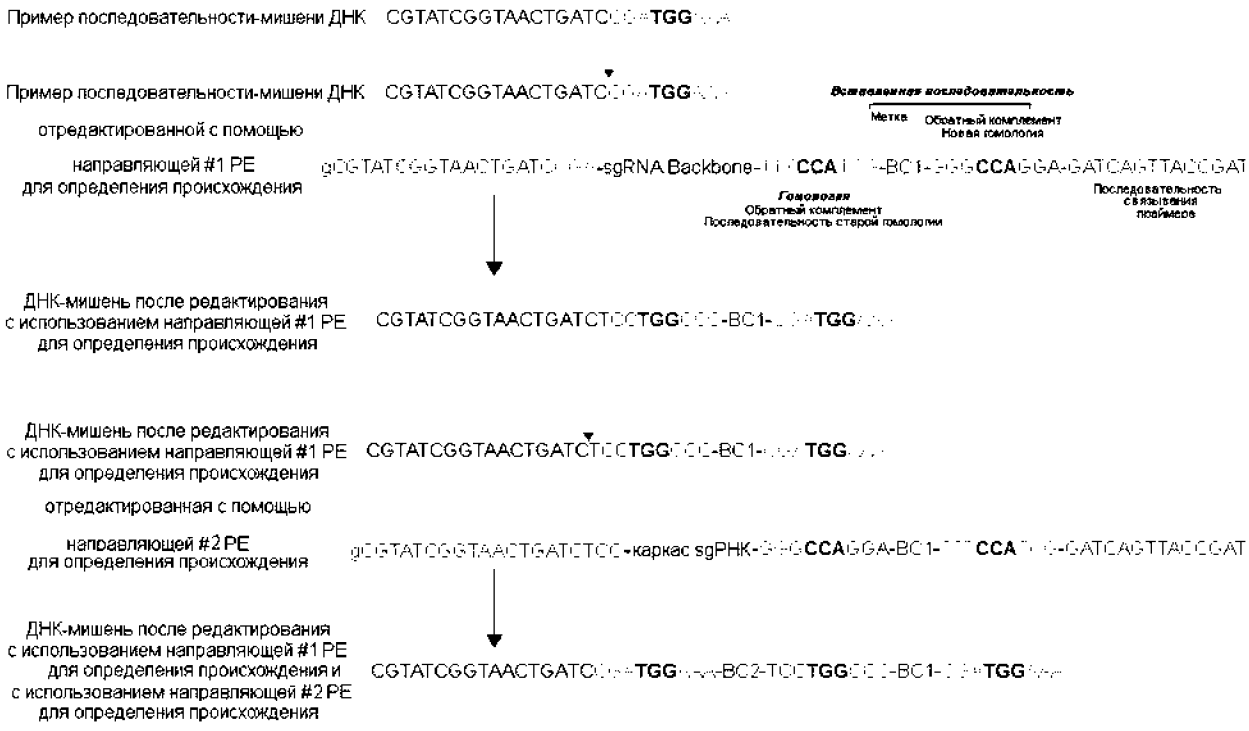
[1443]

Пример последовательности-мишени ДНК	геномная последовательность-CGTATCGTATCGTATCGTATTGG-геномная последовательность
Последовательность-мишень после изменения #1 (внесение CGTAT)	геномная последовательность-CGTATCGTATCGTATCGTATCGTATTGG геномная последовательность
Последовательность-мишень после изменения #2 (внесение CGTAT)	геномная последовательность-CGTATCGTATCGTATCGTATCGTATCGTATTGG-геномная последовательность

[1444] Вставка различных коротких последовательностей будет требовать некоторого количества PEgRNA таким же образом, что и при делециях. Количество записываемых сигналов и размер вставленной последовательности будут определять количество требуемых комбинаций PEgRNA. Одна из трудностей в плане записи нескольких последовательностей в рамках данной системы заключается в различной эффективности каждой PEgRNA при вставке ее соответствующей последовательности.

[1445] Вставка протоспейсеров в качестве индикаторов клеточных сигналов является заманчивой, однако технические сложности могут нарушить эффективность данного подхода. Системы на основе одной PEgRNA будут являться сложными в плане использования, поскольку кассеты PEgRNA будут являться субстратами для самих себя, вызывая вставки в PEgRNA, которые ухудшают эффективность и точность последующих изменений. Та же проблема присущая системам на основе двух или трех PEgRNA, поскольку каждая направляющая является субстратом для другой, что обеспечивает вставку других последовательностей в саму направляющую кассету, что будет приводить к некорректной вставке протоспейсерных последовательностей наряду с некорректным сигналом. Данные системы для вставки на основе

протоспейсеров также сложно представить в плане включения последовательностей меток. Системы с метками на основе одной PEgRNA будут просто перезаписывать используемую метку, приводя к удалению данных, хранящихся в первом изменении. Множество направляющих систем страдают от вставок в другие конструкты экспрессии PEgRNA, что ограничивает их применимость (особенно *in vivo*). Указанные ниже последовательности соответствуют SEQ ID NO: 752, 752, 753, 754, 754, 755, и 756.



[1446]

[1447] Вставка последовательностей гомологии (например последовательностей 3' участка разрыва Cas9) и, в особенности, последовательностей гомологии с ассоциированными метками, представляется особенно полезной стратегией отслеживания происхождения в рамках PE. Данная система обеспечивает устранение проблем, связанных со вставкой протоспейсера, за счет обеспечения того, что последующие раунды редактирования приводят к вставке метки из кассеты PEgRNA, которая не может быть модифицирована в рамках других событий в виде редактирования с использованием PRнРНК в той же клетке. Система мечения является ценной, поскольку несколько меток могут быть ассоциированы с определенным стимулом. Данная система обеспечивает сохранение большей части протоспейсера-мишени, но изменяет основную последовательность, PAM, а также расположенные ниже смежные нуклеотиды. Это обеспечивает возможность соединения нескольких

сигналов с одним локусом редактирования без значительного изменения используемых PEgRNA. Данная стратегия будет обеспечивать возможность множественных вставок меток в ответ на большое количество клеточных стимулов (внутренних или внешних) в отдельном локусе. Она может обеспечивать запись интенсивности, продолжительности и порядка такого количества сигналов, которое соответствует количеству существующих уникальных меток (которые могут быть сконструированы с использованием нескольких N-нуклеотидов для получения 4^N возможных меток, например метка размером 5 nt обеспечит одновременную запись 4^5 или 1024 уникальных сигналов). Данная система может применяться как *in vitro*, так и *in vivo*.

Ссылки, приведенные в Примере 13

[1448] Каждая из следующих ссылок приведена здесь в качеств ссылки.

1. Recording development with single cell dynamic lineage tracing. Aaron McKenna, James A. Gagnon.
2. Whole-organism lineage tracing by combinatorial and cumulative genome editing.Mckenna et al. Science. 2016 Jul 29;353(6298):aaf7907. doi10.1126/science.aaf7907. Epub 2016 May 26.
3. Molecular recording of mammalian embryogenesis. Chan et al. 2019. *Nature*. Jun;570(7759):77-82. doi: 10.1038/s41586-019-1184-5. Epub 2019 May 13.

ПРИМЕР 14 – МОДУЛИРОВАНИЕ БИОМОЛЕКУЛЯРНОЙ АКТИВНОСТИ И/ИЛИ ЛОКАЛИЗАЦИИ С ПОМОЩЬЮ PE

[1449] Субклеточная локализация и связанные с модификацией состояния биомолекул регулируют их активность. Определенные биологические функции, такие как транскрипционный контроль, клеточный метаболизм, и каскады передачи сигналов осторожно управляются в определенных месторасположениях в рамках клетки. Таким образом, модулирование клеточной локализации и связанных с модификацией состояний белков представляет собой потенциальную терапевтическую стратегию для лечения заболеваний. Были разработаны некоторые лекарства для изменения локализации белков-мишеней. Например разработаны ингибиторы фарнезилирования для предотвращения липидирования и нацеливания на мембрану важных онкогенных белков, таких как KRAS. Аналогичным образом, индуцируемое малыми молекулами убиквитинирование белков-мишеней обеспечивает их направление в протеасому для расщепления. Способность привлекать белки к данным и другим уникальным

клеточным компартментам может обеспечивать возможность изменять ряд биологических процессов. Здесь предлагается использование праймерного редактирования (PE) для внесения генетически кодируемых средств, которые обеспечат возможность изменения связанного с модификацией состояния и субклеточного направления биомолекул с использованием генетически кодируемого сигнала (например белков, липидов, сахаров, и нуклеиновых кислот) для терапевтических целей.

[1450] PE представляет собой технологию геномного редактирования, которая обеспечивает вставку, делецию, или замещение коротких последовательностей ДНК в любом геномном локусе, который может служить мишенью для фермента Cas9. С помощью данной технологии в принципе является возможным вносить или удалять важные последовательности, кодирующие ДНК, РНК, или белок, которые изменяют активность данных важных биомолекул. Более конкретно, праймерное редактирование может применяться для внесения мотивов, или сигналов, которые изменяют локализацию или связанные с модификацией свойства биомолекул. Некоторые примеры включают модификацию следующих элементов: аминокислотных последовательностей белка; мотивов для посттрансляционных модификаций; мотивов РНК, которые изменяют фолдинг или локализацию; а также внесение последовательностей ДНК, которые изменяют локальное состояние хроматина или архитектуру окружающей ДНК.

[1451] Одна биомолекула-мишень для PE-опосредованной модификации представляет собой ДНК. Модификации ДНК могут быть реализованы для внесения ряда последовательностей ДНК, которые изменяют доступность локуса-мишени. Доступность хроматина обеспечивает контроль выхода в виде транскрипции генов. Внесение отметок для рекрутирования компактирующих хроматин ферментов должно снижать транскрипционный выход соседних генов, в то время как внесение последовательностей, ассоциированных с открытием хроматина, должно делать участки более доступными и, в свою очередь, повышать транскрипцию. Внесение более сложных мотивов последовательности, которые отражают нативные регуляторные последовательности, должно обеспечивать более точный и биологически чувствительный контроль, чем доступные в настоящее время продукты слияния dCas9 с различным эпигенетическими ферментами для считывания, записи, или стирания — инструменты, которые обычно обеспечивают внесение большого количества одного

типа отметок, которые могут не иметь конкретного биологического предка. Внесение последовательностей, которые обеспечат сближение двух локусов или приведут локусы в контакт с ядерной мембраной, должно также влиять на транскрипционных выход данных локусов, как было продемонстрировано в рамках развивающейся области трехмерной геномной архитектуры.

[1452] Также могут быть реализованы модификации РНК для изменения их активности за счет изменения их клеточной локализации, партнеров по взаимодействию, структурной динамики, или термодинамики фолдинга. Внесение мотивов, которые будут вызывать приостановку трансляции или сдвиг рамки считывания, может изменять содержание иРНК за счет различных механизмов процессирования иРНК. Модифицирование консенсусных сплайсинговых последовательностей также будет изменять содержание и распространение различных видов РНК. Изменение соотношения различных сплайсинговых изоформ будет предсказуемо приводить к изменению соотношения продуктов трансляции белка, при этом это может применяться для изменения многих биологических путей. Например, сдвиг баланса белков репарации митохондриальной и ядерной ДНК будет изменять устойчивость различных видов рака в отношении химиотерапевтических реагентов. Более того, РНК могут быть модифицированы с использованием последовательностей, которые обеспечивают связывание с новыми белками-мишенями. Был разработан ряд РНК аптамеров, которые связываются с клеточными белками с высокой аффинностью. Внесение одного из данных аптамеров может применяться либо для отделения различных видов РНК за счет связывания с белком-мишенью, который будет предотвращать их трансляцию, биологическую активность, или для направления различных видов РНК в определенные субклеточные компартменты. Расщепление биомолекул является другим классом локализационной модификации. Например, метилирование РНК применяется для регулирования РНК в клетке. Консенсусные мотивы для метилирования могут быть введены в РНК-мишени, кодирующие последовательности, с помощью РЕ. РНК также могут быть модифицированы для включения последовательностей, которые направляют механизм несмыслового расщепления или другие пути метаболизма нуклеиновой кислоты для расщепления вида-мишени РНК и изменяют пул РНК в клетке. Кроме того, определенные виды РНК могут быть модифицированы для изменения их состояния агрегации. Последовательности могут быть внесены в отдельный представляющий интерес вид

РНК или несколько видов РНК для получения клубков РНК, которые будут делать их неэффективными субстратами для трансляции или передачи сигнала.

[1453] Модификация белков с использованием посттрансляционной модификации (PTM) также является важным классом биомолекулярных манипуляций, которые могут быть реализованы с помощью РЕ. Как и в случае с отдельными видами РНК, изменение содержания белков в клетке является важной способностью РЕ. Редактирование может быть осуществлено для внесения стоп-кодонов в открытую рамку считывания — это устранит продуцирование полноразмерного продукта подвергнутой редактированию последовательностью ДНК. В качестве альтернативы могут быть внесены пептидные мотивы, которые вызывают изменение степени расщепления белка для белка-мишени. Внесение меток расщепления в тело гена может использоваться для изменения содержания белка в клетке. Более того, внесение дегронов, которые индуцируются малыми молекулами, может обеспечивать временно́й контроль над расщеплением белка. Это может являться важным как для исследований, так и для лечения, поскольку исследователи могут быстро определить являлось ли опосредованное малой молекулой расщепление терапевтического белка определенной мишени жизнеспособной терапевтической стратегией. Мотивы белка также могут быть внесены для изменения субклеточной локализации белка. Аминокислотные мотивы могут быть внесены для предпочтительного направления белков к ряду субклеточных компартментов, включая ядро, митохондрии, клеточную мембрану, пероксисому, лизосому, протеасому, экзосому, и другие.

[1454] Внесение или уничтожение мотивов, модифицированных с помощью механизма PTM, может изменять посттрансляционные модификации белка. Было идентифицировано фосфорилирование, убиквитинилирование, гликозилирование, липидирование (например фарнесилирование, миристоилирование, палмитоилирование, пренилирование, GPI якори), гидроксилирование, метилирование, ацетилирование, кротонилирование, SUMOилирование, образование дисульфидных связей, события в виде расщепления связей боковой цепи, события в виде расщепления полипептидного каркаса (протеолиз), и ряд других PTM белка. Данные PTM изменяют функцию белка, зачастую за счет изменения субклеточной локализации. Действительно, киназы зачастую активируют последующие каскады передачи сигнала за счет событий в виде фосфорилирования. Удаление фосфозита-мишени будет предотвращать трансдукцию сигнала. Способность специфическим в отношении

участка образом устранять или вносить любой PTM мотив при сохранении экспрессии полноразмерного белка будет важным достижением как для исследований, так и для терапевтического использования. Объем вносимых последовательностей и целевое окно PE делают его хорошо подходящим для широкого диапазона PTM модификаций.

[1455] Удаление участков липидирования должно предотвращать направление белков к клеточным мембранам. Заметным ограничением текущих видов лечения, нацеленных на процессы посттрансляционного модифицирования, является их специфичность. Ингибиторы фарнезилтрансферазы интенсивно тестировались на предмет их способности устранять локализацию KRAS на клеточных мембранах. К сожалению, глобальное ингибирование фарнезилирования связано с множеством нецелевых эффектов, которые препятствуют широкому использованию данных малых молекул. Аналогичным образом, специфичное ингибирование протеинкиназ с использованием малых молекул может быть очень затруднено вследствие большого размера генома человека и сходства между различными киназами. PE предлагает потенциальное решение данной проблемы, связанной со специфичностью, поскольку оно обеспечивает ингибирование модификации белка-мишени путем устранения участка модифицирования вместо глобального ингибирования фермента. Например, удаление лапидированного пептидного мотива в KRAS будет представлять собой целевой подход, который может использоваться вместо ингибирования фарнезилтрансферазы. Данный подход является функционально обратным ингибированию активности белка-мишени за счет внесения нацеленного на липид мотива в белок, который не предназначен для связывания с мембраной.

[1456] PE также может применяться для стимулирования событий в виде образования комплексов белок-белок. Белки зачастую функционируют в рамках комплексов для реализации их биологической активности. PE может использоваться для создания или устранения способности белков существовать в рамках таких комплексов. Для устранения событий в виде образования комплексов могут быть внесены аминокислотные замещения или вставки вдоль интерфейса белок:белок для исключения образования комплексов. SSX18 представляет собой белковый компонент комплекса BAF, который является важным комплексом, обеспечивающим ремоделирование гистонов. Мутации в SSX18 вызывают синовиальные саркомы. PE может использоваться для внесения боковых цепей, которые предотвращают связывание SSX18 с его белками-партнерами в рамках комплекса, для предотвращения

его онкогенной активности. PE также может применяться для удаления патогенных мутаций для восстановления активности данного белка, соответствующей дикому типу. В качестве альтернативы, PE может применяться для сохранения белков в рамках их нативного комплекса или для их привлечения для участия во взаимодействиях, которые не относятся к их нативной активности, для ингибирования их активности. Образующиеся комплексы, которые сохраняют одно состояние взаимодействия вместо другого, могут обеспечивать важную терапевтическую модальность. Изменение интерфейсов белок:белок для снижения Kd взаимодействия будет обеспечивать более длительное связывание данных белков. Белковые комплексы могут иметь несколько комплексов передачи сигнала, например как n-myc, запускающий каскады передачи сигнала нейробластомы при заболевании, но в иных случаях принимающий участие в здоровом транскрипционном контроле в других клетках. PE может быть использовано для внесения мутаций, которые запускают ассоциацию n-myc со соответствующими здоровому состоянию партнерами по взаимодействию и снижают его аффинность в отношении онкогенных партнеров по взаимодействию.

Ссылки, приведенные в Примере 14

[1457] Каждая из следующих ссылок приведена здесь в качеств ссылки.

1. Selective Target Protein Degradation via Phthalimide Conjugation. Winter et al. Science. Author manuscript; available in PMC 2016 Jul 8.

[1458] 2. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. Kadoch and Crabtree. Cell. 2013 Mar 28;153(1):71-85. doi: 10.1016/j.cell.2013.02.036.

3. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. Atkins et al. Nucleic Acids Research, Volume 44, Issue 15, 6 September 2016, Pages 7007–7078.

4. Transcriptional Regulation and its Misregulation in Disease. Lee and Young. Cell. Author manuscript; available in PMC 2014 Mar 14.

5. Protein localization in disease and therapy. Mien-Chie Hung, Wolfgang Link Journal of Cell Science 2011 124: 3381-3392.

6. Loss of post-translational modification sites in disease. Li et al. Pac Symp Biocomput. 2010:337-47. PTMD: A Database of Human Disease-associated Post-translational Modifications. Xu et al. Genomics Proteomics Bioinformatics. 2018 Aug;16(4):244-251. Epub 2018 Sep 21.

7. Post-transcriptional gene regulation by mRNA modifications. Zhao et al. Nature Reviews Molecular Cell Biology volume18, pages31–42 (2017).

ПРИМЕР 15 - КОНСТРУИРОВАНИЕ И МОДИФИЦИРОВАНИЕ PEgRNA

[1459] Здесь описана группа конструкций PEgRNA и стратегий, которые могут повышать эффективность праймерного редактирования (PE).

[1460] Праймерное редактирование (PE) представляет собой технологию геномного редактирования, которая может обеспечивать замену, вставку, или удаление определенных последовательностей ДНК в генетическом локусе-мишени с использованием информации, кодируемой в направляющей РНК для праймерного редактирования (PEgRNA). Праймерные редакторы (PE) состоят из программируемого на основе последовательности связывающего ДНК белка с нуклеазной активностью (Cas9), слитого с ферментом в виде обратной транскриптазы (RT). PE образуют комплексы с PEgRNA, которые содержат информацию для нацеливания определенных локусов ДНК в рамках их спейсерных последовательностей, а также информацию с указанием требуемого изменения в сконструированном удлинении, встроенном в стандартный каркас sgРНК. Комплексы PE:PEgRNA связывают и никируют программируемый локус-мишень ДНК, обеспечивая гибридизацию никированной цепи ДНК со сконструированной последовательностью связывания праймера (PBS) PEgRNA. Затем домен обратной транскриптазы копирует кодирующую изменение информацию в часть в виде RT матрицы PEgRNA с использованием никированной геномной ДНК в качестве праймера для полимеризации ДНК. Последующие процессы репарации ДНК обеспечивают встраивание новой синтезированной подвергнутой редактированию цепи ДНК в геномный локус. В то время как универсальность праймерного редактирования является многообещающей в качестве исследовательского инструмента и потенциального способа лечения, существует несколько ограничений в плане эффективности и объема вследствие многоэтапного процесса, требуемого для редактирования. Например, нежелательные структуры РНК, которые могут образовываться в PEgRNA, могут приводить к ингибированию копирования изменений ДНК из PEgRNA в геномный локус. Один потенциальный способ усовершенствования технологии PE заключается в изменении конструкции и модифицировании критического компонента PEgRNA. Усовершенствования конструкции данных PEgRNA вероятно будут необходимы для повышения эффективности PE, а также обеспечения внесения более длинных вставляемых последовательностей в геном.

[1461] Здесь описана группа конструкций PEgRNA, которые, как предполагается, будут повышать эффективность PE. Данные конструкции основаны на ряде ранее описанных подходов для повышения эффективности и/или стабильности sgРНК, также на использовании ряда новых стратегий. Данные улучшения могут относиться к одной или нескольким из ряда различных категорий:

[1462] **(1) Более длинные PEgRNA.** Данная категория связана с улучшенными конструкциями, которые обеспечивают эффективное экспрессирование функциональных PEgRNA из не относящихся к полимеразе III (pol III) промоторов, которые могут обеспечить экспрессию более длинных PEgRNA без обременительных требований в плане последовательности;

[1463] **(2) Усовершенствования ядра.** Данная категория связана с усовершенствованиями ядра, Cas9-связывающего каркаса PEgRNA, которые могут повышать эффективность;

[1464] **(3) Процессирование RT.** Данная категория связана с модификациями PEgRNA, которые повышают процессирование RT, обеспечивая вставку более длинных последовательностей в геномном локусе-мишени; и

[1465] **(4) Концевые мотивы.** Данная категория связана с добавлением мотивов РНК к 5′ и/или 3′-концам PEgRNA, которые повышают стабильность PEgRNA, повышают процессирование RT, предотвращают неправильный фолдинг PEgRNA, или рекрутируют дополнительные факторы, являющиеся важными для геномного редактирования.

[1466] Здесь описан ряд подобных потенциальных конструкций PEgRNA в рамках каждой категории. Некоторые из данных конструкций были ранее описаны для повышения активности sgРНК с Cas9 и отмечены соответствующим образом. Здесь также описана платформа для эволюции PEgRNA для мишеней в виде определенной последовательности, которая будет обеспечивать совершенствование каркаса PEgRNA и повышение активности PE (5). Следует отметить, что данные конструкции также могут легко применяться для усовершенствования PEgRNA, распознаваемых любым Cas9 или его подвергнутым эволюции вариантом.

[1467] (1) Более длинные PEgRNA.

[1468] sgРНК обычно экспрессируются из промотора мяРНК U6. Данный промотор рекрутирует pol III для экспрессии ассоциированной РНК и может применяться для экспрессии коротких РНК, которые остаются в ядре. Однако pol III не является

высокопроцессивной и не способна экспрессировать РНК с длиной более чем несколько сотен нуклеотидов на уровнях, требуемых для эффективного геномного редактирования[183]. Кроме того, pol III может вызывать остановку или терминирование на длинных участках в виде U, что потенциально ограничивает разнообразие цепей, которые могут быть вставлены с помощью PEgRNA. Были проанализированы другие промоторы, которые рекрутируют полимеразу II (например pCMV) или полимеразу I (например промотор мяРНК U1) на их способность экспрессировать более длинные sgРНК[183]. Однако данные промоторы обычно частично транскрибируются, что будет приводить к дополнительной последовательности 5' спейсера в экспрессируемой PEgRNA, что, как было показано, приводит к заметному снижению активности Cas9:sgРНК в зависимости от участка. Кроме того, в то время как транскрибированные из pol III PEgRNA могут просто приводить к терминации на участке из 6-7 U, PEgRNA, транскрибированные из pol II или pol I, будут требовать другого сигнала терминации. Зачастую такие сигналы также приводят к полиаденилированию, которое приведет к нежелательному транспорту PEgRNA из ядра. Аналогичным образом, РНК, экспрессируемые из pol II промоторов, например pCMV, обычно являются 5'-кэпированными, что также приводит к их экспорту из ядра.

[1469] Ранее Rinn и коллеги провели скрининг ряда платформ экспрессии на предмет продуцирования sgРНК, меченных длинными некодирующими РНК (днкРНК)[183]. Данные платформы включают РНК, экспрессированные из pCMV, которые приводят к терминации в элементе ENE из нкРНК MALAT1 человека[184], элементе PAN ENE из KSHV[185], или 3' боксе из sgРНК U1[186]. Следует отметить, что нкРНК MALAT1 и PAN ENE образуют тройные спирали, защищающие polyA-хвост[184, 187]. Ожидается, что помимо обеспечения экспрессии РНК, данные конструкты также могут повышать стабильность РНК (см. раздел iv). Также анализировали использование промотора из sgРНК U1 для обеспечения экспрессии данных более длинных sgРНК[183]. Предполагается, что данные системы экспрессии также обеспечат возможность экспрессии более длинных PEgRNA. Кроме того, был разработан ряд способов для расщепления части pol II промотора, которая будет транскрибирована в виде части PEgRNA, обеспечивая добавление либо саморасщепляемого рибозима, например рибозимов типа "головка молотка" (hammerhead)[188], "пистолет" (pistol)[189], "топор" (hatchet)[189], "шпилька" (hairpin)[190], VS[191], "твистер"[192], или "твистер-систер" (twister sister)[192], или других саморасщепляемых элементов для процессирования

транскрибированной направляющей, или шпильки, которая распознается Csy4[193] и также ведет к процессированию направляющей. Кроме того, существует гипотеза, что встраивание нескольких мотивов ENE может приводить к повышенной экспрессии PEgRNA и стабильности, как ранее было продемонстрировано для РНК KSHV PAN и элемента[185]. Также ожидается, что циркуляризация PEgRNA в форме циркулярной интронной РНК (циРНК) также может приводить к повышенной экспрессии РНК и стабильности, а также ядерной локализации[194].

[1470] Платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и MALAT1 ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTT
CTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGAT
GCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCT
TTGCTTTGACT (SEQ ID NO: 757)

[1471] Платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA

TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAA (SEQ ID NO: 758)

[1472] Платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3xPAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAG
ACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATT
TAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACA
TAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAACACACTGTT
TTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGG
TTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCT
AGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGG
TCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAATCTCTCTGTTTTGGCTGGGTT
TTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTATCCCAG
TGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAAT
TAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGGTCAATGTTTAAT
CCATAAAAAAAAAAAAAAAAAA (SEQ ID NO: 759)

[1473] Платформа экспрессии PEgRNA, состоящая из pCMV, шпильки Csy4, PEgRNA, и 3′ бокса

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC
CTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCG
TGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC
AGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGA
CTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTT
ATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGT
CTGTTTGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTGGTTT
GTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 760)

[1474] Платформа экспрессии PEgRNA, состоящая из pU1, шпильки Csy4, PEgRNA, и 3′ бокса

CTAAGGACCAGCTTCTTTGGGAGAGAACAGACGCAGGGGCGGGAGGGAAAAAG
GGAGAGGCAGACGTCACTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCGGTTGA
GTGGCAGAAAGGCAGACGGGGACTGGGCAAGGCACTGTCGGTGACATCACGGAC
AGGGCGACTTCTATGTAGATGAGGCAGCGCAGAGGCTGCTGCTTCGCCACTTGCT
GCTTCACCACGAAGGAGTTCCCGTGCCCTGGGAGCGGGTTCAGGACCGCTGATCG
GAAGTGAGAATCCCAGCTGTGTGTCAGGGCTGGAAAGGGCTCGGGAGTGCGCGG
GGCAAGTGACCGTGTGTGTAAAGAGTGAGGCGTATGAGGCTGTGTCGGGGCAGA
GGCCCAAGATCTCAGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG
TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAA
AGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTCAGCAAG
TTCAGAGAAATCTGAACTTGCTGGATTTTTGGAGCAGGGAGATGGAATAGGAGCT
TGCTCCGTCCACTCCACGCATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGC
ACCCACTTTCTGGAGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTT
GTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 761)

[1475] (2) Усовершенствования ядра.

[1476] Ядро, представляющее собой Cas9-связывающий каркас PEgRNA, может быть вероятно усовершенствовано для повышения активности PE. Некоторые подобные подходы уже были продемонстрированы. Например, первый элемент связывания каркаса (P1) содержит элемент связывания GTTTT-AAAAC. Было показано, что такие повторы T приводят к остановке pol III и преждевременной терминации РНК транскрипта. Было показано, что разумная мутация одной из пар T-A на пару G-C в этой части P1 повышает активность sgРНК, что свидетельствует о том, что данный подход также может применяться для PEgRNA[195]. Кроме того, также было показано, что повышение длины P1 стимулирует фолдинг sgРНК и ведет к повышенной активности[195], что свидетельствует о том, что оно является другой возможностью повысить активность PEgRNA. Наконец, усовершенствование каркаса PEgRNA за счет направленной эволюции PEgRNA на определенной ДНК-мишени вероятно также приведет к повышенной активности. Это описано в разделе (v).

[1477] PEgRNA, содержащая удлинение P1 длиной 6 nt

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGCTCATGAAAATGAGCTAGCAAG
TTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTC
TGCCATCAAAGCGTGCTCAGTCTGTTTTTTT (SEQ ID NO: 228)

[1478] PEgRNA, содержащая мутацию T-A на G-C в P1

GGCCCAGACTGAGCACGTGAGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC
GTGCTCAGTCTGTTTTTTT (SEQ ID NO: 229)

[1479] (iii) Улучшение процессирования RT за счет модификаций матричного участка PEgRNA

[1480] С увеличением размера вставки, матрицей которого служит PEgRNA, она с большей вероятностью будет расщепляться эндонуклеазами, подвергаться спонтанному гидролизу или фолдингу во вторичные структуры, не способные подвергаться обратной транскрипции за счет RT или нарушающие фолдинг каркаса PEgRNA и последующее связывание Cas9-RT. Соответственно, является вероятным, что для реализации крупных вставок может быть необходимым модифицирование матрицы PEgRNA, например для вставок целых генов. Некоторые стратегии для этого включают встраивание модифицированных нуклеотидов в синтетическую или полусинтетическую PEgRNA, которые делают РНК более устойчивой к расщеплению или гидролизу или менее подверженной включению ингибиторных вторичных структур[196]. Подобные модификации могут включать 8-аза-7-деазагуанозин, который будет снижать содержание вторичной структуры РНК в последовательностях с высоким содержанием G; "заблокированные" нуклеиновые кислоты (LNA), которые снижают расщепление и повышают определенные виды вторичной структуры РНК; модификации в виде 2'-O-метил, 2'-фтор, или 2'-O-метоксиэтокси, которые повышают стабильность РНК. Такие модификации также могут быть включены в другие части PEgRNA для повышения стабильности и активности. В качестве альтернативы или дополнительно, матрица PEgRNA может быть сконструирована таким образом, что она как кодирует требуемый продукт в виде белка, так и с большей вероятностью включает простые вторичные структуры, которые способны подвергаться обратному фолдингу за счет RT. Такие простые структуры будут действовать в качестве термодинамического стока, делая менее вероятным появление более сложных структур, предотвращающих обратную транскрипцию. Наконец, также является возможным разделить матрицу на две отдельные PEgRNA. При такой конструкции PE будет использоваться для инициирования транскрипции, а также рекрутирования отдельной матричной РНК к являющемуся мишенью участку с помощью РНК-связывающего белка, слитого с Cas9, или элемента распознавания РНК на самой PEgRNA, например аптамера MS2. RT может либо напрямую связываться с данной отдельной матричной РНК, или инициировать обратную транскрипцию на исходной PEgRNA перед заменой на вторую матрицу. Такой подход может обеспечивать возможность осуществления длинных вставок за счет как предотвращения неправильного фолдинга PEgRNA при добавлении длинной матрицы, так и за счет

отсутствия необходимости в диссоциации Cas9 из генома для реализации длинных вставок, что вероятно может ингибировать длинные вставки на основе PE.

[1481] (iv) Внесение дополнительных мотивов РНК на 5′ или 3′-концах

[1482] Конструкции PEgRNA также могут быть улучшены ха счет внесения дополнительных мотивов на любом конце РНК. Некоторые подобные мотивы, например PEN ENE из KSHV и ENE из MALAT1, были описаны выше в части (i)[184,185] в качестве возможных способов терминации экспрессии длинных PEgRNA из не относящихся к pol III промоторов. Эти элементы образуют тройные спирали РНК, которые окружают polyA-хвост, что приводит к их удержанию в ядре[184,187]. Однако за счет образования сложных структур на 3′-конце PEgRNA, которые закрывают концевой нуклеотид, данные структуры также вероятно будут помогать предотвращать опосредованное экзонуклеазой расщепление PEgRNA. Другие структурные элементы, вставленные на 3′-конце, также могут повышать стабильность РНК, хотя и без обеспечения терминации из не относящихся к pol III промоторов. Такие мотивы могут включать шпильки или квадруплексы РНК, которые будут закрывать 3′-конец[197], или саморасщепляющиеся рибозимы, например HDV, которые будут приводить к образованию 2’-3′-циклического фосфата на 3′-конце, а также потенциально делать PEgRNA менее подверженной расщеплению экзонуклеазами[198]. Индуцирование циклизации PEgRNA за счет неполного сплайсинга – для образования циРНК – также может повышать стабильность PEgRNA и приводить к удержанию PEgRNA в ядре[194].

[1483] Дополнительные мотивы РНК также могут повышать процессирование RT или повышать активность PEgRNA за счет повышения связывания RT с дуплексом ДНК-РНК. Добавление нативной последовательности, связанной RT в ее родственном ретровирусном геноме, может повышать активность RT[199]. Это может включать нативный участок связывания праймера (PBS), полипуриновый тракт (PPT), или целующиеся петли, принимающие участие в димеризациии ретровирусного генома и инициировании транскрипции[199]. Добавление мотивов димеризации, таких как целующиеся петли или пары тетрапетля GNRA/рецептор тетрапетли [200], на 5′ и 3′-концах PEgRNA также может приводить к эффективной циркуляризации PEgRNA, что повышает стабильность. Кроме того, предполагается, что добавление данных мотивов может обеспечивать физическое разделение спейсера и праймера PEgRNA, предотвращать закрытие спейсера, ухудшающее активность PE. Короткие 5′ удлинения PEgRNA, которые образуют малую опорную шпильку в спейсерном участке, также

могут благоприятным образом конкурировать против участка отжига PEgRNA, связывающего спейсер. Наконец, целующиеся петли также могут использоваться для рекрутирования других матричных РНК в участок генома и обеспечения обмена RT активностью от одной РНК к другой (раздел iii).

[1484] Продукт слияния PEgRNA-HDV

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC
GTGCTCAGTCTGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAA
CATGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 230)

[1485] Целующаяся петля PEgRNA-MMLV

GGTGGGAGACGTCCCACCGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAAT
AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCG
GTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCT
TTTTTT (SEQ ID NO: 231)

[1486] Целующаяся петля PEgRNA-VS рибозим

GAGCAGCATGGCGTCGCTGCTCACGGCCCAGACTGAGCACGTGAGTTTTAGAGCT
AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACC
GAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTCCATCAGTTGACA
CCCTGAGGTTTTTTT (SEQ ID NO: 232)

[1487] PEgRNA-GNRA тетрапетля/рецептор тетрапетли

GCAGACCTAAGTGGUGACATATGGTCTGGGCCCAGACTGAGCACGTGAGTTTTAG
AGCTAUACGTAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTUACGAAGTGG
GACCGAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGCATGCGATT
AGAAATAATCGCATGTTTTTTT (SEQ ID NO: 233)

[1488] PEgRNA матрица с заменой вторичного продукта слияния в виде РНК-HDV

TCTGCCATCAAAGCTGCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCGGCCG
GCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCG
AATGGGACTTTTTTT (SEQ ID NO: 234)

[1489] (v) Эволюция PEgRNA

[1490] Каркас PEgRNA вероятно может быть дополнительно усовершенствован за счет направленной эволюции аналогично тому как были усовершенствованы SpCas9 и редакторы оснований[201]. Направленная эволюция может повышать распознавание PEgRNA Cas9 или подвергнутыми эволюции вариантами Cas9. Кроме того, является возможным, что различные последовательности каркаса PEgRNA будут оптимальными в рамках различных геномных локусов, обеспечивая повышение активности PE в рассматриваемом участке, снижая нецелевые виды активности, или осуществляя оба из данных действий. Наконец, эволюция каркасов PEgRNA, к которым были добавлены другие мотивы РНК, будет практически наверняка повышать активность слитой PEgRNA по сравнению с не подвергавшейся эволюции слитой РНК. Например,

эволюция аллостерических рибозимов, состоящих из аптамеров c-di-GMP-I и рибозимов типа "головка молотка" (hammerhead), приводила к значительному повышению активности[202], что свидетельствует о том, что эволюция также будет повышать активность PEgRNA типа "головка молотка" (hammerhead). Кроме того, в то время как Cas9 в настоящее время в целом не совместим с 5′ удлинением sgРНК, направленная эволюция вероятно обеспечит реализацию мутаций, которые устранят данную несовместимость, обеспечивая возможность использования дополнительных мотивов РНК.

[1491] Как описано здесь, ряд данных подходов уже был описан для применения с комплексами Cas9:sgРНК, однако конструкции для повышения активности PEgRNA не были описаны. Другие стратегии для внесения программируемых мутаций в геном включают редактирование оснований, направленную гомологичную репарацию (HDR), точное опосредуемое микрогомологией соединение концов (MMEJ), или опосредованное транспозазой редактирование. Однако все эти подходы имеют значительные недостатки по сравнению с PE. В то время как современные редакторы оснований являются более эффективными, чем существующие PE, они могут обеспечивать внесение лишь определенных классов геномных мутаций и могут приводить к дополнительным нежелательным нуклеотидным конверсиям в представляющем интерес участке. HDR является реализуемой только в очень малом количестве типов клеток и приводит к сравнительно высоким показателям случайных мутаций в виде вставок и делеций (инделям). Точное MMEJ может приводить к предсказуемой репарации двуцепочечных разрывов, но в целом ограничено внесением делеций, имеет сильную зависимость от участка, и также может иметь сравнительно высокие показатели нежелательных инделей. Было показано, что опосредованное транспозазой редактирование на сегодняшний день функционирует только в бактериях. Таким образом, усовершенствования PE являются вероятно наилучшим путем для терапевтического корректирования большого количества геномных мутаций.

Ссылки, приведенные в Примере 15

[1492] Каждая из следующих ссылок приведена в Примере 15, каждая из которых приведена здесь в качестве ссылки.

1. Schechner, DM, Hacisuleyman E., Younger ST, Rinn JL. Nat Methods 664-70 (2015).

2. Brown JA, et al. Nat Struct Mol Biol 633-40 (2014).

3. Conrad NA and Steitz JA. EMBO J 1831-41 (2005).
4. Bartlett JS, et al. Proc Natl Acad Sci USA 8852-7 (1996).
5. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. Science 1244-7 (2010).
6. Forster AC, Symons RH. Cell. 1987.
7. Weinberg Z, Kim PB, Chen TH, Li S, Harris KA, Lünse CE, Breaker RR. Nat. Chem. Biol.
2015.
8. Feldstein PA, Buzayan JM, Bruening G. Gene 1989.
9. Saville BJ, Collins RA. Cell. 1990.
10. Roth A, Weinberg Z, Chen AG, Kim PG, Ames TD, Breaker RR. Nat Chem Biol. 2013.
11. Borchardt EK, et al. RNA 1921-30 (2015).
12. Zhang Y, et al. Mol Cell 792-806 (2013).
13. Dang Y, et al. Genome Biol 280 (2015).
14. Schaefer M, Kapoor U, and Jantsch MF. Open Biol 170077 (2017).
15. Nahar S, et al. Chem Comm 2377-80 (2018).
16. Gao Y and Zhao Y. J Integr Plant Biol 343-9 (2014).
17. Dubois N, Marquet R, Paillart J, Bernacchi S. Front Microbiol 527 (2018).
18. Costa M and Michel F. EMBO J 1276-85 (1995).
19. Hu JH, et al. Nature 57-63 (2018).
20. Furukawa K, Gu H, Breaker RR. Methods Mol Biol 209-20 (2014).

ПРИМЕР 16 - РАСШИРЕНИЕ ОБЪЕМА НАЦЕЛИВАНИЯ PE С ПОМОЩЬЮ СВЯЗЫВАЮЩИХ ДНК БЕЛКОВ, ОТЛИЧНЫХ ОТ SPCAS9

[1493] Праймерное редактирование (PE) с использованием Cas9 (SpCas9) *Streptococcus pyogenes* может обеспечивать эффективное внесение всех замещений отдельных оснований, вставок, делеций, и их комбинаций в геномных локусах с подходящим образом расположенным смежным с протоспейсером мотивом NGG (PAM), который может быть эффективно связан SpCas9. Описанные здесь способы расширяют способность нацеливания PE за счет увеличения доступных PAM и, таким образом, геномных локусов-мишеней, доступных для эффективного PE. Праймерные редакторы на основе использования РНК-направляемых связывающих ДНК белков, отличных от SpCas9, обеспечивают расширение области нацеливания в плане геномных локусов за счет предоставления доступа к различным PAM. Кроме того, использование РНК-направляемых связывающих ДНК белков с размером менее чем у SpCas9 также

обеспечивает более эффективную вирусную доставку. PE с белками Cas или другими РНК-направляемыми связывающими ДНК белками помимо SpCas9 будет обеспечивать высокоэффективные терапевтические изменения, которые являлись недоступными или неэффективными при использовании PE на основе SpCas9.

[1494] Ожидается, что это будет использоваться в ситуациях, где PE на основе SpCas9 является либо неэффективным вследствие неидеального расположения изменения относительно PAM NGG, или общий размер конструкта на основе SpCas9 препятствует клеточной экспрессии и/или доставке. Определенные связанные с заболеваниями локусы, такие как ген хантингтина, который имеет редкие и неудачно расположенные PAM NGG для SpCas9 рядом с участком-мишенью, могут быть легко нацелены с использованием различных белков Cas в рамках системы PE, например SpCas9-VRQR, который распознает PAM NGA. Белки Cas меньшего размера будут использоваться для получения конструктов PE меньшего размера, которые могут быть упакованы в AAV векторы более эффективным образом, обеспечивая улучшенную доставку в ткани-мишени. На Фиг.61 показан переход к практике праймерного редактирования с использованием CRISPR-Cas на основе *Staphylococcus aureus* в качестве РНК-направляемого связывающего ДНК белка. NT представляет собой контроль без обработки.

[1495] На Фиг.62A-62B продемонстрирована важность протоспейсера для эффективного внесения требуемого изменения в точном месторасположении с помощью праймерного редактирования. Это подчеркивает важность альтернативных PAM и протоспейсеров в качестве новых свойств данной технологии. "н.д." на Фиг.62A означает "не детектирован".

[1496] На Фиг.63 показан переход к практике PE и использованием SpCas9(H840A)-VRQR и SpCas9(H840A)-VRER в качестве РНК-направляемого связывающего ДНК белка в рамках системы праймерного редактора. napDNAbp SpCas9(H840A)-VRQR описан здесь как SEQ ID NO: 87. napDNAbp SpCas9(H840A)-VRER описан здесь как SEQ ID NO: 88. Слитый белок SpCas9(H840A)-VRER-RT MMLV описан здесь как SEQ ID NO: 516, в то время как RT MMLV включает D200N, L603W, T330P, T306K, и W313F замещения по сравнению с RT MMLV дикого типа. Слитый белок SpCas9(H840A)-VRQR-RT MMLV описан здесь как SEQ ID NO: 515, в то время как RT MMLV включает D200N, L603W, T330P, T306K, и W313F замещения по сравнению с RT MMLV дикого типа. Мишенями являются семь различных локусов в геноме

человека: 4 для системы праймерного редактора SpCas9(H840A)-VRQR- RT MMLV и 3 для системы SpCas9(H840A)-VRER-RT MMLV. Аминокислотные последовательности тестируемых конструктов являются следующими:

ПРАЙМЕРНЫЙ РЕДАКТОР SACAS9-RT M-MLV	MKRTADGSEFESPKKKRKVGKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK LLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLA KRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLE RLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFID TYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPE ELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYH DIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKL VPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYG LPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGK ENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHII PRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETF KKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVD TRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFK KERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMF EEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKK PNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLIN KSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGN YLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE EAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIE VNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNL YEVKSKKHPQIIKKGSGGSSGGSSGSETPGTSESATPESSGGSSG GSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGL AVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGIL VPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWR DPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDL ILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQ ICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRR PVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVV ALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAEL IALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSE GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFEPKKKR KV (SEQ ID NO: 660)
ПРАЙМЕРНЫЙ РЕДАКТОР	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR

SPCAS9(H840A)-VRQR-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 661)
ПРАЙМЕРНЫЙ РЕДАКТОР	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR

SPCAS9(H840A)-VRER-ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУСА ЛЕЙКОЗА МЫШЕЙ МОЛОНИ	RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVL DATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESAT PESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLAD FRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYR ASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSE FEPKKKRKV (SEQ ID NO: 662)

[1497] Как показано на Фиг.63, SpCas9(H840A)-VRQR-RT MMLV являлся функциональным в участках PAM, которые включали "AGAG" и "GGAG", и имел некоторую связанную с редактированием активность в последовательностях PAM "GGAT" и "AGAT". SpCas9(H840A)-VRER-RT MMLV являлся функциональным в участках PAM, которые включали "AGCG" и "GGCG", и имел некоторую связанную редактированием активность в "TGCG".

[1498] Эти данные демонстрируют, что праймерное редактирование может быть реализовано с помощью napDNAbp с различной специфичностью в отношении PAM, например описанных здесь вариантов Cas9.

ПРИМЕР 17 – ВСТРАИВАНИЕ УЧАСТКОВ-МИШЕНЕЙ РЕКОМБИНАЗЫ С ПОМОЩЬЮ PE

[1499] В данном Примере описан способ лечения генетического заболевания или получения модифицированных моделей на основе животных или растений с помощью праймерного редактирования (PE) для внесения участков-мишеней рекомбиназы (участков-мишеней SSR) в относящихся к млекопитающим и других геномах с высокой специфичностью и эффективностью.

[1500] В данном Примере описано применение PE для внесения последовательностей распознавания рекомбиназой в имеющие высокую ценность локусы генома человека или другие геномы, которые после подвергания действию специфичной в отношении участка рекомбиназы (рекомбиназ) будут обеспечивать точные и эффективные геномные модификации (Фиг.64). В различных вариантах осуществления, показанных на Фиг.64, PE может применяться для (b) вставки отдельной мишени SSR для использования в качестве участка для геномного встраивания донорской матрицы ДНК. (c) показано, как тандемная вставка участков-мишеней SSR может использоваться для делеции части генома. (d) показано, как тандемная вставка участков-мишеней SSR может использоваться для инверсии части генома. (e) показано, как вставка двух участков-мишеней SSR в двух удаленных хромосомных участках может приводить к хромосомной транслокации. (f) показано, как вставка двух различных участков-мишеней SSR в геноме может использоваться для обмена кассетами из донорской матрицы ДНК. Каждый из данных типов геномных модификаций охватывается при применении PE для вставки мишеней SSR, однако данный перечень также не является ограничивающим.

[1501] При генетическом заболевании возникают многие крупномасштабны изменения, такие как вставки генов, делеции, инверсии, или хромосомные транслокации[1-7]. Кроме того, специализированное и нацеленное манипулирование эукариотическими геномами является важным для исследования заболеваний человека, а также получения трансгенных растений[8,9] или других биотехнологических продуктов. Например, микроделеции хромосом могут приводить к заболеванию, при этом замещение данных делеций путем вставок критических ДНК-элементов может приводить к перманентному облегчению заболевания. Кроме того, заболевания, возникающие вследствие инверсий, изменений количества копий генов, или хромосомных транслокаций, могут подвергаться лечению путем восстановления предшествующей генетической структуры в затронутых клетках. В качестве альтернативы, в растениях и других имеющих высокую ценность эукариотических организмах, используемых в промышленности, внесение рекомбинантной ДНК или нацеленных геномных изменений может приводить к получению усовершенствованных продуктов, например зерновых культур, которые требуют меньшего количества ресурсов или являются устойчивыми к патогенам. Современные технологии для осуществления эффективных крупномасштабных геномных изменений основаны на случайных или стохастических процессах, например использовании транспозонов или ретровирусов, в то время как другие желаемые геномные модификации были реализованы лишь с помощью стратегий на основе гомологичной рекомбинации.

[1502] Один привлекательный класс белков для реализации нацеленного и эффективного геномного модифицирования представляет собой специфические в отношении участка рекомбиназы (SSR). SSR имеют длительную историю использования в качестве инструмента для геномного модифицирования[10-13]. SSR считаются многообещающими инструментами для генной терапии, поскольку они катализируют точное расщепление, обмен цепями, и повторное соединение фрагментов ДНК в определенных мишенях рекомбинации[14] без использования эндогенной репарации двуцепочечных разрывов, которые могут индуцировать индели, транслокации, другие виды перестроения ДНК, или активацию p53[15-18]. Реакции, катализируемые SSR, могут приводить к прямому замещению, вставке, или делеции фрагментов-мишеней ДНК с эффективностью, превосходящей эффективность направленной гомологичной репарации[14,19].

[1503] Несмотря на то, что SSR обеспечивают многие преимущества, они не используются широким образом, поскольку они имеют сильное естественное предпочтение в плане соответствующей им последовательности-мишени. Последовательности распознавания SSR обычно имеют длину ≥ 20 пар оснований и, таким образом, редко встречаются в геноме человека или организмов-моделей. Более того, нативные предпочтения SSR в плане субстратов не являются легкими в плане изменения даже при использовании сложного конструирования в рамках лаборатории или эволюции[20]. Это ограничение преодолевается за счет использования PE для прямого внесения мишеней рекомбиназы в геном или для модификации эндогенных геномных последовательностей, которые естественным образом имеют сходство с мишенями рекомбиназы. Последующее подвергание клетки действию белка в виде рекомбиназы обеспечит точное и эффективное геномное модифицирование, направляемое на основе расположения и ориентации мишени (мишеней) рекомбиназы (Фиг.64).

[1504] PE-опосредованное внесение мишеней рекомбиназы может быть особенно полезным при лечении генетических заболеваний, которые вызываются крупными геномными дефектами, такими как потеря генов, инверсия, или дупликация, или хромосомная транслокация[1-7] (Таблица 6). Например, синдром Вильямса-Бойрена представляет собой нарушение развития, вызванное делецией 24 в хромосоме 721. В настоящее время не существует технологии для эффективной и нацеленной вставки нескольких целых генов в живые клетки (потенциал PE в плане реализации подобной вставки полноразмерных генов в настоящее время оценивается, но еще не был подтвержден); однако опосредуемое рекомбиназой встраивание в мишени, вставленной с помощью PE, является одним из возможных подходов к перманентному излечению данного и других заболеваний. Кроме того, нацеленное внесение последовательностей распознавания рекомбиназой может эффективно обеспечивать различные варианты применения, включая получения трансгенных растений, исследовательских моделей на основе животных, биопроизводственных линий клеток, или других линий модифицированных эукариотических клеток. Например, опосредуемое рекомбиназой геномное модифицирование в трансгенных растениях в PE-специфичных мишенях может обеспечивать преодоление одного из препятствий в плане получения сельскохозяйственных растений с улучшенными свойствами[8,9].

[1505] Был охарактеризован ряд членов семейства SSR, а их последовательности-мишени описаны, включая природные и модифицированные тирозиновые рекомбиназы (Таблица 7), крупные сериновые интегразы (Таблица 8), сериновые резольвазы (Таблица 9), и тирозиновые интегразы (Таблица 10). Модифицированные последовательности-мишени, которые демонстрируют повышенные показатели геномного встраивания, также были описаны для нескольких SSR[22-30]. Помимо природных рекомбиназ, были разработаны программируемые рекомбиназы с различной специфичностью[31-40]. С помощью PE одна или нескольких из данных последовательностей распознавания может быть внесена в геном в определенном месте, например в локус "безопасная гавань"[41-43], в зависимости от требуемого варианта использования.

[1506] Например, внесение одной мишени рекомбиназы в геном будет приводить к интегративной рекомбинации на основе донорской матрицы ДНК (Фиг.64b). Сериновые интегразы, которые функционируют повсеместно в рамках клеток человека, могут являться особенно подходящими для генного встраивания[44,45]. Кроме того, внесение двух мишеней рекомбиназы может приводить к делеции промежуточной последовательности, инверсии промежуточной последовательности, хромосомной транслокации, или обмену кассетами, в зависимости от вида и ориентации мишеней (Фиг.64c-f). За счет выбора эндогенных последовательностей, которые уже близко напоминают мишени рекомбиназы, объем редактирования, необходимого для внесения полной мишени рекомбиназы, будет снижен.

[1507] Наконец, было продемонстрировано, что несколько рекомбиназ встраиваются в геном человека или эукариотические геномы в нативно расположенных псевдосайтах[46-64]. PE может применяться для модификации данных локусов для повышения показателей интеграции в данных природных псевдосайтах или, в качестве альтернативы, для устранения псевдосайтов, которые могут служить нежелательными нецелевыми последовательностями.

[1508] В рамках данного описания описана общая методология внесения последовательностей-мишеней рекомбиназы в эукариотические геномы с помощью PE, варианты использования которого являются практически неограниченными. Реакции геномного редактирования предназначены для применения с "праймерным редактором", который представляет собой химерный продукт слияния CRISPR/белка Cas9 и домена обратной транскриптазы, который использует модифицированную

направляющую РНК для праймерного редактирования (PEgRNA). За счет удлинения инструменты на основе Cas9 и пути направленной гомологичной репарации (HDR) также могут использоваться для внесения мишеней рекомбиназы с помощью ДНК-матриц путем понижения показателей инделей с помощью ряда способов[65-67]. Эксперимент для подтверждения данной концепции в рамках культуры клеток человека показан на Фиг.65.

Таблица 6. Примеры генетических заболеваний, связанных с крупными геномными модификациями.

Заболевание	**Источник**	**Причина**
Трисомия 17p	68	Дупликация генов
Болезнь Шарко-Мари-Тута тип I	69	Дупликация генов
Синдром Смита-Магениса	70	Делеция генов
Синдром Вильямса-Бойрена	21	Делеция генов
Синдром Де ла Шапель	71	Хромосомная
Синдром Дауна (некоторые формы)	72	Хромосомная
Гемофилия A	73	Инверсия генов
Синдром Хантера	74	Инверсия генов

Таблица 7. Тирозиновые рекомбиназы и последовательности мишеней SSR.

Рекомбиназа	**Источник**	**Мишень**	**Название**
Cre	75	ATAACTTCGTATAGCATACATTATACG AAGTTAT (SEQ ID NO: 517)	*loxP*
Dre	76	TAACTTTAAATAATGCCAATTATTTAA AGTTA (SEQ ID NO: 518)	*rox*
VCre	77	TCAATTTCTGAGAACTGTCATTCTCGG AAATTGA (SEQ ID NO: 519)	*loxV*
SCre	77	CTCGTGTCCGATAACTGTAATTATCGG ACATGAT (SEQ ID NO: 520)	*loxS*
Flp	78	GAAGTTCCTATTCTCTAGAAAGTATAG GAACTTC (SEQ ID NO: 521)	*FRT*
B2	79	GAGTTTCATTAAGGAATAACTAATTCC CTAATGAAACTC (SEQ ID NO: 522)	*loxB*
B3	79	GGTTGCTTAAGAATAAGTAATTCTTAA GCAACC (SEQ ID NO: 523)	*loxB3*
Kw	80	ACGAAAAATGGTAAGGAATAGACCAT	

		TCCTTACCATTTTTGGT (SEQ ID NO: 524)	
R	81	TTGATGAAAGAATAACGTATTCTTTCA TCAA (SEQ ID NO: 525)	*RS*
TD1-40	82	GTGCGTCAAATAATAACGTATTATTTG ACACTT (SEQ ID NO: 526)	*TDRS*
Vika	83	AATAGGTCTGAGAACGCCCATTCTCAG ACGTATT (SEQ ID NO: 527)	*vox*
Nigri	84	TGAATGTCCTATAATTACACTTATAGG ACATTCA (SEQ ID NO: 528)	*nox*
Panto	84	GAAACTTTAAATAATAAGTCTTATTTA AAGTTTC (SEQ ID NO: 259)	*pox*
Kd	79	AAACGATATCAGACATTTGTCTGATA ATGCTTCATTATCAGACAAATGTCTG ATATCGTTT (SEQ ID NO: 530)	*loxK*
Fre	85	ATATATACGTATATAGACATATATACG TATATAT (SEQ ID NO: 531)	*loxH*
CreALSHG	86	ATAACTCTATATAATGTATGCTATATA GAGTTAT (SEQ ID NO: 532)	*loxM7*
Tre	87	ACAACATCCTATTACACCCTATATGCC AACATGG (SEQ ID NO: 533)	*loxLTR*
Brec1	12	AACCCACTGCTTAAGCCTCAATAAAGC TTGCCTT (SEQ ID NO: 534)	*loxBTR*
Cre-R3M3	88	GATACAACGTATATACCTTTCTATACG TTGTTTA (SEQ ID NO: 535)	*loxK2*

Таблица 8. Крупные сериновые интегразы и последовательности мишеней SSR.

Интеграза	**Источник**	**Левая мишень**	**Правая мишень**
Bxb1	89	GGTTTGTCTGGTCAACCAC CGCGGTCTCAGTGGTGTAC GGTACAAACC (SEQ ID NO: 536)	GGCTTGTCGACGACGGCGG TCTCCGTCGTCAGGATCAT(S EQ ID NO: 537)
phiC31	90	GTGCCCCAACTGGGGTAAC CTTTGAGTTCTCTCAGTTG GGGG (SEQ ID NO: 538)	TGCGGGTGCCAGGGCGTGC CCTTGGGCTCCCCGGGCGCG TACTCC (SEQ ID NO: 539)
R4	91	TGTTCCCCAAAGCGATACC ACTTGAAGCAGTGGTACTG CTTGTGGGTACA (SEQ ID NO: 540)	GCATGTTCCCCAAAGCGATA CCACTTGAAGCAGTGGTACT GCTTGTGGGTACACTCTGCG GGTG (SEQ ID NO: 541)
phiBT1	92	GGTGCTGGGTTGTTGTCTC TGGACAGTGATCCATGGGA	CAGGTTTTTGACGAAAGTGA TCCAGATGATCCAG (SEQ ID

		AACTACTCAGCACC (SEQ ID NO: 542)	NO: 543)
MJ1 (phiFC1)	93	ATTTTAGGTATATGATTTTGTTTATTAGTGTAAATAACACTATGTACCTAAAAT (SEQ ID NO: 544)	CAAAGGATCACTGAATCAAAAGTATTGCTCATCCACGCGAAA (SEQ ID NO: 545)
MR11	94	TTTGTGCGGAACTACGAACAGTTCATTAATACGAAGTGTACAAACTTCCATACAA (SEQ ID NO: 546)	CGAAAATGTATGGAGGCACTTGTATCAATATAGGATGTATACCTTCGAAGACACTT (SEQ ID NO: 547)
TP901-1	95	GAGTTTTTATTTCGTTTATTTCAATTAAGGTAACTAAAAAACTCCTTTTAAGG (SEQ ID NO: 548)	ATGCCAACACAATTAACATCTCAATCAAGGTAAATGCTTTTTGCTTTTTTTGC (SEQ ID NO: 549)
A118	96	TTCCTCGTTTTCTCTCGTTGGAAGAAGAAGAAACGAGAAA (SEQ ID NO: 550)	TTTCGGATCAAGCTATGAAGGACGCAAAGAGGGAACTAAA(SEQ ID NO: 551)
U153	97	TTCCTCGTTTTCTCTCGTTGGACGGAAACGAATCGAGAAA (SEQ ID NO: 552)	TTTCGGATCAAGCTATGAAGGACGCAAAGAGGGAACTAAA (SEQ ID NO: 553)
phiRV1	98	GTAGTGTATCTCACAGGTCCACGGTTGGCCGTGGACTGCTGAAGAACATTCC (SEQ ID NO: 554)	GAAGGTGTTGGTGCGGGGTTGGCCGTGGTCGAGGTGGGGT (SEQ ID NO: 555)
phi370.1	99	AAAAAAATACAGCGTTTTTCATGTACAACTATACTAGTTGTAGTGCCTAAAA (SEQ ID NO: 556)	TTGTAAAGGAGACTGATAATGGCATGTACAACTATACTCGTCGGTAAAAAGGCA (SEQ ID NO: 557)
TG1	100	TCCAGCCCAACAGTGTTAGTCTTTGCTCTTACCCAGTTGGGCGGGA (SEQ ID NO: 558)	GATCAGCTCCGCGGGCAAGACCTTTCTCCTTCACGGGGTGGAAGGTC (SEQ ID NO: 559)
WB	101	CTAGTTTTAAAGTTGGTTATTAGTTACTGTGATATTTATCACGGTACCCAATAACCAATGAAT (SEQ ID NO: 560)	CGGAAGGTAGCGTCAACGATAGGTGTAACTGTCGTGTTTGTAACGGTACTTCCAACAGCTGGCGCCGCCAC (SEQ ID NO: 561)
BL3	102	CAATGAAAAACTAGGCATGTAGAAGTTGTTTGT (SEQ ID NO: 562)	TTTCCACAGACAACTCACGTGGAGGTAGTCAC (SEQ ID NO: 563)
SprA	103	TGTAGTAAGTATCTTAATATACAGCTTTATCTGTTTTTTAAGATACTTACTACTTT (SEQ ID NO: 564)	CACCCATTGTGTTCACAGGAGATACAGCTTTATCTGTACTGATATTAATGACATGCTG (SEQ ID NO: 565)
phiJoe	104	AGTTGTGGCCATGTGTCCA	ATCTGGATGTGGGTGTCCAT

		TCTGGGGGCAGATGGAGA CGGGGTCACA (SEQ ID NO: 566)	CTGCGGGCAGACGCCGCAG TCGAAGCACGG (SEQ ID NO: 567)
phiK38	105	CCCTAATACGCAAGTCGAT AACTCTCCTGGGAGCGTTG ACAACTTGCGCACCCTGAT CTG (SEQ ID NO: 569)	GAGCGCCGGATCAGGGAGT GGACGGCCTGGGAGCGCTA CACGCTGTGGCTGCGGTCGG TGC (SEQ ID NO: 570)
Int2	105	GCTCATGTATGTGTCTACG CGAGATTCTCGCCCGAGAA CTTCTGCAAGGCACTGCTC TTGGCT (SEQ ID NO: 571)	GGACGGCGCAGAAGGGGAG TAGCTCTTCGCCGGACCGTC GACATACTGCTCAGCTCGTC (SEQ ID NO: 572)
Int3	105	ATGGATAAAAAAATACAG CGTTTTTCATGTACAACTA TACTAGTTGTAGTGCCTAA ATAATGCTT (SEQ ID NO: 573)	GTTTGTAAAGGAGACTGAT AATGGCATGTACAACTATAC TCGTCGGTAAAAAGGCATCT TAT (SEQ ID NO: 574)
Int4	105	AAAAATTACAAAGTTTTCA ACCCTTGATTTGAATTAGC GGTCAAATAATTTGTAATT CGTTT (SEQ ID NO: 575)	TTCCAAAGAGCGCCCAACG CGACCTGAAATTTGAATAA GACTGCTGCTTGTGTAAAGG CGATGATT (SEQ ID NO: 576)
Int7	105	GTGTTATAAACCTGTGTGA GAGTTAAGTTTACATGCCT AACCTTAACTTTTACGCAG GTTCAGCTT(SEQ ID NO: 577)	AGACGAGAAACGTTCCGTC CGTCTGGGTCAGTTGGGCAA AGTTGATGACCGGGTCGTCC GTT (SEQ ID NO: 578)
Int8	105	TTAATAAACTATGGAAGTA TGTACAGTCTTGCAATGTT GAGTGAACAAACTTCCATA ATAAAAT (SEQ ID NO: 579)	CAATCATCAGATAACTATGG CGGCACGTGCATTAACCAC GGTTGTATCCCGTCTAAAGT ACTCGT (SEQ ID NO: 580)
Int9	105	GTGGTTGTTTTTGTTGGAA GTGTGTATCAGGTATCTGC ATAGTTATTCCGAACTTCC AATTA (SEQ ID NO: 581)	TTTATATTGCGAAAAATAAT TGGCGAACGAGGTAACTGG ATACCTCATCCGCCAATTAA AATTTG (SEQ ID NO: 582)
Int10	105	GGAAAATATAAATAATTTT AGTAACCTACATCTCAATC AAGGATAGTAAAACTCTCA CTCTT (SEQ ID NO: 583)	AGCACGCTGATAATCAGCA AGACCACCAACATTTCCACC AATGTAAAAGCTTTAACCTT AGC (SEQ ID NO: 584)
Int11	105	GTTTATATGTTTACTAATA AGACGCTCTCAACCCATAA AGTCTTATTAGTAAACATA TTTCAACT (SEQ ID NO: 585)	ATGGATTTTGCAGATTCCCA GATGCCCCTACAGAAAGAG GTACAAAACATTTATTGGAA TTAATT (SEQ ID NO: 586)
Int12	105	TTTTTGTATGTTAGTTGTGT CACTGGGTAGACCTAAATA GTGACACAACTGCTATTAA AATTTAA (SEQ ID NO: 587)	GTTCGTGGTAACTATGGGTG GTACAGGTGCCACATTAGTT GTACCATTTATGTTTATGTG GTTAAC (SEQ ID NO: 588)

Int13	105	CAATAACGGTTGTATTTGTAGAACTTGACCAGTTGTTTTAGTAACATAAATACAACTCCGAATA (SEQ ID NO: 589)	GCATACATTGTTGTTGTTTTTCCAGATCCAGTTGGTCCTGTAAATATAAGCAATCCATGTGAGT (SEQ ID NO: 590)
LI	106	GTTTAGTATCTCGTTATCTCTCGTTGGAGGGAGAAGAAACGGGATACCAAAA (SEQ ID NO: 591)	TAACTTTTTCGGATCGAGTTATGATGGACGTAAAGAGGGAACAAAGCATCTA (SEQ ID NO: 592)
Peaches	107	TAGTTTCCAATGTTACAGGAACTGCTGGCAGAATCCAACACATTGGAAGTCG (SEQ ID NO: 593)	CGGTCTCCATCGGGATCTGCTGATCGAGCAGCATGCCGACCA (SEQ ID NO: 594)
Bxz2	107	TAACCGCAAGTGTACATCCCTCGGCTGGCCGAGACAAGTACAGTTGCGACAG (SEQ ID NO: 595)	CGGTCTCCATCGGGATCTGCTGATCGAGCAGCATGCCGACCA (SEQ ID NO: 596)
SV1	108	ATGTGGTCCTTTAGATCCACTGACGTGGGTCAGTGTCTCTAAAGGACTCGCG (SEQ ID NO: 597)	CATCAGGGCGGTCAGGCCGTAGATGTGGAAGAAACGGCAGCACGGCGAGGACG (SEQ ID NO: 598)

Таблица 9. Сериновые резольвазы и последовательности мишеней SSR.

Резольваза	**Источник**	**Левая мишень**	**Правая мишень**
Gin	109	CGTTTCCTGTAAACCGAGGTTTTGGATAAACA (SEQ ID NO: 599)	CGTTTCCTGTAAACCGAGGTTTTGGATAATGG (SEQ ID NO: 600)
Cin	110	GAGTTCTCTTAAACCAAGGTTTAGGATTGAAA (SEQ ID NO: 601)	GAGTTCTCTTAAACCAAGGTATTGGATAACAG (SEQ ID NO: 602)
Hin	111	TGGTTCTTGAAAACCAAGGTTTTTGATAAAGC (SEQ ID NO: 603)	AAATTTTCCTTTTTGGAAGGTTTTTGATAACCA (SEQ ID NO: 604)
Min	112	GCCTTCCCCTAAACCAACGTTTTTATGCCGCC (SEQ ID NO: 605)	GCCTTCCCCCAAACCAAGGTAATCAAGAACGC (SEQ ID NO: 606)
Sin	113	TTGTGAAATTTGGGTACACCCTAATCATACAA (SEQ ID NO: 607)	CGTATGATTAGGGTGTATATTAATTT (SEQ ID NO: 608)

Таблица 10. Тирозиновые интегразы и последовательности мишеней.

Интеграза	**Источник**	***attP***	***attB***
HK022	114	CAAATGATTTTATTTTGACTAATA	GCACTTTAGGTGA

		ATGACCTACTTACATTAATTTACT GATAATTAAAGAGATTTTAAATAT ACAACTTATTCACCTAAAGGATG ACAAAA (SEQ ID NO: 609)	AAAAGGTT (SEQ ID NO: 610)
P22	115	CTAAGTGGTTTGGGACAAAAATG GGACATACAAATCTTTGCATCGGT TTGCAAGGCTTTGCATGTCTTTCG AAGATGGGACGTGTGAGCGCAGG TATGACGTGGTATGTGTTGACTTA AAAGGTAGTTCTTATAATTCGTAA TGCGAAGGTCGTAGGTTCGACTCC TATTATCGGCACCAGTTAAATCAA ATACTTACGTATTATTCGTGCCTT CCTTATTTTTACTGTGGGACATAT TTGGGACAGAAGTACCAAAAA (SEQ ID NO: 612)	GCAGCGCATTCGT AATGCGAAGGTCG T (SEQ ID NO: 613)
L5	116	GCGATCCCCATCCGCGACGTGCC AACTAGGTCTCCTCTCGTCGTGAA CAAGGCTACCGGGTTGCAACTCC TGTGCAACTCTCAGGCTTCAACGC GCTTCTACGACCTGCAATTTCTTT CCACTTAGAGGATGCAGCCGAGA GGGGTAAAAACCTATCTTGACCG GCCCATATGTGGTCGGCAGACAC CCATTCTTCCAAACTAGCTACGCG GGTTCGATTCCCGTCGCCCGCTCC GCTGGTCAGAGGGTGTTTTCGCCC TCTGGCCATTTTTCTTTCCAGGGG TCTGCAACTCTTGTGCGACTCTTC TGACCTGGGCATACGCGGTTGCA ACGCATCCCTGATCTGGCTACTTT CGATGCTGACAAACGAATAGAGC CCCCCGCCTGCGCGAACAGACGA GGGGCATTCACA (SEQ ID NO: 614)	GAGCGGGCGACG GGAATCGAACCCG CGTAGCTAGTTTG GAAGA (SEQ ID NO: 615)

Ссылки, приведенные в Примере 17

[1509] Каждая из следующих ссылок приведена в Примере 17, каждая из которых приведена здесь в качестве ссылки.

1. Feuk, L. Inversion variants in the human genome: role in disease and genome architecture. Genome Med 2, 11 (2010).

2. Zhang, F., Gu, W., Hurles, M.E. & Lupski, J.R. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet 10, 451-481 (2009).

3. Shaw, C.J. & Lupski, J.R. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. Hum Mol Genet 13 Spec No 1, R57-64 (2004).

4. Carvalho, C.M., Zhang, F. & Lupski, J.R. Evolution in health and medicine Sackler colloquium: Genomic disorders: a window into human gene and genome evolution. Proc Natl Acad Sci U S A 107 Suppl 1, 1765-1771 (2010).

5. Rowley, J.D. Chromosome translocations: dangerous liaisons revisited. Nat Rev Cancer 1, 245-250 (2001).

6. Aplan, P.D. Causes of oncogenic chromosomal translocation. Trends Genet 22, 46-55 (2006).

7. McCarroll, S.A. & Altshuler, D.M. Copy-number variation and association studies of human disease. Nat Genet 39, S37-42 (2007).

8. Wijnker, E. & de Jong, H. Managing meiotic recombination in plant breeding. Trends Plant Sci 13, 640-646 (2008).

9. Petolino, J.F., Srivastava, V. & Daniell, H. Editing Plant Genomes: a new era of crop improvement. Plant Biotechnol J 14, 435-436 (2016).

10. Smith, M.C.M. Phage-encoded Serine Integrases and Other Large Serine Recombinases. Microbiol Spectr 3 (2015).

11. Meinke, G., Bohm, A., Hauber, J., Pisabarro, M.T. & Buchholz, F. Cre Recombinase and Other Tyrosine Recombinases. Chem Rev 116, 12785-12820 (2016).

12. Karpinski, J. et al. Directed evolution of a recombinase that excises the provirus of most HIV-1 primary isolates with high specificity. Nat Biotechnol 34, 401-409 (2016).

13. Olorunniji, F.J., Rosser, S.J. & Stark, W.M. Site-specific recombinases: molecular machines for the Genetic Revolution. Biochem J 473, 673-684 (2016).

14. Grindley, N.D., Whiteson, K.L. & Rice, P.A. Mechanisms of site-specific recombination. Annu Rev Biochem 75, 567-605 (2006).

15. Lukacsovich, T., Yang, D. & Waldman, A.S. Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. Nucleic Acids Res 22, 5649-5657 (1994).

16. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14, 8096-8106 (1994).

17. Jeggo, P.A. DNA breakage and repair. Adv Genet 38, 185-218 (1998).

18. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. Nat Med 24, 927-930 (2018).

19. Wang, B. et al. Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. Biotechniques 59, 201-202, 204, 206-208 (2015).

20. Bogdanove, A.J., Bohm, A., Miller, J.C., Morgan, R.D. & Stoddard, B.L. Engineering altered protein-DNA recognition specificity. Nucleic Acids Res 46, 4845-4871 (2018).

21. Tassabehji, M. Williams-Beuren syndrome: a challenge for genotype-phenotype correlations. Hum Mol Genet 12 Spec No 2, R229-237 (2003).

22. Araki, K., Araki, M. & Yamamura, K. Targeted integration of DNA using mutant lox sites in embryonic stem cells. Nucleic Acids Res 25, 868-872 (1997).

23. Araki, K., Okada, Y., Araki, M. & Yamamura, K. Comparative analysis of right element mutant lox sites on recombination efficiency in embryonic stem cells. BMC Biotechnol 10, 29 (2010).

24. Thomson, J.G., Rucker, E.B., 3rd & Piedrahita, J.A. Mutational analysis of loxP sites for efficient Cre-mediated insertion into genomic DNA. Genesis 36, 162-167 (2003).

25. Jusiak, B. et al. Comparison of Integrases Identifies Bxb1-GA Mutant as the Most Efficient Site-Specific Integrase System in Mammalian Cells. ACS Synth Biol 8, 16-24 (2019).

26. Xie, F. et al. Adjusting the attB site in donor plasmid improves the efficiency of PhiC31 integrase system. DNA Cell Biol 31, 1335-1340 (2012).

27. Gupta, M., Till, R. & Smith, M.C. Sequences in attB that affect the ability of phiC31 integrase to synapse and to activate DNA cleavage. Nucleic Acids Res 35, 3407-3419 (2007).

28. Kolot, M., Malchin, N., Elias, A., Gritsenko, N. & Yagil, E. Site promiscuity of coliphage HK022 integrase as tool for gene therapy. Gene Ther 22, 602 (2015).

29. Gaj, T., Mercer, A.C., Sirk, S.J., Smith, H.L. & Barbas, C.F., 3rd A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. Nucleic Acids Res 41, 3937-3946 (2013).

30. Chuang, K., Nguyen, E., Sergeev, Y. & Badea, T.C. Novel Heterotypic Rox Sites for Combinatorial Dre Recombination Strategies. G3 (Bethesda) 6, 559-571 (2015).

31. Chaikind, B., Bessen, J.L., Thompson, D.B., Hu, J.H. & Liu, D.R. A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells. Nucleic Acids Res 44, 9758-9770 (2016).

32. Gaj, T., Mercer, A.C., Gersbach, C.A., Gordley, R.M. & Barbas, C.F. Structure-guided reprogramming of serine recombinase DNA sequence specificity. P Natl Acad Sci USA 108, 498-503 (2011).

33. Gaj, T., Sirk, S.J. & Barbas, C.F., 3rd Expanding the scope of site-specific recombinases for genetic and metabolic engineering. Biotechnol Bioeng 111, 1-15 (2014).

34. Akopian, A., He, J., Boocock, M.R. & Stark, W.M. Chimeric recombinases with designed DNA sequence recognition. Proc Natl Acad Sci U S A 100, 8688-8691 (2003).

35. Prorocic, M.M. et al. Zinc-finger recombinase activities in vitro. Nucleic Acids Research 39, 9316-9328 (2011).

36. Gersbach, C.A., Gaj, T., Gordley, R.M., Mercer, A.C. & Barbas, C.F. Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. Nucleic Acids Research 39, 7868-7878 (2011).

37. Sirk, S.J., Gaj, T., Jonsson, A., Mercer, A.C. & Barbas, C.F. Expanding the zinc-finger recombinase repertoire: directed evolution and mutational analysis of serine recombinase specificity determinants. Nucleic Acids Research 42, 4755-4766 (2014).

38. Gaj, T. & Barbas, C.F., 3rd Genome engineering with custom recombinases. Methods Enzymol 546, 79-91 (2014).

39. Olorunniji, F.J., Rosser, S.J. & Marshall Stark, W. Purification and In Vitro Characterization of Zinc Finger Recombinases. Methods Mol Biol 1642, 229-245 (2017).

40. Proudfoot, C., McPherson, A.L., Kolb, A.F. & Stark, W.M. Zinc finger recombinases with adaptable DNA sequence specificity. PLoS One 6, e19537 (2011).

41. Irion, S. et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat Biotechnol 25, 1477-1482 (2007).

42. Sadelain, M., Papapetrou, E.P. & Bushman, F.D. Safe harbours for the integration of new DNA in the human genome. Nat Rev Cancer 12, 51-58 (2012).

43. Pellenz, S. et al. New human chromosomal safe harbor sites for genome engineering with CRISPR/Cas9, TAL effector and homing endonucleases. bioRxiv (2019).

44. Brown, W.R., Lee, N.C., Xu, Z. & Smith, M.C. Serine recombinases as tools for genome engineering. Methods 53, 372-379 (2011).

45. Xu, Z. et al. Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome. BMC Biotechnol 13, 87 (2013).

46. Thyagarajan, B., Guimaraes, M.J., Groth, A.C. & Calos, M.P. Mammalian genomes contain active recombinase recognition sites. Gene 244, 47-54 (2000).

47. Shultz, J.L., Voziyanova, E., Konieczka, J.H. & Voziyanov, Y. A genome-wide analysis of FRT-like sequences in the human genome. PLoS One 6, e18077 (2011).

48. Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S. & Calos, M.P. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. Mol Cell Biol 21, 3926-3934 (2001).

49. Sivalingam, J. et al. Biosafety assessment of site-directed transgene integration in human umbilical cord-lining cells. Mol Ther 18, 1346-1356 (2010).

50. Ortiz-Urda, S. et al. Stable nonviral genetic correction of inherited human skin disease. Nat Med 8, 1166-1170 (2002).

51. Chalberg, T.W. et al. Integration specificity of phage phiC31 integrase in the human genome. J Mol Biol 357, 28-48 (2006).

52. Thyagarajan, B. et al. Creation of engineered human embryonic stem cell lines using phiC31 integrase. Stem Cells 26, 119-126 (2008).

53. Olivares, E.C. et al. Site-specific genomic integration produces therapeutic Factor IX levels in mice. Nat Biotechnol 20, 1124-1128 (2002).

54. Hollis, R.P. et al. Phage integrases for the construction and manipulation of transgenic mammals. Reprod Biol Endocrinol 1, 79 (2003).

55. Held, P.K. et al. In vivo correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery. Mol Ther 11, 399-408 (2005).

56. Ma, H. et al. PhiC31 integrase induces efficient site-specific recombination in the Capra hircus genome. DNA Cell Biol 33, 484-491 (2014).

57. Bi, Y. et al. Pseudo attP sites in favor of transgene integration and expression in cultured porcine cells identified by Streptomyces phage phiC31 integrase. BMC Mol Biol 14, 20 (2013).

58. Ma, Q.W. et al. Identification of pseudo attP sites for phage phiC31 integrase in bovine genome. Biochem Biophys Res Commun 345, 984-988 (2006).

59. Qu, L. et al. Global mapping of binding sites for phic31 integrase in transgenic maden-darby bovine kidney cells using ChIP-seq. Hereditas 156, 3 (2019).

60. Ghahfarokhi, M.K., Dormiani, K., Mohammadi, A., Jafarpour, F. & Nasr-Esfahani, M.H. Blastocyst Formation Rate and Transgene Expression are Associated with Gene Insertion into Safe and Non-Safe Harbors in the Cattle Genome. Sci Rep 7, 15432 (2017).

61. Groth, A.C., Fish, M., Nusse, R. & Calos, M.P. Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics 166, 1775-1782 (2004).

62. Chalberg, T.W., Genise, H.L., Vollrath, D. & Calos, M.P. phiC31 integrase confers genomic integration and long-term transgene expression in rat retina. Invest Ophthalmol Vis Sci 46, 2140-2146 (2005).

63. Keravala, A. et al. A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. Mol Genet Genomics 276, 135-146 (2006).

64. Lei, X., Wang, L., Zhao, G. & Ding, X. Site-specificity of serine integrase demonstrated by the attB sequence preference of BT1 integrase. FEBS Lett 592, 1389-1399 (2018).

65. Chu, V.T. et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33, 543-548 (2015).

66. Yu, C. et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell 16, 142-147 (2015).

67. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature 533, 125 (2016).

68. Martsolf, J.T. et al. Complete trisomy 17p a relatively new syndrome. Ann Genet 31, 172-174 (1988).

69. Bird, T.D. in GeneReviews((R)). (eds. M.P. Adam et al.) (Seattle (WA); 1993).

70. Smith, A.C.M. et al. in GeneReviews((R)). (eds. M.P. Adam et al.) (Seattle (WA); 1993).

71. Dupuy, O. et al. [De La Chapelle syndrome]. Presse Med 30, 369-372 (2001).

72. Jyothy, A. et al. Translocation Down syndrome. Indian J Med Sci 56, 122-126 (2002).

73. Lakich, D., Kazazian, H.H., Jr., Antonarakis, S.E. & Gitschier, J. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. Nat Genet 5, 236-241 (1993).

74. Bondeson, M.L. et al. Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. Hum Mol Genet 4, 615-621 (1995).

75. Abremski, K. & Hoess, R. Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. J Biol Chem 259, 1509-1514 (1984).

76. Sauer, B. & McDermott, J. DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. Nucleic Acids Res 32, 6086-6095 (2004).

77. Suzuki, E. & Nakayama, M. VCre/VloxP and SCre/SloxP: new site-specific recombination systems for genome engineering. Nucleic Acids Res 39, e49 (2011).

78. Sadowski, P.D. The Flp recombinase of the 2-microns plasmid of Saccharomyces cerevisiae. Prog Nucleic Acid Res Mol Biol 51, 53-91 (1995).

79. Nern, A., Pfeiffer, B.D., Svoboda, K. & Rubin, G.M. Multiple new site-specific recombinases for use in manipulating animal genomes. Proc Natl Acad Sci U S A 108, 14198-14203 (2011).

80. Ringrose, L., Angrand, P.O. & Stewart, A.F. The Kw recombinase, an integrase from Kluyveromyces waltii. Eur J Biochem 248, 903-912 (1997).

81. Araki, H. et al. Site-specific recombinase, R, encoded by yeast plasmid pSR1. J Mol Biol 225, 25-37 (1992).

82. Blaisonneau, J., Sor, F., Cheret, G., Yarrow, D. & Fukuhara, H. A circular plasmid from the yeast Torulaspora delbrueckii. Plasmid 38, 202-209 (1997).

83. Karimova, M. et al. Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system. Nucleic Acids Res 41, e37 (2013).

84. Karimova, M., Splith, V., Karpinski, J., Pisabarro, M.T. & Buchholz, F. Discovery of Nigri/nox and Panto/pox site-specific recombinase systems facilitates advanced genome engineering. Sci Rep 6, 30130 (2016).

85. Buchholz, F. & Stewart, A.F. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. Nat Biotechnol 19, 1047-1052 (2001).

86. Santoro, S.W. & Schultz, P.G. Directed evolution of the site specificity of Cre recombinase. Proc Natl Acad Sci U S A 99, 4185-4190 (2002).

87. Sarkar, I., Hauber, I., Hauber, J. & Buchholz, F. HIV-1 proviral DNA excision using an evolved recombinase. Science 316, 1912-1915 (2007).

88. Rufer, A.W. & Sauer, B. Non-contact positions impose site selectivity on Cre recombinase. Nucleic Acids Res 30, 2764-2771 (2002).

89. Kim, A.I. et al. Mycobacteriophage Bxb1 integrates into the Mycobacterium smegmatis groEL1 gene. Mol Microbiol 50, 463-473 (2003).

90. Brown, D.P., Idler, K.B. & Katz, L. Characterization of the genetic elements required for site-specific integration of plasmid pSE211 in Saccharopolyspora erythraea. J Bacteriol 172, 1877-1888 (1990).

91. Matsuura, M. et al. A GENE ESSENTIAL FOR THE SITE-SPECIFIC EXCISION OF ACTINOPHAGE R4 PROPHAGE GENOME FROM THE CHROMOSOME OF A LYSOGEN. The Journal of General and Applied Microbiology 41, 53-61 (1995).

92. Gregory, M.A., Till, R. & Smith, M.C. Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors. J Bacteriol 185, 5320-5323 (2003).

93. Yang, H.Y., Kim, Y.W. & Chang, H.I. Construction of an integration-proficient vector based on the site-specific recombination mechanism of enterococcal temperate phage phiFC1. J Bacteriol 184, 1859-1864 (2002).

94. Rashel, M. et al. A novel site-specific recombination system derived from bacteriophage phiMR11. Biochem Biophys Res Commun 368, 192-198 (2008).

95. Christiansen, B., Johnsen, M.G., Stenby, E., Vogensen, F.K. & Hammer, K. Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. J Bacteriol 176, 1069-1076 (1994).

96. Loessner, M.J., Inman, R.B., Lauer, P. & Calendar, R. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of Listeria monocytogenes: implications for phage evolution. Mol Microbiol 35, 324-340 (2000).

97. Lauer, P., Chow, M.Y., Loessner, M.J., Portnoy, D.A. & Calendar, R. Construction, characterization, and use of two Listeria monocytogenes site-specific phage integration vectors. J Bacteriol 184, 4177-4186 (2002).

98. Bibb, L.A., Hancox, M.I. & Hatfull, G.F. Integration and excision by the large serine recombinase phiRv1 integrase. Mol Microbiol 55, 1896-1910 (2005).

99. Canchaya, C. et al. Genome analysis of an inducible prophage and prophage remnants integrated in the Streptococcus pyogenes strain SF370. Virology 302, 245-258 (2002).

100. Morita, K. et al. The site-specific recombination system of actinophage TG1. FEMS Microbiol Lett 297, 234-240 (2009).

101. Fouts, D.E. et al. Sequencing Bacillus anthracis typing phages gamma and cherry reveals a common ancestry. J Bacteriol 188, 3402-3408 (2006).

102. Kilcher, S., Loessner, M.J. & Klumpp, J. Brochothrix thermosphacta bacteriophages feature heterogeneous and highly mosaic genomes and utilize unique prophage insertion sites. J Bacteriol 192, 5441-5453 (2010).

103. Lazarevic, V. et al. Nucleotide sequence of the Bacillus subtilis temperate bacteriophage SPbetac2. Microbiology 145 (Pt 5), 1055-1067 (1999).

104. Fogg, P.C.M., Haley, J.A., Stark, W.M. & Smith, M.C.M. Genome Integration and Excision by a New Streptomyces Bacteriophage, varphiJoe. Appl Environ Microbiol 83 (2017).

105. Yang, L. et al. Permanent genetic memory with >1-byte capacity. Nat Methods 11, 1261-1266 (2014).

106. Rutherford, K., Yuan, P., Perry, K., Sharp, R. & Van Duyne, G.D. Attachment site recognition and regulation of directionality by the serine integrases. Nucleic Acids Res 41, 8341-8356 (2013).

107. Singh, S., Rockenbach, K., Dedrick, R.M., VanDemark, A.P. & Hatfull, G.F. Cross-talk between diverse serine integrases. J Mol Biol 426, 318-331 (2014).

108. Gupta, N. et al. Cross-talk between cognate and noncognate RpoE sigma factors and Zn(2+)-binding anti-sigma factors regulates photooxidative stress response in Azospirillum brasilense. Antioxid Redox Signal 20, 42-59 (2014).

109. Kahmann, R., Rudt, F., Koch, C. & Mertens, G. G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. Cell 41, 771-780 (1985).

110. Iida, S., Meyer, J., Kennedy, K.E. & Arber, W. A site-specific, conservative recombination system carried by bacteriophage P1. Mapping the recombinase gene cin and the cross-over sites cix for the inversion of the C segment. EMBO J 1, 1445-1453 (1982).

111. Glasgow, A.C., Bruist, M.F. & Simon, M.I. DNA-binding properties of the Hin recombinase. J Biol Chem 264, 10072-10082 (1989).

112. Iida, S. et al. The Min DNA inversion enzyme of plasmid p15B of Escherichia coli 15T-: a new member of the Din family of site-specific recombinases. Mol Microbiol 4, 991-997 (1990).

113. Rowland, S.J., Stark, W.M. & Boocock, M.R. Sin recombinase from Staphylococcus aureus: synaptic complex architecture and transposon targeting. Mol Microbiol 44, 607-619 (2002).

114. Kolot, M., Silberstein, N. & Yagil, E. Site-specific recombination in mammalian cells expressing the Int recombinase of bacteriophage HK022. Mol Biol Rep 26, 207-213 (1999).

115. Cho, E.H., Nam, C.E., Alcaraz, R., Jr. & Gardner, J.F. Site-specific recombination of bacteriophage P22 does not require integration host factor. J Bacteriol 181, 4245-4249 (1999).

116. Lee, M.H., Pascopella, L., Jacobs, W.R., Jr. & Hatfull, G.F. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for Mycobacterium smegmatis,

Mycobacterium tuberculosis, and bacille Calmette-Guerin. Proc Natl Acad Sci U S A 88, 3111-3115 (1991).

ПРИМЕР 18 – ВСТРАИВАНИЕ 3′ ОПОРНОЙ ПЕТЛИ В УЧАСТОК СВЯЗЫВАНИЯ ПРАЙМЕРА (PBS) ПОВЫШАЕТ АКТИВНОСТЬ PEgRNA

[1510] Для дальнейшего повышения активности PE авторы предусмотрели добавление последовательности опорной петли на 3′-конце PEgRNA, имеющей 3′ плечо удлинения. На Фиг.71A приведен пример стандартной PEgRNA SpCas9, имеющей 3′ плечо удлинения (верхняя молекула). 3′ плечо удлинения, в свою очередь, включает RT матрицу (которая включает требуемое изменение) и участок связывания праймера (PBS) на 3′-конце молекулы. Молекула оканчивается последовательностью poly(U), включающей три азотистых основания U (то есть 5′-UUU-3′).

[1511] С другой стороны, в нижней части Фиг.71A показана та же молекула PEgRNA, что и в верхней части Фиг.71A, но где последовательность из 9 азотистых оснований в виде 5′-GAAANNNNN-3′ была вставлена между 3′-концом участка связывания праймера и 5′-концом концевой последовательности poly(U). Данная структура сворачивается обратно на саму себя на 180° с образованием структуры РНК в виде "опорной петли", где последовательности 5′-NNNNN-3′ вставки в виде 9 азотистых оснований отжигаются с комплементарной последовательностью в участке связывания праймера и где часть в виде 5′-GAAA-3′ образует поворот на 180°. Свойства последовательности опорной петли, показанные на Фиг.71A, не ограничивают и не сужают объем возможных опорных петель, которые могут быть использованы вместо нее. Кроме того, последовательность опорной петли будет зависеть от комплементарной последовательности участка связывания праймера. Однако главным является то, что в рамах различных вариантов осуществления последовательность опорной петли может иметь первую часть в виде последовательности, которая образует поворот на 180°, и вторую часть в виде последовательности, которая имеет последовательность, являющуюся комплементарной части участка связывания праймера.

[1512] Без связи с теорией, предполагается, что последовательность опорной петли обеспечивает возможность использования PEgRNA со значительно более длинными участками связывания праймера, чем те, которые были бы возможны в ином случае. В свою очередь, предполагается, что более длинные последовательности PBS повышают

активность PEgRNA PE. Более конкретно, вероятная функция опорной петли заключает в закрытии PBS или, по меньшей мере, минимизации взаимодействия PBS со спейсером. Образование стабильной шпильки между PBS и спейсером может приводить к деактивации PEgRNA. Без опорной петли данное взаимодействие может требовать ограничения длины PBS. Блокирование или минимизация взаимодействия между спейсером и PBS с помощью 3'-концевой опорной петли может приводить к повышению активности PE.

ПРИМЕР 19 – ПРАЙМЕРНОЕ РЕДАКТИРОВАНИЕ С ИСПОЛЬЗОВАНИЕМ АЛЬТЕРНАТИВНЫХ МАТРИЦ НУКЛЕИНОВОЙ КИСЛОТЫ И КОНСТРУКТОВ БЕЛКА РЕДАКТОРА

[1513] До данного примера праймерное редактирование было описано в качестве требующего PEgRNA. Примеры вариантов осуществления, описывающие возможные конфигурации подходящих PEgRNA для использования в рамках праймерного редактирования изображены на Фиг.3A (PEgRNA с 5' плечом удлинения), Фиг.3B (PEgRNA с 3' плечом удлинения), Фиг.3C (PEgRNA с внутренним удлинением), Фиг.3D (PEgRNA с 3' плечом удлинения и включающая участок связывания праймера, матрицу для редактирования, плечо гомологии, и необязательные 3' и 5' модификаторные участки, и участок, обозначенный как матрица для синтеза ДНК), и Фиг.3E (PEgRNA с 5' плечом удлинения и включающая участок связывания праймера, матрицу для редактирования, плечо гомологии, и необязательные 3' и 5' модификаторные участки, и участок, обозначенный как матрица для синтеза ДНК). Кроме того, в разделе Подробное описание изобретения и по всему описанию подробно описана структура и композиция PEgRNA.

[1514] Данный Пример описывает дополнительные варианты конструкции PEgRNA — в некоторых случаях PEgRNA, которые полностью или частично химически синтезируются вне клетки — которые, как предполагается, будут работать совместно с праймерными редакторами согласно данному описанию. Подобные альтернативные конструкции могут улучшать различные аспекты праймерного редактирования, включая вставку более длинных последовательностей ДНК с помощью праймерного редактирования, использование альтернативных полимераз (то есть альтернатив обратной транскриптазе), которые потенциально функционируют с повышенной эффективностью и/или точностью, и применение или рекрутирование альтернативных

и/или дополнительных эффекторных компонентов белков праймерного редактора для улучшения или дополнения праймерного редактирования. Кроме того, использование химически синтезированной PEgRNA может потенциально приводить к продуцированию молекул, которые являются более стабильными и обладают желаемыми свойствами, которые повышают эффективность и возможности праймерного редактирования.

[1515] PEgRNA служит в качестве матрицы нуклеиновой кислоты, которая кодирует требуемую связанную с редактированием генетическую информацию, которая предназначена для внесения в участок-мишень. В одном аспекте PEgRNA создают путем добавления плеча удлинения либо к 5'-концу или 3'-концу sgРНК (то есть как показано в вариантах осуществления на Фиг.3A, 3B, 3D, or 3E) или путем вставки такой же последовательности внутрь sgРНК (например как показано в варианте осуществления на Фиг.3C), при этом плечо удлинения включает матрицу для синтеза ДНК, которая способна кодировать оцДНК-продукт с помощью полимеразы (например обратной транскриптазы) и которая включает представляющую интерес связанную с редактированием генетическую информацию. Плечо удлинения включает участок связывания праймера (PBS) для отжига на napDNAbp-никированной цепи геномной ДНК и матрицу для синтеза ДНК для кодирования подвергаемой редактированию представляющей интерес цепи ДНК, которая становится встроенной в участок-мишень эндогенной ДНК за счет замещения соответствующей цепи ДНК. PEgRNA может быть экспрессирована в клетках из плазмидной ДНК или встроенной в геном кассеты ДНК или она может быть получена вне клеток с помощью *in vitro* транскрипции или с помощью химического синтеза и последующей доставки в клетки. Получение PEgRNA вне клеток, в частности с помощью химического синтеза, обеспечивает возможность значительного модифицирования PEgRNA. Данное изобретение описывает альтернативные конструкции матриц для праймерного редактирования (Фиг.72).
(A) Матрица для синтеза ДНК, экспрессированная в виде отдельных молекул из направляющих РНК (то есть матрица для синтеза ДНК, предоставляемая *in trans* комплексу праймерного редактора (napDNAbp + направляющая РНК).

[1516] В различных описанных здесь вариантах осуществления в рамках праймерного редактирования используется одна PEgRNA, которая служит как программируемые нацеливающие молекулы, так и кодирующая изменение молекула. Данный вариант осуществления изображен на Фиг.72(a), где показана PEgRNA, имеющая 3' плечо

удлинения. Однако в некоторых случаях это может являться вредным, в частности для более сложных молекул PEgRNA, например таких, которые кодируют крупную вставку. Данные РНК могут содержать крупную вторичную структуру, которая нарушает структуру каркаса PEgRNA и взаимодействия с Cas9. В качестве альтернативы, праймерное редактирование может быть осуществлено путем замещения PEgRNA двумя отдельными молекулами РНК: sgРНК и *trans* РНК-матрицей для праймерного редактирования (tPERT), как показано на Фиг.72(b). sgРНК служит для нацеливания Cas9 (или, в более общем случае, napDNAbp) на требуемый геномный участок-мишень, в то время как tPERT используется полимеразой (например обратной транскриптазой) для записи новой последовательности ДНК в локус-мишень.

[1517] В целом, простая экспрессия tPERT приводит к более низкой эффективности редактирования по сравнению с PEgRNA. Однако эффективность *trans* праймерного редактирования может быть повышена путем внесения одного или нескольких РНК-аптамеров MS2 в РНК tPERT наряду с продуктом слияния белка оболочки MS2 (MS2cp) с белком праймерного редактора (для получения MS2cp-Cas9-RT). Это обеспечивает связывание РНК-аптамера MS2 с MS2cp, тем самым обеспечивая солокализацию tPERT (который включает матрицу для синтеза ДНК) с участком редактирования за счет комплекса праймерного редактора. Аптамер MS2 предпочтительно размещается на 3'-конце tPERT для исключения обратной транскрипции последовательности аптамера.

[1518] Несмотря на то, что в данном примере используется способ мечения MS2 (включающий РНК-аптамер MS2 на tPERT, связанной с белком MS2cp, слитым с праймерным редактором), в качестве альтернативы могут использоваться другие системы рекрутирования РНК-белка. Общая предполагаемая здесь концепция заключается в том, что матрица для синтеза ДНК tPERT модифицируется для включения вторичной структуры для рекрутирования РНК (например специализированной шпильки, такой как аптамер MS2), так что tPERT может быть рекрутирована модифицированным слитым белком праймерного редактора, который дополнительно включает РНК-связывающий белок, который специфично распознает и связывает вторичную структуру для рекрутирования РНК на молекуле tPERT. В данной области описаны и другие домены рекрутирования РНК-белка, например в Johansson et al., “RNA recognition by the MS2 phage coat protein,” *Sem Virol.*, 1997, Vol. 8(3): 176-185; Delebecque et al., “Organization of intracellular reactions with rationally

designed RNA assemblies," *Science*, 2011, Vol. 333: 470-474; Mali et al., "Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," *Nat. Biotechnol.*, 2013, Vol.31: 833-838; и Zalatan et al., "Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds," *Cell*, 2015, Vol.160: 339-350, каждый из которых полностью приведен здесь в качестве ссылки. Другие системы включают шпильку PP7, которая специфично рекрутирует белок PCP, и шпильку "com", которая специфично рекрутирует белок Com. См. Zalatan et al. Любая из данных хорошо известных систем рекрутирования может быть использована в рамках *trans* праймерного редактирования, как описано здесь.

[1519] Была протестирована эффективность настоящей *trans* системы праймерного редактирования на основе tPERT. Эффективность вставки His6 (18 bp) на уровне до 20% была достигнута в участке HEK3 в клетках HEK293T при использовании tPERT, содержащей один 3' аптамер MS2, 13 nt участок связывания праймера, и RT матрицу, содержащую последовательность вставки и 34 nt гомологии относительно локуса, наряду с редактором, содержащим MS2cp, слитый с N-концом PE2. См. Фиг.73. Стратегия *trans* праймерного обладает потенциалом в плане устранения трудностей, связанных с конструированием PEgRNA, и может являться более подходящей для более длинных RT матриц для достижения более крупных вставок, делеций, или изменений на бóльших расстояниях от участка разрыва праймерного редактора.

(B) Химически синтезированные PEgRNA с матрицами для синтеза РНК и ДНК

[1520] В химически синтезированных PEgRNA могут использоваться альтернативные матрицы нуклеиновой кислоты (Фиг.72c). Например, синтетическая PEgRNA может быть сконструирована в виде РНК/ДНК гибрида, где спейсерная последовательность и каркас sgРНК состоят из нуклеотидов РНК, а участок связывания праймера и матрица для синтеза (показанная как 3' удлинение на Фиг.72c) состоят из нуклеотидов ДНК. Это может обеспечивать использование ДНК-зависимых ДНК-полимераз вместо обратной транскриптазы в праймерных редакторах. Это также может предотвращать синтез ДНК, матрицей для которой служит последовательность каркаса sgРНК. В рамках других конструкций химические линкеры, состоящие из не относящихся к матрицам нуклеотидов или других подходящих линкерных фрагментов, могут использоваться для связывания матрицы для редактирования нуклеиновой кислоты (состоящей из РНК или ДНК) с каркасом sgРНК. Это может предотвращать непрерывную полимеризацию ДНК каркаса sgРНК и обеспечивать гибкость в

удлинении, которая обеспечивает более эффективный синтез на основе матрицы. Наконец, направленность матрицы для синтеза нуклеиновой кислоты может быть обращена, так что полимеризация ДНК осуществляется по направлению прочь от каркаса sgРНК, а не по направлению к нему.

(C) Рекрутирование ДНК-полимеразы, экспрессируемой *in trans*

[1521] В главном варианте осуществления праймерного редактирования полимераза (например фермент обратная транскриптаза) экспрессируется в виде продукта слияния с napDNAbp (например никазой Cas9). В качестве альтернативы, полимераза (например обратная транскриптаза) может быть экспрессирована *in trans*, при этом ее активность может быть локализована в участке редактирования с использованием систем рекрутирования, таких как РНК-аптамер MS2 и белок оболочки MS2, или других аналогичных систем рекрутирования, известных в данной области. В рамках данной системы PEgRNA модифицируют для включения аптамера MS2 в одну из шпилек каркаса sgРНК, при этом полимераза (например обратная транскриптаза) экспрессируется в виде слитого с MS2cp белка. napDNAbp (например никаза Cas9) также экспрессируется в виде независимого полипептида. Данная система была продемонстрирована с обратной транскриптазой M-MLV дикого типа (Фиг.74) и должна являться пригодной для других вариантов RT. Кроме того, для рекрутирования RT могут быть использованы другие взаимодействия РНК-белок или взаимодействия белок-белок.

К Примеру 19 относятся следующие последовательности:

Последовательности tPERT:

Аптамер MS2 / **RT матрица** / PBS / **Линкер**

5’-MS2_13nt-PBS:

5’GCCAACATGAGGATCACCCATGTCTGCAGGGCC**TGGAGGAAGCAGGGCTTCC TTTCCTCTGCCATCAATGATGGTGATGATGGTG**CGTGCTCAGTCTG – 3’ (SEQ ID NO: 762)

5’-MS2_17nt-PBS:

5’GCCAACATGAGGATCACCCATGTCTGCAGGGCC**TGGAGGAAGCAGGGCTTCC TTTCCTCTGCCATCAATGATGGTGATGATGGTG**CGTGCTCAGTCTGGGCC – 3’ (SEQ ID NO: 773)

3’-MS2_13nt-PBS:

5’**GGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATGATGGTGATGATGGT GCGTGCTCAGTCTG<u>AAATTAACAAATCAA</u>**GCCAACATGAGGATCACCCATGTCT <u>GCAGGGCC</u> – 3’ (SEQ ID NO: 774)

3’-MS2_17nt-PBS:

5’**GGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATGATGGTGATGATGGT GCGTGCTCAGTCTGGGCC<u>AAATTAACAAATCAA</u>**GCCAACATGAGGATCACCCAT <u>GTCTGCAGGGCC</u> – 3’ (SEQ ID NO: 775)

<u>Последовательности PEgRNA MS2:</u>

<u>Спейсер</u> / <u>Аптамер MS2</u> / *Каркас sgPHK* / **RT матрица** / PBS

HEK3_MS2_1
5’**<u>GGCCCAGACTGAGCACGTGA</u>***GTTTTAGAGCTAG*<u>GCCAACATGAGGATCACCCA TGTCTGCAGGGCC</u>*T AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GGACCGAGTCGGTCC*<u>**TCTGCCATCT**</u>CGTGCTCAGTCT – 3’ (SEQ ID NO: 776)

HEK3_MS2_2
5’**<u>GGCCCAGACTGAGCACGTGA</u>***GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC TAGTCCGTTATCAACTTG*<u>GCCAACATGAGGATCACCCATGTCTGCAGGGCC</u>*AAGTG GGACCGAGTCGGTCC*<u>**TCTGCCATCT**</u>CGTGCTCAGTCT – 3’ (SEQ ID NO: 777)

<u>Последовательности белка:</u>

<u>MS2cp</u>-*PE2*
MKRTADGSEFESPKKKRKV<u>GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISS NSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELT IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY</u>SGGSSGGSSGSETPGTSESATP ESSGGSSGGSS*DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN PDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP

*QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRY TSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGG SSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVS IKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQ LREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGL PDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVA LNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKA GAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIY RRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARK AAITETPDTSTLLIENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 778)

MS2cp-*MMLV-RT*
MKRTADGSEFESPKKKRKVGSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISS
NSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELT
IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATP
ESSGGSSGGSS*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRP VQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFE WRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAAT SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMG QPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQ ALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPP CLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSL LQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 779)

MMLV-RT-MS2cp
MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGM GLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLH PTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQ YVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLT EARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLD PVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTH YQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLN VYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKG HSAEARGNRMADQAARKAAITETPDTSTLLIENSSP*SGGSSGGSSGSETPGTSESATPESS GGSSGGSSGSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCS

VRQSSAQNRKYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIV KAMQGLLKDGNPIPSAIAANSGIYSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 780)

ПРИМЕР 20. ДОСТАВКА ПРАЙМЕРНЫХ РЕДАКТОРОВ НА ОСНОВЕ РАЗДЕЛЕННОГО ИНТЕИНА

[1522] Данный пример демонстрирует, что праймерный редактор может быть разделен с использованием интеинов в качестве средства доставки праймерного редактора в клетки с использованием отдельных векторов, при этом каждый вектор кодирует часть слитого белка праймерного редактора. Данный пример основан на слитый слитом белке PE, включающем стандартный SpCas9 (SEQ ID NO: 18). Участки разделения для других белков Cas9 могут соответствовать тому же участку или могут требовать оптимизации для каждого отдельного белка Cas9. В рамках настоящего Примера праймерный редактор был разделен между остатками 1023 и 1024 SpCas9 (SEQ ID NO: 18). Это обозначается как участок разделения "1023/1024".

[1523] Праймерные редакторы (PE) превышают способность AAV в плане упаковывания. Таким образом, было предусмотрено разделение PE путем вставки обеспечивающего транс-сплайсинг интеина *Npu* в остатке 1024 Cas9 *S. pyogenes* в SEQ ID NO: 18, обеспечивающей доставку разделенных SpPE в виде двух отдельных полипептидов, каждый из которых кодируется одним из AAV в рамках двойной системы. Участок разделения 1023/1024 был выбран поскольку (1) он обеспечивает упаковывание в два AAV, в то же время оставляя место для направляющей кассеты (кассет) и минимальных регуляторных элементов, (2) мутация нативного серина в цистеин будет относительно консервативной, (3) данный участок представляет собой гибкую петлю рядом с периферией Cas9, которая на основе стерического предсказания будет обеспечивать реализацию сплайсинга, и (4) Cas9 был успешно изменен в данной цепи с помощью циркулярной пермутации (но ранее не в данном конкретном участке разделения).

[1524] Для определения активности разделенных с использованием *Npu* праймерных редакторов клетки HEK трансфицировали плазмидами, кодирующими разделенные редакторы, с определением того, что они повторяют активность полноразмерного PE3 при анализе с помощью высокопроизводительного секвенирования. Более того, известно, что три нативных амино-концевых остатка в виде *Npu* C-концевого экстеина обеспечивают сплайсинг наиболее эффективным образом, но отличаются от тех, которые нативно фланкируют остаток 1024 Cas9. Замещение данных остатков Cas9

нативными остатками *Npu* может изменять активность праймерного редактора. Таким образом, было определено то, изменяла ли мутация Cas9-нативной последовательности "SEQ" с образованием последовательности "CFN" интеина *Npu* эффективность праймерного редактирования. Остатки "SEQ" обеспечивают праймерное редактирование с эффективностью, аналогичной полноразмерному PE, что свидетельствует о том, что половины разделенного с использованием интеина PE способны ассоциироваться и опосредовать праймерное редактирование. Дополнительное повышение за счет мутации с образованием "CFN" не наблюдалось, что свидетельствует о том, что ассоциирование самостоятельно может являться достаточным для обеспечения активности разделенного PE, как наблюдали ранее в случае разделенных с использованием интеина редакторов оснований.

[1525] Несмотря на то, что ассоциированные не подвергнутые сплайсингу редакторы могут являться активными, нарушения стерических показателей праймерного редактора, связанные с незавершенным сплайсингом в рамках этапов инициации, могут влиять на результаты редактирования. Таким образом, для дальнейших исследования выбран 1024 CFN, поскольку он также повторяет активность полноразмерного PE3.

[1526] Далее приведены аминокислотные последовательности участка разделения 1023/1024.

[1527] Разделение SpPE2 в 1023/1024 N-концевой половине

[1528] MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL

AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAK*CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQW HDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN***SGGS*** *KRTADGSEFEPKKKRKV* (SEQ ID NO: 3875)

[1529] Обозначения: NLS, Cas9, Мутированные остатки, **Линкер**, *Интеин NpuN*, *Интеин NpuC*, **RT**

[1530] Разделение SpPE2 в 1023/1024 С-концевой половина

[1531] MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN*CFNEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR KVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD***SGGSS GGSSGSETPGTSESATPESSGGSSGGSS*TLNIEDEYRLHETSKEPDVSLGSTWLSD FPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLL DQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLL SGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRAL LQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDP VAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLS NARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTR PDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQ RAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIK NKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPD TSTLLIENSSP***SGGS****KRTADGSEFEPKKKRK* SEQ ID NO: 3876)

[1532] Обозначения: NLS, Cas9, Мутированные остатки, **Линкер**, *Интеин NpuN*, *Интеин NpuC*, **RT**

Упаковывание праймерных редакторов в двойной системе на основе AAV

[1533] Праймерные редакторы, разделенные между остатками 1023 и 1024 с использованием обеспечивающего транс-сплайсинг интеина Npu, упаковываются в AAV с высоким титром, сравнимым с титром редакторов оснований с таким же размером генома, полученных таким же образом.

Праймерное редактирование in vivo

[1534] Разделенный SpPE3, доставленный с использованием AAV9 с помощью внутрицеребровентрикулярной инъекции мышам P0, опосредует праймерное редактирование в мозговой ткани *in vivo*. Замещение нуклеотида в положении +5 (то есть на 5 нуклеотидных оснований ниже участка разрыва и в рамках последовательности PAM) в виде G на T приводит в внесению кодирующей мутации Pro>Gln рядом с N-концом DNMT1. Данное изменение демонстрирует осуществимость праймерного редактирования *in vivo* и, как считается, не оказывает какого-либо избирательного давления на подвергнутые редактированию клетки (где "+5" обозначает положение +5 ниже участка разрыва). Протестированные архитектуры AAV включают полноразмерную RT MMLV, а также усеченный вариант без домена РНКазы H. Также был проанализирован усеченный посттрансляционный регуляторный элемент W3, поскольку было показано, что он повышает экспрессию из вирусных кассет in vivo, однако его важность не была протестирована в контексте редакторов оснований или праймерных редакторов. Было обнаружено, что полноразмерная RT превосходит усеченную RT, при этом добавление последовательности W3 повышает активность. Повышение активности редактирования при наличии W3 указывает на то, что экспрессия праймерного редактора является ограниченной *in vivo*, и демонстрирует значительные отличия праймерного редактирования *in vivo* против клеточных культур.

Ссылки

[1535] Oakes, B.L., Fellmann, C. et al. CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification. Cell. 2019 Jan 10;176(1-2):254-267.e16. doi: 10.1016/j.cell.2018.11.052.

ПРИМЕР 21. ОПТИМИЗАЦИЯ ЛИНКЕРА В ФОРМАТЕ PE2

[1536] В рамках данного Примера был сконструирован ряд вариантов праймерных редакторов, все из которых были основаны на PE2. PE2 имеет следующую последовательность и структуру:

[1537] В соответствии с используемым здесь значением "PE2" обозначает комплекс PE, включающий слитый белок, включающий Cas9(H840A) и вариант RT MMLV, имеющий следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)] + требуемая PEgRNA, где продукт слияния PE имеет аминокислотную последовательность SEQ ID NO: 134, которая является следующей:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN**
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD
KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN
ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFD
LAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI
TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHA
ILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV
TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE
DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV
KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV
ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV
SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYD
VRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW
DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF
LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF
LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV
LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT
LIHQSITGLYETRIDLSQLGGD<u>SGGSSGGSSGSETPGTSESATPESSGGSSGGSS</u>*TL*
NIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQY
PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED
IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT
WTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALL
QTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF
LGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTK

PFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG
KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLL
PLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET
EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTS
EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDT
*STLLIENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 134)

ОБОЗНАЧЕНИЯ:
ПОСЛЕДОВАТЕЛЬНОСТЬ ЯДЕРНОЙ ЛОКАЛИЗАЦИИ (NLS) ВЕРХ:(SEQ ID NO: 124), НИЗ: (SEQ ID NO: 133)
CAS9(H840A) (SEQ ID NO: 137)
33-АМИНОКИСЛОТНЫЙ ЛИНКЕР (SEQ ID NO: 127)

[1538] Обратная транскриптаза M-MLV (SEQ ID NO: 139).

[1539] Линкер PE2 представляет собой SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127).

[1540] В данном эксперименте линкер PE2 был замещен одним из следующих замещающих линкеров (или ни одним линкером в одном случае):

Название линкера	Нуклеотидная последовательность	Аминокислотная последовательность
1x SGGS	TCCGGAGGATCT (SEQ ID NO: 3880)	SGGS (SEQ ID NO: 3887)
2x SGGS	TCCGGAGGATCTAGCGGAGGCTCC (SEQ ID NO: 3881)	SGGSSGGS (SEQ ID NO: 3888)
3x SGGS	TCCGGAGGATCTAGCGGAGGCTCCA GCGGAGGCAGC (SEQ ID NO: 3882)	SGGSSGGSSGGS (SEQ ID NO: 3889)
1x XTEN	TCTGGCTCTGAGACACCTGGCACAAG CGAGAGCGCAACACCTGAAAGC (SEQ ID NO: 3883)	SGSETPGTSESATPES (SEQ ID NO: 3890)
Линкер отсутствует		
1x Gly	GGT	G
1x Pro	CCC	P
1x EAAAK	GAAGCAGCTGCTAAA (SEQ ID NO: 3884)	EAAAK (SEQ ID NO: 3891)
2x EAAAK	GAAGCCGCTGCTAAAGAAGCTGCAG CTAAG (SEQ ID NO: 3885)	EAAAKEAAAK (SEQ ID NO: 3892)
3x EAAAK	GAAGCCGCTGCTAAAGAGGCCGCTG CTAAAGAAGCTGCAGCTAAG (SEQ ID NO: 3886)	EAAAKEAAAKEAAAK (SEQ ID NO: 3893)

[1541] На **Фиг.79** показана эффективность редактирования конструктов на основе замещающего линкера. В частности, данные показывают эффективность редактирования конструкта PE2 с текущим линкером (обозначенным как PE2 – белый прямоугольник) по сравнению с различными версиями с линкером, замещенным

последовательностью, как указано в локусах HEK3, EMX1, FANCF, RNF2 для соответствующих PEgRNA для изменений в виде транзиции, трансверсии, вставки, и делеции. Замещающие линкеры обозначены как "1x SGGS", "2x SGGS", "3x SGGS", "1x XTEN", "линкер отсутствует", "1x Gly", "1x Pro", "1x EAAAK", "2x EAAAK", и "3x EAAAK". Эффективность редактирования измеряется в формате гистограммы относительно "контрольной" эффективности редактирования PE2. Линкер PE2 представляет собой SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127). Все редактирование осуществлялось в контексте системы PE3, то есть которая соответствует конструкту для редактирования PE2 плюс добавлению оптимальной вторичной направляющей никирования sgРНК.

[1542] На **Фиг.80** показано, что линкер 1x XTEN обеспечивал повышение эффективности редактирования. Показанный график основан на средней кратной эффективности по сравнению с PE2, свидетельствуя о том, что использование последовательности линкера 1x XTEN повышает эффективность редактирования в среднем в 1,14 раза (n=15).

ПРИМЕР 22. НАПРАВЛЯЮЩИЕ РНК ПРАЙМЕРНОГО РЕДАКТОРА С УЛУЧШЕННОЙ АКТИВНОСТЬЮ

[1543] В различных вариантах осуществления PEgRNA вероятно будут зависеть от участка-мишени ДНК и изменения. То есть последовательность PEgRNA будет зависеть от последовательности-мишени ДНК и конкретного изменения, подвергаемого внесению в нее с помощью праймерного редактирования (например делеции, вставки, инверсии, замещения). Например, в определенных вариантах осуществления при присоединении мотива в виде 3' участка связывания праймера (PBS) PEgRNA линкер между PBS и мотивом предпочтительно предотвращает стерическое сталкивание с доменом полимеразы (например обратной транскриптазы) слитого белка PE. Однако природа данного линкера может являться различной для каждого участка. Например, при использовании одного линкера для каждого участка это может искусственно превращать 13 nt PBS в 16 nt PBS за счет произвольного спаривания со спейсерной последовательностью. Аналогичным образом, линкер может сам связываться с PBS, что приводит к его закрытию и потенциально снижению активности. Поэтому, в отличие от редакторов на основе белка, таких как PE или BE4, вариант в виде одной последовательности линкера, соединяющей два элемента, может

являться неэффективным в рамках каждого конструкта, но будет частично зависеть от последовательности ДНК-мишени и представляющего интерес изменения.

[1544] Частично основываясь на информации, приведенной в Примере 15 (Конструирование и модифицирование PEgRNA), в рамках данного Примера были сконструированы и затем протестированы эффекты различных структурных модификаций PEgRNA на функционирование редактирования, помимо других аспектов.

Экспрессия PEgRNA из не относящихся к pol III промоторов

[1545] Было протестировано множество систем экспрессии PEgRNA на их способность продуцировать PEgRNA с использованием вставки последовательностей длиной 102 нуклеотида из FKBP в качестве считываемого фрагмента.

[1546] Транскрипция PEgRNA может обеспечиваться обычным конститутивным промотором, например промотором U6. Несмотря на то, что промотор U6 в большинстве случаев является эффективным в плане обеспечения транскрипции PEgRNA, промотор U6 не является особенно эффективным в плане обеспечения транскрипции более длинных PEgRNA или РНК с большим содержанием U. Последовательности РНК с большим содержанием U вызывают преждевременную терминацию транскрипции. В рамках данного Примера сравниваются результаты редактирования направляющих, экспрессированных из промотора CMV или промотора U1 с промотором U6. Данные промоторы требуют другой последовательности терминатора, например MASC ENE или PAN ENE, как показано ниже. Повышение в плане редактирования наблюдали при использовании системы pCMV/MASC-ENE, однако эти направляющие приводили к неполной вставке последовательности, в то время как при использовании промотора U6 наблюдали полную вставку при более низких уровнях редактирования. См. Фиг.81. Эти данные свидетельствуют о вероятности того, что для длинных вставок могут применяться альтернативные системы экспрессии.

[1547] Нуклеотидная последовательность систем экспрессии pCMV/MASC-ENE является следующей (в направлении от 5’ к 3’) (с указанием мотива жирным шрифтом непосредственно перед участком, который он обозначает):

[1548] -промотор pCMV -

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAG
TTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC

CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAAT
GGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCA
GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC
ATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA
TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA
GGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATC-**петля Csy4**-
GTTCACTGCCGTATAGGCAG-**спейсер**-GGCCCAGACTGAGCACGTGA-**каркас**-
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGGACCGAGTCGGTCC-**матрица**-
TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCA-**вставка**-
AAATTTCTTTCCATCTTCAAGCATCCCGGTGTAGTGCACCACGCAGGTCTGGCCG
CGCTTGGGGAAGGTGCGCCCGTCTCCTGGGGAGATGGTTTCCACCTGCACTCC-
PBS-CGTGCTCAGTCTG-**линкер**-TTT-**MASC ENE**-
TAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTTCTCAGG
TTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGATGCTGGT
GGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCTTTGCTTT
GACT-**произвольная плазмидная последовательность**-
TTTTTTTAAGCTTGGGCCGCTCGAGGTAGCAGC-**промотор Ubc**-
GGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCCTCCTCACGGCGAG
CGCTGCCACGTCAGACGAAGGGCGCAGGAGCGTTCCTGATCCTTCCGCCCGGAC
GCTCAGGACAGCGGCCCGCTGCTCATAAGACTCGGCCTTAGAACCCCAGTATCA
GCAGAAGGACATTTTAGGACGGGACTTGGGTGACTCTAGGGCACTGGTTTTCTTT
CCAGAGAGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAG
GGATCTCCGTGGGGCGGTGAACGCCGATGATTATATAAGGACGCGCCGGGTGTG
GCACAGCTAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATC
GCTGTGATCGTCACTTGGTGAGTTGCGGGCTGCTGGGCTGGCCGGGGCTTTCGTG
GCCGCCGGGCCGCTCGGTGGGACGGAAGCGTGTGGAGAGACCGCCAAGGGCTGT
AGTCTGGGTCCGCGAGCAAGGTTGCCCTGAACTGGGGGTTGGGGGGAGCGCACA
AAATGGCGGCTGTTCCCGAGTCTTGAATGGAAGACGCTTGTAAGGCGGGCTGTG
AGGTCGTTGAAACAAGGTGGGGGGCATGGTGGGCGGCAAGAACCCAAGGTCTTG

AGGCCTTCGCTAATGCGGGAAAGCTCTTATTCGGGTGAGATGGGCTGGGGCACC
ATCTGGGGACCCTGACGTGAAGTTTGTCACTGACTGGAGAACTCGGGTTTGTCGT
CTGGTTGCGGGGGCGGCAGTTATGCGGTGCCGTTGGGCAGTGCACCCGTACCTTT
GGGAGCGCGCGCCTCGTCGTGTCGTGACGTCACCCGTTCTGTTGGCTTATAATGC
AGGGTGGGGCCACCTGCCGGTAGGTGTGCGGTAGGCTTTTCTCCGTCGCAGGACG
CAGGGTTCGGGCCTAGGGTAGGCTCTCCTGAATCGACAGGCGCCGGACCTCTGGT
GAGGGGAGGGATAAGTGAGGCGTCAGTTTCTTTGGTCGGTTTTATGTACCTATCT
TCTTAAGTAGCTGAAGCTCCGGTTTTGAACTATGCGCTCGGGGTTGGCGAGTGTG
TTTTGTGAAGTTTTTTAGGCACCTTTTGAAATGTAATCATTTGGGTCAATATGTAA
TTTTCAGTGTTAGACTAGTAAATTGTCCGCTAAATTCTGGCCGTTTTTGGCTTTTT
TGTTAGACAGGATCCCCGGGTACCGGTCGCCACC-**Csy4 и NLS-**
ATGGGCTCTTTTACTATGGACCACTACCTGGATATTAGACTGAGACCTGACCCTG
AGTTCCCACCCGCCCAGCTGATGAGCGTGCTGTTCGGCAAGCTGCACCAGGCCCT
GGTGGCACAGGGAGGCGACCGGATCGGCGTGAGCTTCCCCGACCTGGATGAGAG
CAGATCCAGGCTGGGAGAGCGCCTGAGGATCCACGCATCCGCCGACGATCTGCG
CGCCCTGCTGGCCCGGCCATGGCTGGAGGGCCTGCGCGACCACCTGCAGTTTGG
AGAGCCAGCAGTGGTGCCACACCCTACCCCATACAGGCAGGTGTCCAGGGTGCA
GGCAAAGTCTAACCCTGAGCGGCTGCGGAGAAGGCTGATGCGCCGGCACGATCT
GTCTGAGGAGGAGGCCAGAAAGAGGATCCCCGACACCGTGGCCAGAACACTGG
ATCTGCCTTTCGTGACCCTGCGGAGCCAGAGCACAGGCCAGCACTTCAGACTGTT
TATCAGGCACGGCCCACTGCAGGTGACAGCCGAGGAAGGAGGATTCACTTGTTA
CGGACTGTCTAAAGGAGGATTCGTGCCCTGGTTCAGCAGCCTGAGGCCTCCTAAG
AAGAAGAGGAAGGTTTAA-**терминатор SV40-**
TGATCATAATCAAGCCATATCACATCTGTAGAGGTTTACTTGCTTTAAAAAACCT
CCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACT
TGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCAC
AAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATG
TATCTTATCATGTCTGGATCTGC.

Обозначения:

[Промотор pCMV] – связывает pol II РНК-полимеразу

[Петля Csy4] – связана белком Csy4, приводит к расщеплению 3’ петли. Требуется, поскольку часть в виде [промотор CMV] транскрибируется, и если данная

последовательность присоединена к 5’ нРНК, то она уменьшит/устранит активность (ранее известно).

[Спейсерная последовательность] PEgRNA

[Каркас PEgRNA]

[Матрица для синтеза ДНК]

[Изменение в виде вставки (108 nt от FKBP)]

[Участок связывания праймера]

[Линкер] (высоковариабельный) – соединяет PBS и терминаторный элемент

[Терминатор транскрипции MASC ENE] - транскрипция данного элемента приводит к терминации транскрипции; при этом polyA-хвост кодируется и затем секвестируется элементом ENE

[Произвольная последовательность]

[Промотор Ubc] – требуется для экспрессии белка Csy4

[Белок Csy4 и NLS] – требуются для процессирования 5'-конца направляющей. Также могут использоваться другие стратегии, которые не требуют экспрессии крупного белка (например рибозим-опосредованное расщепление спейсера), но они потребуют более индивидуальной настройки для различных спейсерных последовательностей, поэтому была использована эта.

[Терминатор SV40] – для терминации белка Csy4.

Усовершенствования каркаса PEgRNA

[1549] Также был протестирован ряд структурных модификаций каркаса нРНК, ни одна из которых не показала значительного повышения эффективности редактирования (см. Фиг.82, 3.30.13-3.30.9 по оси X, в отличие от 3.30).Однако эти данные имеют два нюанса, которые стоит отметить. Во-первых, данная направляющая уже работала достаточно хорошо, при этом было бы лучше протестировать менее эффективную направляющую. Во-вторых, в клетках HEK трансфекция является достаточно эффективной, при этом было отмечено, что количество трансфицированной направляющей РНК является значительно превышенным по сравнению с тем, которое требуется (снижение количества в 4-8 раз не оказывает влияния на редактирование). Данные улучшения могут быть видны только в других типах клеток, где эффективность трансфекции является более низкой, или при

использовании менее эффективных направляющих. Многие из данных изменений предшествуют повышению активности sgРНК в других линиях клеток.

[1550] Последовательности конструктов на Фиг.82 являются следующими:

Последовательность PEgRNA HEK3.30: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3873]
Последовательность PEgRNA HEK3.30.0: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUUUU [SEQ ID NO: 3874]
Последовательность PEgRNA HEK3.30.1: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-[нет] [SEQ ID NO: 3875]
Последовательность PEgRNA HEK3.30.2: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGCUCGAGGCGGAAACGCCUCGAGCUUUU [SEQ ID NO: 3876]
Последовательность PEgRNA HEK3.30.2b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGCUCGAGGCGGAAACGCCUCGAGC [SEQ ID NO: 3877]
Последовательность PEgRNA HEK3.30.3: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGCUCGAGGCGUACGCGAAAGCGUACGCCUCGAGCUUUU [SEQ ID NO: 3878]
Последовательность PEgRNA HEK3.30.3b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGCUCGAGGCGUACGCGAAAGCGUACGCCUCGAGC [SEQ ID NO: 3879]
Последовательность PEgRNA HEK3.30.5: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-

UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGCUCGAGGCGUACGCCCGAUGAAAAUCGGGCGUACGCCUCGAGCUUUU [SEQ ID NO: 3880]
Последовательность PEgRNA HEK3.30.5a: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUUGGGGUUGGGGUUGGGGUUGGGGUUUU [SEQ ID NO: 3881]
Последовательность PEgRNA HEK3.30.5b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUGGUGGUGGUGGUUUU [SEQ ID NO: 3882]
Последовательность PEgRNA HEK3.30.13: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGCGAAAGCUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3883]
Последовательность PEgRNA HEK3.30.15: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGCUCGAAAGAGCUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3884]
Последовательность PEgRNA HEK3.30.15: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGCUCAUGAAAAUGAGCUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3885]
Последовательность PEgRNA HEK3.30.16: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGCUCAUCCGAAAGGAUGAGCUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3886]
Последовательность PEgRNA HEK3.30.17: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGCUCAUCCUGGAAACAGGAUGAGCUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3887]
Последовательность PEgRNA HEK3.30.18: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUGAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3888]
Последовательность PEgRNA HEK3.30.19: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUGAGAGCUAGCUCAUGAAAAUGAGCUAGCAAGUUUAAAUAAGGCUAGUC

CGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3889]
Последовательность PEgRNA HEK3.56: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAAAGCUUCGACCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3890]
Последовательность PEgRNA HEK3.56.1a: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGGCGAAAGCCUCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3891]
Последовательность PEgRNA HEK3.56.1b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGACGAAAGCCUCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3892]
Последовательность PEgRNA HEK3.56.1c: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGGCGAAAGCCCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3892]
Последовательность PEgRNA HEK3.56.2a: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGAUGCGAAAGCAUCUCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3893]
Последовательность PEgRNA HEK3.56.2b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGAUGCGAAAGCACCUCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3894]
Последовательность PEgRNA HEK3.56.2c: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGAUGCGAAAGCAUCCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3895]
Последовательность PEgRNA HEK3.56.3a: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGACAUGCGAAAGCAUGUCUCGUGCUCAGUCUG-Концевой мотив-

UUUU [SEQ ID NO: 3896]
Последовательность PEgRNA HEK3.56.3b: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGACAUGCGAAAGCAGGCCCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3897]
Последовательность PEgRNA HEK3.56.3c: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-Матрица и PBS-UCUGCCAUCAGACAUGCGAAAGCAUGUCUCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3898]
Последовательность PEgRNA HEK3.56.4a: спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAUACGUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUUACGAAGUGGGACCGAGUCGGUCC-Матрица и PBS-UCUGCCAUCAAAGCUUCGACCGUGCUCAGUCUG-Концевой мотив-UUUU [SEQ ID NO: 3899]
Последовательность PEgRNA HEK3.56.4b: 5' мотив-GCAGACCUAAGUGGUGACAUAUGGUCUG-spacer-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAUACGUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUUACGAAGUGGGACCGAGUCGGUCC-Матрица и PBS-Концевой мотив-UUUU [SEQ ID NO: 3922]
Последовательность PEgRNA HEK3.56.4c: 5' мотив-GCAGACCUAAGUGGUGACAUAUGGUCUG-спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAUACGUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUUACGAAGUGGGACCGAGUCGGUCC-Матрица и PBS--Концевой мотив-UUUU [SEQ ID NO: 3900]
Последовательность PEgRNA HEK3.56.4d: 5' мотив-GCAGACCUAAGUGGUGACAUAUGGUCUG-спейсер-GGCCCAGACUGAGCACGUGA-каркас-GUUUUAGAGCUAUACGUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUUACGAAGUGGGACCGAGUCGGUCC-Матрица и PBS-Концевой мотив-UUUU [SEQ ID NO: 3901]

Следует отметить, что при отсутствии концевого мотива или при наличии концевого мотива, который не оканчивается в виде последовательности из нескольких U, транскрипция терминировалась с помощью следующего рибозима HDV: GGCCGGCAUGGUCCCAGCCUCCUCGCUGGCGCCGGCUGGGCAACAUGCUUCGGCAUGGCGAAUGGGAC [SEQ ID NO: 3923]

Дополнительные мотивы РНК

[1551] См. подробное описание некоторых мотивов на Фиг.82, например рибозима HDV 3' PEgRNA, или вставки в виде G-квадруплекса, удлинений P1, шпилек матрицы, и обеспечивающих циркуляризацию за счет добавления тетрапетли, которые могут быть внесены в PEgRNA для повышения ее эффективности.

[1552] В частности, в рамках данного Примера было протестировано влияние внесения мотива тРНК 3' участка связывания праймера. Данный элемент был выбран в связи с несколькими потенциальными функциями:

[1553] (1) мотив нРНК является очень стабильным мотивом РНК и поэтому потенциально может снижать расщепление PEgRNA;

[1554] (2) RT MMLV использует пролил-тРНК в качестве праймера при конвертации вирусного генома в ДНК во время транскрипции, поэтому ожидалось, что тот же кэп может быть связан RT с повышением связывания PEgRNA за счет PE, а также повышением стабильности РНК и возвращением PBS обратно в непосредственную близость от геномного участка, что потенциально также повышает активность.

[1555] В рамках данных конструктов был удлинен P1 тРНК (см. Фиг.84). P1 обозначает первую шпильку/элемент в виде спаренных оснований тРНК (см. Фиг.84). Предполагалось, что это является необходимым для предотвращения опосредуемого РНКазой P расщепления тРНК 5' P1, что приводит к его удалению из PEgRNA.

[1556] В рамках данной конструкции использовали пролил-тРНК (кодон CGG) с удлиненным P1 и коротким 3 nt линкером между тРНК и PBS. Было протестировано множество конструкций тРНК, а также была протестирована эффективность редактирования по сравнению с PEgRNA без тРНК-кэпа – см. сравнительные данные на Фиг.83 (изображен эксперимент на основе PE, нацеленный на редактирование гена HEK3, в частности нацеленный на внесение вставки длиной 10 nt в положении +1 относительно участка разрыва и использование PE3), Фиг.85 (изображен эксперимент на основе PE, нацеленный на редактирование гена FANCF, в частности нацеленный на конверсию G на T в положении +5 относительно участка разрыва и использование конструкта PE3), и Фиг.86 (изображен эксперимент на основе PE, нацеленный на редактирование гена HEK3, в частности нацеленный на вставку метки FLAG длиной 71 nt в положении +1 относительно участка разрыва и использование конструкта PE3). Модифицированные на основе тРНК PEgRNA тестировали против контроля в виде немодифицированной PEgRNA.

[1557] UGG/CGG обозначает использованный кодон, при этом число обозначает длину добавленного удлинения P1, "длинный" обозначает линкер длиной 8 nt, отсутствие обозначения обозначает линкер длиной 3 nt.

[1558] Данные свидетельствуют о том, что внесение тРНК может обеспечивать применение более коротких PBS, что вероятно будет приводить к дополнительному повышению активности. В случае RNF2 является возможным/вероятным, что использованный линкер приводил к повышению связывания PBS со спейсером и возникающему вследствие этого снижению активности.

[1559] Некоторые из использованных последовательностей:

[1560] Вставка +1 FLAG-метки в HEK3, пролил-нРНК{UGG} P1 удлинение 5 nt, линкер 3 nt

[1561] GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAA AUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGA GGAAGCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAU CCGUGCUCAGUCUGUCUGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCU UCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3902]

[1562] Конверсия +5 G на T в FANCF пролил-тРНК{CGG} P1 удлинение 5 nt, линкер 3 nt

[1563] GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAA AUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAA AAGCGAUCAAGGUGCUGCAGAAGGGAUCUGGCGGGGCUCGUUGGUCUAGGGG UAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCC CCGCCUUUU [SEQ ID NO: 3903]

[1564] 10 nt вставка +1 в HEK3, пролил-тРНК{UGG} P1 удлинение 5 nt, линкер 3 nt

[1565] GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAA AUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGACCGAGUCGGUCCUCUG CCAUCAAAGCUUCGACCGUGCUCAGUCUUCUGCUCGAGGCGGGGCUCGUUGGU CUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGG ACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3904]

[1566] Последовательности, указанные в рамках данных, приведенных на Фиг.85 и 86, являются следующими:

Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS-

GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-частичный CGG тРНК линкер 8- UCUCUCUCUGGUCUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGG UUCAAUUUU [SEQ ID NO: 3905]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 5 линкер 3- UCUGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAG GUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3906]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 5 линкер 8- UCUCUCUCGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGC GAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3907]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 8 линкер 3- UCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGA GAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3908]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 8 линкер 8- UCUCUCUCCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGG UGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3909]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 11 линкер 3- UCUGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUG CGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3910]
Последовательность PEgRNA FANCF +5 G на T: спейсер, каркас матрицы и PBS- GGAAUCCCUUCUGCAGCACCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAAAAGC GAUCAAGGUGCUGCAGAAGGGA-UGG P1 удлинение 11 линкер 8- UCUCUCUCGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUG GGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3911]
Последовательность PEgRNA HEK3 +1 10 nt вставка: спейсер, каркас матрицы и PBS-

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTCTGCCATCAAAG CTTCGACCGTGCTCAGTCTG-UGG P1 удлинение 5 линкер 3- UCUGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAG GUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3913
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGGAA GCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAUCCGU GCUCAGUCUG-частичный CGG тРНК линкер 8- UCUCUCUCUGGUCUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGG UUCAAUUUU [SEQ ID NO: 3914]
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGGAA GCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAUCCGU GCUCAGUCUG-UGG P1 удлинение 5 линкер 3- UCUGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAG GUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3915]
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGGAA GCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAUCCGU GCUCAGUCUG-UGG P1 удлинение 5 линкер 8- UCUCUCUCGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGC GAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUUUU [SEQ ID NO: 3916]
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGG AAGCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAU CCGUGCUCAGUCUG-UGG P1 удлинение 8 линкер 8- UCUCUCUCCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGG GUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3917]
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGGAA GCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAUCCGU GCUCAGUCUG-UGG P1 удлинение 11 линкер 8- UCUCUCUCGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUG GGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3918]
Последовательность PEgRNA HEK3 +1 вставка FLAG: спейсер, каркас матрицы и PBS- GGCCCAGACUGAGCACGUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUGGAGGAA GCAGGGCUUCCUUUCCUCUGCCAUCACUUAUCGUCGUCAUCCUUGUAAUCCGU GCUCAGUCUG-UGG P1 удлинение 14 линкер 8-

UCUCUCUCGGUGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCACCUUUU [SEQ ID NO: 3919]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{CGG}-5-UCUGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCC [SEQ ID NO: 3920]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{CGG}-8-UCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3921]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{CGG}-11-UCUGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUCGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3922]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{Lys}-5-UCUGGCGGGCCCGGAUAGCUCAGUCGGUAGAGCAUCAGACUUUUAAUCUGAGGGUCCAGGGUUCAAGUCCCUGUUCGGGCCCGCCUUUU [SEQ ID NO: 3923]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{Lys}-8-UCUCGAGGCGGGCCCGGAUAGCUCAGUCGGUAGAGCAUCAGACUUUUAAUCUGAGGGUCCAGGGUUCAAGUCCCUGUUCGGGCCCGCCUCGUUUU [SEQ ID NO: 3924]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{UGG}-8-UCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3925]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{UGG}-11-UCUGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID

NO: 3926]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{UGG}-8-более длинный линкер-UCUCUCUCCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGUUUU [SEQ ID NO: 3927]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{UGG}-11-более длинный линкер-UCUCUCUCGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCUUUU [SEQ ID NO: 3928]
Последовательность PEgRNA RNF2 +1 C на A: спейсер, каркас матрицы и PBS-GUCAUCUUAGUCAUUACCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGAACACCUCAUGUAAUGACUAAGAUG-тРНК-Pro{UGG}-14-более длинный линкер-UCUCUCUCGGUGCUCGAGGCGGGGCUCGUUGGUCUAGGGGUAUGAUUCUCGCUUUGGGUGCGAGAGGUCCCGGGUUCAAAUCCCGGACGAGCCCCGCCUCGAGCACCUUUU [SEQ ID NO: 3929]

ПРИМЕР 23. ПРИМЕНЕНИЕ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ ДЛЯ КОРРЕКТИРОВАНИЯ CDKL5 И СЕРПОВИДНОКЛЕТОЧНОЙ АНЕМИИ

Применение PE для конструирования модели CDKL5 с мутацией 1412delA на основе мышей

[1567] Связанное с дефицитом CDKL5 заболевание (CDD) представляет собой нейродегенеративное заболевание, которое наиболее часто вызывается спонтанными мутациями в циклин-зависимом киназоподобном 5 гене. Симптомы проявляются в раннем детстве и включают судороги, нерегулярный сон, желудочно-кишечный стресс, и задержку развития. Некоторые мутации, которые вызывают CDD, включая 1412delA, не могут быть скорректированы с помощью редактирования оснований. Однако праймерное редактирование имеет потенциал в плане точного корректирования изменений в виде основание-на-основание, делеций, и вставок. В плане мутации 1412delA в рамках данного Примера были сконструированы и протестированы PEgRNA, способные обеспечивать вставку мутации в попытке определить линию нейронных клеток мыши (N2A), несущих мутацию. Это обеспечит возможность проведения широкого скрининга для поиска потенциально терапевтических PEgRNA,

которые корректируют данную мутацию. Конечная цель заключается в возможности проведения анализа модели CDD на основе мышей с мутацией 1412delA для анализа терапевтического эффекта. Современные модели CDD на основе мышей не имеют гуманизированного аллеля, однако оптимизация PEgRNA в настоящее время проводится также в клетках HEK293T. На Фиг.87 и 88 показаны результаты первого скрининга в клетках N2A, где PEgRNA обеспечивает внесение 1412Adel, с подробной информацией о длине участка связывания праймера (PBS) и длине матрицы в виде обратной транскриптазы (RT). (Показаны с/без инделей).

Применение PE для лечения серповидноклеточной анемии (SCA)

[1568] Серповидноклеточная анемия (SCA) представляет собой рецессивное заболевание крови, вызываемое мутацией в виде глутамат на валин в положении 6 в гене β-глобина. Результатом являются серповидные красные кровяные клетки, являющиеся плохими транспортерами кислорода и подверженные агрегации. Связанные с агрегацией симптомы могут угрожать жизни. Ранее в лаборатории доктора Liu удалось показать как внесение, так и корректирование локуса SCA в клетках HEK293T с использованием праймерного редактирования путем трансфекции ДНК-плазмидой. Поскольку гемопоэтические стволовые клетки (HSC) являются сложными в плане редактирования на основе трансфекции ДНК-плазмидой, в данном Примере тестировали систему PE3 в HSC с нуклеофекцией белка и иРНК. На Фиг.89 показаны результаты редактирования в промежуточном локусе в гене β-глобина и в HEK3 в здоровых HSC при различной концентрации редактора относительно PEgRNA и никирующей нРНК.

Протокол нуклеофекции иРНК:

[1569] Протокол был усовершенствован путем корректирования соотношения редактора и направляющих (соотношение [редактор] / [направляющая] или соотношение редактор:направляющая).

[1570] Последовательность никирующего направляющего протоспейсера являлась следующей: CCTTGATACCAACCTGCCCA

[1571] Последовательность PEgRNA являлась следующей: CATGGTGCACCTGACTCCTGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC

TAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAGACTTCTCTTCAG GAGTCAGGTGCACTTT

[1572] 1. Разморозить клетки CD34+ и предварительно стимулировать в среде X-Vivo15 цитокинами (SCF, Flt3 и TPO) в течение 24 часов. Плотность высевания 1 млн клеток/мл.

[1573] 2. На следующий день (спустя 24 часа) аспирировать клетки CD34+ (1 x 105) и промыть один раз PBS (центрифугирование осуществлять при 300 g течение 10 минут при комнатной температуре) и ресуспендировать в раствор P3 (Lonza) (20 мкл на одно условие).

[1574] 3. В вытяжном шкафу соединить 2 мкг иРНК SpPE2, 1 мкг соединенной PEgRNA и sgРНК в стерильном ПЦР стрипе на льду. Довести объем до 20 мкл (клетки в растворе P3 + смесь иРНК и нРНК).

[1575] 4. Внести клеточную суспензию в 20 мкл Lonza 4D 16-луночный стрип с помощью пипетки и подвергнуть электропорации с помощью программы DS130.

[1576] 5. Подождать 10-15 минут после электропорации и добавить 80 мкл среды X-Vivo10+цитокин в клеточную суспензию и перенести в предварительно нагретую среду X-Vivo10 с цитокином.

[1577] 6. Спустя 72 часа после электропорации собрать клетки и проверить выделение клеток и осуществить генотипирование в рамках общей популяции, а также отсортированных популяций CD34+ и CD34+90+.

<u>Белок CDKL5 дикого типа (кат. # NP_001032420) (изоформа 1 – человек)</u>
MKIPNIGNVM NKFEILGVVG EGAYGVVLKC RHKETHEIVA IKKFKDSEEN EEVKETTLRE LKMLRTLKQE NIVELKEAFR RRGKLYLVFE YVEKNMLELL EEMPNGVPPE KVKSYIYQLI KAIHWCHKND IVHRDIKPEN LLISHNDVLK LCDFGFARNL SEGNNANYTE YVATRWYRSP ELLLGAPYGK SVDMWSVGCI LGELSDGQPL FPGESEIDQL FTIQKVLGPL PSEQMKLFYS NPRFHGLRFP AVNHPQSLER RYLGILNSVL LDLMKNLLKL DPADRYLTEQ CLNHPTFQTQ

RLLDRSPSRS AKRKPYHVES STLSNRNQAG KSTALQSHHR SNSKDIQNLS
VGLPRADEGL PANESFLNGN LAGASLSPLH TKTYQASSQP GSTSKDLTNN
NIPHLLSPKE AKSKTEFDFN IDPKPSEGPG TKYLKSNSRS QQNRHSFMES
SQSKAGTLQP NEKQSRHSYI DTIPQSSRSP SYRTKAKSHG ALSDSKSVSN
LSEARAQIAE PSTSRYFPSS CLDLNSPTSP TPTRHSDTRT LLSPSGRNNR
NEGTLDSRRT TTRHSKTMEE LKLPEHMDSS HSHSLSAPHE SFSYGLGYTS
PFSSQQRPHR HSMYVTRDKV RAKGLDGSLS IGQGMAARAN SLQLLSPQPG
EQLPPEMTVA RSSVKETSRE GTSSFHTRQK SEGGVYHDPH SDDGTAPKEN
RHLYNDPVPR RVGSFYRVPS PRPDNSFHEN NVSTRVSSLP SESSSGTNHS
KRQPAFDPWK SPENISHSEQ LKEKEKQGFF RSMKKKKKKS QTVPNSDSPD
LLTLQKSIHS ASTPSSRPKE WRPEKISDLQ TQSQPLKSLR KLLHLSSASN
HPASSDPRFQ PLTAQQTKNS FSEIRIHPLS QASGGSSNIR QEPAPKGRPA
LQLPDGGCDG RRQRHHSGPQ DRRFMLRTTE QQGEYFCCGD PKKPHTPCVP
NRALHRPISS PAPYPVLQVR GTSMCPTLQV RGTDAFSCPT QQSGFSFFVR
HVMREALIHR AQVNQAALLT YHENAALTGK

<u>Белок CDKL5 дикого типа (кат. # NP_001310218) (изоформа 2 – человек)</u>

MKIPNIGNVM NKFEILGVVG EGAYGVVLKC RHKETHEIVA IKKFKDSEEN
EEVKETTLRE LKMLRTLKQE NIVELKEAFR RRGKLYLVFE YVEKNMLELL
EEMPNGVPPE KVKSYIYQLI KAIHWCHKND IVHRDIKPEN LLISHNDVLK
LCDFGFARNL SEGNNANYTE YVATRWYRSP ELLLGAPYGK SVDMWSVGCI
LGELSDGQPL FPGESEIDQL FTIQKVLGPL PSEQMKLFYS NPRFHGLRFP
AVNHPQSLER RYLGILNSVL LDLMKNLLKL DPADRYLTEQ CLNHPTFQTQ
RLLDRSPSRS AKRKPYHVES STLSNRNQAG KSTALQSHHR SNSKDIQNLS
VGLPRADEGL PANESFLNGN LAGASLSPLH TKTYQASSQP GSTSKDLTNN
NIPHLLSPKE AKSKTEFDFN IDPKPSEGPG TKYLKSNSRS QQNRHSFMES
SQSKAGTLQP NEKQSRHSYI DTIPQSSRSP SYRTKAKSHG ALSDSKSVSN
LSEARAQIAE PSTSRYFPSS CLDLNSPTSP TPTRHSDTRT LLSPSGRNNR
NEGTLDSRRT TTRHSKTMEE LKLPEHMDSS HSHSLSAPHE SFSYGLGYTS
PFSSQQRPHR HSMYVTRDKV RAKGLDGSLS IGQGMAARAN SLQLLSPQPG
EQLPPEMTVARSSVKETSRE GTSSFHTRQK SEGGVYHDPH SDDGTAPKEN
RHLYNDPVPR RVGSFYRVPS PRPDNSFHEN NVSTRVSSLP SESSSGTNHS

KRQPAFDPWK SPENISHSEQ LKEKEKQGFF RSMKKKKKKS QTVPNSDSPD LLTLQKSIHS ASTPSSRPKE WRPEKISDLQ TQSQPLKSLR KLLHLSSASN HPASSDPRFQ PLTAQQTKNS FSEIRIHPLS QASGGSSNIR QEPAPKGRPA LQLPGQMDPG WHVSSVTRSA TEGPSYSEQL GAKSGPNGHP YNRTNRSRMP NLNDLKETAL

ПРИМЕР 24. ПРИМЕНЕНИЕ ПРАЙМЕРНОГО РЕДАКТИРОВАНИЯ ДЛЯ НАЦЕЛИВАНИЯ НА ПАТОГЕННЫЕ АЛЛЕЛИ AOL1 В КАЧЕСТВЕ ЛЕЧЕНИЯ НЕДИАБЕТИЧЕСКОГО ХРОНИЧЕСКОГО ЗАБОЛЕВАНИЯ ПОЧЕК

[1578] В рамках данного Примера сконструированы PEgRNA, которые способны нацеливаться на патогенные аллели APOL1 для применения в рамках праймерного редактирования для лечения или снижения вероятности развития заболевания почек.

[1579] Терминальная стадия почечной недостаточности (ESKD) представляет собой растущую проблему, которая в настоящее время затрагивает более полумиллиона человек в Соединенных Штатах. Стоимость ухода за пациентами с ESKD в настоящее время составляет более 40 миллиардов долларов в год. В США вероятность того, что у субъектов африканского происхождения разовьется ESKD, в 4-5 раз выше, чем у американцев неафриканского происхождения. Данные факты соответствуют расхождению в виде 12-13% населения США африканского происхождения и 40% подвергаемых диализу пациентов в США, которые являются афроамериканцами. Эпидемия факторов риска, связанных с заболеванием почек, таких как ожирение и метаболический синдром, свидетельствует о том, что масштаб данной проблемы будет только увеличиваться.

[1580] Специфических видов лечения для подавляющего большинства прогрессирующих заболеваний почек не существует. Некоторые типы прогрессирования хронических заболевания почек могут быть замедлены за счет контролирования кровяного давления с помощью специальных агентов, однако нефрологи не могут точно предсказать то, у каких пациентов будет возникать ответ. Более того, в то время как успешное лечение обычно замедляет прогрессирование, оно не обеспечивает предотвращение заболевания и не останавливает прогрессирование заболевания.

[1581] Недавно было определено, что определенные генетические варианты, которые изменяют последовательность белка Аполипопротеин L1 (APOL1), ассоциированы с

прогрессирующим заболеванием почек. Является неожиданным то, что связанные с заболеванием почек варианты APOL1 оказывают значительное влияние на многие различные типы заболевания почек, включая ассоциированную с гипертензией терминальную стадию почечной недостаточности (H-ESRD), очаговый сегментарный гломерулосклероз (FSGS), и ВИЧ-ассоциированную нефропатию (HIVAN). Субъекты с данными вариантами аллелей APOL1 имеют 7-30-кратно повышенный риск развития заболевания почек. На основе высокой частоты данных связанных с риском аллелей APOL1, более 3,5 миллионов афроамериканцев вероятно имеют связанный с высоким риском генотип в плане APOL1. Афроамериканцы без связанного с высоким риском генотипа имеют малый риск по сравнению с американцами европейского происхождения.

[1582] Несмотря на свидетельство того, что варианты гена APOL1 вызывают заболевание почек, очень мало известно о биологических свойствах его продукта, APOL1, или его роли в почке. APOL1 имеет определенную роль в плане устойчивости к трипаносомам, а варианты G1 и G2 похоже стали обыденными в Африке, поскольку они придает защиту против тех форм трипаносом, которые вызывают африканскую сонную болезнь.

[1583] Все еще существует потребность в способах лечения заболеваний почек у пациентов с одним или несколькими связанными с риском аллелями APOL1, которые вызывают высокую заболеваемость и смертность с высокими экономическим потерями в рамках данной и других популяций субъектов

[1584] В рамках данного Примера предусмотрены три примера вариантов конструкции PEgRNA, основанные на определенной служащей в качестве примера последовательности-мишени, которые могут использоваться в рамках праймерного редактирования для корректирования дефективных аллелей APOL1.

[1585] PEgRNA 1

[1586] Конструирование PEgRNA для аллеля APOL1 rs73885319 (p.S342G). Это соответствует G→A коррекции у затронутых субъектов. Последовательность-мишень представляет собой 5-GGAGTC**AAGCTCACGGATGTGGCCCC**TGTA(G на A)GCTTCTTTCTTGTG CTGGATGTAGTCTACCT-3.

[1587] Протоспейсер (обозначен выше с помощью жирного шрифта) представляет собой AAGCTCACGGATGTGGCCCC. Выбранный PE включает SaCas9(D10A).

[1588] Участок связывания праймера может представлять собой:

[1589] GTGGCCCC

[1590] TGTGGCCCC

[1591] ATGTGGCCCC

[1592] GATGTGGCCCC

[1593] GGATGTGGCCCC

[1594] CGGATGTGGCCCC

[1595] ACGGATGTGGCCCC

[1596] CACGGATGTGGCCCC

[1597] TCACGGATGTGGCCCC

[1598] CTCACGGATGTGGCCCC.

[1599] RT матрицы могут представлять собой:

[1600] AAGAAGCTTACA

[1601] AAAGAAGCTTACA

[1602] GAAAGAAGCTTACA

[1603] AGAAAGAAGCTTACA

[1604] AAGAAAGAAGCTTACA

[1605] CAAGAAAGAAGCTTACA

[1606] ACAAGAAAGAAGCTTACA

[1607] CACAAGAAAGAAGCTTACA

[1608] GCACAAGAAAGAAGCTTACA

[1609] AGCACAAGAAAGAAGCTTACA

[1610] CAGCACAAGAAAGAAGCTTACA

[1611] CCAGCACAAGAAAGAAGCTTACA

[1612] TCCAGCACAAGAAAGAAGCTTACA

[1613] ATCCAGCACAAGAAAGAAGCTTACA.

[1614] Никирующая матрица может представлять собой

GCTTTGATTCGTACACGAGG.

[1615] PEgRNA 2

[1616] Конструирование PEgRNA для аллеля APOL1 rs60910145. Это соответствует G→T коррекции у затронутых субъектов.

[1617] Протоспейсер представляет собой GCTGGAGGAGAAGCTAAACA.

Выбранный PE включает SpCas9(D10A)-NG.

[1618] Участок связывания праймера может представлять собой:

[1619] GAAGCTAA

[1620] AGAAGCTAA

[1621] GAGAAGCTAA

[1622] GGAGAAGCTAA

[1623] AGGAGAAGCTAA

[1624] GAGGAGAAGCTAA

[1625] GGAGGAGAAGCTAA

[1626] TGGAGGAGAAGCTAA

[1627] CTGGAGGAGAAGCTAA

[1628] GCTGGAGGAGAAGCTAA.

[1629] RT матрица может представлять собой (не может заканчиваться на C):

[1630] AGAATGT

[1631] GAGAATGT

[1632] TGAGAATGT

[1633] TTGAGAATGT

[1634] GTTGAGAATGT

[1635] TGTTGAGAATGT

[1636] TTGTTGAGAATGT

[1637] ATTGTTGAGAATGT

[1638] TATTGTTGAGAATGT

[1639] TTATTGTTGAGAATGT

[1640] ATTATTGTTGAGAATGT

[1641] AATTATTGTTGAGAATGT

[1642] TAATTATTGTTGAGAATGT

[1643] ATAATTATTGTTGAGAATGT.

[1644] Никирующие матрицы могут представлять собой:

[1645] CCTGTGGTCACAGTTCTTGG

[1646] CCACAGGGCAGGGCAGCCAC.

[1647] PEgRNA 3

[1648] Конструирование PEgRNA для аллеля APOL1 rs71785313. Это соответствует вставке следующим образом: ATTCTCAACAA[вставка: TAATTA]TAAGATTC.

[1649] Протоспейсер может представлять собой: TCTCAACAATAAGATTCTGC

[1650] PE включает SaKKH-PE2.

[1651] Участок связывания праймера может представлять собой:

[1652] TTCTCAAC

[1653] ATTCTCAAC

[1654] CATTCTCAAC

[1655] ACATTCTCAAC

[1656] AACATTCTCAAC

[1657] AAACATTCTCAAC

[1658] TAAACATTCTCAAC

[1659] CTAAACATTCTCAAC

[1660] GCTAAACATTCTCAAC

[1661] AGCTAAACATTCTCAAC.

[1662] RT матрица может представлять собой:

[1663] AATCTTATAATTATT

[1664] GAATCTTATAATTATT

[1665] AGAATCTTATAATTATT

[1666] CAGAATCTTATAATTATT

[1667] GCAGAATCTTATAATTATT

[1668] TGCAGAATCTTATAATTATT

[1669] CTGCAGAATCTTATAATTATT

[1670] CCTGCAGAATCTTATAATTATT

[1671] GCCTGCAGAATCTTATAATTATT

[1672] CGCCTGCAGAATCTTATAATTATT

[1673] CCGCCTGCAGAATCTTATAATTATT

[1674] TCCGCCTGCAGAATCTTATAATTATT

[1675] GTCCGCCTGCAGAATCTTATAATTATT

[1676] GGTCCGCCTGCAGAATCTTATAATTATT.

[1677] Никирующие матрицы могут представлять собой:

[1678] CCTGTGGTCACAGTTCTTGG

[1679] CCACAGGGCAGGGCAGCCAC.

ДРУГИЕ ССЫЛКИ, ПРИВЕДЕННЫЕ В РАМКАХ ДАННОГО ОПИСАНИЯ

Каждая из следующих ссылок полностью приведена здесь в качестве ссылки.

1. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

2. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).

3. Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* **168**, 20–36 (2017).

4. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).

5. Nishida, K. *et al.* Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, aaf8729 (2016).

6. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

7. ClinVar, July 2019.

8. Dunbar, C. E. *et al.* Gene therapy comes of age. *Science* **359**, eaan4672 (2018).

9. Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* **21**, 121–131 (2015).

10. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **9**, 1911 (2018).

11. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).

12. Kleinstiver, B. P. *et al.* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).

13. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).

14. Nishimasu, H. *et al.* Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259–1262 (2018).

15. Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**, a012740 (2013).

16. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016).

17. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).

18. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).

19. Ihry, R. J. *et al.* p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).

20. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).

21. Srivastava, M. *et al.* An Inhibitor of Nonhomologous End-Joining Abrogates Double-Strand Break Repair and Impedes Cancer Progression. *Cell* **151**, 1474–1487 (2012).

22. Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).

23. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).

24. Kim, Y. B. *et al.* Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* **35**, 371–376

(2017).

25. Li, X. *et al.* Base editing with a Cpf1–cytidine deaminase fusion. *Nat. Biotechnol.* **36**, 324–327 (2018).

26. Gehrke, J. M. *et al.* An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* (2018). doi:10.1038/nbt.4199

27. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 1 (2018). doi:10.1038/s41576-018-0059-1.

28. Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. *Annu. Rev. Genet.* **35**, 501–538 (2001).

29. Zimmerly, S., Guo, H., Perlman, P. S. & Lambowltz, A. M. Group II intron mobility occurs by target DNA- primed reverse transcription. *Cell* **82**, 545–554 (1995).

30. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).

31. Feng, Q., Moran, J. V., Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).

32. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* **343**, 1247997 (2014).

33. Jiang, F. *et al.* Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* aad8282 (2016). doi:10.1126/science.aad8282

34. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

35. Tang, W., Hu, J. H. & Liu, D. R. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat. Commun.* **8**, 15939 (2017).

36. Shechner, D. M., Hacisuleyman, E., Younger, S. T. & Rinn, J. L. Multiplexable,

locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664–670 (2015).

37. Anders, C. & Jinek, M. Chapter One - *In vitro* Enzymology of Cas9. in *Methods in Enzymology* (eds. Doudna, J. A. & Sontheimer, E. J.) **546**, 1–20 (Academic Press, 2014).

38. Briner, A. E. *et al.* Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality. *Mol. Cell* **56**, 333–339 (2014).

39. Nowak, C. M., Lawson, S., Zerez, M. & Bleris, L. Guide RNA engineering for versatile Cas9 functionality. *Nucleic Acids Res.* **44**, 9555–9564 (2016).

40. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).

41. Mohr, S. *et al.* Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA* **19**, 958–970 (2013).

42. Stamos, J. L., Lentzsch, A. M. & Lambowitz, A. M. Structure of a Thermostable Group II Intron Reverse Transcriptase with Template-Primer and Its Functional and Evolutionary Implications. *Mol. Cell* **68**, 926- 939.e4 (2017).

43. Zhao, C. & Pyle, A. M. Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nat. Struct. Mol. Biol.* **23**, 558–565 (2016).

44. Zhao, C., Liu, F. & Pyle, A. M. An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA* **24**, 183–195 (2018).

45. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

46. Liu, Y., Kao, H.-I. & Bambara, R. A. Flap endonuclease 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* **73**, 589–615 (2004).

47. Krokan, H. E. & Bjørås, M. Base Excision Repair. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).

48. Kelman, Z. PCNA: structure, functions and interactions. *Oncogene* **14**, 629–640 (1997).

49. Choe, K. N. & Moldovan, G.-L. Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol. Cell* **65**, 380–392 (2017).

50. Li, X., Li, J., Harrington, J., Lieber, M. R. & Burgers, P. M. Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. *J. Biol. Chem.* **270**, 22109–22112 (1995).

51. Tom, S., Henricksen, L. A. & Bambara, R. A. Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J. Biol. Chem.* **275**, 10498–10505 (2000).

52. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).

53. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).

54. Dahlman, J. E. *et al.* Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat. Biotechnol.* **33**, 1159–1161 (2015).

55. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).

56. Tsai, S. Q. *et al.* CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).

57. Schek N, Cooke C, Alwine JC. Molecular and Cellular Biology. (1992).

58. Gil A, Proudfoot NJ. Cell. (1987).

59. Zhao, B.S., Roundtree, I.A., He, C. Nat Rev Mol Cell Biol. (2017).

60. Rubio, M.A.T., Hopper, A.K. Wiley Interdiscip Rev RNA (2011).

61. Shechner, D.M., Hacisuleyman E., Younger, S.T., Rinn, J.L. Nat Methods. (2015).

62. Paige, J.S., Wu, K.Y., Jaffrey, S.R. Science (2011).

63. Ray D., … Hughes TR. Nature (2013).

64. Chadalavada, D. M., Cerrone-Szakal, A. L., Bevilacqua, P.C. RNA (2007).

65. Forster AC, Symons RH. Cell. (1987).

66. Weinberg Z, Kim PB, Chen TH, Li S, Harris KA, Lünse CE, Breaker RR. Nat. Chem. Biol. (2015).

67. Feldstein PA, Buzayan JM, Bruening G. Gene (1989).

68. Saville BJ, Collins RA. Cell. (1990).

69. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. Nature (2004).

70. Roth A, Weinberg Z, Chen AG, Kim PG, Ames TD, Breaker RR. Nat Chem Biol. (2013).

71. Choudhury R, Tsai YS, Dominguez D, Wang Y, Wang Z. Nat Commun. (2012).

72. MacRae IJ, Doudna JA. Curr Opin Struct Biol. (2007).

73. Bernstein E, Caudy AA, Hammond SM, Hannon GJ Nature (2001).

74. Filippov V, Solovyev V, Filippova M, Gill SS. Gene (2000).

75. Cadwell RC and Joyce GF. PCR Methods Appl. (1992).

76. McInerney P, Adams P, and Hadi MZ. Mol Biol Int. (2014).

77. Esvelt KM, Carlson JC, and Liu DR. Nature. (2011).

78. Naorem SS, Hin J, Wang S, Lee WR, Heng X, Miller JF, Guo H. Proc Natl Acad Sci USA (2017).

79. Martinez MA, Vartanian JP, Wain-Hobson S. Proc Natl Acad Sci USA (1994).

80. Meyer AJ, Ellefson JW, Ellington AD. Curr Protoc Mol Biol. (2014).

81. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Nature. (2009).

82. Nyerges Á et al. Proc Natl Acad Sci USA. (2016).

83. Mascola JR, Haynes BF. Immunol Rev. (2013).

84. X. Wen, K. Wen, D. Cao, G. Li, R. W. Jones, J. Li, S. Szu, Y. Hoshino, L. Yuan, Inclusion of a universal tetanus toxoid CD4(+) T cell epitope P2 significantly enhanced the immunogenicity of recombinant rotavirus ΔVP8* subunit parenteral vaccines. Vaccine 32, 4420-4427 (2014).

85. G. Ada, D. Isaacs, Carbohydrate-protein conjugate vaccines. Clin Microbiol Infect 9, 79-85 (2003).

86. E. Malito, B. Bursulaya, C. Chen, P. L. Surdo, M. Picchianti, E. Balducci, M. Biancucci, A. Brock, F. Berti, M. J. Bottomley, M. Nissum, P. Costantino, R. Rappuoli, G. Spraggon, Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. Proceedings of the National Academy of Sciences 109, 5229 (2012).

87. J. de Wit, M. E. Emmelot, M. C. M. Poelen, J. Lanfermeijer, W. G. H. Han, C. van Els, P. Kaaijk, The Human CD4(+) T Cell Response against Mumps Virus Targets a Broadly Recognized Nucleoprotein Epitope. J Virol 93, (2019).

88. M. May, C. A. Rieder, R. J. Rowe, Emergent lineages of mumps virus suggest the need for a polyvalent vaccine. Int J Infect Dis 66, 1-4 (2018).

89. M. Ramamurthy, P. Rajendiran, N. Saravanan, S. Sankar, S. Gopalan, B. Nandagopal, Identification of immunogenic B-cell epitope peptides of rubella virus E1 glycoprotein towards development of highly specific immunoassays and/or vaccine. Conference Abstract, (2019).

90. U. S. F. Tambunan, F. R. P. Sipahutar, A. A. Parikesit, D. Kerami, Vaccine Design for H5N1 Based on B- and T-cell Epitope Predictions. Bioinform Biol Insights 10, 27-35 (2016).

91. Asante, EA. et. al. "A naturally occurring variant of the human prion protein completely prevents prion disease". Nature. (2015).

92. Crabtree, G. R. & Schreiber, S. L. Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem. Sci.* **21**, 418–22 (1996).

93. Liu, J. *et al.* Calcineurin Is a Common Target of A and FKBP-FK506 Complexes. *Cell* **66**, 807–815 (1991).

94. Keith, C. T. *et al.* A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* **369**, 756–758 (2003).

95. Spencer, D. M., Wandless, T. J., Schreiber, S. L. S. & Crabtree, G. R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019–24 (1993).

96. Pruschy, M. N. *et al.* Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chem. Biol.* **1**, 163–172 (1994).

97. Spencer, D. M. *et al.* Functional analysis of Fas signaling *in vivo* using synthetic inducers of dimerization. *Curr. Biol.* **6**, 839–847 (1996).

98. Belshaw, P. J., Spencer, D. M., Crabtree, G. R. & Schreiber, S. L. Controlling programmed cell death with a cyclophilin-cyclosporin-based chemical inducer of dimerization. *Chem. Biol.* **3**, 731–738 (1996).

99. Yang, J. X., Symes, K., Mercola, M. & Schreiber, S. L. Small-molecule control of insulin and PDGF receptor signaling and the role of membrane attachment. *Curr. Biol.* **8**, 11–18 (1998).

100. Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc. Natl. Acad. Sci.* **93**, 4604–4607 (2002).

101. Stockwell, B. R. & Schreiber, S. L. Probing the role of homomeric and heteromeric receptor interactions in TGF-β signaling using small molecule dimerizers. *Curr. Biol.* **8**, 761–773 (2004).

102. Spencer, D. M., Graef, I., Austin, D. J., Schreiber, S. L. & Crabtree, G. R. A general strategy for producing conditional alleles of Src-like tyrosine kinases. *Proc. Natl. Acad. Sci.* **92**, 9805–9809 (2006).

103. Holsinger, L. J., Spencer, D. M., Austin, D. J., Schreiber, S. L. & Crabtree, G. R. Signal transduction in T lymphocytes using a conditional allele of Sos. *Proc. Natl. Acad. Sci.* **92**, 9810–9814 (2006).

104. Myers, M. G. Insulin Signal Transduction and the IRS Proteins. *Annu. Rev. Pharmacol. Toxicol.* **36**, 615–658 (1996).
105. Watowich, S. S. The erythropoietin receptor: Molecular structure and hematopoietic signaling pathways. *J. Investig. Med.* **59**, 1067–1072 (2011).
106. Blau, C. A., Peterson, K. R., Drachman, J. G. & Spencer, D. M. A proliferation switch for genetically modified cells. *Proc. Natl. Acad. Sci.* **94**, 3076–3081 (2002).
107. Clackson, T. *et al.* Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc. Natl. Acad. Sci.* **95**, 10437–10442 (1998).
108. Diver, S. T. & Schreiber, S. L. Single-step synthesis of cell-permeable protein dimerizers that activate signal transduction and gene expression. *J. Am. Chem. Soc.* **119**, 5106–5109 (1997).
109. Guo, Z. F., Zhang, R. & Liang, F. Sen. Facile functionalization of FK506 for biological studies by the thiol-ene 'click' reaction. *RSC Adv.* **4**, 11400–11403 (2014).
110. Robinson, D. R., Wu, Y.-M. & Lin, S.-F. The protein tyrosine kinase family of the human genome. *Oncogene* **19**, 5548–5557 (2000).
111. Landrum, M. J. et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 44, D862–D868 (2016).
112. Jinek, M. et al. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821 (2012).
113. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339, 819–823 (2013).
114. Mali, P. et al. RNA-Guided Human Genome Engineering via Cas9. Science 339, 823–826 (2013).
115. Yang, H. et al. One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering. Cell 154, 1370–1379 (2013).
116. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 24, 1012–1019 (2014).
117. Orlando, S. J. et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res. 38, e152–e152 (2010).
118. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. 33, 187–197 (2015).

119. Suzuki, K. et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature 540, 144–149 (2016).
120. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nat. Biotechnol. 36, 765–771 (2018).
121. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nat. Med. 24, 927–930 (2018).
122. Ihry, R. J. et al. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells.
Nat. Med. 24, 939–946 (2018).
123. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the end game: DNA double- strand break repair pathway choice. Mol. Cell 47, 497–510 (2012).
124. Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. Nat. Med. 21, 121–131 (2015).
125. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature 533, 125–129 (2016).
126. Chu, V. T. et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9- induced precise gene editing in mammalian cells. Nat. Biotechnol. 33, 543–548 (2015).
127. Maruyama, T. et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat. Biotechnol. 33, 538–542 (2015).
128. Rees, H. A., Yeh, W.-H. & Liu, D. R. Development of hRad51–Cas9 nickase fusions that mediate HDR without double-stranded breaks. Nat. Commun. 10, 1–12 (2019).
129. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420– 424 (2016).
130. Gaudelli, N. M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551, 464–471 (2017).
131. Gao, X. et al. Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. Nature 553, 217–221 (2018).

132. Ingram, V. M. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. Nature 178, 792–794 (1956).
133. Myerowitz, R. & Costigan, F. C. The major defect in Ashkenazi Jews with Tay-Sachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase. J. Biol. Chem. 263, 18587–18589 (1988).
134. Zielenski, J. Genotype and Phenotype in Cystic Fibrosis. Respiration 67, 117–133 (2000).
135. Mead, S. et al. A Novel Protective Prion Protein Variant that Colocalizes with Kuru Exposure. N. Engl. J. Med. 361, 2056–2065 (2009).
136. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322, 1843–1845 (2008).
137. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes.
Science 315, 1709–1712 (2007).
138. Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. Annu. Rev. Biophys. 46, 505–529 (2017).
139. Hille, F. et al. The Biology of CRISPR-Cas: Backward and Forward. Cell 172, 1239–1259 (2018).
140. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non- LTR retrotransposition. Cell 72, 595–605 (1993).
141. Liu, Y., Kao, H.-I. & Bambara, R. A. Flap endonuclease 1: a central component of DNA metabolism. Annu. Rev. Biochem. 73, 589–615 (2004).
142. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat. Rev. Genet. 19, 770 (2018).
143. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat. Biotechnol. 34, 339–344 (2016).
144. Qi, L. S. et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152, 1173–1183 (2013).
145. Shechner, D. M., Hacisuleyman, E., Younger, S. T. & Rinn, J. L. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. Nat. Methods 12, 664–670 (2015).

146. Tang, W., Hu, J. H. & Liu, D. R. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. Nat. Commun. 8, 15939 (2017).
147. Jinek, M. et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. Science 343, 1247997 (2014).
148. Nishimasu, H. et al. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell 156, 935–949 (2014).
149. Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9–guide RNA complex preorganized for target DNA recognition. Science 348, 1477–1481 (2015).
150. Baranauskas, A. et al. Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. Protein Eng. Des. Sel. 25, 657–668 (2012).
151. Gerard, G. F. et al. The role of template-primer in protection of reverse transcriptase from thermal inactivation. Nucleic Acids Res. 30, 3118–3129 (2002).
152. Arezi, B. & Hogrefe, H. Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. Nucleic Acids Res. 37, 473–481 (2009).
153. Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M. & Gerard, G. F. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. Nucleic Acids Res. 16, 265–277 (1988).
154. Shen, M. W. et al. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature 563, 646–651 (2018).
155. Thuronyi, B. W. et al. Continuous evolution of base editors with expanded target compatibility and improved activity. Nat. Biotechnol. (2019). doi:10.1038/s41587-019- 0193-0
156. Kim, Y. B. et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat. Biotechnol. 35, 371–376 (2017).
157. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. (2018). doi:10.1038/nbt.4172

158. Komor, A. C. et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. Sci. Adv. 3, eaao4774 (2017).
159. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495 (2016).
160. Zuo, E. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science 364, 289–292 (2019).
161. Jin, S. et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science 364, 292–295 (2019).
162. Kim, D., Kim, D., Lee, G., Cho, S.-I. & Kim, J.-S. Genome-wide target specificity of CRISPR RNA-guided adenine base editors. Nat. Biotechnol. 37, 430–435 (2019).
163. Grünewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. Nature 569, 433–437 (2019).
164. Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. Nature 571, 275–278 (2019).
165. Rees, H. A., Wilson, C., Doman, J. L. & Liu, D. R. Analysis and minimization of cellular RNA editing by DNA adenine base editors. Sci. Adv. 5, eaax5717 (2019).
166. Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. Annu. Rev. Genet. 35, 501–538 (2001).
167. Griffiths, D. J. Endogenous retroviruses in the human genome sequence. Genome Biol.
2, REVIEWS1017 (2001).
168. Berkhout, B., Jebbink, M. & Zsíros, J. Identification of an Active Reverse Transcriptase Enzyme Encoded by a Human Endogenous HERV-K Retrovirus. J. Virol. 73, 2365–2375 (1999).
169. Halvas, E. K., Svarovskaia, E. S. & Pathak, V. K. Role of Murine Leukemia Virus Reverse Transcriptase Deoxyribonucleoside Triphosphate-Binding Site in Retroviral Replication and In Vivo Fidelity. J. Virol. 74, 10349–10358 (2000).
170. Dever, D. P. et al. CRISPR/Cas9 Beta-globin Gene Targeting in Human Hematopoietic Stem Cells. Nature 539, 384–389 (2016).
171. Park, S. H. et al. Highly efficient editing of the β-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. Nucleic Acids Res. doi:10.1093/nar/gkz475

172. Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. Annu. Rev. Neurosci. 24, 519–550 (2001).
173. Asante, E. A. et al. A naturally occurring variant of the human prion protein completely prevents prion disease. Nature 522, 478–481 (2015).
174. Anzalone, A. V., Lin, A. J., Zairis, S., Rabadan, R. & Cornish, V. W. Reprogramming eukaryotic translation with ligand-responsive synthetic RNA switches. Nat. Methods 13, 453–458 (2016).
175. Badran, A. H. et al. Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. Nature 533, 58–63 (2016).
176. Anders, C. & Jinek, M. Chapter One - In Vitro Enzymology of Cas9. in Methods in Enzymology (eds. Doudna, J. A. & Sontheimer, E. J.) 546, 1–20 (Academic Press, 2014).
177. Pirakitikulr, N., Ostrov, N., Peralta-Yahya, P. & Cornish, V. W. PCRless library mutagenesis via oligonucleotide recombination in yeast. Protein Sci. Publ. Protein Soc. 19, 2336–2346 (2010).
178. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. 37, 224–226 (2019).
179. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. 33, 187–197 (2015).
180. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off- target effects. Nature 529, 490–495 (2016).
181. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. (2018). doi:10.1038/nbt.4172
182. Baranauskas, A. et al. Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. Protein Eng. Des. Sel. 25, 657–668 (2012).
183. Schechner, DM, Hacisuleyman E., Younger ST, Rinn JL. Nat Methods 664-70 (2015).
184. Brown JA, et al. Nat Struct Mol Biol 633-40 (2014).
185. Conrad NA and Steitz JA. EMBO J 1831-41 (2005).
186. Bartlett JS, et al. Proc Natl Acad Sci USA 8852-7 (1996).
187. Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA. Science 1244-7 (2010).
188. Forster AC, Symons RH. Cell. 1987.

189. Weinberg Z, Kim PB, Chen TH, Li S, Harris KA, Lünse CE, Breaker RR. Nat. Chem. Biol. 2015.
190. Feldstein PA, Buzayan JM, Bruening G. Gene 1989.
191. Saville BJ, Collins RA. Cell. 1990.
192. Roth A, Weinberg Z, Chen AG, Kim PG, Ames TD, Breaker RR. Nat Chem Biol. 2013.
193. Borchardt EK, et al. RNA 1921-30 (2015).
194. Zhang Y, et al. Mol Cell 792-806 (2013).
195. Dang Y, et al. Genome Biol 280 (2015).
196. Schaefer M, Kapoor U, and Jantsch MF. Open Biol 170077 (2017).
197. Nahar S, et al. Chem Comm 2377-80 (2018).
198. Gao Y and Zhao Y. J Integr Plant Biol 343-9 (2014).
199. Dubois N, Marquet R, Paillart J, Bernacchi S. Front Microbiol 527 (2018).
200. Costa M and Michel F. EMBO J 1276-85 (1995).
201. Hu JH, et al. Nature 57-63 (2018).
202. Furukawa K, Gu H, Breaker RR. Methods Mol Biol 209-20 (2014).
203. Zettler, J., Schütz, V. & Mootz, H. D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. FEBS Lett. 583, 909–914 (2009).
204. Kügler, S., Kilic, E. & Bähr, M. Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. Gene Ther. 10, 337–347 (2003).
205. de Felipe, P., Hughes, L. E., Ryan, M. D. & Brown, J. D. Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. J. Biol. Chem. 278, 11441–11448 (2003).
206. Levy, J. M. & Nicoll, R. A. Membrane-associated guanylate kinase dynamics reveal regional and developmental specificity of synapse stability. J. Physiol. 595, 1699–1709 (2017).
207. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323 (2011).
208. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47–e47 (2015).

ВАРИАНТЫ ОСУЩЕСТВЛЕНИЯ

[1680] Приведенные ниже варианты осуществления находятся в рамках объема настоящего изобретения. Более того, изобретение охватывает все вариации, комбинации, и пермутации данных вариантов осуществления, в которых одно или несколько ограничений, элементов, пунктов, и описательных элементов из одного или нескольких перечисленных вариантов осуществления внесены в другой перечисленный вариант осуществления в рамках данного раздела. Например, любой перечисленный вариант осуществления, который зависит от другого варианта осуществления, может быть модифицирован для включения одного или нескольких ограничений, указанных в любом другом перечисленном варианте осуществления в рамках данного раздела, который зависит от того же первоначального варианта осуществления. Если элементы представлены в виде перечней, например в формате группы Маркуша, каждая подгруппа данных элементов также включается, при этом любой элемент (элементы) могут быть удалены из группы. Следует понимать, что, в целом, при обозначении изобретения или аспектов изобретения в качестве включающих определенные элементы и/или свойства, определенные варианты осуществления изобретения или аспекты изобретения состоят или состоят по существу из подобных элементов и/или свойств. Следует также отметить, что термины "включающий" и "содержащий" являются открытыми и допускают включение дополнительных элементов или этапов. При указании диапазонов их границы включаются. Более того, если иное не указано или не следует из контекста и понимания среднего специалиста в данной области, значения, которые выражены в виде диапазонов, могут соответствовать любому конкретному значению или поддиапазону в рамках указанных диапазонов в различных вариантах осуществления изобретения до десятой доли единицы нижней границы диапазона, если из контекста явно не следует иное.

ГРУППА 1. ВАРИАНТЫ ОСУЩЕСТВЛЕНИЯ 1-212

1. Слитый белок, включающий программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу.

2. Слитый белок по варианту осуществления 1, где слитый белок способен осуществлять геномное редактирование с помощью мишень-инициированной обратной транскрипции в присутствии удлиненной направляющей РНК.

3. Слитый белок по варианту осуществления 1, где napDNAbp имеет никазную активность.

4. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

5. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

6. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9)..

7. Слитый белок по варианту осуществления 1, где napDNAbp выбран из группы, состоящей из: Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2C3, и Argonaute, и необязательно имеет никазную активность.

8. Слитый белок по варианту осуществления 1, где слитый белок при образовании комплекса с удлиненной направляющей РНК способен связываться с последовательностью-мишенью ДНК.

9. Слитый белок по варианту осуществления 8, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

10. Слитый белок по варианту осуществления 8, где связывание слитого белка в виде комплекса с удлиненной направляющей РНК образует R-петлю.

11. Слитый белок по варианту осуществления 10, где R-петля включает (i) РНК-ДНК гибрид, включающий удлиненную направляющую РНК и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

12. Слитый белок по варианту осуществления 11, где комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности обратной транскриптазы, имеющей свободный 3′-конец.

13. Слитый белок по варианту осуществления 2, где удлиненная направляющая РНК включает (a) направляющую РНК и (b) удлинение РНК на 5′ или 3′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

14. Слитый белок по варианту осуществления 13, где удлинение РНК включает (i) последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

15. Слитый белок по варианту осуществления 14, где последовательность матрицы обратной транскрипции кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

16. Слитый белок по варианту осуществления 13, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

17. Слитый белок по варианту осуществления 15, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения.

18. Слитый белок по варианту осуществления 15, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5′-конец.

19. Слитый белок по варианту осуществления 18, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

20. Слитый белок по варианту осуществления 18, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

21. Слитый белок по варианту осуществления 14, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM, или где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

22. Слитый белок по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности SEQ ID NO: 2.

23. Слитый белок по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

24. Слитый белок по варианту осуществления 1, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

25. Слитый белок по варианту осуществления 1, где обратная транскриптаза включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

26. Слитый белок по варианту осуществления 1, где обратная транскриптаза представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

27. Слитый белок по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-[обратная транскриптаза]-COOH; или NH_2-[обратная транскриптаза]-[napDNAbp]-COOH, при этом каждый экземпляр "]-[" обозначает наличие необязательной последовательности линкера.

28. Слитый белок по варианту осуществления 27, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 37-47.

29. Слитый белок по варианту осуществления 14, где требуемое нуклеотидное изменение представляет собой изменение одного нуклеотида, вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

30. Слитый белок по варианту осуществления 29, где вставка или делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22 , по меньшей мере, 23 , по меньшей мере, 24 , по меньшей мере, 25, по меньшей мере, 26, по меньшей мере, 27, по меньшей мере, 28, по меньшей мере, 29, по меньшей мере, 30, по меньшей мере, 31, по меньшей мере, 32, по меньшей мере, 33, по меньшей мере, 34, по меньшей мере, 35, по меньшей мере, 36, по меньшей мере, 37, по меньшей мере, 38, по меньшей мере, 39, по меньшей мере, 40, по меньшей мере, 41, по меньшей мере, 42, по меньшей мере, 43, по меньшей мере, 44, по меньшей мере, 45, по меньшей мере, 46, по меньшей мере, 47, по меньшей мере, 48, по меньшей мере, 49, или, по меньшей мере, 50.

31. Удлиненная направляющая РНК, включающая направляющую РНК и, по меньшей мере, одно удлинение РНК.

32. Удлиненная направляющая РНК по варианту осуществления 1, где удлинение РНК расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

33. Удлиненная направляющая РНК по варианту осуществления 31, где удлиненная направляющая РНК способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

34. Удлиненная направляющая РНК по варианту осуществления 33, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

35. Удлиненная направляющая РНК по варианту осуществления 31, где указанное, по меньшей мере, одно удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

36. Удлиненная направляющая РНК по варианту осуществления 35, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

37. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по

меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

38. Удлиненная направляющая РНК по варианту осуществления 35, где участок связывания праймера обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

39. Удлиненная направляющая РНК по варианту осуществления 35, где необязательная последовательность линкера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

40. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

41. Удлиненная направляющая РНК по варианту осуществления 40, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

42. Удлиненная направляющая РНК по варианту осуществления 41, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

43. Удлиненная направляющая РНК по варианту осуществления 41, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

44. Удлиненная направляющая РНК по варианту осуществления 31, включающая нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

45. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

46. Удлиненная направляющая РНК по варианту осуществления 35, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

47. Удлиненная направляющая РНК по варианту осуществления 35, где необязательная последовательность линкера имеет длину, по меньшей мере, 1 нуклеотид, или, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, или, по меньшей мере, 15 нуклеотидов.

48. Комплекс, включающий слитый белок по любому одному из вариантов осуществления 1-30 и удлиненную направляющую РНК.

49. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК включает направляющую РНК и удлинение РНК на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

50. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

51. Комплекс по варианту осуществления 50, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

52. Комплекс по варианту осуществления 49, где указанное, по меньшей мере, одно удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

53. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК включает нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

54. Комплекс по варианту осуществления 52, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

55. Комплекс по варианту осуществления 52, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

56. Комплекс, включающий napDNAbp и удлиненную направляющую РНК.

57. Комплекс по варианту осуществления 56, где napDNAbp представляет собой никазу Cas9.

58. Комплекс по варианту осуществления 56, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную

последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 2.

59. Комплекс по варианту осуществления 57, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

60. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК включает направляющую РНК и удлинение РНК на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

61. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК способна направлять napDNAbp к последовательности-мишени ДНК.

62. Комплекс по варианту осуществления 61, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом спейсерная последовательность гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

63. Комплекс по варианту осуществления 61, где удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

64. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК включает нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

65. Комплекс по варианту осуществления 63, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

66. Комплекс по варианту осуществления 63, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

67. Комплекс по варианту осуществления 63, где необязательная последовательность линкера имеет длину, по меньшей мере, 1 нуклеотид, или, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, или, по меньшей мере, 15 нуклеотидов.

68. Полинуклеотид, кодирующий слитый белок по любому одному из вариантов осуществления 1-30.

69. Вектор, включающий полинуклеотид по варианту осуществления 68.

70. Клетка, включающая слитый белок по любому из вариантов осуществления 1-30 и удлиненную направляющую РНК, связанную с napDNAbp слитого белка.

71. Клетка, включающая комплекс по любому одному из вариантов осуществления 48-67.

72. Фармацевтическая композиция, включающая: (i) слитый белок по любому из вариантов осуществления 1-30, комплекс по вариантам осуществления 48-67, полинуклеотид по варианту осуществления 68, или вектор по варианту осуществления 69; и (ii) фармацевтически приемлемый наполнитель.

73. Фармацевтическая композиция, включающая: (i) комплекс по вариантам осуществления 48-67, (ii) обратную транскриптазу, поступающую in trans; и (iii) фармацевтически приемлемый наполнитель.

74. Набор, включающий конструкт нуклеиновой кислоты, включающий: (i) последовательность нуклеиновой кислоты, кодирующую слитый белок по любому одному из вариантов осуществления 1-30; и (ii) промотор, который обеспечивает экспрессию последовательности (i).

75. Способ внесения требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК, при этом способ включает:

(i) контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и удлиненную направляющую РНК, при этом слитый белок включает napDNAbp и обратную транскриптазу, и при этом удлиненная направляющая РНК включает последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение;

(ii) никирование последовательности двуцепочечной ДНК в не являющейся мишенью цепи, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3´-конец;

(iii) гибридизацию 3` конца свободной одноцепочечной ДНК с последовательностью матрицы обратной транскрипции, тем самым обеспечивая праймирование домена обратной транскриптазы;

(iv) полимеризацию цепи ДНК с 3´-конца, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение;

(v) замещение цепи эндогенной ДНК, примыкающей к участку разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

76. Способ по варианту осуществления 75, где этап замещения (v) включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

77. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой замещение одного нуклеотида, делецию, или вставку.

78. Способ по варианту осуществления 77, где замещение одного нуклеотида представляет собой транзицию или трансверсию.

79. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на

G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

80. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

81. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

82. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена.

83. Способ по варианту осуществления 82, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса.

84. Способ по варианту осуществления 82, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

85. Способ по варианту осуществления 76, где napDNAbp представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9.

86. Способ по варианту осуществления 76, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2.

87. Способ по варианту осуществления 76, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

88. Способ по варианту осуществления 76, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

89. Способ по варианту осуществления 76, где домен обратной транскриптазы включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

90. Способ по варианту осуществления 76, где удлиненная направляющая РНК включает удлинение РНК на 3'- или 5'-концах или во внутримолекулярном положении в направляющей РНК, при этом удлинение РНК включает последовательность матрицы обратной транскрипции.

91. Способ по варианту осуществления 90, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

92. Способ по варианту осуществления 76, где удлиненная направляющая РНК имеет нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 18-36.

93. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий:

(i) контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и направляющей РНК, которая нацеливает napDNAbp на локус-мишень, при этом направляющая РНК включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение;

(ii) образование свободного 3′-конца в цепи ДНК в локусе-мишени;

(iii) гибридизацию свободного 3′-конца с последовательность RT матрицы для праймирования обратной транскрипции;

(iv) синтезирование одноцепочечного ДНК-флэпа, включающего указанное, по меньшей мере, одно требуемое нуклеотидное изменение, на основе последовательности RT матрицы с помощью обратной транскриптазы;

(v) и встраивание указанного, по меньшей мере, одного требуемого нуклеотидного изменения в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

94. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают транзицию.

95. Способ по варианту осуществления 94, где транзиция выбрана из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

96. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают трансверсию.

97. Способ по варианту осуществления 96, где трансверсия выбрана из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

98. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

99. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

100. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают корректирование связанного с заболеванием гена.

101. Способ по варианту осуществления 100, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса.

102. Способ по варианту осуществления 100, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

103. Способ по варианту осуществления 93, где napDNAbp представляет собой имеющий нуклеазную активность Cas9 или его вариант.

104. Способ по варианту осуществления 93, где napDNAbp представляет собой не имеющий нуклеазную активность Cas9 (dCas9), или никазу Cas9 (nCas9), или их вариант.

105. Способ по варианту осуществления 93, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2.

106. Способ по варианту осуществления 93, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

107. Способ по варианту осуществления 93, где обратная транскриптаза вносится in trans.

108. Способ по варианту осуществления 93, где napDNAbp включает продукт слияния с обратной транскриптазой.

109. Способ по варианту осуществления 93, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

110. Способ по варианту осуществления 93, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

111. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает никирование цепи ДНК нуклеазой.

112. Способ по варианту осуществления 111, где нуклеаза представляет собой napDNAbp и поступает в виде слитого домена napDNAbp или поступает in trans.

113. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает контактирование цепи ДНК с химическим агентом.

114. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает внесение ошибки репликации.

115. Способ по варианту осуществления 93, где этап контактирования молекулы ДНК с napDNAbp и направляющей РНК образует R-петлю.

116. Способ по варианту осуществления 115, где цепь ДНК, в которой образован свободный 3'-конец, образована в R-петле.

117. Способ по варианту осуществления 93, где направляющая РНК включает удлиненную часть, которая включает последовательность матрицы в виде обратной транскриптазы (RT).

118. Способ по варианту осуществления 117, где удлиненная часть расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

119. Способ по варианту осуществления 93, где направляющая РНК дополнительно включает участок связывания праймера.

120. Способ по варианту осуществления 93, где направляющая РНК дополнительно включает спейсерную последовательность.

121. Способ по варианту осуществления 93, где последовательность RT матрицы является гомологичной относительно соответствующей эндогенной ДНК.

122. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-

инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение; и (c) встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

123. Способ по варианту осуществления 122, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

124. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

125. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

126. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

127. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

128. Полинуклеотид, кодирующий удлиненную направляющую РНК по любому одному из вариантов осуществления 31-47.

129. Вектор, включающий полинуклеотид по варианту осуществления 128.

130. Клетка, включающая вектор по варианту осуществления 129.

131. Слитый белок по любому из вариантов осуществления 1-30, где обратная транскриптаза представляет собой подверженную ошибкам обратную транскриптазу.

132. Способ мутагенеза молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp), и подверженной ошибкам обратной транскриптазой, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения мутагенизированной одноцепочечной ДНК; и (c) встраивание мутагенизированной одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

133. Способ по варианту осуществления 132, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

134. Способ по варианту осуществления 132, где napDNAbp представляет собой никазу Cas9 (nCas9).

135. Способ по варианту осуществления 132, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

136. Способ по варианту осуществления 132, где направляющая РНК включает SEQ ID NO: 222.

137. Способ по варианту осуществления 132, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

138. Способ замещения мутации в виде экспансии тринуклеотидных повторов в молекуле-мишени ДНК здоровой последовательностью, включающей здоровое количество тринуклеотидных повторов, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим

программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую замещающую последовательность; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей замещающую последовательность; и (c) встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

139. Способ по варианту осуществления 138, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

140. Способ по варианту осуществления 138, где napDNAbp представляет собой никазу Cas9 (nCas9).

141. Способ по варианту осуществления 138, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

142. Способ по варианту осуществления 138, где направляющая РНК включает SEQ ID NO: 222.

143. Способ по варианту осуществления 138, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

144. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов ассоциирована с болезнью Хантингтона, синдромом ломкой Х-хромосомы, или атаксией Фридрейха.

145. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов CAG.

146. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов GAA.

147. Способ внесения функционального фрагмента в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и функциональный фрагмент.

148. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой пептидную метку.

149. Способ по варианту осуществления 148, где пептидная метка представляет собой аффинную метку, метку солюбилизации, хроматографическую метку, эпитопную метку, или флуоресцентную метку.

150. Способ по варианту осуществления 148, где пептидная метка выбрана из группы, состоящей из: AviTag (SEQ ID NO: 245); C-метки (SEQ ID NO: 246); калмодулиновой метки (SEQ ID NO: 247); полиглутаматной метки (SEQ ID NO: 248); E-метки (SEQ ID NO: 249); FLAG-метки (SEQ ID NO: 250); HA-метки (SEQ ID NO: 251); His-метки (SEQ ID NO: 252-262); Myc-метки (SEQ ID NO: 263); NE-метки (SEQ ID NO: 264); Rho1D4-метки (SEQ ID NO: 265); S-метки (SEQ ID NO: 266); SBP-метки (SEQ ID NO: 267); Softag-1 (SEQ ID NO: 268); Softag-2 (SEQ ID NO: 269); Spot-метки (SEQ ID NO: 270); Strep-метки (SEQ ID NO: 271); метки TC (SEQ ID NO: 272); метки Ty (SEQ ID NO: 273); метки V5 (SEQ ID NO: 274); VSV-метки (SEQ ID NO: 275); и метки Xpress (SEQ ID NO: 276).

151. Способ по варианту осуществления 148, где пептидная метка выбрана из группы, состоящей из: эпитопа AU1 (SEQ ID NO: 278); эпитопа AU5 (SEQ ID NO: 279); эпитопа бактериофага T7 (T7-метка) (SEQ ID NO: 280); метки на основе вируса

Bluetongue (B-метки) (SEQ ID NO: 281); эпитопа E2 (SEQ ID NO: 282); аффинной гистидиновой метки (HAT) (SEQ ID NO: 283); эпитопа HSV (SEQ ID NO: 284); полиаргинина (Arg-метки) (SEQ ID NO: 285); полиаспартата (Asp метки) (SEQ ID NO: 286); полифенилаланина (Phe-метки) (SEQ ID NO: 287); S1-метки (SEQ ID NO: 288); S-метки (SEQ ID NO: 289); и VSV-G (SEQ ID NO: 290).

152. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой иммуноэпитоп.

153. Способ по варианту осуществления 152, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 398); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 2 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

154. Способ по варианту осуществления 147, где функциональный фрагмент изменяет локализацию представляющего интерес белка.

155. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой метку расщепления, так что обеспечивается изменение скорости расщепления представляющего интерес белка.

156. Способ по варианту осуществления 155, где метка расщепления.

157. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой домен связывания малой молекулы.

158. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

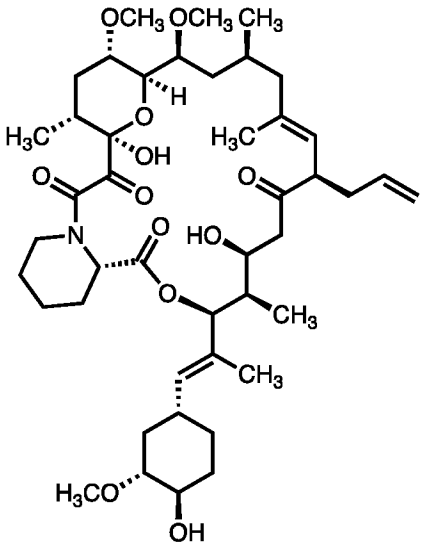
159. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

160. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 494.

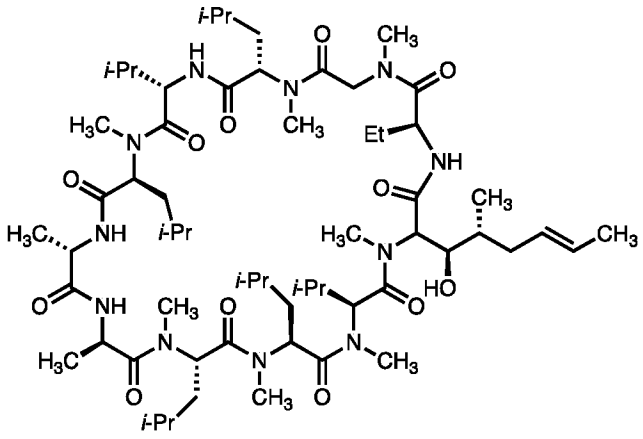
161. Способ по варианту осуществления 157, где домен связывания малой молекулы вносится в два или более представляющих интерес белка.

162. Способ по варианту осуществления 161, где указанные два или более представляющие интерес белка могут подвергаться димеризации при контактировании с малой молекулой.

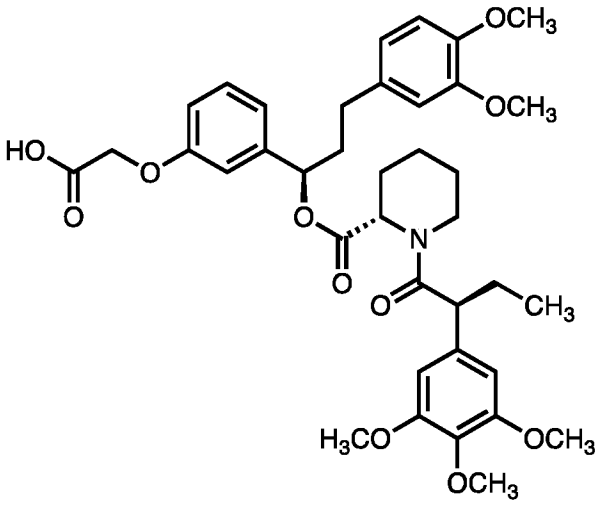
163. Способ по варианту осуществления 157, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из:



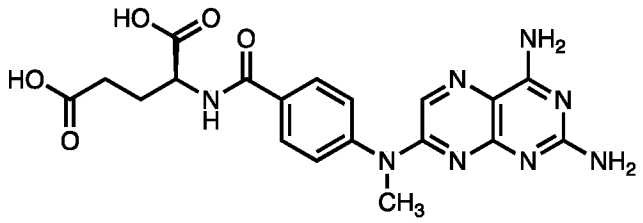
;



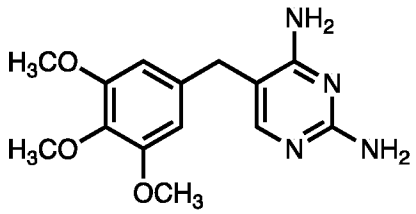
;



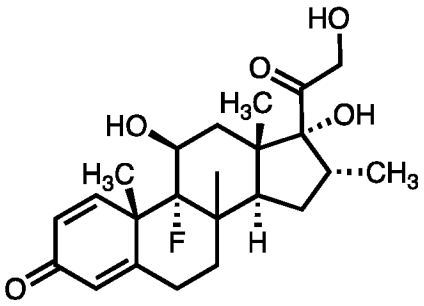
;



;



; и



.

164. Способ внесения иммуноэпитопа в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и иммуноэпитоп.

165. Способ по варианту осуществления 164, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 398); иммуноэпитопа паротита 1 (SEQ ID NO: 400);

иммуноэпитопа паротита 2 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

166. Способ по варианту осуществления 164, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

167. Способ по варианту осуществления 164, где napDNAbp представляет собой никазу Cas9 (nCas9).

168. Способ по варианту осуществления 164, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

169. Способ по варианту осуществления 164, где направляющая РНК включает SEQ ID NO: 222.

170. Способ внесения домена димеризации на основе малой молекулы в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую домен димеризации на основе малой молекулы; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и домен димеризации на основе малой молекулы.

171. Способ по варианту осуществления 170, дополнительно включающий осуществление способа в отношении второго представляющего интерес белка.

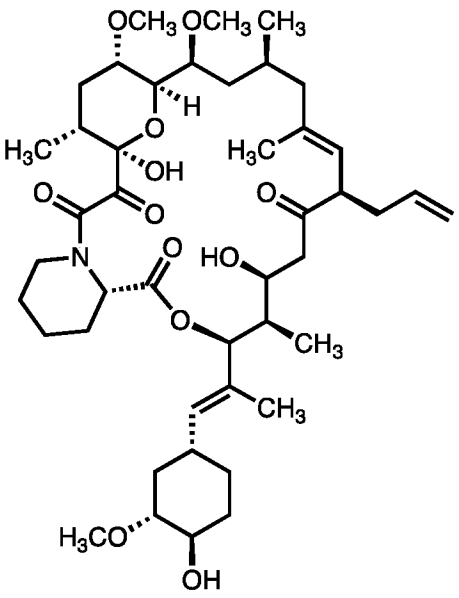
172. Способ по варианту осуществления 171, где первый представляющий интерес белок и второй представляющий интерес белок подвергаются димеризации в присутствии малой молекулы, которая связывается с доменом димеризации на каждом из указанных белков.

173. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

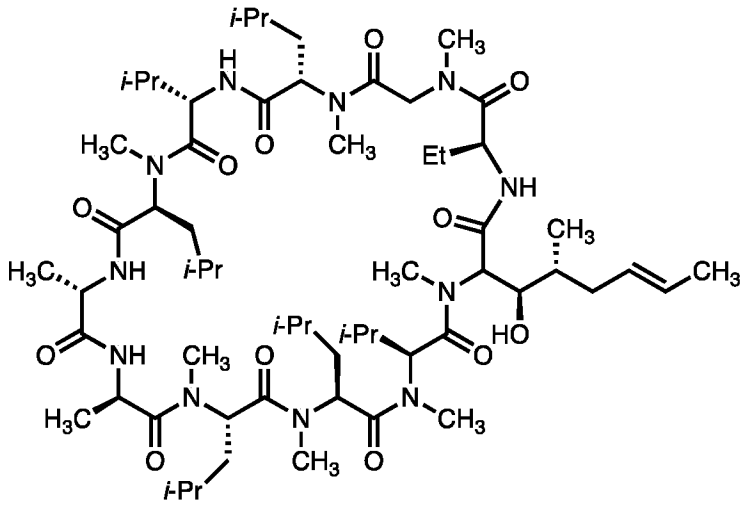
174. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

175. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 492-494.

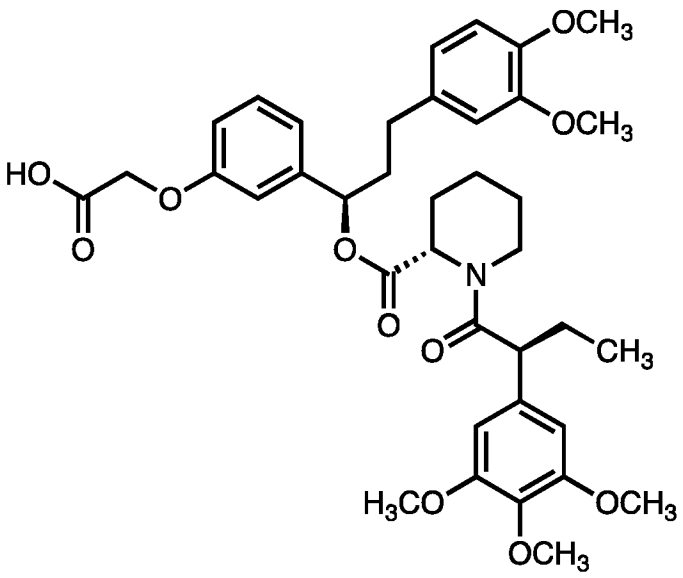
176. Способ по варианту осуществления 170, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из:



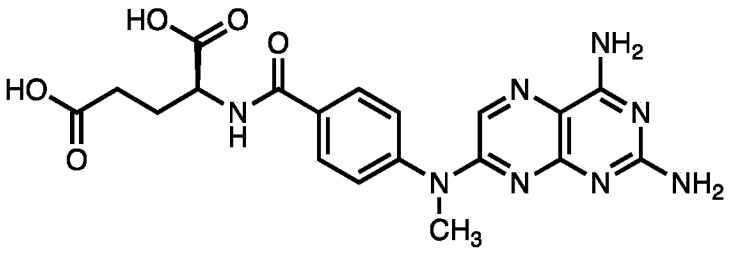
;



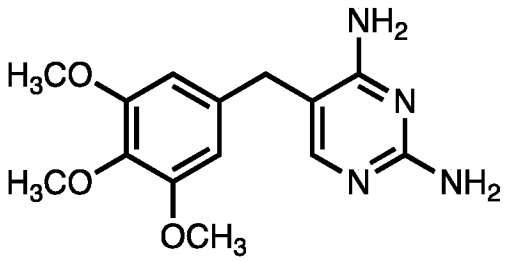
;



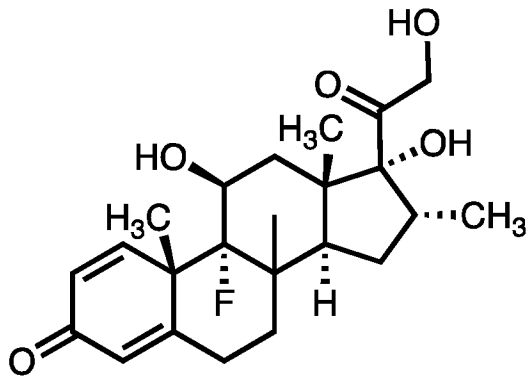
;



;



; и



.

177. Способ по варианту осуществления 170, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

178. Способ по варианту осуществления 170, где napDNAbp представляет собой никазу Cas9 (nCas9).

179. Способ по варианту осуществления 170, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

180. Способ по варианту осуществления 170, где направляющая РНК включает SEQ ID NO: 222.

181. Способ внесения пептидной метки или эпитопа в белок с помощью праймерного редактирования, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с конструктом праймерного редактора, выполненного с возможностью вносить в нее вторую нуклеотидную последовательность, кодирующую пептидную метку, для получения рекомбинантной нуклеотидной последовательности, так что пептидная метка и белок экспрессируются из рекомбинантной нуклеотидной последовательности в виде слитого белка.

182. Способ по варианту осуществления 181, где пептидная метка применяется для очистки и/или детектирования белка.

183. Способ по варианту осуществления 181, где пептидная метка представляет собой полигистидин (например HHHHHH), FLAG (например DYKDDDDK), V5 (например GKPIPNPLLGLDST), GCN4, HA (например YPYDVPDYA), Myc (например EQKLISEED), GST….и так далее.

184. Способ по варианту осуществления 181, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 245-290.

185. Способ по варианту осуществления 181, где пептидная метка слита с белком с помощью линкера.

186. Способ по варианту осуществления 181, где слитый белок имеет следующую структуру: [белок]-[пептидная метка] или [пептидная метка]-[белок], где “]-[“ обозначает необязательный линкер.

187. Способ по варианту осуществления 181, где линкер имеет аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

188. Способ по варианту осуществления 181, где конструкт праймерного редактора включает PEgRNA, включающую нуклеотидную последовательность согласно SEQ ID NO: 18-25.

189. Способ по варианту осуществления 181, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

190. Способ по варианту осуществления 181, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

191. Способ предотвращения или остановки прогрессирования прионного заболевания путем внесения одной или нескольких защитных мутаций в PRNP, кодируемых нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей защитную мутацию; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует PRNP, включающий защитную мутацию и являющийся устойчивым к неправильному фолдингу.

192. Способ по варианту осуществления 191, где прионное заболевание представляет собой прионное заболевание человека.

193. Способ по варианту осуществления 191, где прионное заболевание представляет собой прионное заболевание животного.

194. Способ по варианту осуществления 192, где прионное заболевание представляет собой болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, или куру.

195. Способ по варианту осуществления 192, где прионное заболевание представляет собой коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных.

196. Способ по варианту осуществления 191, где аминокислотная последовательность PRNP дикого типа представляет собой SEQ ID NO: 291-292.

197. Способ по варианту осуществления 191, где способ приводит к получению модифицированной аминокислотной последовательности PRNP, выбранной из группы, состоящей из SEQ ID NO: 293-323, при этом указанный модифицированный белок PRNP устойчив к неправильному фолдингу.

198. Способ по варианту осуществления 191, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

199. Способ по варианту осуществления 191, где napDNAbp представляет собой никазу Cas9 (nCas9).

200. Способ по варианту осуществления 191, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

201. Способ по варианту осуществления 191, где направляющая РНК включает SEQ ID NO: 222.

202. Способ внесения рибонуклеотидного мотива или метки в представляющую интерес РНК, кодируемую нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором,

включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую рибонуклеотидный мотив или метку; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей рибонуклеотидный мотив или метку; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует представляющую интерес модифицированную РНК, включающую рибонуклеотидный мотив или метку.

203. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка представляет собой фрагмент для детектирования.

204. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка влияет на уровень экспрессии представляющей интерес РНК.

205. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка влияет на транспорт или субклеточную локализацию представляющей интерес РНК.

206. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка выбрана из группы, состоящей из SV40 типа 1, SV40 типа 2, SV40 типа 3, hGH, BGH, rbGlob, TK, MALAT1 ENE-mascРНК, KSHV PAN ENE, Sm-бокса/бокса мяРНК U1, 3’ бокса мяРНК U1, тРНК-лизина, аптамера broccoli, аптамера spinach, аптамера mango, рибозима HDV, и m6A.

207. Способ по варианту осуществления 202, где PEgRNA включает SEQ ID NO: 18-25.

208. Способ по варианту осуществления 202, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

209. Способ по варианту осуществления 202, где napDNAbp представляет собой никазу Cas9 (nCas9).

210. Способ по варианту осуществления 202, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

211. Способ внесения или удаления функционального фрагмента в представляющем интерес белке, кодируемом нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент или его удаление; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент или его удаление; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует модифицированный белок, включающий представляющий интерес белок и функциональный фрагмент или его удаление, при этом функциональный фрагмент изменяет связанное с модификацией состояние или связанное с локализацией состояние белка.

212. Способ по варианту осуществления 211, где функциональный фрагмент изменяет связанное с фосфорилированием, убиквитинилированием, гликозилированием, липидированием, гидроксилированием, метилированием, ацетилированием, кротонилированием, SUMOилированием состояние представляющего интерес белка.

ГРУППА 2. ВАРИАНТЫ ОСУЩЕСТВЛЕНИЯ 213-424

213. Слитый белок, включающий программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу.

214. Слитый белок по варианту осуществления 213, где слитый белок способен осуществлять геномное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA).

215. Слитый белок по варианту осуществления 213, где napDNAbp имеет никазную активность.

216. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой белок Cas9 или его вариант.

217. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

218. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой никазу Cas9 (nCas9)..

219. Слитый белок по варианту осуществления 213, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

220. Слитый белок по варианту осуществления 213, где слитый белок при образовании комплекса с PEgRNA способен связываться с последовательностью-мишенью ДНК.

221. Слитый белок по варианту осуществления 220, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

222. Слитый белок по варианту осуществления 220, где связывание слитого белка в виде комплекса с PEgRNA образует R-петлю.

223. Слитый белок по варианту осуществления 222, где R-петля включает (i) РНК-ДНК гибрид, включающий PEgRNA и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

224. Слитый белок по варианту осуществления 223, где комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности, имеющей свободный 3′-конец.

225. Слитый белок по варианту осуществления 214, где PEgRNA включает (a) направляющую РНК и (b) плечо удлинения на 5′ или 3′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

226. Слитый белок по варианту осуществления 225, где плечо удлинения включает (i) последовательность матрицы для синтеза ДНК, включающую требуемое нуклеотидное изменение, и (ii) участок связывания праймера.

227. Слитый белок по варианту осуществления 226, где последовательность матрицы для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

228. Слитый белок по варианту осуществления 225, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

229. Слитый белок по варианту осуществления 227, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения.

230. Слитый белок по варианту осуществления 227, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5'-конец.

231. Слитый белок по варианту осуществления 230, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

232. Слитый белок по варианту осуществления 230, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

233. Слитый белок по варианту осуществления 226, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM, или где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

234. Слитый белок по варианту осуществления 213, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности SEQ ID NO: 2.

235. Слитый белок по варианту осуществления 213, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

236. Слитый белок по варианту осуществления 213, где полимераза представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

237. Слитый белок по варианту осуществления 213, где полимераза представляет собой обратную транскриптазу, включающую аминокислотную последовательность,

имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

238. Слитый белок по варианту осуществления 213, где полимераза представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

239. Слитый белок по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-[полимераза]-COOH; или NH_2-[полимераза]-[napDNAbp]- COOH, где каждый экземпляр "]-[" указывает на наличие последовательности необязательного линкера.

240. Слитый белок по варианту осуществления 239, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 37-47.

241. Слитый белок по варианту осуществления 226, где требуемое нуклеотидное изменение представляет собой изменение одного нуклеотида, вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

242. Слитый белок по варианту осуществления 241, где вставка или делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22 , по меньшей мере, 23 , по меньшей мере, 24 , по меньшей мере, 25, по меньшей мере, 26, по меньшей мере, 27, по меньшей мере, 28, по меньшей мере, 29, по меньшей мере, 30, по меньшей мере, 31, по меньшей мере, 32, по меньшей мере, 33, по меньшей мере, 34, по меньшей мере, 35, по меньшей мере, 36, по меньшей мере, 37, по меньшей мере, 38, по меньшей мере, 39, по меньшей мере, 40, по меньшей мере, 41, по меньшей мере, 42, по меньшей мере, 43, по меньшей мере, 44, по меньшей мере, 45, по меньшей мере, 46, по меньшей мере, 47, по меньшей мере, 48, по меньшей мере, 49, или, по меньшей мере, 50.

243. PEgRNA, включающая направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты, включающее матрицу для синтеза ДНК.

244. PEgRNA по варианту осуществления 241, где плечо удлинения нуклеиновой кислоты расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном расположении в направляющей РНК и где плечо удлинения нуклеиновой кислоты представляет собой ДНК или РНК.

245. PEgRNA по варианту осуществления 242, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

246. PEgRNA по варианту осуществления 245, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

247. PEgRNA по варианту осуществления 243, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

248. PEgRNA по варианту осуществления 247, где плечо удлинения нуклеиновой кислоты имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

249. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7

нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

250. PEgRNA по варианту осуществления 247, где участок связывания праймера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

251. PEgRNA по варианту осуществления 243, дополнительно включающая, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

252. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

253. PEgRNA по варианту осуществления 252, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

254. PEgRNA по варианту осуществления 253, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

255. PEgRNA по варианту осуществления 253, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

256. PEgRNA по варианту осуществления 243, включающая нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

257. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

258. PEgRNA по варианту осуществления 247, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

259. PEgRNA по варианту осуществления 251, где указанная, по меньшей мере, одна дополнительная структура расположена на 3'- или 5'-конце PEgRNA.

260. Комплекс, включающий слитый белок по любому одному из вариантов осуществления 213-242 и PEgRNA.

261. Комплекс по варианту осуществления 260, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

262. Комплекс по варианту осуществления 260, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

263. Комплекс по варианту осуществления 262, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

264. Комплекс по варианту осуществления 261, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

265. Комплекс по варианту осуществления 260, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную

последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

266. Комплекс по варианту осуществления 264, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

267. Комплекс по варианту осуществления 264, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

268. Комплекс, включающий napDNAbp и PEgRNA.

269. Комплекс по варианту осуществления 268, где napDNAbp представляет собой никазу Cas9.

270. Комплекс по варианту осуществления 268, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2.

271. Комплекс по варианту осуществления 268, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

272. Комплекс по варианту осуществления 268, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

273. Комплекс по варианту осуществления 268, где PEgRNA способна направлять napDNAbp к последовательности-мишени ДНК.

274. Комплекс по варианту осуществления 272, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом спейсерная последовательность PEgRNA гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

275. Комплекс по варианту осуществления 273, где плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

276. Комплекс по варианту осуществления 269, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 18-36.

277. Комплекс по варианту осуществления 276, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

278. Комплекс по варианту осуществления 276, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

279. Комплекс по варианту осуществления 276, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

280. Полинуклеотид, кодирующий слитый белок по любому одному из вариантов осуществления 213-242.

281. Вектор, включающий полинуклеотид по варианту осуществления 280.

282. Клетка, включающая слитый белок по любому из вариантов осуществления 213-242 и PEgRNA, связанную с napDNAbp слитого белка.

283. Клетка, включающая комплекс по любому одному из вариантов осуществления 260-279.

284. Фармацевтическая композиция, включающая: (i) слитый белок по любому из вариантов осуществления 213-242, комплекс по вариантам осуществления 260-279, полинуклеотид по варианту осуществления 68, или вектор по варианту осуществления 69; и (ii) фармацевтически приемлемый наполнитель.

285. Фармацевтическая композиция, включающая: (i) комплекс по вариантам осуществления 260-279, (ii) полимеразу, поступающую in trans; и (iii) фармацевтически приемлемый наполнитель.

286. Набор, включающий конструкт нуклеиновой кислоты, включающий: (i) последовательность нуклеиновой кислоты, кодирующую слитый белок по любому одному из вариантов осуществления 213-242; и (ii) промотор, который обеспечивает экспрессию последовательности (i).

287. Способ внесения требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК, при этом способ включает:

(i) контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и PEgRNA, при этом слитый белок включает napDNAbp и полимеразу, и при этом PEgRNA включает матрицу для синтеза ДНК, включающую требуемое нуклеотидное изменение и участок связывания праймера;

(ii) никирование последовательности двуцепочечной ДНК, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3´-конец;

(iii) гибридизирование 3'-конца свободной одноцепочечной ДНК с участком связывания праймера, тем самым обеспечивая праймирование полимеразы;

(iv) полимеризацию цепи ДНК с 3´-конца, гибридизированного с участком связывания праймера, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение и являющегося комплементарным матрице для синтеза ДНК;

(v) замещение цепи эндогенной ДНК, примыкающей к участку разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

288. Способ по варианту осуществления 287, где этап замещения (v) включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

289. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой замещение одного нуклеотида, делецию, или вставку.

290. Способ по варианту осуществления 289, где замещение одного нуклеотида представляет собой транзицию или трансверсию.

291. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

292. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

293. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

294. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена.

295. Способ по варианту осуществления 294, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого

комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

296. Способ по варианту осуществления 294, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

297. Способ по варианту осуществления 287, где napDNAbp представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9.

298. Слитый белок по варианту осуществления 287, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности SEQ ID NO: 2.

299. Способ по варианту осуществления 287, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

300. Способ по варианту осуществления 287, где полимераза представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

301. Способ по варианту осуществления 287, где полимераза представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

302. Способ по варианту осуществления 287, где PEgRNA включает плечо удлинения нуклеиновой кислоты на 3'- или 5'-концах или во внутримолекулярном

положении в направляющей РНК, при этом плечо удлинения включает последовательность матрицы для синтеза ДНК и участок связывания праймера.

303. Способ по варианту осуществления 302, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

304. Способ по варианту осуществления 287, где PEgRNA имеет нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 18-36.

305. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий:

(i) контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и PEgRNA, которая нацеливает napDNAbp на локус-мишень, при этом PEgRNA включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение и участок связывания праймера;

(ii) образование свободного 3′-конца в цепи ДНК в локусе-мишени;

(iii) гибридизацию свободного 3′-конца с участком связывания праймера для праймирования обратной транскрипции;

(iv) синтезирование одноцепочечного ДНК-флэпа, включающего указанное, по меньшей мере, одно требуемое нуклеотидное изменение, на основе последовательности RT матрицы с помощью обратной транскриптазы;

(v) и встраивание указанного, по меньшей мере, одного требуемого

нуклеотидного изменения в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

306. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают транзицию.

307. Способ по варианту осуществления 306, где транзиция выбрана из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

308. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают трансверсию.

309. Способ по варианту осуществления 308, где трансверсия выбрана из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

310. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (9) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

311. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

312. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают корректирование связанного с заболеванием гена.

313. Способ по варианту осуществления 312, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из:

дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

314. Способ по варианту осуществления 312, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

315. Способ по варианту осуществления 305, где napDNAbp представляет собой имеющий нуклеазную активность Cas9 или его вариант.

316. Способ по варианту осуществления 305, где napDNAbp представляет собой не имеющий нуклеазную активность Cas9 (dCas9), или никазу Cas9 (nCas9), или их вариант.

317. Способ по варианту осуществления 305, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 2.

318. Способ по варианту осуществления 305, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 2-10.

319. Способ по варианту осуществления 305, где обратная транскриптаза вносится in trans.

320. Способ по варианту осуществления 305, где napDNAbp включает продукт слияния с обратной транскриптазой.

321. Способ по варианту осуществления 305, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 11-17.

322. Способ по варианту осуществления 305, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 11-17.

323. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает никирование цепи ДНК нуклеазой.

324. Способ по варианту осуществления 323, где нуклеаза поступает *in trans*.

325. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает контактирование цепи ДНК с химическим агентом.

326. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает внесение ошибки репликации.

327. Способ по варианту осуществления 305, где этап контактирования молекулы ДНК с napDNAbp и направляющей РНК образует R-петлю.

328. Способ по варианту осуществления 327, где цепь ДНК, в которой образован свободный 3'-конец, образована в R-петле.

329. Способ по варианту осуществления 315, где PEgRNA включает плечо удлинения, которое включает последовательность матрицы в виде обратной транскриптазы (RT) и участок связывания праймера.

330. Способ по варианту осуществления 329, где плечо удлинения расположено на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

331. Способ по варианту осуществления 305, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей

из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования РНК-белка.

332. Способ по варианту осуществления 305, где PEgRNA дополнительно включает плечо гомологии.

333. Способ по варианту осуществления 305, где последовательность RT матрицы является гомологичной относительно соответствующей эндогенной ДНК.

334. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции матрицы RT для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение; и (c) встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

335. Способ по варианту осуществления 334, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

336. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

337. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

338. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

339. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

340. Полинуклеотид, кодирующий PEgRNA по любому одному из вариантов осуществления 243-259.

341. Вектор, включающий полинуклеотид по варианту осуществления 340.

342. Клетка, включающая вектор по варианту осуществления 341.

343. Слитый белок по варианту осуществления 213, где полимераза представляет собой подверженную ошибкам обратную транскриптазу.

344. Способ мутагенеза молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и подверженную ошибкам обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения мутагенизированной одноцепочечной ДНК; и (c) встраивание мутагенизированной одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

345. Способ по любому предыдущему варианту осуществления, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

346. Способ по любому предыдущему варианту осуществления, где napDNAbp представляет собой никазу Cas9 (nCas9).

347. Способ по варианту осуществления 344, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

348. Способ по варианту осуществления 344, где направляющая РНК включает SEQ ID NO: 26-36.

349. Способ по варианту осуществления 344, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

350. Способ замещения мутации в виде экспансии тринуклеотидных повторов в молекуле-мишени ДНК здоровой последовательностью, включающей здоровое количество тринуклеотидных повторов, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК, включающую замещающую последовательность и участок связывания праймера; (b) осуществление праймерного редактирования для получения одноцепочечной ДНК, включающей замещающую последовательность; и (c) встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

351. Способ по варианту осуществления 350, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

352. Способ по варианту осуществления 350, где napDNAbp представляет собой никазу Cas9 (nCas9).

353. Способ по варианту осуществления 350, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

354. Способ по варианту осуществления 350, где направляющая РНК включает SEQ ID NO: 26-36.

355. Способ по варианту осуществления 350, где этап (b) осуществления праймерного редактирования включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать полимеразу за счет отжига с участком связывания праймера на направляющей РНК.

356. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов ассоциирована с болезнью Хантингтона, синдромом ломкой Х-хромосомы, или атаксией Фридрейха.

357. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов CAG.

358. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов GAA.

359. Способ внесения функционального фрагмента в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и функциональный фрагмент.

360. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой пептидную метку.

361. Способ по варианту осуществления 360, где пептидная метка представляет собой аффинную метку, метку солюбилизации, хроматографическую метку, эпитопную или иммуноэпитопную метку, или флуоресцентную метку.

362. Способ по варианту осуществления 360, где пептидная метка выбрана из группы, состоящей из: AviTag (SEQ ID NO: 245); C-метки (SEQ ID NO: 246); калмодулиновой метки (SEQ ID NO: 247); полиглутаматной метки (SEQ ID NO: 248); E-метки (SEQ ID NO: 249); FLAG-метки (SEQ ID NO: 250); HA-метки (SEQ ID NO: 251); His-метки (SEQ ID NO: 252-262); Myc-метки (SEQ ID NO: 263); NE-метки (SEQ ID NO: 264); Rho1D4-метки (SEQ ID NO: 265); S-метки (SEQ ID NO: 266); SBP-метки (SEQ ID NO: 267); Softag-1 (SEQ ID NO: 268); Softag-2 (SEQ ID NO: 269); Spot-метки (SEQ ID NO: 270); Strep-метки (SEQ ID NO: 271); метки TC (SEQ ID NO: 272); метки Ty (SEQ ID NO: 273); метки V5 (SEQ ID NO: 274); VSV-метки (SEQ ID NO: 275); и метки Xpress (SEQ ID NO: 276).

363. Способ по варианту осуществления 360, где пептидная метка выбрана из группы, состоящей из: эпитопа AU1 (SEQ ID NO: 278); эпитопа AU5 (SEQ ID NO: 279); эпитопа бактериофага T7 (T7-метка) (SEQ ID NO: 280); метки на основе вируса Bluetongue (B-метки) (SEQ ID NO: 281); эпитопа E2 (SEQ ID NO: 282); аффинной гистидиновой метки (HAT) (SEQ ID NO: 283); эпитопа HSV (SEQ ID NO: 284); полиаргинина (Arg-метки) (SEQ ID NO: 285); полиаспартата (Asp метки) (SEQ ID NO: 286); полифенилаланина (Phe-метки) (SEQ ID NO: 287); S1-метки (SEQ ID NO: 288); S-метки (SEQ ID NO: 289); и VSV-G (SEQ ID NO: 290).

364. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой иммуноэпитоп.

365. Способ по варианту осуществления 364, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 398); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 2 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

366. Способ по варианту осуществления 359, где функциональный фрагмент изменяет локализацию представляющего интерес белка.

367. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой метку расщепления, так что обеспечивается изменение скорости расщепления представляющего интерес белка.

368. Способ по варианту осуществления 367, где метка расщепления приводит к расщеплению помеченного белка.

369. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой домен связывания малой молекулы.

370. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

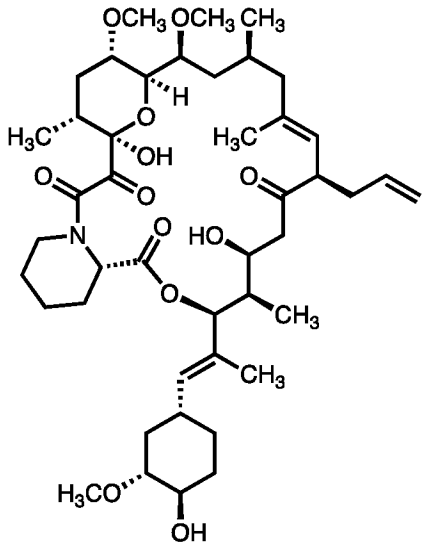
371. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

372. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 492-494.

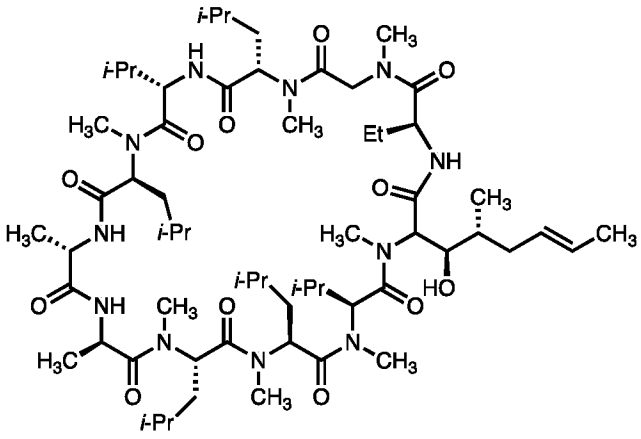
373. Способ по варианту осуществления 359, где домен связывания малой молекулы вносится в два или более представляющих интерес белка.

374. Способ по варианту осуществления 373, где указанные два или более представляющие интерес белка могут подвергаться димеризации при контактировании с малой молекулой.

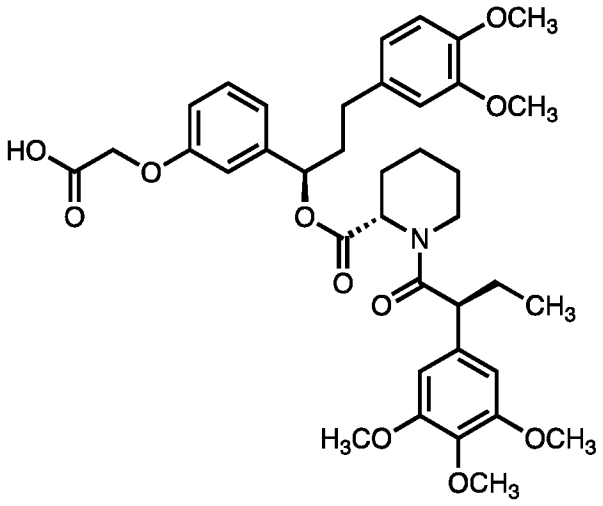
375. Способ по варианту осуществления 369, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из:



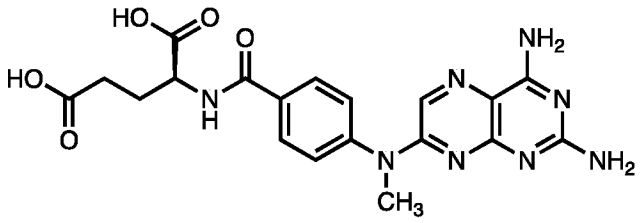
;



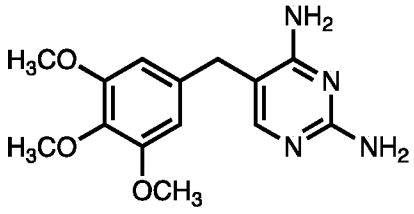
;



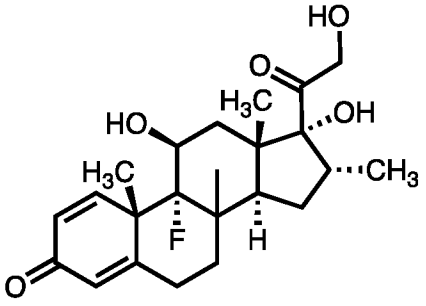
;



;



; и



.

376. Способ внесения иммуноэпитопа в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности

ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и иммуноэпитоп.

377. Способ по варианту осуществления 376, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 398); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 2 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

378. Способ по варианту осуществления 376, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

379. Способ по варианту осуществления 376, где napDNAbp представляет собой никазу Cas9 (nCas9).

380. Способ по варианту осуществления 376, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

381. Способ по варианту осуществления 376, где PEgRNA включает SEQ ID NO: 26-36.

382. Способ внесения домена димеризации на основе малой молекулы в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным

редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую домен димеризации на основе малой молекулы; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и домен димеризации на основе малой молекулы.

383. Способ по варианту осуществления 382, дополнительно включающий осуществление способа в отношении второго представляющего интерес белка.

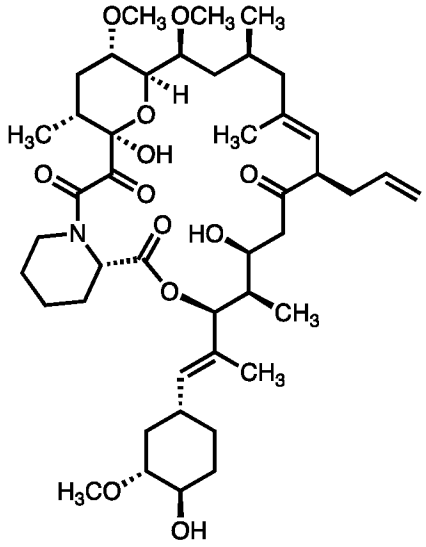
384. Способ по варианту осуществления 383, где первый представляющий интерес белок и второй представляющий интерес белок подвергаются димеризации в присутствии малой молекулы, которая связывается с доменом димеризации на каждом из указанных белков.

385. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

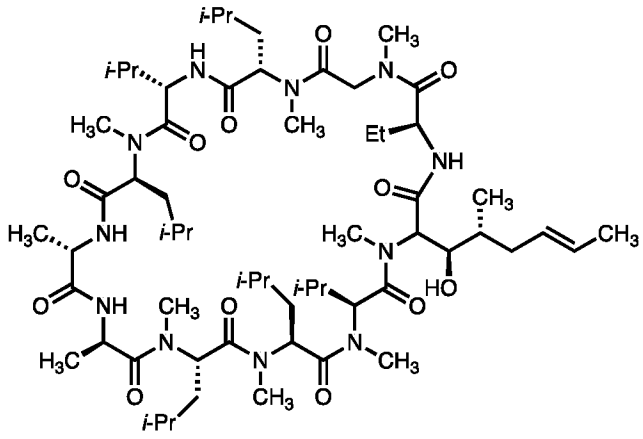
386. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

387. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 492-494.

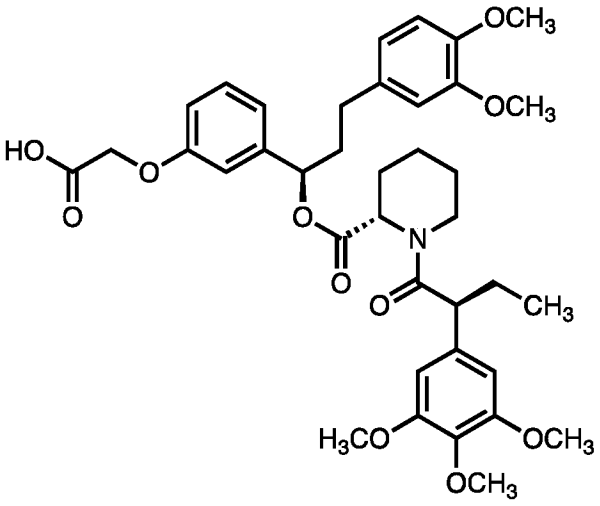
388. Способ по варианту осуществления 382, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из:



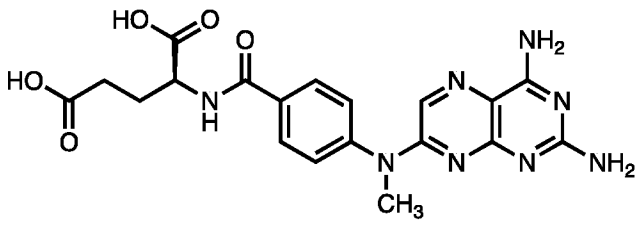
;



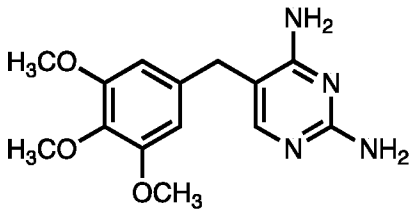
;



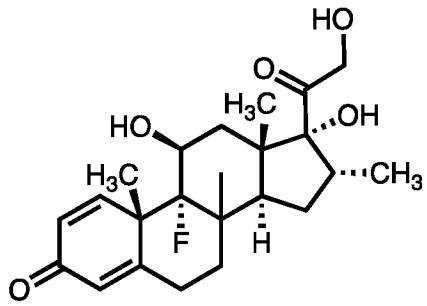
;



;



; и



.

389. Способ по варианту осуществления 382, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

390. Способ по варианту осуществления 382, где napDNAbp представляет собой никазу Cas9 (nCas9).

391. Способ по варианту осуществления 382, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

392. Способ по варианту осуществления 382, где PEgRNA включает SEQ ID NO: 26-36.

393. Способ внесения пептидной метки или эпитопа в белок с помощью праймерного редактирования, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с конструктом праймерного редактора, выполненного с возможностью вносить в нее вторую нуклеотидную последовательность, кодирующую пептидную метку, для получения рекомбинантной нуклеотидной последовательности, так что пептидная метка и белок экспрессируются из рекомбинантной нуклеотидной последовательности в виде слитого белка.

394. Способ по варианту осуществления 383, где пептидная метка применяется для очистки и/или детектирования белка.

395. Способ по варианту осуществления 383, где пептидная метка представляет собой полигистидин (например HHHHHH), FLAG (например DYKDDDDK), V5 (например GKPIPNPLLGLDST), GCN4, HA (например YPYDVPDYA), Myc (например EQKLISEED), или GST.

396. Способ по варианту осуществления 383, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 245-290.

397. Способ по варианту осуществления 383, где пептидная метка слита с белком с помощью линкера.

398. Способ по варианту осуществления 383, где слитый белок имеет следующую структуру: [белок]-[пептидная метка] или [пептидная метка]-[белок], где "]-[" обозначает необязательный линкер.

399. Способ по варианту осуществления 383, где линкер имеет аминокислотную последовательность SEQ ID NO: 37-47.

400. Способ по варианту осуществления 383, где конструкт праймерного редактора включает PEgRNA, включающую нуклеотидную последовательность согласно SEQ ID NO: 18-25.

401. Способ по варианту осуществления 383, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

402. Способ по варианту осуществления 383, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

403. Способ предотвращения или остановки прогрессирования прионного заболевания путем внесения одной или нескольких защитных мутаций в PRNP, кодируемых нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей защитную мутацию; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует PRNP, включающий защитную мутацию и являющийся устойчивым к неправильному фолдингу.

404. Способ по варианту осуществления 403, где прионное заболевание представляет собой прионное заболевание человека.

405. Способ по варианту осуществления 403, где прионное заболевание представляет собой прионное заболевание животного.

406. Способ по варианту осуществления 404, где прионное заболевание представляет собой болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, или куру.

407. Способ по варианту осуществления 403, где прионное заболевание представляет собой коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных.

408. Способ по варианту осуществления 403, где аминокислотная последовательность PRNP дикого типа представляет собой SEQ ID NO: 291-292.

409. Способ по варианту осуществления 403, где способ приводит к получению модифицированной аминокислотной последовательности PRNP, выбранной из группы, состоящей из SEQ ID NO: 293-323, при этом указанный модифицированный белок PRNP устойчив к неправильному фолдингу.

410. Способ по варианту осуществления 403, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

411. Способ по варианту осуществления 403, где napDNAbp представляет собой никазу Cas9 (nCas9).

412. Способ по варианту осуществления 403, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

413. Способ по варианту осуществления 403, где PEgRNA включает SEQ ID NO: 26-36.

414. Способ внесения рибонуклеотидного мотива или метки в представляющую интерес РНК, кодируемую нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую рибонуклеотидный мотив или метку; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей рибонуклеотидный мотив или метку; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает

получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует представляющую интерес модифицированную РНК, включающую рибонуклеотидный мотив или метку.

415. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка представляет собой фрагмент для детектирования.

416. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка влияет на уровень экспрессии представляющей интерес РНК.

417. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка влияет на транспорт или субклеточную локализацию представляющей интерес РНК.

418. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка выбрана из группы, состоящей из SV40 типа 1, SV40 типа 2, SV40 типа 3, hGH, BGH, rbGlob, TK, MALAT1 ENE-mascРНК, KSHV PAN ENE, Sm-бокса/бокса мяРНК U1, 3’ бокса мяРНК U1, тРНК-лизина, аптамера broccoli, аптамера spinach, аптамера mango, рибозима HDV, и m6A.

419. Способ по варианту осуществления 414, где PEgRNA включает SEQ ID NO: 18-25.

420. Способ по варианту осуществления 414, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

421. Способ по варианту осуществления 414, где napDNAbp представляет собой никазу Cas9 (nCas9).

422. Способ по варианту осуществления 414, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

423. Способ внесения или удаления функционального фрагмента в представляющем интерес белке, кодируемом нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок

(napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент или его удаление; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент или его удаление; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует модифицированный белок, включающий представляющий интерес белок и функциональный фрагмент или его удаление, при этом функциональный фрагмент изменяет связанное с модификацией состояние или связанное с локализацией состояние белка.

424. Способ по варианту осуществления 423, где функциональный фрагмент изменяет связанное с фосфорилированием, убиквитинилированием, гликозилированием, липидированием, гидроксилированием, метилированием, ацетилированием, кротонилированием, SUMOилированием состояние представляющего интерес белка.

ГРУППА A. СЛИТЫЕ БЕЛКИ, НАПРАВЛЯЮЩИЕ, И СПОСОБЫ

[1681] Вариант осуществления 1. Слитый белок, включающий программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу.

[1682] Вариант осуществления 2. Слитый белок по варианту осуществления 1, где слитый белок способен осуществлять геномное редактирование с помощью мишень-инициированной обратной транскрипции в присутствии удлиненной направляющей РНК.

[1683] Вариант осуществления 3. Слитый белок по варианту осуществления 1, где napDNAbp имеет никазную активность.

[1684] Вариант осуществления 4. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

[1685] Вариант осуществления 5. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[1686] Вариант осуществления 6. Слитый белок по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1687] Вариант осуществления 7. Слитый белок по варианту осуществления 1, где napDNAbp выбран из группы, состоящей из: Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2C3, и Argonaute, и необязательно имеет никазную активность.

[1688] Вариант осуществления 8. Слитый белок по варианту осуществления 1, где слитый белок при образовании комплекса с удлиненной направляющей РНК способен связываться с последовательностью-мишенью ДНК.

[1689] Вариант осуществления 9. Слитый белок по варианту осуществления 8, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[1690] Вариант осуществления 10. Слитый белок по варианту осуществления 8, где связывание слитого белка в виде комплекса с удлиненной направляющей РНК образует R-петлю.

[1691] Вариант осуществления 11. Слитый белок по варианту осуществления 10, где R-петля включает (i) РНК-ДНК гибрид, включающий удлиненную направляющую РНК и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

[1692] Вариант осуществления 12. Слитый белок по варианту осуществления 11, где комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности обратной транскриптазы, имеющей свободный 3´-конец.

[1693] Вариант осуществления 13. Слитый белок по варианту осуществления 2, где удлиненная направляющая РНК включает (a) направляющую РНК и (b) удлинение РНК на 5´ или 3´-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1694] Вариант осуществления 14. Слитый белок по варианту осуществления 13, где удлинение РНК включает (i) последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

[1695] Вариант осуществления 15. Слитый белок по варианту осуществления 14, где последовательность матрицы обратной транскрипции кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся

смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[1696] Вариант осуществления 16. Слитый белок по варианту осуществления 13, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[1697] Вариант осуществления 17. Слитый белок по варианту осуществления 15, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения.

[1698] Вариант осуществления 18. Слитый белок по варианту осуществления 15, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5′-конец.

[1699] Вариант осуществления 19. Слитый белок по варианту осуществления 18, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

[1700] Вариант осуществления 20. Слитый белок по варианту осуществления 18, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[1701] Вариант осуществления 21. Слитый белок по варианту осуществления 14, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM, или где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,

24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

[1702] Вариант осуществления 22. Слитый белок по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности SEQ ID NO: 18.

[1703] Вариант осуществления 23. Слитый белок по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1704] [0001]Слитый белок по варианту осуществления 1, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89.

[1705] Вариант осуществления 24. Слитый белок по варианту осуществления 1, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1706] [0002]Слитый белок по варианту осуществления 1, где обратная транскриптаза включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 89.

[1707] Вариант осуществления 25. Слитый белок по варианту осуществления 1, где обратная транскриптаза включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1708] Вариант осуществления 26. Слитый белок по варианту осуществления 1, где обратная транскриптаза представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

[1709] Вариант осуществления 27. Слитый белок по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-

[обратная транскриптаза]-COOH; или NH_2-[обратная транскриптаза]-[napDNAbp]-COOH, при этом каждый экземпляр "]-[" обозначает наличие необязательной последовательности линкера.

[1710] Вариант осуществления 28. Слитый белок по варианту осуществления 27, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[1711] Вариант осуществления 29. Слитый белок по варианту осуществления 14, где требуемое нуклеотидное изменение представляет собой изменение одного нуклеотида, вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

[1712] Вариант осуществления 30. Слитый белок по варианту осуществления 29, где вставка или делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22 , по меньшей мере, 23 , по меньшей мере, 24 , по меньшей мере, 25, по меньшей мере, 26, по меньшей мере, 27, по меньшей мере, 28, по меньшей мере, 29, по меньшей мере, 30, по меньшей мере, 31, по меньшей мере, 32, по меньшей мере, 33, по меньшей мере, 34, по меньшей мере, 35, по меньшей мере, 36, по меньшей мере, 37, по меньшей мере, 38, по меньшей мере, 39, по меньшей мере, 40, по меньшей мере, 41, по меньшей мере, 42, по меньшей мере, 43, по меньшей мере, 44, по меньшей мере, 45, по меньшей мере, 46, по меньшей мере, 47, по меньшей мере, 48, по меньшей мере, 49, или, по меньшей мере, 50.

[1713] Вариант осуществления 31. Удлиненная направляющая РНК, включающая направляющую РНК и, по меньшей мере, одно удлинение РНК.

[1714] Вариант осуществления 32. Удлиненная направляющая РНК по варианту осуществления 1, где удлинение РНК расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1715] Вариант осуществления 33. Удлиненная направляющая РНК по варианту осуществления 31, где удлиненная направляющая РНК способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[1716] Вариант осуществления 34. Удлиненная направляющая РНК по варианту осуществления 33, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.
[1717] Вариант осуществления 35. Удлиненная направляющая РНК по варианту осуществления 31, где указанное, по меньшей мере, одно удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.
[1718] Вариант осуществления 36. Удлиненная направляющая РНК по варианту осуществления 35, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.
[1719] Вариант осуществления 37. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.
[1720] Вариант осуществления 38. Удлиненная направляющая РНК по варианту осуществления 35, где участок связывания праймера обратной транскрипции имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере,

12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[1721] Вариант осуществления 39. Удлиненная направляющая РНК по варианту осуществления 35, где необязательная последовательность линкера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[1722] Вариант осуществления 40. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[1723] Вариант осуществления 41. Удлиненная направляющая РНК по варианту осуществления 40, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

[1724] Вариант осуществления 42. Удлиненная направляющая РНК по варианту осуществления 41, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

[1725] Вариант осуществления 43. Удлиненная направляющая РНК по варианту осуществления 41, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[1726] Вариант осуществления 44. Удлиненная направляющая РНК по варианту осуществления 31, включающая нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из

SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810.

[1727] Вариант осуществления 45. Удлиненная направляющая РНК по варианту осуществления 35, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентичная эндогенной ДНК-мишени.

[1728] Вариант осуществления 46. Удлиненная направляющая РНК по варианту осуществления 35, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

[1729] Вариант осуществления 47. Удлиненная направляющая РНК по варианту осуществления 35, где необязательная последовательность линкера имеет длину, по меньшей мере, 1 нуклеотид, или, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, или, по меньшей мере, 15 нуклеотидов.

[1730] Вариант осуществления 48. Комплекс, включающий слитый белок по любому одному из вариантов осуществления 1-30 и удлиненную направляющую РНК.

[1731] Вариант осуществления 49. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК включает направляющую РНК и удлинение РНК на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1732] Вариант осуществления 50. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[1733] Вариант осуществления 51. Комплекс по варианту осуществления 50, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[1734] Вариант осуществления 52. Комплекс по варианту осуществления 49, где указанное, по меньшей мере, одно удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

[1735] Вариант осуществления 53. Комплекс по варианту осуществления 48, где удлиненная направляющая РНК включает нуклеотидную последовательность SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810.

[1736] Вариант осуществления 54. Комплекс по варианту осуществления 52, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, имеющую, по меньшей мере, 80%, или 85%, или 90%, или 95%, или 99% идентичность последовательности относительно эндогенной ДНК-мишени.

[1737] Вариант осуществления 55. Комплекс по варианту осуществления 52, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

[1738] Вариант осуществления 56. Комплекс, включающий napDNAbp и удлиненную направляющую РНК.

[1739] Вариант осуществления 57. Комплекс по варианту осуществления 56, где napDNAbp представляет собой никазу Cas9.

[1740] Вариант осуществления 58. Комплекс по варианту осуществления 56, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно SEQ ID NO: 18.

[1741] Вариант осуществления 59. Комплекс по варианту осуществления 57, где napDNAbp включает аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1742] Вариант осуществления 60. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК включает направляющую РНК и удлинение РНК на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1743] Вариант осуществления 61. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК способна направлять napDNAbp к последовательности-мишени ДНК.

[1744] Вариант осуществления 62. Комплекс по варианту осуществления 61, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом спейсерная последовательность гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[1745] Вариант осуществления 63. Комплекс по варианту осуществления 61, где удлинение РНК включает (i) последовательность матрицы обратной транскрипции, (ii) участок связывания праймера обратной транскрипции, и (iii) необязательно последовательность линкера.

[1746] Вариант осуществления 64. Комплекс по варианту осуществления 57, где удлиненная направляющая РНК включает нуклеотидную последовательность SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810.

[1747] Вариант осуществления 65. Комплекс по варианту осуществления 63, где последовательность матрицы обратной транскрипции включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентичная эндогенной ДНК-мишени.

[1748] Вариант осуществления 66. Комплекс по варианту осуществления 63, где участок связывания праймера обратной транскрипции гибридизируется со свободным 3'-концом разрезанной ДНК.

[1749] Вариант осуществления 67. Комплекс по варианту осуществления 63, где необязательная последовательность линкера имеет длину, по меньшей мере, 1 нуклеотид, или, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, или, по меньшей мере, 15 нуклеотидов.

[1750] Вариант осуществления 68. Полинуклеотид, кодирующий слитый белок по любому одному из вариантов осуществления 1-30.

[1751] Вариант осуществления 69. Вектор, включающий полинуклеотид по варианту осуществления 68.

[1752] Вариант осуществления 70. Клетка, включающая слитый белок по любому из вариантов осуществления 1-30 и удлиненную направляющую РНК, связанную с napDNAbp слитого белка.

[1753] Вариант осуществления 71. Клетка, включающая комплекс по любому одному из вариантов осуществления 48-67.

[1754] Вариант осуществления 72. Фармацевтическая композиция, включающая: (i) слитый белок по любому из вариантов осуществления 1-30, комплекс по вариантам осуществления 48-67, полинуклеотид по варианту осуществления 68, или вектор по варианту осуществления 69; и (ii) фармацевтически приемлемый наполнитель.

[1755] Вариант осуществления 73. Фармацевтическая композиция, включающая: (i) комплекс по вариантам осуществления 48-67, (ii) обратную транскриптазу, поступающую in trans; и (iii) фармацевтически приемлемый наполнитель.

[1756] Вариант осуществления 74. Набор, включающий конструкт нуклеиновой кислоты, включающий: (i) последовательность нуклеиновой кислоты, кодирующую слитый белок по любому одному из вариантов осуществления 1-30; и (ii) промотор, который обеспечивает экспрессию последовательности (i).

[1757] Вариант осуществления 75. Способ внесения требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК, при этом способ включает:

[1758] (i) контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и удлиненную направляющую РНК, при этом слитый белок включает napDNAbp и обратную транскриптазу, и при этом удлиненная направляющая РНК включает последовательность матрицы обратной транскрипции, включающую требуемое нуклеотидное изменение;

[1759] (ii) никирование последовательности двуцепочечной ДНК в не являющейся мишенью цепи, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3′-конец;

[1760] (iii) гибридизацию 3`конца свободной одноцепочечной ДНК с последовательностью матрицы обратной транскрипции, тем самым обеспечивая праймирование домена обратной транскриптазы;

[1761] (iv) полимеризацию цепи ДНК с 3′-конца, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение;

[1762] (v) замещение цепи эндогенной ДНК, примыкающей к участку разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

[1763] Вариант осуществления 76. Способ по варианту осуществления 75, где этап замещения (v) включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

[1764] Вариант осуществления 77. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой замещение одного нуклеотида, делецию, или вставку.

[1765] Вариант осуществления 78. Способ по варианту осуществления 77, где замещение одного нуклеотида представляет собой транзицию или трансверсию.

[1766] Вариант осуществления 79. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[1767] Вариант осуществления 80. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[1768] Вариант осуществления 81. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение представляет собой вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[1769] Вариант осуществления 82. Способ по варианту осуществления 76, где требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена.

[1770] Вариант осуществления 83. Способ по варианту осуществления 82, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса.

[1771] Вариант осуществления 84. Способ по варианту осуществления 82, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[1772] Вариант осуществления 85. Способ по варианту осуществления 76, где napDNAbp представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9.

[1773] Вариант осуществления 86. Способ по варианту осуществления 76, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[1774] Вариант осуществления 87. Способ по варианту осуществления 76, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1775] Вариант осуществления 88. Способ по варианту осуществления 76, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1776] Вариант осуществления 89. Способ по варианту осуществления 76, где домен обратной транскриптазы включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной

из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1777] Вариант осуществления 90. Способ по варианту осуществления 76, где удлиненная направляющая РНК включает удлинение РНК на 3'- или 5'-концах или во внутримолекулярном положении в направляющей РНК, при этом удлинение РНК включает последовательность матрицы обратной транскрипции.

[1778] Вариант осуществления 91. Способ по варианту осуществления 90, где удлинение РНК имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[1779] Вариант осуществления 92. Способ по варианту осуществления 76, где удлиненная направляющая РНК имеет нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 394, 429-442, 641-649, 678-692, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3556, 3628-3698, и 3755-3810.

[1780] Вариант осуществления 93. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий:

[1781] (i) контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и направляющей РНК, которая нацеливает napDNAbp на локус-мишень, при этом направляющая РНК включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение;

[1782] (ii) образование свободного 3′-конца в цепи ДНК в локусе-мишени;

[1783] (iii) гибридизацию свободного 3′-конца с последовательность RT матрицы для праймирования обратной транскрипции;

[1784] (iv) синтезирование одноцепочечного ДНК-флэпа, включающего указанное, по меньшей мере, одно требуемое нуклеотидное изменение, на основе последовательности RT матрицы с помощью обратной транскриптазы;

[1785] (v) и встраивание указанного, по меньшей мере, одного требуемого нуклеотидного изменения в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

[1786] Вариант осуществления 94. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают транзицию.

[1787] Вариант осуществления 95. Способ по варианту осуществления 94, где транзиция выбрана из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[1788] Вариант осуществления 96. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают трансверсию.

[1789] Вариант осуществления 97. Способ по варианту осуществления 96, где трансверсия выбрана из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[1790] Вариант осуществления 98. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[1791] Вариант осуществления 99. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[1792] Вариант осуществления 100. Способ по варианту осуществления 93, где указанное одно или несколько изменений в нуклеотидной последовательности включают корректирование связанного с заболеванием гена.

[1793] Вариант осуществления 101. Способ по варианту осуществления 100, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса.

[1794] Вариант осуществления 102. Способ по варианту осуществления 100, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[1795] Вариант осуществления 103. Способ по варианту осуществления 93, где napDNAbp представляет собой имеющий нуклеазную активность Cas9 или его вариант.

[1796] Вариант осуществления 104. Способ по варианту осуществления 93, где napDNAbp представляет собой не имеющий нуклеазную активность Cas9 (dCas9), или никазу Cas9 (nCas9), или их вариант.

[1797] Вариант осуществления 105. Способ по варианту осуществления 93, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[1798] Вариант осуществления 106. Способ по варианту осуществления 93, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1799] Вариант осуществления 107. Способ по варианту осуществления 93, где обратная транскриптаза вносится in trans.

[1800] Вариант осуществления 108. Способ по варианту осуществления 93, где napDNAbp включает продукт слияния с обратной транскриптазой.

[1801] Вариант осуществления 109. Способ по варианту осуществления 93, где обратная транскриптаза включает любую одну из аминокислотных

последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1802] Вариант осуществления 110. Способ по варианту осуществления 93, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1803] Вариант осуществления 111. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает никирование цепи ДНК нуклеазой.

[1804] Вариант осуществления 112. Способ по варианту осуществления 111, где нуклеаза представляет собой napDNAbp и поступает в виде слитого домена napDNAbp или поступает in trans.

[1805] Вариант осуществления 113. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает контактирование цепи ДНК с химическим агентом.

[1806] Вариант осуществления 114. Способ по варианту осуществления 93, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает внесение ошибки репликации.

[1807] Вариант осуществления 115. Способ по варианту осуществления 93, где этап контактирования молекулы ДНК с napDNAbp и направляющей РНК образует R-петлю.

[1808] Вариант осуществления 116. Способ по варианту осуществления 115, где цепь ДНК, в которой образован свободный 3'-конец, расположена в R-петле.

[1809] Вариант осуществления 117. Способ по варианту осуществления 93, где направляющая РНК включает удлиненную часть, которая включает последовательность матрицы в виде обратной транскриптазы (RT).

[1810] Вариант осуществления 118. Способ по варианту осуществления 117, где удлиненная часть расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[1811] Вариант осуществления 119. Способ по варианту осуществления 93, где направляющая РНК дополнительно включает участок связывания праймера.

[1812] Вариант осуществления 120. Способ по варианту осуществления 93, где направляющая РНК дополнительно включает спейсерную последовательность.

[1813] Вариант осуществления 121. Способ по варианту осуществления 93, где последовательность RT матрицы является гомологичной относительно соответствующей эндогенной ДНК.

[1814] Вариант осуществления 122. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение; и (c) встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[1815] Вариант осуществления 123. Способ по варианту осуществления 122, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[1816] Вариант осуществления 124. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

[1817] Вариант осуществления 125. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[1818] Вариант осуществления 126. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[1819] Вариант осуществления 127. Способ по варианту осуществления 122, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C,

(5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[1820] Вариант осуществления 128. Полинуклеотид, кодирующий удлиненную направляющую РНК по любому одному из вариантов осуществления 31-47.

[1821] Вариант осуществления 129. Вектор, включающий полинуклеотид по варианту осуществления 128.

[1822] Вариант осуществления 130. Клетка, включающая вектор по варианту осуществления 129.

[1823] Вариант осуществления 131. Слитый белок по любому из вариантов осуществления 1-30, где обратная транскриптаза представляет собой подверженную ошибкам обратную транскриптазу.

[1824] Вариант осуществления 132. Способ мутагенеза молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и подверженную ошибкам обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения мутагенизированной одноцепочечной ДНК; и (c) встраивание мутагенизированной одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[1825] Вариант осуществления 133. Способ по варианту осуществления 132, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1826] Вариант осуществления 134. Способ по варианту осуществления 132, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1827] Вариант осуществления 135. Способ по варианту осуществления 132, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1828] Вариант осуществления 136. Способ по варианту осуществления 132, где направляющая РНК включает SEQ ID NO: 222.

[1829] Вариант осуществления 137. Способ по варианту осуществления 132, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

[1830] Вариант осуществления 138. Способ замещения мутации в виде экспансии тринуклеотидных повторов в молекуле-мишени ДНК здоровой последовательностью, включающей здоровое количество тринуклеотидных повторов, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую замещающую последовательность; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей замещающую последовательность; и (c) встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[1831] Вариант осуществления 139. Способ по варианту осуществления 138, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1832] Вариант осуществления 140. Способ по варианту осуществления 138, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1833] Вариант осуществления 141. Способ по варианту осуществления 138, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1834] Вариант осуществления 142. Способ по варианту осуществления 138, где направляющая РНК включает SEQ ID NO: 222.

[1835] Вариант осуществления 143. Способ по варианту осуществления 138, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

[1836] Вариант осуществления 144. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов ассоциирована с болезнью Хантингтона, синдромом ломкой Х-хромосомы, или атаксией Фридрейха.

[1837] Вариант осуществления 145. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов CAG.

[1838] Вариант осуществления 146. Способ по варианту осуществления 138, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов GAA.

[1839] Вариант осуществления 147. Способ внесения функционального фрагмента в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и функциональный фрагмент.

[1840] Вариант осуществления 148. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой пептидную метку.

[1841] Вариант осуществления 149. Способ по варианту осуществления 148, где пептидная метка представляет собой аффинную метку, метку солюбилизации, хроматографическую метку, эпитопную метку, или флуоресцентную метку.

[1842] Вариант осуществления 150. Способ по варианту осуществления 148, где пептидная метка выбрана из группы, состоящей из: AviTag (SEQ ID NO: 245); C-метки (SEQ ID NO: 246); калмодулиновой метки (SEQ ID NO: 247); полиглутаматной метки (SEQ ID NO: 248); E-метки (SEQ ID NO: 249); FLAG-метки (SEQ ID NO: 2); HA-метки (SEQ ID NO: 5); His-метки (SEQ ID NO: 252-262); Myc-метки (SEQ ID NO: 6); NE-метки (SEQ ID NO: 264); Rho1D4-метки (SEQ ID NO: 265); S-метки (SEQ ID NO: 266); SBP-метки (SEQ ID NO: 267); Softag-1 (SEQ ID NO: 268); Softag-2 (SEQ ID NO: 269); Spot-метки (SEQ ID NO: 270); Strep-метки (SEQ ID NO: 271); метки TC (SEQ ID NO:

272); метки Ty (SEQ ID NO: 273); метки V5 (SEQ ID NO: 3); VSV-метки (SEQ ID NO: 275); и метки Xpress (SEQ ID NO: 276).

[1843] Вариант осуществления 151. Способ по варианту осуществления 148, где пептидная метка выбрана из группы, состоящей из: эпитопа AU1 (SEQ ID NO: 278); эпитопа AU5 (SEQ ID NO: 279); эпитопа бактериофага T7 (T7-метка) (SEQ ID NO: 280); метки на основе вируса Bluetongue (B-метки) (SEQ ID NO: 281); эпитопа E2 (SEQ ID NO: 282); аффинной гистидиновой метки (HAT) (SEQ ID NO: 283); эпитопа HSV (SEQ ID NO: 284); полиаргинина (Arg-метки) (SEQ ID NO: 285); полиаспартата (Asp метки) (SEQ ID NO: 286); полифенилаланина (Phe-метки) (SEQ ID NO: 287); S1-метки (SEQ ID NO: 288); S-метки (SEQ ID NO: 266); и VSV-G (SEQ ID NO: 275).

[1844] Вариант осуществления 152. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой иммуноэпитоп.

[1845] Вариант осуществления 153. Способ по варианту осуществления 152, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 398); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 3 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

[1846] Вариант осуществления 154. Способ по варианту осуществления 147, где функциональный фрагмент изменяет локализацию представляющего интерес белка.

[1847] Вариант осуществления 155. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой метку расщепления, так что обеспечивается изменение скорости расщепления представляющего интерес белка.

[1848] Вариант осуществления 156. Способ по варианту осуществления 155, где метка расщепления включает аминокислотную последовательность. кодирующую метки расщепления, как описано здесь.

[1849] Вариант осуществления 157. Способ по варианту осуществления 147, где функциональный фрагмент представляет собой домен связывания малой молекулы.

[1850] Вариант осуществления 158. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

[1851] Вариант осуществления 159. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

[1852] Вариант осуществления 160. Способ по варианту осуществления 157, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 492-494.

[1853] Вариант осуществления 161. Способ по варианту осуществления 157, где домен связывания малой молекулы вносится в два или более представляющих интерес белка.

[1854] Вариант осуществления 162. Способ по варианту осуществления 161, где указанные два или более представляющие интерес белка могут подвергаться димеризации при контактировании с малой молекулой.

[1855] Вариант осуществления 163. Способ по варианту осуществления 157, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из соединений, описанных в Варианте осуществления 163 Группы 1. .

[1856] Вариант осуществления 164. Способ внесения иммуноэпитопа в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая

кодирует слитый белок, включающий представляющий интерес белок и иммуноэпитоп.

[1857] Вариант осуществления 165. Способ по варианту осуществления 164, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 3 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

[1858] Вариант осуществления 166. Способ по варианту осуществления 164, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1859] Вариант осуществления 167. Способ по варианту осуществления 164, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1860] Вариант осуществления 168. Способ по варианту осуществления 164, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1861] Вариант осуществления 169. Способ по варианту осуществления 164, где направляющая РНК включает SEQ ID NO: 222.

[1862] Вариант осуществления 170. Способ внесения домена димеризации на основе малой молекулы в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую домен димеризации на основе малой молекулы; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной

последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и домен димеризации на основе малой молекулы.

[1863] Вариант осуществления 171. Способ по варианту осуществления 170, дополнительно включающий осуществление способа в отношении второго представляющего интерес белка.

[1864] Вариант осуществления 172. Способ по варианту осуществления 171, где первый представляющий интерес белок и второй представляющий интерес белок подвергаются димеризации в присутствии малой молекулы, которая связывается с доменом димеризации на каждом из указанных белков.

[1865] Вариант осуществления 173. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

[1866] Вариант осуществления 174. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

[1867] Вариант осуществления 175. Способ по варианту осуществления 170, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 490 и 493-494.

[1868] Вариант осуществления 176. Способ по варианту осуществления 170, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из соединений, описанных в Варианте осуществления 163 Группы 1.

[1869] Вариант осуществления 177. Способ по варианту осуществления 170, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1870] Вариант осуществления 178. Способ по варианту осуществления 170, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1871] Вариант осуществления 179. Способ по варианту осуществления 170, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1872] Вариант осуществления 180. Способ по варианту осуществления 170, где направляющая РНК включает SEQ ID NO: 222.

[1873] Вариант осуществления 181. Способ внесения пептидной метки или эпитопа в белок с помощью праймерного редактирования, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с конструктом праймерного редактора, выполненного с возможностью вносить в нее вторую нуклеотидную последовательность, кодирующую пептидную метку, для получения рекомбинантной нуклеотидной последовательности, так что пептидная метка и белок экспрессируются из рекомбинантной нуклеотидной последовательности в виде слитого белка.

[1874] Вариант осуществления 182. Способ по варианту осуществления 181, где пептидная метка применяется для очистки и/или детектирования белка.

[1875] Вариант осуществления 183. Способ по варианту осуществления 181, где пептидная метка представляет собой полигистидин (например HHHHHH) (SEQ ID NO: 252-262), FLAG (например DYKDDDDK) (SEQ ID NO: 2), V5 (например GKPIPNPLLGLDST) (SEQ ID NO: 3), GCN4, HA (например YPYDVPDYA) (SEQ ID NO: 5), Myc (например EQKLISEED) (SEQ ID NO: 6), GST….и так далее

[1876] Вариант осуществления 184. Способ по варианту осуществления 181, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 1-6, 245-249, 252-262, 264-273, 275-276, 281, 278-288, и 622.

[1877] Вариант осуществления 185. Способ по варианту осуществления 181, где пептидная метка слита с белком с помощью линкера.

[1878] Вариант осуществления 186. Способ по варианту осуществления 181, где слитый белок имеет следующую структуру: [белок]-[пептидная метка] или [пептидная метка]-[белок], где "]-[" обозначает необязательный линкер.

[1879] Вариант осуществления 187. Способ по варианту осуществления 181, где линкер имеет аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[1880] Вариант осуществления 188. Способ по варианту осуществления 181, где конструкт праймерного редактора включает PEgRNA, включающую нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[1881] Вариант осуществления 189. Способ по варианту осуществления 181, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер

комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[1882] Вариант осуществления 190. Способ по варианту осуществления 181, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[1883] Вариант осуществления 191. Способ предотвращения или остановки прогрессирования прионного заболевания путем внесения одной или нескольких защитных мутаций в PRNP, кодируемых нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей защитную мутацию; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует PRNP, включающий защитную мутацию и являющийся устойчивым к неправильному фолдингу.

[1884] Вариант осуществления 192. Способ по варианту осуществления 191, где прионное заболевание представляет собой прионное заболевание человека.

[1885] Вариант осуществления 193. Способ по варианту осуществления 191, где прионное заболевание представляет собой прионное заболевание животного.

[1886] Вариант осуществления 194. Способ по варианту осуществления 192, где прионное заболевание представляет собой болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, или куру.

[1887] Вариант осуществления 195. Способ по варианту осуществления 192, где прионное заболевание представляет собой коровью губчатую энцефалопатию (BSE или

"коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных.

[1888] Вариант осуществления 196. Способ по варианту осуществления 191, где аминокислотная последовательность PRNP дикого типа представляет собой SEQ ID NO: 291-292.

[1889] Вариант осуществления 197. Способ по варианту осуществления 191, где способ приводит к получению модифицированной аминокислотной последовательности PRNP, выбранной из группы, состоящей из SEQ ID NO: 293-309, и 311-323, при этом указанный модифицированный белок PRNP устойчив к неправильному фолдингу.

[1890] Вариант осуществления 198. Способ по варианту осуществления 191, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1891] Вариант осуществления 199. Способ по варианту осуществления 191, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1892] Вариант осуществления 200. Способ по варианту осуществления 191, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1893] Вариант осуществления 201. Способ по варианту осуществления 191, где направляющая РНК включает SEQ ID NO: 222.

[1894] Вариант осуществления 202. Способ внесения рибонуклеотидного мотива или метки в представляющую интерес РНК, кодируемую нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую рибонуклеотидный мотив или метку; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей рибонуклеотидный мотив или метку; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной

последовательности-мишени, которая кодирует представляющую интерес модифицированную РНК, включающую рибонуклеотидный мотив или метку.

[1895] Вариант осуществления 203. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка представляет собой фрагмент для детектирования.

[1896] Вариант осуществления 204. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка влияет на уровень экспрессии представляющей интерес РНК.

[1897] Вариант осуществления 205. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка влияет на транспорт или субклеточную локализацию представляющей интерес РНК.

[1898] Вариант осуществления 206. Способ по варианту осуществления 202, где рибонуклеотидный мотив или метка выбрана из группы, состоящей из SV40 типа 1, SV40 типа 2, SV40 типа 3, hGH, BGH, rbGlob, TK, MALAT1 ENE-mascРНК, KSHV PAN ENE, Sm-бокса/бокса мяРНК U1, 3’ бокса мяРНК U1, тРНК-лизина, аптамера broccoli, аптамера spinach, аптамера mango, рибозима HDV, и m6A.

[1899] Вариант осуществления 207. Способ по варианту осуществления 202, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[1900] Вариант осуществления 208. Способ по варианту осуществления 202, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[1901] Вариант осуществления 209. Способ по варианту осуществления 202, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1902] Вариант осуществления 210. Способ по варианту осуществления 202, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1903] Вариант осуществления 211. Способ внесения или удаления функционального фрагмента в представляющем интерес белке, кодируемом нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) PEgRNA, включающей матрицу для редактирования,

кодирующую функциональный фрагмент или его удаление; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент или его удаление; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует модифицированный белок, включающий представляющий интерес белок и функциональный фрагмент или его удаление, при этом функциональный фрагмент изменяет связанное с модификацией состояние или связанное с локализацией состояние белка.

[1904] Вариант осуществления 212. Способ по варианту осуществления 211, где функциональный фрагмент изменяет связанное с фосфорилированием, убиквитинилированием, гликозилированием, липидированием, гидроксилированием, метилированием, ацетилированием, кротонилированием, SUMOилированием состояние представляющего интерес белка.

[1905] Вариант осуществления 213. Слитый белок, включающий программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу.

[1906] Вариант осуществления 214. Слитый белок по варианту осуществления 213, где слитый белок способен осуществлять геномное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA).

[1907] Вариант осуществления 215. Слитый белок по варианту осуществления 213, где napDNAbp имеет никазную активность.

[1908] Вариант осуществления 216. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой белок Cas9 или его вариант.

[1909] Вариант осуществления 217. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[1910] Вариант осуществления 218. Слитый белок по варианту осуществления 213, где napDNAbp представляет собой никазу Cas9 (nCas9).

[1911] Вариант осуществления 219. Слитый белок по варианту осуществления 213, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[1912] Вариант осуществления 220. Слитый белок по варианту осуществления 213, где слитый белок при образовании комплекса с PEgRNA способен связываться с последовательностью-мишенью ДНК.

[1913] Вариант осуществления 221. Слитый белок по варианту осуществления 220, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[1914] Вариант осуществления 222. Слитый белок по варианту осуществления 220, где связывание слитого белка в виде комплекса с PEgRNA образует R-петлю.

[1915] Вариант осуществления 223. Слитый белок по варианту осуществления 222, где R-петля включает (i) РНК-ДНК гибрид, включающий PEgRNA и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

[1916] Вариант осуществления 224. Слитый белок по варианту осуществления 223, где комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности, имеющей свободный 3′-конец.

[1917] Вариант осуществления 225. Слитый белок по варианту осуществления 214, где PEgRNA включает (a) направляющую РНК и (b) плечо удлинения на 5′ или 3′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1918] Вариант осуществления 226. Слитый белок по варианту осуществления 225, где плечо удлинения включает (i) последовательность матрицы для синтеза ДНК, включающую требуемое нуклеотидное изменение, и (ii) участок связывания праймера.

[1919] Вариант осуществления 227. Слитый белок по варианту осуществления 226, где последовательность матрицы для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[1920] Вариант осуществления 228. Слитый белок по варианту осуществления 225, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей

мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[1921] Вариант осуществления 229. Слитый белок по варианту осуществления 227, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения.

[1922] Вариант осуществления 230. Слитый белок по варианту осуществления 227, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5'-конец.

[1923] Вариант осуществления 231. Слитый белок по варианту осуществления 230, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

[1924] Вариант осуществления 232. Слитый белок по варианту осуществления 230, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[1925] Вариант осуществления 233. Слитый белок по варианту осуществления 226, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM, или где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

[1926] Вариант осуществления 234. Слитый белок по варианту осуществления 213, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%,

95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности SEQ ID NO: 18.

[1927] Вариант осуществления 235. Слитый белок по варианту осуществления 213, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1928] Вариант осуществления 236. Слитый белок по варианту осуществления 213, где полимераза представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1929] Вариант осуществления 237. Слитый белок по варианту осуществления 213, где полимераза представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1930] Вариант осуществления 238. Слитый белок по варианту осуществления 213, где полимераза представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

[1931] Вариант осуществления 239. Слитый белок по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-[полимераза]-COOH; или NH_2-[полимераза]-[napDNAbp]-COOH, где каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера.

[1932] Вариант осуществления 240. Слитый белок по варианту осуществления 239, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[1933] Вариант осуществления 241. Слитый белок по варианту осуществления 226, где требуемое нуклеотидное изменение представляет собой изменение одного нуклеотида, вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

[1934] Вариант осуществления 242. Слитый белок по варианту осуществления 241, где вставка или делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22 , по меньшей мере, 23 , по меньшей мере, 24 , по меньшей мере, 25, по меньшей мере, 26, по меньшей мере, 27, по меньшей мере, 28, по меньшей мере, 29, по меньшей мере, 30, по меньшей мере, 31, по меньшей мере, 32, по меньшей мере, 33, по меньшей мере, 34, по меньшей мере, 35, по меньшей мере, 36, по меньшей мере, 37, по меньшей мере, 38, по меньшей мере, 39, по меньшей мере, 40, по меньшей мере, 41, по меньшей мере, 42, по меньшей мере, 43, по меньшей мере, 44, по меньшей мере, 45, по меньшей мере, 46, по меньшей мере, 47, по меньшей мере, 48, по меньшей мере, 49, или, по меньшей мере, 50.

[1935] Вариант осуществления 243. PEgRNA, включающая направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты, включающее матрицу для синтеза ДНК.

[1936] Вариант осуществления 244. PEgRNA по варианту осуществления 241, где плечо удлинения нуклеиновой кислоты расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК и где плечо удлинения нуклеиновой кислоты представляет собой ДНК или РНК.

[1937] Вариант осуществления 245. PEgRNA по варианту осуществления 242, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[1938] Вариант осуществления 246. PEgRNA по варианту осуществления 245, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[1939] Вариант осуществления 247. PEgRNA по варианту осуществления 243, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

[1940] Вариант осуществления 248. PEgRNA по варианту осуществления 247, где плечо удлинения нуклеиновой кислоты имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[1941] Вариант осуществления 249. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[1942] Вариант осуществления 250. PEgRNA по варианту осуществления 247, где участок связывания праймера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[1943] Вариант осуществления 251. PEgRNA по варианту осуществления 243, дополнительно включающая, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования РНК-белка.

[1944] Вариант осуществления 252. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной

относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[1945] Вариант осуществления 253. PEgRNA по варианту осуществления 252, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

[1946] Вариант осуществления 254. PEgRNA по варианту осуществления 253, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

[1947] Вариант осуществления 255. PEgRNA по варианту осуществления 253, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[1948] Вариант осуществления 256. PEgRNA по варианту осуществления 243, включающая нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[1949] Вариант осуществления 257. PEgRNA по варианту осуществления 247, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[1950] Вариант осуществления 258. PEgRNA по варианту осуществления 247, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[1951] Вариант осуществления 259. PEgRNA по варианту осуществления 251, где указанная, по меньшей мере, одна дополнительная структура расположена на 3'- или 5'-конце PEgRNA.

[1952] Вариант осуществления 260. Комплекс, включающий слитый белок по любому одному из вариантов осуществления 213-242 и PEgRNA.

[1953] Вариант осуществления 261. Комплекс по варианту осуществления 260, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1954] Вариант осуществления 262. Комплекс по варианту осуществления 260, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[1955] Вариант осуществления 263. Комплекс по варианту осуществления 262, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[1956] Вариант осуществления 264. Комплекс по варианту осуществления 261, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

[1957] Вариант осуществления 265. Комплекс по варианту осуществления 260, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[1958] Вариант осуществления 266. Комплекс по варианту осуществления 264, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[1959] Вариант осуществления 267. Комплекс по варианту осуществления 264, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[1960] Вариант осуществления 268. Комплекс, включающий napDNAbp и PEgRNA.

[1961] Вариант осуществления 269. Комплекс по варианту осуществления 268, где napDNAbp представляет собой никазу Cas9.

[1962] Вариант осуществления 270. Комплекс по варианту осуществления 268, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[1963] Вариант осуществления 271. Комплекс по варианту осуществления 268, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1964] Вариант осуществления 272. Комплекс по варианту осуществления 268, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[1965] Вариант осуществления 273. Комплекс по варианту осуществления 268, где PEgRNA способна направлять napDNAbp к последовательности-мишени ДНК.

[1966] Вариант осуществления 274. Комплекс по варианту осуществления 272, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом спейсерная последовательность PEgRNA гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[1967] Вариант осуществления 275. Комплекс по варианту осуществления 273, где плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

[1968] Вариант осуществления 276. Комплекс по варианту осуществления 269, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[1969] Вариант осуществления 277. Комплекс по варианту осуществления 276, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[1970] Вариант осуществления 278. Комплекс по варианту осуществления 276, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[1971] Вариант осуществления 279. Комплекс по варианту осуществления 276, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

[1972] Вариант осуществления 280. Полинуклеотид, кодирующий слитый белок по любому одному из вариантов осуществления 213-242.

[1973] Вариант осуществления 281. Вектор, включающий полинуклеотид по варианту осуществления 280.

[1974] Вариант осуществления 282. Клетка, включающая слитый белок по любому из вариантов осуществления 213-242 и PEgRNA, связанную с napDNAbp слитого белка.

[1975] Вариант осуществления 283. Клетка, включающая комплекс по любому одному из вариантов осуществления 260-279.

[1976] Вариант осуществления 284. Фармацевтическая композиция, включающая: (i) слитый белок по любому из вариантов осуществления 213-242, комплекс по вариантам осуществления 260-279, полинуклеотид по варианту осуществления 68, или вектор по варианту осуществления 69; и (ii) фармацевтически приемлемый наполнитель.

[1977] Вариант осуществления 285. Фармацевтическая композиция, включающая: (i) комплекс по вариантам осуществления 260-279, (ii) полимеразу, поступающую in trans; и (iii) фармацевтически приемлемый наполнитель.

[1978] Вариант осуществления 286. Набор, включающий конструкт нуклеиновой кислоты, включающий: (i) последовательность нуклеиновой кислоты, кодирующую слитый белок по любому одному из вариантов осуществления 213-242; и (ii) промотор, который обеспечивает экспрессию последовательности (i).

[1979] Вариант осуществления 287. Способ внесения требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК, при этом способ включает:

[1980] (i) контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и PEgRNA, при этом слитый белок включает napDNAbp и полимеразу, и при этом PEgRNA включает матрицу для синтеза ДНК, включающую требуемое нуклеотидное изменение и участок связывания праймера;

[1981] (ii) никирование последовательности двуцепочечной ДНК, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3′-конец;

[1982] (iii) гибридизирование 3'-конца свободной одноцепочечной ДНК с участком связывания праймера, тем самым обеспечивая праймирование полимеразы;

[1983] (iv) полимеризацию цепи ДНК с 3′-конца, гибридизированного с участком связывания праймера, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение и являющегося комплементарным матрице для синтеза ДНК;

[1984] (v) замещение цепи эндогенной ДНК, примыкающей к участку разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

[1985] Вариант осуществления 288. Способ по варианту осуществления 287, где этап замещения (v) включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

[1986] Вариант осуществления 289. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой замещение одного нуклеотида, делецию, или вставку.

[1987] Вариант осуществления 290. Способ по варианту осуществления 289, где замещение одного нуклеотида представляет собой транзицию или трансверсию.

[1988] Вариант осуществления 291. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[1989] Вариант осуществления 292. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[1990] Вариант осуществления 293. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение представляет собой вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[1991] Вариант осуществления 294. Способ по варианту осуществления 288, где требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена.

[1992] Вариант осуществления 295. Способ по варианту осуществления 294, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

[1993] Вариант осуществления 296. Способ по варианту осуществления 294, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[1994] Вариант осуществления 297. Способ по варианту осуществления 287, где napDNAbp представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9.

[1995] Вариант осуществления 298. Слитый белок по варианту осуществления 287, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или

аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности SEQ ID NO: 18.

[1996] Вариант осуществления 299. Способ по варианту осуществления 287, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[1997] Вариант осуществления 300. Слитый белок по варианту осуществления 287, где полимераза представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1998] Вариант осуществления 301. Слитый белок по варианту осуществления 287, где полимераза представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[1999] Вариант осуществления 302. Способ по варианту осуществления 287, где PEgRNA включает плечо удлинения нуклеиновой кислоты на 3'- или 5'-концах или во внутримолекулярном положении в направляющей РНК, при этом плечо удлинения включает последовательность матрицы для синтеза ДНК и участок связывания праймера.

[2000] Вариант осуществления 303. Способ по варианту осуществления 302, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23

нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[2001] Вариант осуществления 304. Способ по варианту осуществления 287, где PEgRNA имеет нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2002] Вариант осуществления 305. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий:

[2003] (i) контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и PEgRNA, которая нацеливает napDNAbp на локус-мишень, при этом PEgRNA включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение и участок связывания праймера;

[2004] (ii) образование свободного 3′-конца в цепи ДНК в локусе-мишени;

[2005] (iii) гибридизацию свободного 3′-конца с участком связывания праймера для праймирования обратной транскрипции;

[2006] (iv) синтезирование одноцепочечного ДНК-флэпа, включающего указанное, по меньшей мере, одно требуемое нуклеотидное изменение, на основе последовательности RT матрицы с помощью обратной транскриптазы;

[2007] (v) и встраивание указанного, по меньшей мере, одного требуемого нуклеотидного изменения в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

[2008] Вариант осуществления 306. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают транзицию.

[2009] Вариант осуществления 307. Способ по варианту осуществления 306, где транзиция выбрана из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[2010] Вариант осуществления 308. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают трансверсию.

[2011] Вариант осуществления 309. Способ по варианту осуществления 308, где трансверсия выбрана из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[2012] Вариант осуществления 310. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (9) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2013] Вариант осуществления 311. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[2014] Вариант осуществления 312. Способ по варианту осуществления 305, где указанное одно или несколько изменений в нуклеотидной последовательности включают корректирование связанного с заболеванием гена.

[2015] Вариант осуществления 313. Способ по варианту осуществления 312, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

[2016] Вариант осуществления 314. Способ по варианту осуществления 312, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[2017] Вариант осуществления 315. Способ по варианту осуществления 305, где napDNAbp представляет собой имеющий нуклеазную активность Cas9 или его вариант.

[2018] Вариант осуществления 316. Способ по варианту осуществления 305, где napDNAbp представляет собой не имеющий нуклеазную активность Cas9 (dCas9), или никазу Cas9 (nCas9), или их вариант.

[2019] Вариант осуществления 317. Способ по варианту осуществления 305, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 18.

[2020] Вариант осуществления 318. Способ по варианту осуществления 305, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2021] Вариант осуществления 319. Способ по варианту осуществления 305, где обратная транскриптаза вносится in trans.

[2022] Вариант осуществления 320. Способ по варианту осуществления 305, где napDNAbp включает продукт слияния с обратной транскриптазой.

[2023] Вариант осуществления 321. Способ по варианту осуществления 305, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2024] Вариант осуществления 322. Способ по варианту осуществления 305, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2025] Вариант осуществления 323. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает никирование цепи ДНК нуклеазой.

[2026] Вариант осуществления 324. Способ по варианту осуществления 323, где нуклеаза поступает in trans.

[2027] Вариант осуществления 325. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает контактирование цепи ДНК с химическим агентом.

[2028] Вариант осуществления 326. Способ по варианту осуществления 305, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает внесение ошибки репликации.

[2029] Вариант осуществления 327. Способ по варианту осуществления 305, где этап контактирования молекулы ДНК с napDNAbp и направляющей РНК образует R-петлю.

[2030] Вариант осуществления 328. Способ по варианту осуществления 327, где цепь ДНК, в которой образован свободный 3'-конец, расположена в R-петле.

[2031] Вариант осуществления 329. Способ по варианту осуществления 315, где PEgRNA включает плечо удлинения, которое включает последовательность матрицы в виде обратной транскриптазы (RT) и участок связывания праймера.

[2032] Вариант осуществления 330. Способ по варианту осуществления 329, где плечо удлинения расположено на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[2033] Вариант осуществления 331. Способ по варианту осуществления 305, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования РНК-белка.

[2034] Вариант осуществления 332. Способ по варианту осуществления 305, где PEgRNA дополнительно включает плечо гомологии.

[2035] Вариант осуществления 333. Способ по варианту осуществления 305, где последовательность RT матрицы является гомологичной относительно соответствующей эндогенной ДНК.

[2036] Вариант осуществления 334. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление

мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение; и (c) встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[2037] Вариант осуществления 335. Способ по варианту осуществления 334, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[2038] Вариант осуществления 336. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

[2039] Вариант осуществления 337. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[2040] Вариант осуществления 338. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[2041] Вариант осуществления 339. Способ по варианту осуществления 334, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2042] Вариант осуществления 340. Полинуклеотид, кодирующий PEgRNA по любому одному из вариантов осуществления 243-259.

[2043] Вариант осуществления 341. Вектор, включающий полинуклеотид по варианту осуществления 340.

[2044] Вариант осуществления 342. Клетка, включающая вектор по варианту осуществления 341.

[2045] Вариант осуществления 343. Слитый белок по варианту осуществления 213, где полимераза представляет собой подверженную ошибкам обратную транскриптазу.

[2046] Вариант осуществления 344. Способ мутагенеза молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и подверженную ошибкам обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение; (b) осуществление мишень-инициированной обратной транскрипции RT матрицы для получения мутагенизированной одноцепочечной ДНК; и (c) встраивание мутагенизированной одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[2047] Вариант осуществления 345. Способ по любому предыдущему варианту осуществления, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2048] Вариант осуществления 346. Способ по любому предыдущему варианту осуществления, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2049] Вариант осуществления 347. Способ по варианту осуществления 344, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-25.

[2050] Вариант осуществления 348. Способ по варианту осуществления 344, где направляющая РНК включает SEQ ID NO: 222.

[2051] Вариант осуществления 349. Способ по варианту осуществления 344, где этап (b) осуществления мишень-инициированной обратной транскрипции включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать обратную транскрипцию за счет отжига с участком связывания праймера на направляющей РНК.

[2052] Вариант осуществления 350. Способ замещения мутации в виде экспансии тринуклеотидных повторов в молекуле-мишени ДНК здоровой последовательностью, включающей здоровое количество тринуклеотидных повторов, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК,

включающую замещающую последовательность и участок связывания праймера; (b) осуществление праймерного редактирования для получения одноцепочечной ДНК, включающей замещающую последовательность; и (c) встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[2053] Вариант осуществления 351. Способ по варианту осуществления 350, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2054] Вариант осуществления 352. Способ по варианту осуществления 350, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2055] Вариант осуществления 353. Способ по варианту осуществления 350, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2056] Вариант осуществления 354. Способ по варианту осуществления 350, где направляющая РНК включает SEQ ID NO: 222.

[2057] Вариант осуществления 355. Способ по варианту осуществления 350, где этап (b) осуществления праймерного редактирования включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать полимеразу за счет отжига с участком связывания праймера на направляющей РНК.

[2058] Вариант осуществления 356. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов ассоциирована с болезнью Хантингтона, синдромом ломкой Х-хромосомы, или атаксией Фридрейха.

[2059] Вариант осуществления 357. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов CAG.

[2060] Вариант осуществления 358. Способ по варианту осуществления 350, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов GAA.

[2061] Вариант осуществления 359. Способ внесения функционального фрагмента в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой

ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и функциональный фрагмент.

[2062] Вариант осуществления 360. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой пептидную метку.

[2063] Вариант осуществления 361. Способ по варианту осуществления 360, где пептидная метка представляет собой аффинную метку, метку солюбилизации, хроматографическую метку, эпитопную или иммуноэпитопную метку, или флуоресцентную метку.

[2064] Вариант осуществления 362. Способ по варианту осуществления 360, где пептидная метка выбрана из группы, состоящей из: AviTag (SEQ ID NO: 245); C-метки (SEQ ID NO: 246); калмодулиновой метки (SEQ ID NO: 247); полиглутаматной метки (SEQ ID NO: 248); E-метки (SEQ ID NO: 249); FLAG-метки (SEQ ID NO: 2); HA-метки (SEQ ID NO: 5); His-метки (SEQ ID NO: 252-262); Myc-метки (SEQ ID NO: 6); NE-метки (SEQ ID NO: 264); Rho1D4-метки (SEQ ID NO: 265); S-метки (SEQ ID NO: 266); SBP-метки (SEQ ID NO: 267); Softag-1 (SEQ ID NO: 268); Softag-2 (SEQ ID NO: 269); Spot-метки (SEQ ID NO: 270); Strep-метки (SEQ ID NO: 271); метки TC (SEQ ID NO: 272); метки Ty (SEQ ID NO: 273); метки V5 (SEQ ID NO: 3); VSV-метки (SEQ ID NO: 275); и метки Xpress (SEQ ID NO: 276).

[2065] Вариант осуществления 363. Способ по варианту осуществления 360, где пептидная метка выбрана из группы, состоящей из: эпитопа AU1 (SEQ ID NO: 278); эпитопа AU5 (SEQ ID NO: 279); эпитопа бактериофага T7 (T7-метка) (SEQ ID NO: 280); метки на основе вируса Bluetongue (B-метки) (SEQ ID NO: 281); эпитопа E2 (SEQ ID NO: 282); аффинной гистидиновой метки (HAT) (SEQ ID NO: 283); эпитопа HSV (SEQ ID NO: 284); полиаргинина (Arg-метки) (SEQ ID NO: 285); полиаспартата (Asp метки) (SEQ ID NO: 286); полифенилаланина (Phe-метки) (SEQ ID NO: 287); S1-метки (SEQ ID NO: 288); S-метки (SEQ ID NO: 266); и VSV-G (SEQ ID NO: 275).

[2066] Вариант осуществления 364. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой иммуноэпитоп.

[2067] Вариант осуществления 365. Способ по варианту осуществления 364, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 3 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

[2068] Вариант осуществления 366. Способ по варианту осуществления 359, где функциональный фрагмент изменяет локализацию представляющего интерес белка.

[2069] Вариант осуществления 367. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой метку расщепления, так что обеспечивается изменение скорости расщепления представляющего интерес белка.

[2070] Вариант осуществления 368. Способ по варианту осуществления 367, где метка расщепления приводит к расщеплению помеченного белка.

[2071] Вариант осуществления 369. Способ по варианту осуществления 359, где функциональный фрагмент представляет собой домен связывания малой молекулы.

[2072] Вариант осуществления 370. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

[2073] Вариант осуществления 371. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

[2074] Вариант осуществления 372. Способ по варианту осуществления 359, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 490 и 493-494.

[2075] Вариант осуществления 373. Способ по варианту осуществления 359, где домен связывания малой молекулы вносится в два или более представляющих интерес белка.

[2076] Вариант осуществления 374. Способ по варианту осуществления 373, где указанные два или более представляющие интерес белка могут подвергаться димеризации при контактировании с малой молекулой.

[2077] Вариант осуществления 375. Способ по варианту осуществления 369, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из соединений, описанных в Варианте осуществления 163 Группы 1.

[2078] Вариант осуществления 376. Способ внесения иммуноэпитопа в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и иммуноэпитоп.

[2079] Вариант осуществления 377. Способ по варианту осуществления 376, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 3 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом

гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

[2080] Вариант осуществления 378. Способ по варианту осуществления 376, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2081] Вариант осуществления 379. Способ по варианту осуществления 376, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2082] Вариант осуществления 380. Способ по варианту осуществления 376, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2083] Вариант осуществления 381. Способ по варианту осуществления 376, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2084] Вариант осуществления 382. Способ внесения домена димеризации на основе малой молекулы в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую домен димеризации на основе малой молекулы; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и домен димеризации на основе малой молекулы.

[2085] Вариант осуществления 383. Способ по варианту осуществления 382, дополнительно включающий осуществление способа в отношении второго представляющего интерес белка.

[2086] Вариант осуществления 384. Способ по варианту осуществления 383, где первый представляющий интерес белок и второй представляющий интерес белок

подвергаются димеризации в присутствии малой молекулы, которая связывается с доменом димеризации на каждом из указанных белков.

[2087] Вариант осуществления 385. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

[2088] Вариант осуществления 386. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

[2089] Вариант осуществления 387. Способ по варианту осуществления 382, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 490 и 493-494.

[2090] Вариант осуществления 388. Способ по варианту осуществления 382, где малая молекула представляет собой димер малой молекулы, выбранной из группы, состоящей из соединений, описанных в Варианте осуществления 163 Группы 1.

[2091] Вариант осуществления 389. Способ по варианту осуществления 382, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2092] Вариант осуществления 390. Способ по варианту осуществления 382, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2093] Вариант осуществления 391. Способ по варианту осуществления 382, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2094] Вариант осуществления 392. Способ по варианту осуществления 382, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2095] Вариант осуществления 393. Способ внесения пептидной метки или эпитопа в белок с помощью праймерного редактирования, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с конструктом праймерного редактора, выполненного с возможностью вносить в нее вторую нуклеотидную последовательность, кодирующую пептидную метку, для получения рекомбинантной нуклеотидной последовательности, так что пептидная метка и белок экспрессируются из рекомбинантной нуклеотидной последовательности в виде слитого белка.

[2096] Вариант осуществления 394. Способ по варианту осуществления 383, где пептидная метка применяется для очистки и/или детектирования белка.

[2097] Вариант осуществления 395. Способ по варианту осуществления 383, где пептидная метка представляет собой полигистидин (например HHHHHH) (SEQ ID NO: 252-262), FLAG (например DYKDDDDK) (SEQ ID NO: 2), V5 (например GKPIPNPLLGLDST) (SEQ ID NO: 3), GCN4, HA (например , YPYDVPDYA) (SEQ ID NO: 5), Myc (например EQKLISEED) (SEQ ID NO: 6), или GST.

[2098] Вариант осуществления 396. Способ по варианту осуществления 383, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 1-6, 245-249, 252-262, 264-273, 275-276, 281, 278-288, и 622.

[2099] Вариант осуществления 397. Способ по варианту осуществления 383, где пептидная метка слита с белком с помощью линкера.

[2100] Вариант осуществления 398. Способ по варианту осуществления 383, где слитый белок имеет следующую структуру: [белок]-[пептидная метка] или [пептидная метка]-[белок], где "]-[" обозначает необязательный линкер.

[2101] Вариант осуществления 399. Способ по варианту осуществления 383, где линкер имеет аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2102] Вариант осуществления 400. Способ по варианту осуществления 383, где конструкт праймерного редактора включает PEgRNA, включающую нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2103] Вариант осуществления 401. Способ по варианту осуществления 383, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[2104] Вариант осуществления 402. Способ по варианту осуществления 383, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[2105] Вариант осуществления 403. Способ предотвращения или остановки прогрессирования прионного заболевания путем внесения одной или нескольких защитных мутаций в PRNP, кодируемых нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей защитную мутацию; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует PRNP, включающий защитную мутацию и являющийся устойчивым к неправильному фолдингу.

[2106] Вариант осуществления 404. Способ по варианту осуществления 403, где прионное заболевание представляет собой прионное заболевание человека.

[2107] Вариант осуществления 405. Способ по варианту осуществления 403, где прионное заболевание представляет собой прионное заболевание животного.

[2108] Вариант осуществления 406. Способ по варианту осуществления 404, где прионное заболевание представляет собой болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, или куру.

[2109] Вариант осуществления 407. Способ по варианту осуществления 403, где прионное заболевание представляет собой коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных.

[2110] Вариант осуществления 408. Способ по варианту осуществления 403, где аминокислотная последовательность PRNP дикого типа представляет собой SEQ ID NO: 291-292.

[2111] Вариант осуществления 409. Способ по варианту осуществления 403, где способ приводит к получению модифицированной аминокислотной

последовательности PRNP, выбранной из группы, состоящей из SEQ ID NO: 293-309, и 311-323, при этом указанный модифицированный белок PRNP устойчив к неправильному фолдингу.

[2112] Вариант осуществления 410. Способ по варианту осуществления 403, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2113] Вариант осуществления 411. Способ по варианту осуществления 403, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2114] Вариант осуществления 412. Способ по варианту осуществления 403, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2115] Вариант осуществления 413. Способ по варианту осуществления 403, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2116] Вариант осуществления 414. Способ внесения рибонуклеотидного мотива или метки в представляющую интерес РНК, кодируемую нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую рибонуклеотидный мотив или метку; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей рибонуклеотидный мотив или метку; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует представляющую интерес модифицированную РНК, включающую рибонуклеотидный мотив или метку.

[2117] Вариант осуществления 415. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка представляет собой фрагмент для детектирования.

[2118] Вариант осуществления 416. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка влияет на уровень экспрессии представляющей интерес РНК.

[2119] Вариант осуществления 417. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка влияет на транспорт или субклеточную локализацию представляющей интерес РНК.

[2120] Вариант осуществления 418. Способ по варианту осуществления 414, где рибонуклеотидный мотив или метка выбрана из группы, состоящей из SV40 типа 1, SV40 типа 2, SV40 типа 3, hGH, BGH, rbGlob, TK, MALAT1 ENE-mascРНК, KSHV PAN ENE, Sm-бокса/бокса мяРНК U1, 3' бокса мяРНК U1, тРНК-лизина, аптамера broccoli, аптамера spinach, аптамера mango, рибозима HDV, и m6A.

[2121] Вариант осуществления 419. Способ по варианту осуществления 414, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2122] Вариант осуществления 420. Способ по варианту осуществления 414, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2123] Вариант осуществления 421. Способ по варианту осуществления 414, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2124] Вариант осуществления 422. Способ по варианту осуществления 414, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2125] Вариант осуществления 423. Способ внесения или удаления функционального фрагмента в представляющем интерес белке, кодируемом нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент или его удаление; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент или его удаление; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует модифицированный белок, включающий представляющий интерес белок и

функциональный фрагмент или его удаление, при этом функциональный фрагмент изменяет связанное с модификацией состояние или связанное с локализацией состояние белка.

[2126] Вариант осуществления 424. Способ по варианту осуществления 423, где функциональный фрагмент изменяет связанное с фосфорилированием, убиквитинилированием, гликозилированием, липидированием, гидроксилированием, метилированием, ацетилированием, кротонилированием, SUMOилированием состояние представляющего интерес белка.

[2127] Вариант осуществления 425. Слитый белок, включающий домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен, включающий активность РНК-зависимой ДНК-полимеразы.

[2128] Вариант осуществления 426. Слитый белок по варианту осуществления 425, где слитый белок способен осуществлять геномное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA) для внесения требуемого нуклеотидного изменения в последовательность-мишень.

[2129] Вариант осуществления 427. Слитый белок по варианту осуществления 425, где домен napDNAbp имеет никазную активность.

[2130] Вариант осуществления 428. Слитый белок по варианту осуществления 425, где домен napDNAbp представляет собой белок Cas9 или его вариант.

[2131] Вариант осуществления 429. Слитый белок по варианту осуществления 425, где домен napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2132] Вариант осуществления 430. Слитый белок по варианту осуществления 425, где домен napDNAbp представляет собой никазу Cas9 (nCas9).

[2133] Вариант осуществления 431. Слитый белок по варианту осуществления 425, где домен napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2134] Вариант осуществления 432. Слитый белок по варианту осуществления 425, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2135] Вариант осуществления 433. Слитый белок по варианту осуществления 425, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766, и необязательно где домен, включающий РНК-зависимую ДНК-полимеразу является подверженным ошибкам.

[2136] Вариант осуществления 434. Слитый белок по варианту осуществления 425, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

[2137] Вариант осуществления 435. Слитый белок по варианту осуществления 425, где слитый белок при образовании комплекса с PEgRNA способен связываться с последовательностью-мишенью ДНК.

[2138] Вариант осуществления 436. Слитый белок по варианту осуществления 435, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[2139] Вариант осуществления 437. Слитый белок по варианту осуществления 435, где связывание слитого белка в виде комплекса с PEgRNA образует R-петлю.

[2140] Вариант осуществления 438. Слитый белок по варианту осуществления 437, где R-петля включает (i) РНК-ДНК гибрид, включающий PEgRNA и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

[2141] Вариант осуществления 439. Слитый белок по варианту осуществления 437, где мишень или комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности, имеющей свободный 3′-конец.

[2142] Вариант осуществления 440. Слитый белок по варианту осуществления 439, где участок разрыва расположен выше последовательности PAM на цепи-мишени.

[2143] Вариант осуществления 441. Слитый белок по варианту осуществления 439, где участок разрыва расположен выше последовательности PAM на не являющейся мишенью цепи.

[2144] Вариант осуществления 442. Слитый белок по варианту осуществления 439, где участок разрыва расположен -1, -2, -3, -4, -5, -6, -7, -8, или -9 относительно 5'-конца последовательности PAM.

[2145] Вариант осуществления 443. Слитый белок по варианту осуществления 426, где PEgRNA включает направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты.

[2146] Вариант осуществления 444. Слитый белок по варианту осуществления 443, где плечо удлинение расположено на 5' или 3'-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[2147] Вариант осуществления 445. Слитый белок по варианту осуществления 443, где плечо удлинения включает (i) последовательность матрицы для синтеза ДНК, включающую требуемое нуклеотидное изменение, и (ii) участок связывания праймера.

[2148] Вариант осуществления 446. Слитый белок по варианту осуществления 445, где последовательность матрицы для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[2149] Вариант осуществления 447. Слитый белок по варианту осуществления 443, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 26 нуклеотидов, по меньшей мере, 27 нуклеотидов, по меньшей мере, 28 нуклеотидов, по меньшей мере, 29 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 31 нуклеотидов, по меньшей мере, 32 нуклеотидов, по меньшей мере, 33 нуклеотидов, по меньшей мере, 34 нуклеотидов, по меньшей мере, 35 нуклеотидов, по меньшей мере, 36 нуклеотидов, по меньшей мере, 37 нуклеотидов, по меньшей мере, 38 нуклеотидов, по меньшей мере, 39 нуклеотидов,

по меньшей мере, 40 нуклеотидов, по меньшей мере, 41 нуклеотидов, по меньшей мере, 42 нуклеотидов, по меньшей мере, 43 нуклеотидов, по меньшей мере, 44 нуклеотидов, по меньшей мере, 45 нуклеотидов, по меньшей мере, 46 нуклеотидов, по меньшей мере, 47 нуклеотидов, по меньшей мере, 48 нуклеотидов, по меньшей мере, 49 нуклеотидов, или, по меньшей мере, 50 нуклеотидов.

[2150] Вариант осуществления 448. Слитый белок по варианту осуществления 443, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения в цепь-мишень.

[2151] Вариант осуществления 449. Слитый белок по варианту осуществления 443, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5'-конец.

[2152] Вариант осуществления 450. Слитый белок по варианту осуществления 446, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

[2153] Вариант осуществления 451. Слитый белок по варианту осуществления 446, где последовательность эндогенной ДНК, имеющая 5'-конец, подвергается эксцизии флэп-эндонуклеазой.

[2154] Вариант осуществления 452. Слитый белок по варианту осуществления 448, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения в не являющуюся мишенью цепь, тем самым обеспечивая образование требуемого продукта.

[2155] Вариант осуществления 453. Слитый белок по варианту осуществления 449, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM.

[2156] Вариант осуществления 454. Слитый белок по варианту осуществления 449, где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81,

82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

[2157] Вариант осуществления 455. Слитый белок по варианту осуществления 425, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности SEQ ID NO: 18.

[2158] Вариант осуществления 456. Слитый белок по варианту осуществления 425, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2159] Вариант осуществления 457. Слитый белок по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-COOH; или NH_2-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-[napDNAbp]-COOH, где каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера.

[2160] Вариант осуществления 458. Слитый белок по варианту осуществления 457, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2161] Вариант осуществления 459. Слитый белок по варианту осуществления 425, где требуемое нуклеотидное изменение представляет собой изменение одного нуклеотида, вставку одного или нескольких нуклеотидов, или делецию одного или нескольких нуклеотидов.

[2162] Вариант осуществления 460. Слитый белок по варианту осуществления 459, где вставка или делеция имеет длину, по меньшей мере, 1, по меньшей мере, 2, по меньшей мере, 3, по меньшей мере, 4, по меньшей мере, 5, по меньшей мере, 6, по меньшей мере, 7, по меньшей мере, 8, по меньшей мере, 9, по меньшей мере, 10, по меньшей мере, 11, по меньшей мере, 12, по меньшей мере, 13, по меньшей мере, 14, по меньшей мере, 15, по меньшей мере, 16, по меньшей мере, 17, по меньшей мере, 18, по меньшей мере, 19, по меньшей мере, 20, по меньшей мере, 21, по меньшей мере, 22 , по меньшей мере, 23 , по меньшей мере, 24 , по меньшей мере, 25, по меньшей мере, 26,

по меньшей мере, 27, по меньшей мере, 28, по меньшей мере, 29, по меньшей мере, 30, по меньшей мере, 31, по меньшей мере, 32, по меньшей мере, 33, по меньшей мере, 34, по меньшей мере, 35, по меньшей мере, 36, по меньшей мере, 37, по меньшей мере, 38, по меньшей мере, 39, по меньшей мере, 40, по меньшей мере, 41, по меньшей мере, 42, по меньшей мере, 43, по меньшей мере, 44, по меньшей мере, 45, по меньшей мере, 46, по меньшей мере, 47, по меньшей мере, 48, по меньшей мере, 49, или, по меньшей мере, 50.

[2163] Вариант осуществления 461. Комплекс, включающий слитый белок по любому из вариантов осуществления 425-460 и PEgRNA, где PEgRNA направляет слитый белок к последовательности-мишени ДНК для праймерного редактирования.

[2164] Вариант осуществления 462. Комплекс по варианту осуществления 461, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[2165] Вариант осуществления 463. Комплекс по варианту осуществления 462, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[2166] Вариант осуществления 464. Комплекс по варианту осуществления 463, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь, при этом направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[2167] Вариант осуществления 465. Комплекс по варианту осуществления 464, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты включает (i) матрицу для синтеза ДНК и (ii) участок связывания праймера.

[2168] Вариант осуществления 466. Комплекс по варианту осуществления 464, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2169] Вариант осуществления 467. Комплекс по варианту осуществления 465, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[2170] Вариант осуществления 468. Комплекс по варианту осуществления 465, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[2171] Вариант осуществления 469. Комплекс по варианту осуществления 461, где napDNAbp представляет собой никазу Cas9.

[2172] Вариант осуществления 470. Комплекс по варианту осуществления 461, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно SEQ ID NO: 18.

[2173] Вариант осуществления 471. Комплекс по варианту осуществления 461, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2174] Вариант осуществления 472. Комплекс по варианту осуществления 461, где PEgRNA включает направляющую РНК и плечо удлинения нуклеиновой кислоты на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК.

[2175] Вариант осуществления 473. Комплекс по варианту осуществления 465, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[2176] Вариант осуществления 474. Комплекс по варианту осуществления 465, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[2177] Вариант осуществления 475. Комплекс по варианту осуществления 462, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

[2178] Вариант осуществления 476. Полинуклеотид, кодирующий слитый белок по любому одному из вариантов осуществления 425-461.

[2179] Вариант осуществления 477. Полинуклеотид, кодирующий PEgRNA по любому одному из указанных выше вариантов осуществления.

[2180] Вариант осуществления 478. Вектор, включающий полинуклеотид по варианту осуществления 476, где экспрессия слитого белка контролируется промотором.

[2181] Вариант осуществления 479. Вектор, включающий полинуклеотид по варианту осуществления 477, где экспрессия PEgRNA контролируется промотором.

[2182] Вариант осуществления 480. Вектор по варианту осуществления 479, где промотор представляет собой промотор U6.

[2183] Вариант осуществления 481. Вектор по варианту осуществления 479, где промотор представляет собой промотор CMV.

[2184] Вариант осуществления 482. Вектор по варианту осуществления 480, где PEgRNA встроена для удаления одного или нескольких повторяющихся кластеров в виде T в плече удлинения для повышения эффективности транскрипции на основе промотора U6.

[2185] Вариант осуществления 483. Вектор по варианту осуществления 482, где указанный один или несколько повторяющихся кластеров в виде T, которые удаляются, включают, по меньшей мере, 3 T, по меньшей мере, 4 T, по меньшей мере, 5 T, по меньшей мере, 6 T, по меньшей мере, 7 T, по меньшей мере, 8 T, по меньшей мере, 9 T, по меньшей мере, 10 T, по меньшей мере, 11 T, по меньшей мере, 12 T, по меньшей мере, 13 T, по меньшей мере, 14 T, по меньшей мере, 15 T, по меньшей мере, 16 T, по меньшей мере, 17 T, по меньшей мере, 18 T, по меньшей мере, 19 T, или, по меньшей мере, 20 T.

[2186] Вариант осуществления 484. Клетка, включающая слитый белок по любому из вариантов осуществления 425-460 и PEgRNA, связанную с napDNAbp слитого белка.

[2187] Вариант осуществления 485. Клетка, включающая комплекс по любому одному из вариантов осуществления 461-475.

[2188] Вариант осуществления 486. Фармацевтическая композиция, включающая: (i) слитый белок по любому из вариантов осуществления 425-460, комплекс по вариантам осуществления 461-475, полинуклеотид по вариантам

осуществления 476-477, или вектор по вариантам осуществления 478-483; и (ii) фармацевтически приемлемый наполнитель.

[2189] Вариант осуществления 487. Фармацевтическая композиция, включающая: (i) комплекс по вариантам осуществления 461-475, (ii) полимеразу, поступающую in trans; и (iii) фармацевтически приемлемый наполнитель.

[2190] Вариант осуществления 488. Набор для праймерного редактирования, включающий: (i) молекулу нуклеиновой кислоты, кодирующую слитый белок по любому одному из вариантов осуществления 425-460; и (ii) молекулу нуклеиновой кислоты, кодирующую PEgRNA, которая способна направлять слитый белок к участку-мишени ДНК, где молекула нуклеиновой кислоты (i) и (ii) может содержаться в одном конструкте ДНК или разных конструктах ДНК.

[2191] Вариант осуществления 489. Набор по варианту осуществления 488, где молекула нуклеиновой кислоты (i) дополнительно включает промотор, который обеспечивает экспрессию слитого белка.

[2192] Вариант осуществления 490. Набор по варианту осуществления 488, где молекула нуклеиновой кислоты (ii) дополнительно включает промотор, который обеспечивает экспрессию PEgRNA.

[2193] Вариант осуществления 491. Набор по варианту осуществления 490, где промотор представляет собой промотор U6.

[2194] Вариант осуществления 492. Набор по варианту осуществления 490, где промотор представляет собой промотор CMV.

[2195] Вариант осуществления 493. Слитый белок по варианту осуществления 457, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 174 (1x SGGS), 3888 (2x SGGS), 3889 (3x SGGS), 3890 (1x XTEN), 3891 (1x EAAAK), 3892 (2x EAAAK), и 3893 (3x EAAAK).

ГРУППА B. PE НАПРАВЛЯЮЩИЕ И СПОСОБЫ КОНСТРУИРОВАНИЯ

[2196] Вариант осуществления 1. PEgRNA, включающая направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты, включающее матрицу для синтеза ДНК.

[2197] Вариант осуществления 2. PEgRNA по варианту осуществления 1, где плечо удлинения нуклеиновой кислоты расположено на 3′- или 5′-конце направляющей РНК

или во внутримолекулярном положении в направляющей РНК и где плечо удлинения нуклеиновой кислоты представляет собой ДНК или РНК.

[2198] Вариант осуществления 3. PEgRNA по варианту осуществления 1, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[2199] Вариант осуществления 4. PEgRNA по варианту осуществления 3, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[2200] Вариант осуществления 5. PEgRNA по варианту осуществления 3, где направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[2201] Вариант осуществления 6. PEgRNA по варианту осуществления 1, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты дополнительно включает участок связывания праймера.

[2202] Вариант осуществления 7. PEgRNA по варианту осуществления 1, где плечо удлинения нуклеиновой кислоты имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 26 нуклеотидов, по меньшей мере, 27 нуклеотидов, по меньшей мере, 28 нуклеотидов, по меньшей мере, 29 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 31 нуклеотидов, по меньшей мере, 32 нуклеотидов, по меньшей мере, 33 нуклеотидов, по меньшей мере, 34 нуклеотидов, по меньшей мере, 35 нуклеотидов, по меньшей мере, 36 нуклеотидов, по меньшей мере, 37 нуклеотидов, по меньшей мере, 38 нуклеотидов, по меньшей мере, 39 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 41 нуклеотидов, по меньшей мере, 42 нуклеотидов, по меньшей мере, 43 нуклеотидов, по меньшей мере, 44 нуклеотидов, по меньшей мере, 45 нуклеотидов, по меньшей мере, 46 нуклеотидов, по

меньшей мере, 47 нуклеотидов, по меньшей мере, 48 нуклеотидов, по меньшей мере, 49 нуклеотидов, или, по меньшей мере, 50 нуклеотидов.

[2203] Вариант осуществления 8. PEgRNA по варианту осуществления 1, где матрица для синтеза ДНК имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[2204] Вариант осуществления 9. PEgRNA по варианту осуществления 6, где участок связывания праймера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[2205] Вариант осуществления 10. PEgRNA по варианту осуществления 1, дополнительно включающая, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из тРНК, линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования РНК-белка.

[2206] Вариант осуществления 11. PEgRNA по варианту осуществления 1, где матрица для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[2207] Вариант осуществления 12. PEgRNA по варианту осуществления 11, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

[2208] Вариант осуществления 13. PEgRNA по варианту осуществления 11, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

[2209] Вариант осуществления 14. PEgRNA по варианту осуществления 13, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[2210] Вариант осуществления 15. PEgRNA по варианту осуществления 1, включающая нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2211] Вариант осуществления 16. PEgRNA по варианту осуществления 1, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[2212] Вариант осуществления 17. PEgRNA по варианту осуществления 6, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[2213] Вариант осуществления 18. PEgRNA по варианту осуществления 10, где указанная, по меньшей мере, одна дополнительная структура расположена на 3'- или 5'-конце PEgRNA.

[2214] Вариант осуществления 19. PEgRNA по варианту осуществления 10, где линкер включает нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2215] Вариант осуществления 20. PEgRNA по варианту осуществления 10, где структура типа "стебель-петля" включает нуклеотидную последовательность, выбранную из описанных здесь структур типа "стебель-петля".

[2216] Вариант осуществления 21. PEgRNA по варианту осуществления 10, где шпилька включает нуклеотидную последовательность, выбранную из описанных здесь шпилек.

[2217] Вариант осуществления 22. PEgRNA по варианту осуществления 10, где опорная петля включает нуклеотидную последовательность, выбранную из описанных здесь опорных петель.

[2218] Вариант осуществления 23. PEgRNA по варианту осуществления 10, где аптамер включает нуклеотидную последовательность, выбранную из описанных здесь аптамеров.

[2219] Вариант осуществления 24. PEgRNA по варианту осуществления 10, где домен рекрутирования РНК-белка включает нуклеотидную последовательность, выбранную из описанных здесь доменов рекрутирования РНК-белка.

[2220] Вариант осуществления 25. Способ конструирования PEgRNA для применения в рамках праймерного редактирования для внесения требуемого нуклеотидного изменения в нуклеотидную последовательность-мишень, где указанная PEgRNA включает спейсер, ядро нРНК, плечо удлинения, и где указанное плечо удлинения включает участок связывания праймера и матрицу для синтеза ДНК, при этом указанный способ включает:

[2221] (i) выбор требуемого участка-мишени редактирования в нуклеотидной последовательности-мишени;

[2222] (ii) определение окружающей нуклеотидной последовательности, расположенной выше и ниже участка-мишени редактирования;

[2223] (iii) поиск предполагаемых участков смежного с протоспейсером мотива (PAM) в окружающей нуклеотидной последовательности, расположенных с рядом с требуемым участком-мишенью редактирования;

[2224] (iv) идентификацию соответствующих участков разрыва для каждого предполагаемого участка PAM;

[2225] (v) конструирование спейсера;

[2226] (vi) конструирование ядра нРНК;

[2227] (vii) конструирование плеча удлинения; и

[2228] (viii) конструирование полной PEgRNA путем конкатенирования спейсера, ядра нРНК, и плеча удлинения.

[2229] Вариант осуществления 26. Способ по варианту осуществления 25, где этап (i) выбора требуемого участка-мишени редактирования включает выбор вызывающей заболевание мутации.

[2230] Вариант осуществления 27. Способ по варианту осуществления 26, где вызывающая заболевание мутация ассоциирована с заболеванием, выбранным из группы, состоящей из: рака, аутоиммунных заболеваний, неврологических

заболеваний, кожных заболеваний. респираторных заболеваний, и сердечных заболеваний.

[2231] Вариант осуществления 28. Способ по варианту 25, где этап (ii) определения окружающей нуклеотидной последовательности, расположенной выше и ниже участка-мишени редактирования, включает определение приблизительно 50-55 пар оснований (bp), приблизительно 55-60 bp, приблизительно 60-65 bp, приблизительно 65-70 bp, приблизительно 70-75 bp, приблизительно 75-80 bp, приблизительно 80-85 bp, приблизительно 85-90 bp, приблизительно 90-95 bp, приблизительно 95-100 bp, приблизительно 100-105 bp, приблизительно 105-110 bp, приблизительно 110-125 bp, приблизительно 125-130 bp, приблизительно 130-135 bp, приблизительно 135-140 bp, приблизительно 140-145 bp, приблизительно 145-150 bp, приблизительно 150-155 bp, приблизительно 155-160 bp, приблизительно 160-165 bp, приблизительно 165-170 bp, приблизительно 170-175 bp, приблизительно 175-180 bp, приблизительно 180-185 bp, приблизительно 185-190 bp, приблизительно 190-195 bp, приблизительно 195-200 bp, приблизительно 200-205 bp, приблизительно 205-210 bp, приблизительно 210-215 bp, приблизительно 215-220 bp, приблизительно 220-225 bp, приблизительно 225-230 bp, приблизительно 230-235 bp, приблизительно 235-240 bp, приблизительно 240-245 bp, или приблизительно 245-250 bp участка, который включает требуемый участок-мишень редактирования.

[2232] Вариант осуществления 29. Способ по варианту осуществления 28, где требуемый участок-мишень редактирования расположен приблизительно на равном расстоянии от каждого конца окружающей нуклеотидной последовательности.

[2233] Вариант осуществления 30. Способ по варианту осуществления 25, где на этапе (iii) предполагаемые участки PAM расположены рядом с требуемым участком-мишенью редактирования.

[2234] Вариант осуществления 31. Способ по варианту осуществления 25, где на этапе (iii) предполагаемые участки PAM включают участки, связанные с участками разрыва, расположенные на расстоянии менее чем 30 нуклеотидов от участка-мишени редактирования или менее чем 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, или 2 нуклеотидов от участка-мишени редактирования.

[2235] Вариант осуществления 32. Способ по варианту осуществления 25, где на этапе (iii) предполагаемые участки PAM включают участки, связанные с участками разрыва, расположенные на расстоянии более чем 30 нуклеотидов от участка-мишени

редактирования или более чем 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84 ,85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов от участка-мишени редактирования.

[2236] Вариант осуществления 33. Способ по варианту осуществления 25, где на этапе (iii) предполагаемые участки PAM связаны с одним соответствующим napDNAbp, который связывается с указанными участками PAM.

[2237] Вариант осуществления 34. Способ по варианту осуществления 33, где предполагаемый участки PAM и их соответствующие napDNAb выбраны из любой из следующих групп: (a) SpCas9 согласно SEQ ID NO: 18-25 и 87-88 и NGG; (b) Sp VQR nCas9; и (c) NGAN.

[2238] Вариант осуществления 35. Способ по варианту осуществления 25, где этап (v) конструирования спейсера включает определение комплементарной нуклеотидной последовательности протоспейсерной последовательности, связанной с каждым предполагаемым PAM.

[2239] Вариант осуществления 36. Способ по варианту осуществления 25, где этап (vi) конструирования ядра нРНК в контексте каждого предполагаемого PAM включает выбор последовательности ядра нРНК, которая способна связываться с napDNAbp, который связан с каждым из указанных предполагаемых PAM.

[2240] Вариант осуществления 37. Способ по варианту осуществления 25, где этап (vii) конструирования плеча удлинения включает конструирование (a) матрицы для синтеза ДНК, включающей представляющее интерес изменение, и (b) участок связывания праймера.

[2241] Вариант осуществления 38. Способ по варианту осуществления 37, где конструирование участка связывания праймера включает (a) идентификацию ДНК-праймера на PAM-содержащей цепи нуклеотидной последовательности-мишени, где 3'-конец ДНК-праймера представляет собой первый нуклеотид выше участка разрыва, связанный с участком PAM, и (b) конструирование комплемента ДНК-праймера, где указанный комплемент образует участок связывания праймера.

[2242] Вариант осуществления 39. Способ по варианту осуществления 38, где участок связывания праймера имеет длину от 8 до 15 нуклеотидов.

[2243] Вариант осуществления 40. Способ по варианту осуществления 38, где участок связывания праймера имеет длину 12-13 нуклеотидов, если ДНК-праймер содержит около 40-60% содержимого в виде GC.

[2244] Вариант осуществления 41. Способ по варианту осуществления 38, где участок связывания праймера имеет длину 14-15 нуклеотидов, если ДНК-праймер содержит менее чем около 40% содержимого в виде GC.

[2245] Вариант осуществления 42. Способ по варианту осуществления 38, где участок связывания праймера имеет длину 8-11 нуклеотидов, если ДНК-праймер содержит более чем около 60% содержимого в виде GC.

[2246] Вариант осуществления 43. Способ праймерного редактирования, включающий контактирование последовательности-мишени ДНК с PEgRNA по любому из вариантов осуществления 1-24 и слитым белком праймерного редактора, включающим napDNAbp и домен, имеющий активность РНК-зависимой ДНК-полимеразы.

[2247] Вариант осуществления 44. Способ по варианту осуществления 43, где napDNAbp имеет никазную активность.

[2248] Вариант осуществления 45. Способ по варианту осуществления 43, где napDNAbp представляет собой белок Cas9 или его вариант.

[2249] Вариант осуществления 46. Способ по варианту осуществления 43, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2250] Вариант осуществления 47. Способ по варианту осуществления 43, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2251] Вариант осуществления 48. Способ по варианту осуществления 43, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2252] Вариант осуществления 49. Способ по варианту осуществления 43, где домен, включающий активность РНК-зависимой ДНК-полимеразы представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2253] Вариант осуществления 50. Способ по варианту осуществления 43, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу, включающую аминокислотную последовательность,

имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2254] Вариант осуществления 51. Способ по варианту осуществления 43, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

ГРУППА C. PE КОМПЛЕКСЫ

[2255] Вариант осуществления 1. Комплекс для праймерного редактирования, включающий:

[2256] (i) слитый белок, включающий домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен, включающий активность РНК-зависимой ДНК-полимеразы; и

[2257] (ii) направляющую РНК для праймерного редактирования (PEgRNA).

[2258] Вариант осуществления 2. Комплекс по варианту осуществления 1, где слитый белок способен осуществлять праймерное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA) для внесения требуемого нуклеотидного изменения в последовательность-мишень.

[2259] Вариант осуществления 3. Комплекс по варианту осуществления 11, где napDNAbp имеет никазную активность.

[2260] Вариант осуществления 4. Комплекс по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

[2261] Вариант осуществления 5. Комплекс по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2262] Вариант осуществления 6. Комплекс по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2263] Вариант осуществления 7. Комплекс по варианту осуществления 1, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2264] Вариант осуществления 8. Комплекс по варианту осуществления 1, где домен, включающий активность РНК-зависимой ДНК-полимеразы представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2265] Вариант осуществления 9. Комплекс по варианту осуществления 1, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2266] Вариант осуществления 10. Комплекс по варианту осуществления 1, где домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

[2267] Вариант осуществления 11. Комплекс по варианту осуществления 1, где слитый белок при образовании комплекса с PEgRNA способен связываться с последовательностью-мишенью ДНК.

[2268] Вариант осуществления 12. Комплекс по варианту осуществления 1, где PEgRNA включает направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты, включающее матрицу для синтеза ДНК.

[2269] Вариант осуществления 13. Комплекс по варианту осуществления 12, где плечо удлинения нуклеиновой кислоты расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК и где плечо удлинения нуклеиновой кислоты представляет собой ДНК или РНК.

[2270] Вариант осуществления 14. Комплекс по варианту осуществления 12, где PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

[2271] Вариант осуществления 15. Комплекс по варианту осуществления 14, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[2272] Вариант осуществления 16. Комплекс по варианту осуществления 12, где направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

[2273] Вариант осуществления 17. Комплекс по варианту осуществления 12, где указанное, по меньшей мере, одно плечо удлинения нуклеиновой кислоты дополнительно включает участок связывания праймера.

[2274] Вариант осуществления 18. Комплекс по варианту осуществления 1, где плечо удлинения нуклеиновой кислоты имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 26 нуклеотидов, по меньшей мере, 27 нуклеотидов, по меньшей мере, 28 нуклеотидов, по меньшей мере, 29 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 31 нуклеотидов, по меньшей мере, 32 нуклеотидов, по меньшей мере, 33 нуклеотидов, по меньшей мере, 34 нуклеотидов, по меньшей мере, 35 нуклеотидов, по меньшей мере, 36 нуклеотидов, по меньшей мере, 37 нуклеотидов, по меньшей мере, 38 нуклеотидов, по меньшей мере, 39 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 41 нуклеотидов, по меньшей мере, 42 нуклеотидов, по меньшей мере, 43 нуклеотидов, по меньшей мере, 44 нуклеотидов, по меньшей мере, 45 нуклеотидов, по меньшей мере, 46 нуклеотидов, по меньшей мере, 47 нуклеотидов, по меньшей мере, 48 нуклеотидов, по меньшей мере, 49 нуклеотидов, или, по меньшей мере, 50 нуклеотидов.

[2275] Вариант осуществления 19. Комплекс по варианту осуществления 12, где матрица для синтеза ДНК имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по

меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[2276] Вариант осуществления 20. Комплекс по варианту осуществления 17, где участок связывания праймера имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

[2277] Вариант осуществления 21. Комплекс по варианту осуществления 12, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

[2278] Вариант осуществления 22. Комплекс по варианту осуществления 12, где матрица для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

[2279] Вариант осуществления 23. Комплекс по варианту осуществления 22, где одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и где эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

[2280] Вариант осуществления 24. Комплекс по варианту осуществления 23, где эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

[2281] Вариант осуществления 25. Комплекс по варианту осуществления 23, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

[2282] Вариант осуществления 26. Комплекс по варианту осуществления 12, где PEgRNA включает нуклеотидную последовательность SEQ ID NO: 89-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по

меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2283] Вариант осуществления 27. Комплекс по варианту осуществления 12, где матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

[2284] Вариант осуществления 28. Комплекс по варианту осуществления 17, где участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

[2285] Вариант осуществления 29. Комплекс по варианту осуществления 21, где указанная, по меньшей мере, одна дополнительная структура расположена на 3'- или 5'-конце PEgRNA.

[2286] Вариант осуществления 30. Комплекс по варианту осуществления 29, где линкер включает нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2287] Вариант осуществления 31. Комплекс по варианту осуществления 29, где структура типа "стебель-петля" включает нуклеотидную последовательность, выбранную из описанных здесь структур типа "стебель-петля".

[2288] Вариант осуществления 32. Комплекс по варианту осуществления 29, где шпилька включает нуклеотидную последовательность, выбранную из описанных здесь шпилек.

[2289] Вариант осуществления 33. Комплекс по варианту осуществления 29, где опорная петля включает нуклеотидную последовательность, выбранную из описанных здесь опорных петель.

[2290] Вариант осуществления 34. Комплекс по варианту осуществления 29, где аптамер включает нуклеотидную последовательность, выбранную из описанных здесь аптамеров.

[2291] Вариант осуществления 35. Комплекс по варианту осуществления 29, где домен рекрутирования РНК-белка включает нуклеотидную последовательность, выбранную из описанных здесь доменов рекрутирования РНК-белка.

[2292] Вариант осуществления 36. Комплекс по варианту осуществления 1, где последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

[2293] Вариант осуществления 37. Комплекс по варианту осуществления 36, где R-петля включает (i) РНК-ДНК гибрид, включающий PEgRNA и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

[2294] Вариант осуществления 38. Комплекс по варианту осуществления 37, где мишень или комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности, имеющей свободный 3′-конец.

[2295] Вариант осуществления 39. Комплекс по варианту осуществления 38, где участок разрыва расположен выше последовательности PAM на цепи-мишени.

[2296] Вариант осуществления 40. Комплекс по варианту осуществления 38, где участок разрыва расположен выше последовательности PAM на не являющейся мишенью цепи.

[2297] Вариант осуществления 41. Комплекс по варианту осуществления 38, где участок разрыва расположен -1, -2, -3, -4, -5, -6, -7, -8, или -9 относительно 5'-конца последовательности PAM.

[2298] Вариант осуществления 42. Комплекс по варианту осуществления 22, где одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения в цепь-мишень.

[2299] Вариант осуществления 43. Комплекс по варианту осуществления 22, где одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5'-конец.

[2300] Вариант осуществления 44. Комплекс по варианту осуществления 22, где последовательность эндогенной ДНК, имеющая 5′-конец, подвергается эксцизии клеткой.

[2301] Вариант осуществления 45. Комплекс по варианту осуществления 44, где последовательность эндогенной ДНК, имеющая 5'-конец, подвергается эксцизии флэп-эндонуклеазой.

[2302] Вариант осуществления 46. Комплекс по варианту осуществления 43, где клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого

нуклеотидного изменения в не являющуюся мишенью цепь, тем самым обеспечивая образование требуемого продукта.

[2303] Вариант осуществления 47. Комплекс по варианту осуществления 46, где требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM.

[2304] Вариант осуществления 48. Комплекс по варианту осуществления 449, где требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

[2305] Вариант осуществления 49. Комплекс по любому одному из предыдущих вариантов осуществления, где слитый белок включает структуру NH_2-[napDNAbp]-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-COOH; или NH_2-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-[napDNAbp]-COOH, где каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера.

[2306] Вариант осуществления 50. Комплекс по варианту осуществления 49, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2307] Вариант осуществления 51. Комплекс по варианту осуществления 1, где слитый белок дополнительное включает линкер, который соединяет napDNAbp и домен, включающий активность РНК-зависимой ДНК-полимеразы.

[2308] Вариант осуществления 52. Комплекс по варианту осуществления 51, где последовательность линкера включает аминокислотную последовательность SEQ ID NO: 3887 (1x SGGS), 3888 (2x SGGS), 3889 (3x SGGS), 3890 (1x XTEN), 3891 (1x EAAAK), 3892 (2x EAAAK), и 3893 (3x EAAAK).

ГРУППА D. СПОСОБ РЕ ДЛЯ КОРРЕКТИРОВАНИЯ МУТАЦИЙ

[2309] Вариант осуществления 1. Способ внесения требуемого нуклеотидного изменения в двуцепочечную последовательности ДНК, при этом способ включает контактирование последовательности двуцепочечной ДНК с комплексом, включающим слитый белок и PEgRNA, при этом слитый белок включает napDNAbp и полимеразу, и при этом PEgRNA включает матрицу для синтеза ДНК, включающую требуемое нуклеотидное изменение и участок связывания праймера;

[2310] тем самым обеспечивая никирование последовательности двуцепочечной ДНК, тем самым обеспечивая создание свободной одноцепочечной ДНК, имеющей 3′-конец;

[2311] тем самым обеспечивая гибридизирование 3′-конца свободной одноцепочечной ДНК с участком связывания праймера, тем самым обеспечивая праймирование полимеразы;

[2312] тем самым обеспечивая полимеризацию цепи ДНК с 3′-конца, гибридизированного с участком связывания праймера, тем самым обеспечивая создание одноцепочечного ДНК-флэпа, включающего требуемое нуклеотидное изменение и являющегося комплементарным матрице для синтеза ДНК;

[2313] тем самым обеспечивая замещение цепи эндогенной ДНК, примыкающей к участку разрыва, одноцепочечным ДНК-флэпом, тем самым обеспечивая внесение требуемого нуклеотидного изменения в последовательность двуцепочечной ДНК.

[2314] Вариант осуществления 2. Способ по варианту замещения 1, где замещение цепи эндогенной ДНК включает: (i) гибридизацию одноцепочечного ДНК-флэпа с цепью эндогенной ДНК, примыкающей к участку разрыва, для создания несовпадения последовательности; (ii) эксцизию цепи эндогенной ДНК; и (iii) репарацию несовпадения для образования требуемого продукта, включающего требуемое нуклеотидное изменение в обеих цепях ДНК.

[2315] Вариант осуществления 3. Способ по варианту осуществления 1, где требуемое нуклеотидное изменение представляет собой замещение одного нуклеотида, делецию, или вставку.

[2316] Вариант осуществления 4. Способ по варианту осуществления 3, где замещение одного нуклеотида представляет собой транзицию или трансверсию.

[2317] Вариант осуществления 5. Способ по варианту осуществления 1, где требуемое нуклеотидное изменение представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T

на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[2318] Вариант осуществления 6. Способ по варианту осуществления 1, где требуемое нуклеотидное изменение обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2319] Вариант осуществления 7. Способ по варианту осуществления 1, где требуемое нуклеотидное изменение представляет собой вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[2320] Вариант осуществления 8. Способ по варианту осуществления 1, где требуемое нуклеотидное изменение обеспечивает корректирование связанного с заболеванием гена.

[2321] Вариант осуществления 9. Способ по варианту осуществления 8, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

[2322] Вариант осуществления 10. Способ по варианту осуществления 8, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[2323] Вариант осуществления 11. Способ по варианту осуществления 1, где napDNAbp представляет собой не имеющий нуклеазной активности Cas9 (dCas9), никазу Cas9 (nCas9), или имеющий нуклеазную активность Cas9.

[2324] Вариант осуществления 12. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[2325] Вариант осуществления 13. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2326] Вариант осуществления 14. Способ по варианту осуществления 1, где полимераза представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739- 741, и 766.

[2327] Вариант осуществления 15. Способ по варианту осуществления 1, где полимераза представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739- 741, и 766.

[2328] Вариант осуществления 16. Способ по варианту осуществления 1, где PEgRNA включает плечо удлинения нуклеиновой кислоты на 3'- или 5'-концах или во внутримолекулярном положении в направляющей РНК, при этом плечо удлинения включает последовательность матрицы для синтеза ДНК и участок связывания праймера.

[2329] Вариант осуществления 17. Способ по варианту осуществления 16, где плечо удлинения имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, или, по меньшей мере, 25 нуклеотидов.

[2330] Вариант осуществления 18. Способ по варианту осуществления 1, где PEgRNA имеет нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2331] Вариант осуществления 19. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени, включающий контактирование молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и PEgRNA, которая нацеливает napDNAbp на локус-мишень, при этом PEgRNA включает последовательность матрицы в виде обратной транскриптазы (RT), включающую, по меньшей мере, одно требуемое нуклеотидное изменение и участок связывания праймера;

[2332] тем самым обеспечивая образование свободного 3′-конца в цепи ДНК в локусе-мишени;

[2333] тем самым обеспечивая гибридизацию свободного 3′-конца с участком связывания праймера для праймирования обратной транскрипции;

[2334] тем самым обеспечивая синтезирование одноцепочечного ДНК-флэпа, включающего указанное, по меньшей мере, одно требуемое нуклеотидное изменение, на основе последовательности RT матрицы с помощью обратной транскриптазы;

[2335] тем самым обеспечивая встраивание указанного, по меньшей мере, одного требуемого нуклеотидного изменения в соответствующую эндогенную ДНК, тем самым обеспечивая внесение одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени.

[2336] Вариант осуществления 20. Способ по варианту осуществления 19, где указанное одно или несколько изменений в нуклеотидной последовательности включают транзицию.

[2337] Вариант осуществления 21. Способ по варианту осуществления 19, где транзиция выбрана из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[2338] Вариант осуществления 22. Способ по варианту осуществления 19, где указанное одно или несколько изменений в нуклеотидной последовательности включают трансверсию.

[2339] Вариант осуществления 23. Способ по варианту осуществления 22, где трансверсия выбрана из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[2340] Вариант осуществления 24. Способ по варианту осуществления 19, где указанное одно или несколько изменений в нуклеотидной последовательности включают изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (9) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2341] Вариант осуществления 25. Способ по варианту осуществления 19, где указанное одно или несколько изменений в нуклеотидной последовательности включают вставку или делецию 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, или 25 нуклеотидов.

[2342] Вариант осуществления 26. Способ по варианту осуществления 19, где указанное одно или несколько изменений в нуклеотидной последовательности включают корректирование связанного с заболеванием гена.

[2343] Вариант осуществления 27. Способ по варианту осуществления 26, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

[2344] Вариант осуществления 28. Способ по варианту осуществления 26, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[2345] Вариант осуществления 29. Способ по варианту осуществления 19, где napDNAbp представляет собой имеющий нуклеазную активность Cas9 или его вариант.

[2346] Вариант осуществления 30. Способ по варианту осуществления 19, где napDNAbp представляет собой не имеющий нуклеазную активность Cas9 (dCas9), или никазу Cas9 (nCas9), или их вариант.

[2347] Вариант осуществления 31. Способ по варианту осуществления 19, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[2348] Вариант осуществления 32. Способ по варианту осуществления 19, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2349] Вариант осуществления 33. Способ по варианту осуществления 19, где обратная транскриптаза вносится in trans.

[2350] Вариант осуществления 34. Способ по варианту осуществления 19, где napDNAbp включает продукт слияния с обратной транскриптазой.

[2351] Вариант осуществления 35. Способ по варианту осуществления 19, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2352] Вариант осуществления 36. Способ по варианту осуществления 19, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2353] Вариант осуществления 37. Способ по варианту осуществления 19, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает никирование цепи ДНК нуклеазой.

[2354] Вариант осуществления 38. Способ по варианту осуществления 37, где нуклеаза представляет собой napDNAbp и поступает в виде слитого домена napDNAbp или поступает in trans.

[2355] Вариант осуществления 39. Способ по варианту осуществления 19, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает контактирование цепи ДНК с химическим агентом.

[2356] Вариант осуществления 40. Способ по варианту осуществления 19, где этап образования свободного 3'-конца в цепи ДНК в локусе-мишени включает внесение ошибки репликации.

[2357] Вариант осуществления 41. Способ по варианту осуществления 19, где этап контактирования молекулы ДНК с napDNAbp и направляющей РНК образует R-петлю.

[2358] Вариант осуществления 42. Способ по варианту осуществления 41, где цепь ДНК, в которой образован свободный 3'-конец, расположена в R-петле.

[2359] Вариант осуществления 43. Способ по варианту осуществления 19, где PEgRNA включает плечо удлинения, которое включает последовательность матрицы в виде обратной транскриптазы (RT) и участок связывания праймера.

[2360] Вариант осуществления 44. Способ по варианту осуществления 43, где плечо удлинения расположено на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[2361] Вариант осуществления 45. Способ по варианту осуществления 19, где PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования РНК-белка.

[2362] Вариант осуществления 46. Способ по варианту осуществления 19, где PEgRNA дополнительно включает плечо гомологии.

[2363] Вариант осуществления 47. Способ по варианту осуществления 19, где последовательность RT матрицы является гомологичной относительно соответствующей эндогенной ДНК.

[2364] Вариант осуществления 48. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение;

[2365] тем самым обеспечивая осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение;

[2366] тем самым обеспечивая встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени за счет процесса репарации и/или репликации ДНК.

[2367] Вариант осуществления 49. Способ по варианту осуществления 48, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[2368] Вариант осуществления 50. Способ по варианту осуществления 48, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

[2369] Вариант осуществления 51. Способ по варианту осуществления 48, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[2370] Вариант осуществления 52. Способ по варианту осуществления 48, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[2371] Вариант осуществления 53. Способ по варианту осуществления 48, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2372] Вариант осуществления 54. Способ замещения мутации в виде экспансии тринуклеотидных повторов в молекуле-мишени ДНК здоровой последовательностью, включающей здоровое количество тринуклеотидных повторов, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК,

включающую замещающую последовательность и участок связывания праймера; (b) осуществление праймерного редактирования для получения одноцепочечной ДНК, включающей замещающую последовательность; и (c) встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[2373] Вариант осуществления 55. Способ по варианту осуществления 54, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2374] Вариант осуществления 56. Способ по варианту осуществления 54, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2375] Вариант осуществления 57. Способ по варианту осуществления 54, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2376] Вариант осуществления 58. Способ по варианту осуществления 54, где направляющая РНК включает SEQ ID NO: 222.

[2377] Вариант осуществления 59. Способ по варианту осуществления 54, где этап (b) осуществления праймерного редактирования включает получение 3'-концевой последовательности связывания праймера в локусе-мишени, которая способна праймировать полимеразу за счет отжига с участком связывания праймера на направляющей РНК.

[2378] Вариант осуществления 60. Способ по варианту осуществления 54, где мутация в виде экспансии тринуклеотидных повторов ассоциирована с болезнью Хантингтона, синдромом ломкой Х-хромосомы, или атаксией Фридрейха.

[2379] Вариант осуществления 61. Способ по варианту осуществления 54, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов CAG.

[2380] Вариант осуществления 62. Способ по варианту осуществления 54, где мутация в виде экспансии тринуклеотидных повторов включает повторяющуюся единицу из триплетов GAA.

[2381] Вариант осуществления 63. Способ внесения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в локусе-мишени с помощью мишень-инициированной обратной транскрипции, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и

обратную транскриптазу, и (ii) направляющей РНК, включающей RT матрицу, включающую требуемое нуклеотидное изменение;

[2382] тем самым обеспечивая осуществление мишень-инициированной обратной транскрипции RT матрицы для получения одноцепочечной ДНК, включающей требуемое нуклеотидное изменение;

[2383] тем самым обеспечивая встраивание требуемого нуклеотидного изменения в молекулу ДНК в локусе-мишени за счет процесса репарации и/или репликации ДНК.

[2384] Вариант осуществления 64. Способ по варианту осуществления 63, где RT матрица расположена на 3'-конце направляющей РНК, 5'-конце направляющей РНК, или во внутримолекулярном положении в направляющей РНК.

[2385] Вариант осуществления 65. Способ по варианту осуществления 63, где требуемое нуклеотидное изменение включает транзицию, трансверсию, вставку, или делецию, или любую их комбинацию.

[2386] Вариант осуществления 66. Способ по варианту осуществления 63, где требуемое нуклеотидное изменение включает транзицию, выбранную из группы, состоящей из: (a) T на C; (b) A на G; (c) C на T; и (d) G на A.

[2387] Вариант осуществления 67. Способ по варианту осуществления 63, где требуемое нуклеотидное изменение включает трансверсию, выбранную из группы, состоящей из: (a) T на A; (b) T на G; (c) C на G; (d) C на A; (e) A на T; (f) A на C; (g) G на C; и (h) G на T.

[2388] Вариант осуществления 68. Способ по варианту осуществления 63, где требуемое нуклеотидное изменение включает изменение (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2389] Вариант осуществления 69. Способ предотвращения или остановки прогрессирования прионного заболевания путем внесения одной или нескольких защитных мутаций в PRNP, кодируемых нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a)

контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент;

[2390] тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, кодирующей защитную мутацию;

[2391] тем самым обеспечивая встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени за счет процесса репарации и/или репликации ДНК;

[2392] где способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует PRNP, включающий защитную мутацию и являющийся устойчивым к неправильному фолдингу.

[2393] Вариант осуществления 70. Способ по варианту осуществления 69, где прионное заболевание представляет собой прионное заболевание человека.

[2394] Вариант осуществления 71. Способ по варианту осуществления 69, где прионное заболевание представляет собой прионное заболевание животного.

[2395] Вариант осуществления 72. Способ по варианту осуществления 69, где прионное заболевание представляет собой болезнь Крейтцфельдта-Якоба (CJD), вариант болезни Крейтцфельдта-Якоба, синдром Герстмана-Штрауслера-Шейнкера, фатальную семейную бессонницу, или куру.

[2396] Вариант осуществления 73. Способ по варианту осуществления 69, где прионное заболевание представляет собой коровью губчатую энцефалопатию (BSE или "коровье бешенство"), хроническое истощающее заболевание (CWD), скрапи, трансмиссивную энцефалопатию норок, губчатую энцефалопатию кошек, и губчатую энцефалопатию копытных.

[2397] Вариант осуществления 74. Способ по варианту осуществления 69, где аминокислотная последовательность PRNP дикого типа представляет собой SEQ ID NO: 291-292.

[2398] Вариант осуществления 75. Способ по варианту осуществления 69, где способ приводит к получению модифицированной аминокислотной последовательности PRNP, выбранной из группы, состоящей из SEQ ID NO: 293-309, и 311-323, при этом указанный модифицированный белок PRNP устойчив к неправильному фолдингу.

[2399] Вариант осуществления 76. Способ по варианту осуществления 69, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2400] Вариант осуществления 77. Способ по варианту осуществления 69, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2401] Вариант осуществления 78. Способ по варианту осуществления 69, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2402] Вариант осуществления 79. Способ по варианту осуществления 69, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2403] Вариант осуществления 80. Способ лечения связанного с дефицитом CDKL5 заболевания путем корректирования мутации в циклин-зависимом киназоподобном 5 гене (CDKL5) в нуклеотидной последовательности-мишени с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, которая корректирует мутацию в CDKL5;

[2404] тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, кодирующей изменение;

[2405] тем самым обеспечивая встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени за счет процесса репарации и/или репликации ДНК;

[2406] где способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует скорректированный ген CDKL5.

[2407] Вариант осуществления 81. Способ по варианту осуществления 80, где мутация в CDKL5 представляет собой 1412delA.

ГРУППА E. PE СПОСОБЫ МОДИФИЦИРОВАНИЯ СТРУКТУРЫ БЕЛКА/ФУНКЦИИ И/ИЛИ МУТАГЕНЕЗА

[2408] Вариант осуществления 1. Способ мутагенеза молекулы ДНК в локусе-мишени с помощью праймерного редактирования, при этом способ включает: (a) контактирование молекулы ДНК в локусе-мишени с (i) слитым белком, включающим

программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и подверженный ошибкам полимер (например подверженную ошибкам обратную транскриптазу), и (ii) направляющей РНК, включающей матрицу для редактирования, включающую требуемое нуклеотидное изменение; тем самым обеспечивая полимеризацию одноцепочечной ДНК, матрицей для которой служит матрица для редактирования; и встраивание одноцепочечной ДНК в молекулу ДНК в локусе-мишени с помощью процесса репарации и/или репликации ДНК.

[2409] Вариант осуществления 2. Способ по любому предыдущему варианту осуществления, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2410] Вариант осуществления 3. Способ по любому предыдущему варианту осуществления, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2411] Вариант осуществления 4. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2412] Вариант осуществления 5. Способ по варианту осуществления 1, где направляющая РНК включает SEQ ID NO: 222.

[2413] Вариант осуществления 6. Способ внесения иммуноэпитопа в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент;

[2414] тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп;

[2415] тем самым обеспечивая встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени за счет процесса репарации и/или репликации ДНК;

[2416] где способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и иммуноэпитоп.

[2417] Вариант осуществления 7. Способ по варианту осуществления 6, где иммуноэпитоп выбран из группы, состоящей из: столбнячного токсоида (SEQ ID NO: 396); мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630); иммуноэпитопа паротита 1 (SEQ ID NO: 400); иммуноэпитопа паротита 3 (SEQ ID NO: 402); иммуноэпитопа паротита 3 (SEQ ID NO: 404); вируса рубеллы (SEQ ID NO: 406); гемагглутинина (SEQ ID NO: 408); нейраминидазы (SEQ ID NO: 410); TAP1 (SEQ ID NO: 412); TAP2 (SEQ ID NO: 414); эпитопов гемагглутинина к HLA класса I (SEQ ID NO: 416); эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418); эпитопов гемагглутинина к HLA класса II (SEQ ID NO: 420); эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422); HLA класса I и класса II, связанного с эпитопом гемагглутинина H5N1 (SEQ ID NO: 424); HLA класса I и класса II, связанного с эпитопом нейраминидазы H5N1 (SEQ ID NO: 426).

[2418] Вариант осуществления 8. Способ по варианту осуществления 6, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2419] Вариант осуществления 9. Способ по варианту осуществления 6, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2420] Вариант осуществления 10. Способ по варианту осуществления 6, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2421] Вариант осуществления 11. Способ по варианту осуществления 6, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2422] Вариант осуществления 12. Способ внесения домена димеризации на основе малой молекулы в представляющий интерес белок, кодируемый нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую доме димеризации на основе малой молекулы;

[2423] тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, кодирующей иммуноэпитоп;

[2424] тем самым обеспечивая встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени за счет процесса репарации и/или репликации ДНК;

[2425] где способ обеспечивает получение модифицированной нуклеотидной последовательности-мишени, которая кодирует слитый белок, включающий представляющий интерес белок и домен димеризации на основе малой молекулы.

[2426] Вариант осуществления 13. Способ по варианту осуществления 12, дополнительно включающий осуществление способа в отношении второго представляющего интерес белка.

[2427] Вариант осуществления 14. Способ по варианту осуществления 13, где первый представляющий интерес белок и второй представляющий интерес белок подвергаются димеризации в присутствии малой молекулы, которая связывается с доменом димеризации на каждом из указанных белков.

[2428] Вариант осуществления 15. Способ по варианту осуществления 12, где домен связывания малой молекулы представляет собой FKBP12 согласно SEQ ID NO: 488.

[2429] Вариант осуществления 16. Способ по варианту осуществления 12, где домен связывания малой молекулы представляет собой FKBP12-F36V согласно SEQ ID NO: 489.

[2430] Вариант осуществления 17. Способ по варианту осуществления 12, где домен связывания малой молекулы представляет собой циклофилин согласно SEQ ID NO: 490 и 493-494.

[2431] Вариант осуществления 18. Способ по варианту осуществления 12, где малая молекула представляет собой димер малой молекулы, как описано здесь.

[2432] Вариант осуществления 19. Способ по варианту осуществления 12, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2433] Вариант осуществления 20. Способ по варианту осуществления 12, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2434] Вариант осуществления 21. Способ по варианту осуществления 12, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2435] Вариант осуществления 22. Способ по варианту осуществления 12, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2436] Вариант осуществления 23. Способ внесения пептидной метки или эпитопа в белок с помощью праймерного редактирования, включающий: контактирование нуклеотидной последовательности-мишени, кодирующей белок, с конструктом праймерного редактора, выполненного с возможностью вносить в нее вторую нуклеотидную последовательность, кодирующую пептидную метку, для получения рекомбинантной нуклеотидной последовательности, так что пептидная метка и белок экспрессируются из рекомбинантной нуклеотидной последовательности в виде слитого белка.

[2437] Вариант осуществления 24. Способ по варианту осуществления 23, где пептидная метка применяется для очистки и/или детектирования белка.

[2438] Вариант осуществления 25. Способ по варианту осуществления 23, где пептидная метка представляет собой полигистидин (например HHHHHH) (SEQ ID NO: 252-262), FLAG (например DYKDDDDK) (SEQ ID NO: 2), V5 (например GKPIPNPLLGLDST) (SEQ ID NO: 3), GCN4, HA (например , YPYDVPDYA) (SEQ ID NO: 5), Myc (например EQKLISEED) (SEQ ID NO: 6), или GST.

[2439] Вариант осуществления 26. Способ по варианту осуществления 23, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 1-6, 245-249, 252-262, 264-273, 275-276, 281, 278-288, и 622.

[2440] Вариант осуществления 27. Способ по варианту осуществления 23, где пептидная метка слита с белком с помощью линкера.

[2441] Вариант осуществления 28. Способ по варианту осуществления 23, где слитый белок имеет следующую структуру: [белок]-[пептидная метка] или [пептидная метка]-[белок], где "]-[" обозначает необязательный линкер.

[2442] Вариант осуществления 29. Способ по варианту осуществления 23, где линкер имеет аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

[2443] Вариант осуществления 30. Способ по варианту осуществления 23, где конструкт праймерного редактора включает PEgRNA, включающую нуклеотидную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2444] Вариант осуществления 31. Способ по варианту осуществления 23, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен

нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[2445] Вариант осуществления 32. Способ по варианту осуществления 23, где PEgRNA включает спейсер, ядро нРНК, и плечо удлинения, при этом спейсер комплементарен нуклеотидной последовательности-мишени, а плечо удлинения включает матрицу в виде обратной транскриптазы, которая кодирует пептидную метку.

[2446] Вариант осуществления 33. Способ внесения или удаления функционального фрагмента в представляющем интерес белке, кодируемом нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую функциональный фрагмент или его удаление; (b) полимеризацию одноцепочечной последовательности ДНК, кодирующей функциональный фрагмент или его удаление; и (c) встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение рекомбинантной нуклеотидной последовательности-мишени, которая кодирует модифицированный белок, включающий представляющий интерес белок и функциональный фрагмент или его удаление, при этом функциональный фрагмент изменяет связанное с модификацией состояние или связанное с локализацией состояние белка.

[2447] Вариант осуществления 34. Способ по варианту осуществления 33, где функциональный фрагмент изменяет связанное с фосфорилированием, убиквитинилированием, гликозилированием, липидированием, гидроксилированием, метилированием, ацетилированием, кротонилированием, или SUMOилированием состояние представляющего интерес белка.

ГРУППА F. СПОСОБЫ ДОСТАВКИ PE И КОМПОЗИЦИИ

[2448] Вариант осуществления 1. Полипептид, включающий N-концевую половину или C-концевую половину слитого белка праймерного редактора.

[2449] Вариант осуществления 2. Полипептид по варианту осуществления 1, где слитый белок праймерного редактора включает домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен полимеразы.

[2450] Вариант осуществления 3. Полипептид по варианту осуществления 1, где слитый белок праймерного редактора способен осуществлять геномное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA).

[2451] Вариант осуществления 4. Полипептид по варианту осуществления 2, где napDNAbp представляет собой белок Cas9 или его вариант.

[2452] Вариант осуществления 5. Полипептид по варианту осуществления 2, где napDNAbp представляет собой нуклеазу с никазной активностью.

[2453] Вариант осуществления 6. Полипептид по варианту осуществления 2, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2454] Вариант осуществления 7. Полипептид по варианту осуществления 2, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2455] Вариант осуществления 8. Полипептид по варианту осуществления 1, где полипептид образован путем разделения слитого белка праймерного редактора в участке разделения.

[2456] Вариант осуществления 9. Полипептид по варианту осуществления 8, где участок разделения представляет собой пептидную связь в домене napDNAbp.

[2457] Вариант осуществления 10. Полипептид по варианту осуществления 8, где участок разделения представляет собой пептидную связь в домене полимеразы.

[2458] Вариант осуществления 11. Полипептид по варианту осуществления 8, где участок разделения представляет собой пептидную связь в линкере между доменом napDNAbp и доменом полимеразы.

[2459] Вариант осуществления 12. Полипептид по варианту осуществления 9, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и 13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18 и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 ad 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и

49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 SEQ ID NO: 18 (стандартный SpCas9), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента SpCas9 согласно SEQ ID NO: 18.

[2460] Вариант осуществления 13. Полипептид по варианту осуществления 9, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и 13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18 и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 и 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и 49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, или 600-667 SEQ ID NO: 89 (стандартная обратная транскриптаза, RT M-MLV), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента обратной транскриптазы согласно SEQ ID NO: 89.

[2461] Вариант осуществления 14. Полипептид по варианту осуществления 1, где полипептид представляет собой N-концевую половину слитого белка праймерного редактора.

[2462] Вариант осуществления 15. Полипептид по варианту осуществления 1, где полипептид представляет собой C-концевую половину слитого белка праймерного редактора.

[2463] Вариант осуществления 16. Нуклеотидная последовательность, кодирующая полипептид по любому из вариантов осуществления 1-15 и необязательно PEgRNA.

[2464] Вариант осуществления 17. Вирусный геном, включающий нуклеотидную последовательность, кодирующую полипептид по любому из вариантов осуществления 1-15 и необязательно PEgRNA.

[2465] Вариант осуществления 18. Вирусный геном по варианту осуществления 17, где нуклеотидная последовательность дополнительно включает последовательность промотора, подходящую для экспрессии полипептида по любому из вариантов осуществления 1-15.

[2466] Вариант осуществления 19. Вирусный геном по варианту осуществления 17, где нуклеотидная последовательность дополнительно включает последовательность, кодирующую PEgRNA.

[2467] Вариант осуществления 20. Вирусный геном, включающий геном, включающий нуклеотидную последовательность, кодирующую полипептид по любому из вариантов осуществления 1-15 и необязательно PEgRNA.

[2468] Вариант осуществления 21. Вирусная частица по варианту осуществления 20, где вирусная частица представляет собой аденовирусную частицу, адено-ассоциированную вирусную частицу, или лентивирусную частицу.

[2469] Вариант осуществления 22. Вирусная частица по варианту осуществления 20, где полипептид, кодируемый геномом, представляет собой N-концевую половину слитого белка праймерного редактора.

[2470] Вариант осуществления 23. Вирусная частица по варианту осуществления 20, где полипептид, кодируемый геномом, представляет собой C-концевую половину слитого белка праймерного редактора.

[2471] Вариант осуществления 24. Фармацевтическая композиция, включающая вирусную частицу по любому из вариантов осуществления 20-23 и фармацевтический наполнитель.

[2472] Вариант осуществления 25. Фармацевтическая композиция, включающая вирусную частицу по варианту осуществления 22 (кодирующую N-концевую половину) и фармацевтический наполнитель.

[2473] Вариант осуществления 26. Фармацевтическая композиция, включающая вирусную частицу по варианту осуществления 23 (кодирующую C-концевую половину) и фармацевтический наполнитель.

[2474] Вариант осуществления 27. Комплекс рибонуклеопротеина (RNP), включающий нуклеотидную последовательность, кодирующую полипептид по любому из вариантов осуществления 1-15 и необязательно PEgRNA.

[2475] Вариант осуществления 28. Комплекс рибонуклеопротеина (RNP) по варианту осуществления 27, где полипептид, кодируемый геномом, представляет собой N-концевую половину слитого белка праймерного редактора.

[2476] Вариант осуществления 29. Комплекс рибонуклеопротеина (RNP) по варианту осуществления 27, где полипептид, кодируемый геномом, представляет собой C-концевую половину слитого белка праймерного редактора.

[2477] Вариант осуществления 30. Фармацевтическая композиция, включающая комплекс рибонуклеопротеина (RNP) по любому из вариантов осуществления 27-29 и фармацевтический наполнитель.

[2478] Вариант осуществления 31. Фармацевтическая композиция, включающая комплекс рибонуклеопротеина (RNP) по варианту осуществления 28 (кодирующий N-концевую половину) и фармацевтический наполнитель.

[2479] Вариант осуществления 32. Фармацевтическая композиция, включающая комплекс рибонуклеопротеина (RNP) по варианту осуществления 29 (кодирующий C-концевую половину) и фармацевтический наполнитель.

[2480] Вариант осуществления 33. Фармацевтическая композиция, включающая первую AAV частицу и вторую AAV частицу, где первый AAV вектор экспрессирует N-концевую половину слитого белка праймерного редактора, а второй AAV вектор экспрессирует C-концевую половину слитого белка праймерного редактора, где N-концевая половина и C-концевая половина объединены в клетке для восстановления праймерного редактора.

[2481] Вариант осуществления 34. Фармацевтическая композиция по варианту осуществления 33, где первая или вторая AAV частица также экспрессирует PEgRNA, которая нацеливает восстановленный праймерный редактор на участок-мишень ДНК.

[2482] Вариант осуществления 35. Фармацевтическая композиция по варианту осуществления 33, где слитый белок праймерного редактора включает домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен полимеразы.

[2483] Вариант осуществления 36. Фармацевтическая композиция по варианту осуществления 33, где слитый белок праймерного редактора способен осуществлять праймерное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA).

[2484] Вариант осуществления 37. Фармацевтическая композиция по варианту осуществления 35, где napDNAbp представляет собой белок Cas9 или его вариант.

[2485] Вариант осуществления 38. Фармацевтическая композиция по варианту осуществления 35, где napDNAbp представляет собой нуклеазу с никазной активностью.

[2486] Вариант осуществления 39. Фармацевтическая композиция по варианту осуществления 35, где napDNAbp представляет собой имеющий нуклеазную

активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2487] Вариант осуществления 40. Фармацевтическая композиция по варианту осуществления 35, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2488] Вариант осуществления 41. Фармацевтическая композиция по варианту осуществления 33, где N-концевая и C-концевая половина образованы путем разделения слитого белка праймерного редактора в участке разделения.

[2489] Вариант осуществления 42. Фармацевтическая композиция по варианту осуществления 41, где участок разделения представляет собой пептидную связь в домене napDNAbp.

[2490] Вариант осуществления 43. Фармацевтическая композиция по варианту осуществления 41, где участок разделения представляет собой пептидную связь в домене полимеразы.

[2491] Вариант осуществления 44. Фармацевтическая композиция по варианту осуществления 41, где участок разделения представляет собой пептидную связь в линкере.

[2492] Вариант осуществления 45. Фармацевтическая композиция по варианту осуществления 41, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и 13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18 и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 ad 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и 49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 SEQ ID NO: 18 (стандартный SpCas9), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента SpCas9 согласно SEQ ID NO: 18.

[2493] Вариант осуществления 46. Фармацевтическая композиция по варианту осуществления 41, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и

13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18 и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 и 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и 49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, или 600-667 SEQ ID NO: 89 (стандартная обратная транскриптаза, RT M-MLV), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента обратной транскриптазы согласно SEQ ID NO: 89.

[2494] Вариант осуществления 47. Фармацевтическая композиция по варианту осуществления 33, где N-концевая половина слитого белка праймерного редактора имеет аминокислотную последовательность, кодирующую N-концевые слитые белки праймерного редактора, как описано здесь.

[2495] Вариант осуществления 48. Фармацевтическая композиция по варианту осуществления 33, где C-концевая половина слитого белка праймерного редактора имеет аминокислотную последовательность, кодирующую C-концевые слитые белки праймерного редактора, как описано здесь.

[2496] Вариант осуществления 49. Способ доставки слитого белка праймерного редактора в клетку, включающий трансфекцию клетки первой AAV частицей и второй AAV частицей, где первый AAV вектор экспрессирует N-концевую половину слитого белка праймерного редактора, а второй AAV вектор экспрессирует C-концевую половину слитого белка праймерного редактора, где N-концевая половина и C-концевая половина объединены в клетке для восстановления слитого белка праймерного редактора.

[2497] Вариант осуществления 50. Способ по варианту осуществления 49, где первая или вторая AAV частица также экспрессирует PEgRNA, которая нацеливает восстановленный праймерный редактор на участок-мишень ДНК.

[2498] Вариант осуществления 51. Способ по варианту осуществления 49, где слитый белок праймерного редактора включает домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен полимеразы.

[2499] Вариант осуществления 52. Способ по варианту осуществления 49, где слитый белок праймерного редактора способен осуществлять праймерное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA).

[2500] Вариант осуществления 53. Способ по варианту осуществления 51, где napDNAbp представляет собой белок Cas9 или его вариант.

[2501] Вариант осуществления 54. Способ по варианту осуществления 53, где napDNAbp представляет собой нуклеазу с никазной активностью.

[2502] Вариант осуществления 55. Способ по варианту осуществления 53, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2503] Вариант осуществления 56. Способ по варианту осуществления 53, где napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

[2504] Вариант осуществления 57. Способ по варианту осуществления 49, где N-концевая и C-концевая половина образованы путем разделения слитого белка праймерного редактора в участке разделения.

[2505] Вариант осуществления 58. Способ по варианту осуществления 57, где участок разделения представляет собой пептидную связь в домене napDNAbp.

[2506] Вариант осуществления 59. Способ по варианту осуществления 57, где участок разделения представляет собой пептидную связь в домене полимеразы.

[2507] Вариант осуществления 60. Способ по варианту осуществления 57, где участок разделения представляет собой пептидную связь в линкере.

[2508] Вариант осуществления 61. Способ по варианту осуществления 57, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и 13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18 и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 и 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и 49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, или 1300-1368 SEQ ID NO: 18 (стандартный SpCas9), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента SpCas9 согласно SEQ ID NO: 18.

[2509] Вариант осуществления 62. Способ по варианту осуществления 57, где участок разделения расположен в пептидной связи между остатками 1 и 2, 2 и 3, 3 и 4, 4 и 5, 5 и 6, 6 и 7, 7 и 8, 8 и 9, 9 и 10, 10 и 11, 11 и 12, 12 и 13, 13 и 14, 14 и 15, 16 и 17, 17 и 18, 18

и 19, 19 и 20, 20 и 21, 21 и 22, 22 и 23, 23 и 24, 24 и 25, 25 и 26, 26 и 27, 27 и 28, 28 и 29, 29 и 30, 30 и 31, 31 и 32, 32 и 33, 33 и 34, 34 и 35, 35 и 36, 36 и 37, 37 и 38, 38 и 39, 39 и 40, 40 и 41, 41 и 42, 42 и 43, 43 и 44, 44 и 45, 45 и 46, 46 и 47, 47 и 48, 48 и 49, 49 и 50, или между любыми двумя остатками между остатками 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-600, или 600-667 SEQ ID NO: 89 (стандартная обратная транскриптаза, RT M-MLV), или между любыми двумя эквивалентными аминокислотными остатками гомолога или эквивалента обратной транскриптазы согласно SEQ ID NO: 89.

[2510] Вариант осуществления 63. Способ по варианту осуществления 49, где N-концевая половина слитого белка праймерного редактора имеет аминокислотную последовательность, кодирующую N-концевые слитые белки праймерного редактора, как описано здесь.

[2511] Вариант осуществления 64. Способ по варианту осуществления 49, где C-концевая половина слитого белка праймерного редактора имеет аминокислотную последовательность, кодирующую C-концевые слитые белки праймерного редактора, как описано здесь.

[2512] Вариант осуществления 65. Способ по варианту осуществления 49, где первая AAV частица включает рекомбинантный геном AAV, включающий нуклеотидную последовательность, которая кодирует первый компонент праймерного редактора.

[2513] Вариант осуществления 66. Способ по варианту осуществления 49, где вторая AAV частица включает рекомбинантный геном AAV, включающий нуклеотидную последовательность, которая кодирует второй компонент праймерного редактора.

[2514] Вариант осуществления 67. Способ по варианту осуществления 49, где этап трансфекции осуществляется in vivo.

[2515] Вариант осуществления 68. Способ по варианту осуществления 49, где этап трансфекции осуществляется ex vivo.

[2516] Вариант осуществления 69. Способ по варианту осуществления50, где участок-мишень ДНК представляет собой связанный с заболеванием ген.

[2517] Вариант осуществления 70. Способ по варианту осуществления 69, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1 антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана;

нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; связанного с тринуклеотидными повторами заболевания; прионного заболевания; и болезни Тея-Сакса.

[2518] Вариант осуществления 71. Способ по варианту осуществления 69, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[2519] Вариант осуществления 72. Способ по варианту осуществления 51, где домен программируемого ДНК-связывающего белка (napDNAbp).

[2520] Вариант осуществления 73. Способ по варианту осуществления 51, где домен полимеразы представляет собой обратную транскриптазу.

[2521] Вариант осуществления 74. Способ по варианту осуществления 73, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2522] Вариант осуществления 75. Способ по варианту осуществления 73, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2523] Вариант осуществления 76. Способ по варианту осуществления 73, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[2524] Вариант осуществления 77. Способ по варианту осуществления 73, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

Вариант осуществления 78. Полипептид по варианту осуществления 8, где участок разделения расположен между 1023 и 1024 SEQ ID NO: 18 или в соответствующем положении в аминокислотной последовательности, имеющей, по меньшей мере, 85%,

по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 99%, или, по меньшей мере, 99,5% идентичность последовательности относительно SEQ ID NO: 18.

ГРУППА G. СПОСОБЫ РЕ ДЛЯ МОДИФИЦИРОВАНИЯ СТРУКТУРЫ/ФУНКЦИИ РНК

[2525] Вариант осуществления 1. Способ внесения рибонуклеотидного мотива или метки в представляющую интерес РНК, кодируемую нуклеотидной последовательностью-мишенью, с помощью праймерного редактирования, при этом способ включает: (a) контактирование нуклеотидной последовательности-мишени с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, кодирующую рибонуклеотидный мотив или метку; тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, кодирующей рибонуклеотидный мотив или метку; и встраивание одноцепочечной последовательности ДНК вместо соответствующей эндогенной цепи в нуклеотидной последовательности-мишени с помощью процесса репарации и/или репликации ДНК, при этом способ обеспечивает получение нуклеотидной последовательности-мишени, которая кодирует представляющую интерес модифицированную РНК, включающую рибонуклеотидный мотив или метку.

[2526] Вариант осуществления 2. Способ по варианту осуществления 1, где рибонуклеотидный мотив или метка представляет собой фрагмент для детектирования.

[2527] Вариант осуществления 3. Способ по варианту осуществления 1, где рибонуклеотидный мотив или метка влияет на уровень экспрессии представляющей интерес РНК.

[2528] Вариант осуществления 4. Способ по варианту осуществления 1, где рибонуклеотидный мотив или метка влияет на транспорт или субклеточную локализацию представляющей интерес РНК.

[2529] Вариант осуществления 5. Способ по варианту осуществления 1, где рибонуклеотидный мотив или метка выбрана из группы, состоящей из SV40 типа 1, SV40 типа 2, SV40 типа 3, hGH, BGH, rbGlob, TK, MALAT1 ENE-mascРНК, KSHV PAN ENE, Sm-бокса/бокса мяРНК U1, 3’ бокса мяРНК U1, тРНК-лизина, аптамера broccoli, аптамера spinach, аптамера mango, рибозима HDV, и m6A.

[2530] Вариант осуществления 6. Способ по варианту осуществления 1, где PEgRNA включает SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777 (см. Таблицу).

[2531] Вариант осуществления 7. Способ по варианту осуществления 1, где слитый белок включает аминокислотную последовательность PE1, PE2, или PE3.

[2532] Вариант осуществления 8. Способ по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2533] Вариант осуществления 9. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

ГРУППА H. СПОСОБЫ PE ДЛЯ ПОЛУЧЕНИЯ БИБЛИОТЕК ГЕНОВ

[2534] Вариант осуществления 1. Способ конструирования программируемой библиотеки мутантных генов с помощью праймерного редактирования, при этом способ включает:

[2535] (a) контактирование библиотеки нуклеотидных последовательностей-мишеней, каждая из которых включает один или несколько генетических локусов-мишеней, с (i) праймерным редактором, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для редактирования, включающую последовательность, имеющую, по меньшей мере, одно генетическое изменение относительно одного или нескольких генетических локусов-мишеней;

[2536] тем самым обеспечивая полимеризацию одноцепочечной последовательности ДНК, матрицей для которой служит матрица для редактирования; и

[2537] встраивание одноцепочечной последовательности ДНК вместо одного или нескольких генетических локусов-мишеней за счет репарации и/или репликации ДНК, тем самым обеспечивая встраивание указанного, по меньшей мере, одного генетического изменения в генетические локусы-мишени нуклеотидных последовательностей-мишеней указанной библиотеки.

[2538] Вариант осуществления 2. Способ по варианту осуществления 1, где библиотека представляет собой библиотеку плазмид.

[2539] Вариант осуществления 3. Способ по варианту осуществления 1, где библиотека представляет собой библиотеку фагов.

[2540] Вариант осуществления 4. Способ по варианту осуществления 1, где один или несколько генетических локусов-мишеней включают участок, кодирующий белок.

[2541] Вариант осуществления 5. Способ по варианту осуществления 1, где один или несколько генетических локусов-мишеней включают участок, кодирующий мотив вторичной структуры белка.

[2542] Вариант осуществления 6. Способ по варианту осуществления 5, где мотив вторичной структуры представляет собой альфа-спираль.

[2543] Вариант осуществления 7. Способ по варианту осуществления 5, где мотив вторичной структуры представляет собой бета-лист.

[2544] Вариант осуществления 8. Способ по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

[2545] Вариант осуществления 9. Способ по варианту осуществления 1, где napDNAbp представляет собой нуклеазу с никазной активностью.

[2546] Вариант осуществления 10. Способ по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2547] Вариант осуществления 11. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18.

[2548] Вариант осуществления 12. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2549] Вариант осуществления 13. Способ по варианту осуществления 1, где домен полимеразы представляет собой обратную транскриптазу.

[2550] Вариант осуществления 14. Способ по варианту осуществления 13, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2551] Вариант осуществления 15. Способ по варианту осуществления 14, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей

мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2552] Вариант осуществления 16. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение в матрице для редактирования представляет собой вставку.

[2553] Вариант осуществления 17. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение в матрице для редактирования представляет собой делецию.

[2554] Вариант осуществления 18. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение в матрице для редактирования представляет собой замещение.

[2555] Вариант осуществления 19. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение представляет собой вставку одного или нескольких кодонов.

[2556] Вариант осуществления 20. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение представляет собой делецию одного или нескольких кодонов.

[2557] Вариант осуществления 21. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение представляет собой вставку стоп-кодона.

[2558] Вариант осуществления 22. Способ по варианту осуществления 1, где указанное, по меньшей мере, одно генетическое изменение представляет собой конверсию кодона, не являющегося стоп-кодоном, в стоп-кодон.

[2559] Вариант осуществления 23. Способ по варианту осуществления 1, где способ осуществляется одновременно, по меньшей мере, в 2, или 3, или 4, или 5, или 6, или 7, или 8, или 9, или 10, или в 10-100, или в 100-200, или в 200-300, или в 300-400, или в 400-500 генетических локусах-мишенях в каждой нуклеотидной последовательности-мишени библиотеки.

[2560] Вариант осуществления 24. Способ по варианту осуществления 1, где способ осуществляется во время PACE или PANCE эволюции и где каждый случай встраивания указанного, по меньшей мере, одного генетического изменения в

генетические локусы-мишени нуклеотидных последовательностей-мишеней указанной библиотеки также обеспечивает внесение новой последовательности-мишени.

ГРУППА I. СПОСОБЫ РЕ ДЛЯ НЕЦЕЛЕВОГО ДЕТЕКТИРОВАНИЯ

[2561] Вариант осуществления 1. Способ оценки нецелевого редактирования с помощью праймерного редактора, при этом способ включает:

[2562] (a) контактирование нуклеотидной последовательности-мишени, имеющей участок редакторивания, с (i) слитым белком праймерного редактора, включающим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полимеразу, и (ii) PEgRNA, включающей матрицу для синтеза ДНК, которая кодирует детектируемую последовательность;

[2563] где PEgRNA образует комплекс со слитым белком и направляет указанный слитый белок к участку редактирования и, в случае наличия, к одному или нескольким нецелевым участкам;

[2564] и где слитый белок праймерного редактора обеспечивает внесение детектируемой последовательности в участке редактирования и, в случае наличия, в одном или нескольких нецелевых участках;

[2565] (b) определение нуклеотидной последовательности участка редактирования и одного или нескольких нецелевых участков.

[2566] Вариант осуществления 2. Способ по варианту осуществления 1, где нуклеотидная последовательность-мишень представляет собой геном.

[2567] Вариант осуществления 3. Способ по варианту осуществления 1, где этап контактирования осуществляется in vitro.

[2568] Вариант осуществления 4. Способ по варианту осуществления 1, где этап контактирования осуществляется in vivo.

[2569] Вариант осуществления 5. Способ по варианту осуществления 1, где участок редактирования представляет собой мутацию в связанном с заболеванием гене.

[2570] Вариант осуществления 6. Способ по варианту осуществления 5, где мутация представляет собой замещение одного основания, вставку, делецию, или инверсию.

[2571] Вариант осуществления 7. Способ по варианту осуществления 1, где связанный с заболеванием ген связан с моногенетическим заболеванием, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA); дефицита альфа-1

антитрипсина; муковисцидоза; мышечной дистрофии Дюшенна; галактоземии; гемохроматоза; болезни Хантингтона; болезни кленового сиропа; синдрома Марфана; нейрофиброматоза 1 типа; врожденной пахионихии; фенилкеотнурии; тяжелого комбинированного иммунодефицита; серповидноклеточной анемии; синдрома Смита-Лемли-Опица; и болезни Тея-Сакса.

[2572] Вариант осуществления 8. Способ по варианту осуществления 1, где связанный с заболеванием ген связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака, и ожирения.

[2573] Вариант осуществления 9. Способ по варианту осуществления 1, где слитый белок имеет аминокислотную последовательность SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно SEQ ID NO: 101-104, 181-183, 223-244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

[2574] Вариант осуществления 10. Способ по варианту осуществления 1, где napDNAbp представляет собой Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute, или вариант Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute.

[2575] Вариант осуществления 11. Способ по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

[2576] Вариант осуществления 12. Способ по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2577] Вариант осуществления 13. Способ по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2578] Вариант осуществления 14. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 18.

[2579] Вариант осуществления 15. Способ по варианту осуществления 1, где napDNAbp представляет собой SpCas9 дикого типа или его вариант согласно любой одной из аминокислотных последовательностей 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотной последовательности, имеющей, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2580] Вариант осуществления 16. Способ по варианту осуществления 1, где napDNAbp представляет собой ортолог SpCas9.

[2581] Вариант осуществления 17. Способ по варианту осуществления 1, где napDNAbp представляет собой любую одну из аминокислотных последовательностей 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2582] Вариант осуществления 18. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2583] Вариант осуществления 19. Способ по варианту осуществления 1, где полимераза включает активность РНК-зависимой ДНК-полимеразы.

[2584] Вариант осуществления 20. Способ по варианту осуществления 1, где полимераза представляет собой обратную транскриптазу.

[2585] Вариант осуществления 21. Способ по варианту осуществления 1, где обратная транскриптаза представляет собой встречающуюся в природе обратную транскриптазу дикого типа, имеющую аминокислотную последовательность любой одной из SEQ ID NO: 89 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой из SEQ ID NO: 89.

[2586] Вариант осуществления 22. Способ по варианту осуществления 1, где обратная транскриптаза представляет собой вариант обратной транскриптазы, имеющей аминокислотную последовательность любой одной из SEQ ID NO: 89-100, 105-122,

128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2587] Вариант осуществления 23. Способ по варианту осуществления 1, где обратная транскриптаза включает любую одну из аминокислотных последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2588] Вариант осуществления 24. Способ по варианту осуществления 1, где обратная транскриптаза включает аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2589] Вариант осуществления 25. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой вставку, по меньшей мере, 1, или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2590] Вариант осуществления 26. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой делецию, по меньшей мере, 1,

или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2591] Вариант осуществления 27. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение азотистого основания.

[2592] Вариант осуществления 28. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой переходную мутацию.

[2593] Вариант осуществления 29. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой трансверсивную мутацию.

[2594] Вариант осуществления 30. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение одного нуклеотида, выбранное из группы, состоящей из: (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[2595] Вариант осуществления 31. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение одного нуклеотида, которое обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары

оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2596] Вариант осуществления 32. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой последовательность в виде метки.

[2597] Вариант осуществления 33. Способ по варианту осуществления 1, где этап (d) определения нуклеотидной последовательности в целевых и нецелевых участках включает (i) фрагментирование нуклеотидной последовательности-мишени для образования фрагментов, (ii) присоединение адаптерных последовательностей к концам фрагментов, (iii) ПЦР амплификацию ампликона с помощью пары праймеров, при этом один праймер отжигается на адаптерной последовательности, присоединенной к одному концу фрагментов, а другой праймер отжигается на адаптерной последовательности, вставленной с помощью праймерного редактирования, расположенной во фрагменте, и (iv) секвенирование ампликона для определения месторасположения изменения.

ГРУППА J. СПОСОБЫ PE ДЛЯ ЗАПИСИ КЛЕТОЧНЫХ ДАННЫХ

[2598] Вариант осуществления 1. Способ записи клеточного события с помощью праймерного редактирования, при этом способ включает: (A) внесение в клетку одного или нескольких конструктов, кодирующих (i) слитый белок праймерного редактора, включающий napDNAbp и РНК-зависимую ДНК-полимеразу, и (ii) PEgRNA, где экспрессия слитого белка и/или PEgRNA индуцируется за счет наступления клеточного события и где экспрессия слитого белка и/или PEgRNA приводит к праймерному редактированию участка-мишени редактирования в геноме клетки для внесения детектируемой последовательности, и (B) идентификацию детектируемой последовательности, тем самым обеспечивая идентификацию наступления клеточного события.

[2599] Вариант осуществления 2. Способ по варианту осуществления 1, где праймерное редактирование этапа (A) одновременно обеспечивает внесение нового

участка-мишени редактирования, так что запись клеточного события может осуществляться итеративным образом.

[2600] Вариант осуществления 3. Способ по варианту осуществления 1, где PEgRNA включает матрицу для редактирования, которая кодирует детектируемую последовательность.

[2601] Вариант осуществления 4. Способ по варианту осуществления 3, где матрица для редактирования дополнительно кодирует новый участок-мишень редактирования.

[2602] Вариант осуществления 5. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой вставку, по меньшей мере, 1, или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2603] Вариант осуществления 6. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой делецию, по меньшей мере, 1, или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по

меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2604] Вариант осуществления 7. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение азотистого основания.

[2605] Вариант осуществления 8. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой переходную мутацию.

[2606] Вариант осуществления 9. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой трансверсивную мутацию.

[2607] Вариант осуществления 10. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение одного нуклеотида, где замещение одного нуклеотида представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[2608] Вариант осуществления 11. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой замещение одного нуклеотида, которое обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2609] Вариант осуществления 12. Способ по варианту осуществления 1, где детектируемая последовательность представляет собой последовательность в виде метки.

[2610] Вариант осуществления 13. Способ по варианту осуществления 1, где длина детектируемой последовательности увеличивается с течением времени в результате итеративной вставки детектируемой последовательности для каждого наступления клеточного события.

[2611] Вариант осуществления 14. Способ по варианту осуществления 1, где этап детектирования включает секвенирование подвергнутого редактированию участка-мишени или ампликона подвергнутого редактированию участка-мишени.

[2612] Вариант осуществления 15. Способ по варианту осуществления 1, где napDNAbp представляет собой Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute, или вариант Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute.

[2613] Вариант осуществления 16. Способ по варианту осуществления 1, где napDNAbp представляет собой белок Cas9 или его вариант.

[2614] Вариант осуществления 17. Способ по варианту осуществления 1, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2615] Вариант осуществления 18. Способ по варианту осуществления 1, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2616] Вариант осуществления 19. Способ по варианту осуществления 1, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 2 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 18.

[2617] Вариант осуществления 20. Способ по варианту осуществления 1, где napDNAbp представляет собой SpCas9 дикого типа или его вариант согласно любой одной из аминокислотных последовательностей 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотной последовательности, имеющей, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2618] Вариант осуществления 21. Способ по варианту осуществления 1, где napDNAbp представляет собой ортолог SpCas9.

[2619] Вариант осуществления 22. Способ по варианту осуществления 1, где napDNAbp представляет собой любую одну из аминокислотных последовательностей

18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2620] Вариант осуществления 23. Способ по варианту осуществления 1, где РНК-зависимая ДНК-полимераза представляет собой обратную транскриптазу.

[2621] Вариант осуществления 24. Способ по варианту осуществления 23, где обратная транскриптаза представляет собой встречающуюся в природе обратную транскриптазу дикого типа, имеющую аминокислотную последовательность любой одной из SEQ ID NO: 89 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой из SEQ ID NO: 89.

[2622] Вариант осуществления 25. Способ по варианту осуществления 23, где обратная транскриптаза представляет собой вариант обратной транскриптазы, имеющей аминокислотную последовательность любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2623] Вариант осуществления 26. Способ по варианту осуществления 1, где слитый белок включает аминокислотную последовательность любой одной из SEQ ID NO: 123 и 134 (PE1, PE2), или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно SEQ ID NO: 123 and 134 (PE1, PE2).

[2624] Вариант осуществления 27. Способ по варианту осуществления 1, где один или несколько конструктов, которые кодируют слитый белок праймерного редактора и/или PEgRNA, дополнительно включают один или несколько промоторов, которые индуцируются при наступлении клеточного события.

[2625] Вариант осуществления 28. Способ по варианту осуществления 1, где клеточное событие соответствует стимулу, получаемому клеткой.

[2626] Вариант осуществления 29. Способ по варианту осуществления 28, где стимул представляет собой малую молекулу, белок, пептид, аминокислоту, метаболит,

неорганическую молекулу, органометаллическую молекулу, органическую молекулу, лекарство или лекарство-кандидат, сахар, липид, металл, нуклеиновую кислоту, молекулу, полученную при активации эндогенного или экзогенного каскада передачи сигнала, свет, температуру, звук, давление, механическое давление, давление вследствие сдвига, или вирус или другой микроорганизм, изменение pH, или изменение в плане состояния окисления/восстановления.

[2627] Вариант осуществления 30. Плазмида для записи клеточных данных для записи клеточного события с помощью праймерного редактора, включающая:

[2628] i. слитый белок, включающий (a) последовательность нуклеиновой кислоты, кодирующую программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp), и (b) РНК-зависимую ДНК-полимеразу, при этом указанный слитый белок функционально связан с первым промотором;

[2629] ii. последовательность нуклеиновой кислоты, кодирующую направляющую РНК праймерного редактора (PEgRNA), функционально связанную со вторым промотором, где PEgRNA комплементарна последовательности-мишени; и

[2630] iii. ориджин репликации;

[2631] где, по меньшей мере, один из промоторов представляет собой индуцируемый промотор и где PEgRNA связывается с napDNAbp при условиях, которые индуцируют эксперссию PEgRNA и экспрессию napDNAbp, являющуюся достаточной для внесения детектируемой последовательности в участке-мишени детектирования.

[2632] Вариант осуществления 31. Плазмида для записи клеточных данных по варианту осуществления 30, где индуцируемый промотор индуцируется клеточным событием.

[2633] Вариант осуществления 32. Плазмида для записи клеточных данных по варианту осуществления 30, где клеточное событие соответствует стимулу, получаемому клеткой.

[2634] Вариант осуществления 33. Плазмида для записи клеточных данных по варианту осуществления 32, где стимул представляет собой малую молекулу, белок, пептид, аминокислоту, метаболит, неорганическую молекулу, органометаллическую молекулу, органическую молекулу, лекарство или лекарство-кандидат, сахар, липид, металл, нуклеиновую кислоту, молекулу, полученную при активации эндогенного или экзогенного каскада передачи сигнала, свет, температуру, звук, давление, механическое давление, давление вследствие сдвига, или вирус или другой

микроорганизм, изменение pH, или изменение в плане состояния окисления/восстановления.

[2635] Вариант осуществления 34. Плазмида для записи клеточных данных по варианту осуществления 30, где первый и второй промотор являются одинаковыми.

[2636] Вариант осуществления 35. Плазмида для записи клеточных данных по варианту осуществления 30, где первый и второй промотор являются различающимися.

[2637] Вариант осуществления 36. Плазмида для записи клеточных данных по варианту осуществления 30, где, по меньшей мере, один индуцируемый промотор представляет собой индуцируемый ангидротетрациклином промотор, индуцируемый IPTG промотор, индуцируемый рамнозой промотор, или индуцируемый арабинозой промотор.

[2638] Вариант осуществления 37. Плазмида для записи клеточных данных по варианту осуществления 30, где первый и второй промотор является конститутивным промотором.

[2639] Вариант осуществления 38. Плазмида для записи клеточных данных по варианту осуществления 37, где конститутивный промотор представляет собой промотор Lac, промотор цитомегаловируса (CMV), конститутивный промотор полимеразы III, или промотор UBC.

[2640] Вариант осуществления 39. Плазмида для записи клеточных данных по варианту осуществления 30, где ориджин репликации включает последовательность ориджина репликации pSC101, pMB1, pBR322, ColE1, или p15A.

[2641] Вариант осуществления 40. Плазмида для записи клеточных данных по варианту осуществления 30, где PEgRNA включает матрицу для редактирования, которая кодирует детектируемую последовательность.

[2642] Вариант осуществления 41. Плазмида для записи клеточных данных по варианту осуществления 40, где матрица для редактирования дополнительно кодирует новый участок-мишень редактирования.

[2643] Вариант осуществления 42. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой вставку, по меньшей мере, 1, или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей

мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2644] Вариант осуществления 43. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой делецию, по меньшей мере, 1, или, по меньшей мере, 2, или, по меньшей мере, 3, или, по меньшей мере, 4, или, по меньшей мере, 5, или, по меньшей мере, 6, или, по меньшей мере, 7, или, по меньшей мере, 8, или, по меньшей мере, 9, или, по меньшей мере, 10, или, по меньшей мере, 11, или, по меньшей мере, 12, или, по меньшей мере, 13, или, по меньшей мере, 14, или, по меньшей мере, 15, или, по меньшей мере, 16, или, по меньшей мере, 17, или, по меньшей мере, 18, или, по меньшей мере, 19, или, по меньшей мере, 20, или, по меньшей мере, 21, или, по меньшей мере, 22, или, по меньшей мере, 23, или, по меньшей мере, 24, или, по меньшей мере, 25, или, по меньшей мере, 26, или, по меньшей мере, 27, или, по меньшей мере, 28, или, по меньшей мере, 29, или, по меньшей мере, 30, или, по меньшей мере, 31, или, по меньшей мере, 32, или, по меньшей мере, 33, или, по меньшей мере, 34, или, по меньшей мере, 35, или, по меньшей мере, 40, или, по меньшей мере, 50, или, по меньшей мере, 60, или, по меньшей мере, 70, или, по меньшей мере, 80, или, по меньшей мере, 90, или, по меньшей мере, 100 азотистых оснований.

[2645] Вариант осуществления 44. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой замещение азотистого основания.

[2646] Вариант осуществления 45. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой переходную мутацию.

[2647] Вариант осуществления 46. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой трансверсивную мутацию.

[2648] Вариант осуществления 47. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой замещение одного нуклеотида, где замещение одного нуклеотида представляет собой (1) замещение G на T, (2) замещение G на A, (3) замещение G на C, (4) замещение T на G, (5) замещение T на A, (6) замещение T на C, (7) замещение C на G, (8) замещение C на T, (9) замещение C на A, (10) замещение A на T, (11) замещение A на G, или (12) замещение A на C.

[2649] Вариант осуществления 48. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой замещение одного нуклеотида, которое обеспечивает конвертирование (1) пары оснований G:C на пару оснований T:A, (2) пары оснований G:C на пару оснований A:T, (3) пары оснований G:C на пару оснований C:G, (4) пары оснований T:A на пару оснований G:C, (5) пары оснований T:A на пару оснований A:T, (6) пары оснований T:A на пару оснований C:G, (7) пары оснований C:G на пару оснований G:C, (8) пары оснований C:G на пару оснований T:A, (9) пары оснований C:G на пару оснований A:T, (10) пары оснований A:T на пару оснований T:A, (11) пары оснований A:T на пару оснований G:C, или (12) пары оснований A:T на пару оснований C:G.

[2650] Вариант осуществления 49. Плазмида для записи клеточных данных по варианту осуществления 30, где детектируемая последовательность представляет собой последовательность в виде метки.

[2651] Вариант осуществления 50. Плазмида для записи клеточных данных по варианту осуществления 30, где длина детектируемой последовательности увеличивается с течением времени в результате итеративной вставки детектируемой последовательности для каждого наступления клеточного события.

[2652] Вариант осуществления 51. Плазмида для записи клеточных данных по варианту осуществления 30, где этап детектирования включает секвенирование подвергнутого редактированию участка-мишени или ампликона подвергнутого редактированию участка-мишени.

[2653] Вариант осуществления 52. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой Cas9, Cas12e, Cas12d,

Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute, или вариант Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, или Argonaute.

[2654] Вариант осуществления 53. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой Cas9 или его вариант.

[2655] Вариант осуществления 54. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

[2656] Вариант осуществления 55. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой никазу Cas9 (nCas9).

[2657] Вариант осуществления 56. Способ по варианту осуществления 30, где napDNAbp включает аминокислотную последовательность SEQ ID NO: 18 или аминокислотную последовательность, которая, по меньшей мере, на 80%, 85%, 90%, 95%, 98%, или 99% идентична SEQ ID NO: 18.

[2658] Вариант осуществления 57. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой SpCas9 дикого типа или его вариант согласно любой одной из аминокислотных последовательностей 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотной последовательности, имеющей, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2659] Вариант осуществления 58. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой ортолог SpCas9.

[2660] Вариант осуществления 59. Плазмида для записи клеточных данных по варианту осуществления 30, где napDNAbp представляет собой любую одну из аминокислотных последовательностей 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 18-88, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, и 482-487.

[2661] Вариант осуществления 60. Плазмида для записи клеточных данных по варианту осуществления 30, где РНК-зависимая ДНК-полимераза представляет собой обратную транскриптазу.

[2662] Вариант осуществления 61. Плазмида для записи клеточных данных по варианту осуществления 30, где обратная транскриптаза представляет собой встречающуюся в природе обратную транскриптазу дикого типа, имеющую аминокислотную последовательность любой одной из SEQ ID NO: 89 или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность относительно аминокислотной последовательности любой из SEQ ID NO: 89.

[2663] Вариант осуществления 62. Плазмида для записи клеточных данных по варианту осуществления 30, где обратная транскриптаза представляет собой вариант обратной транскриптазы, имеющей аминокислотную последовательность любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766, или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

[2664] Вариант осуществления 63. Плазмида для записи клеточных данных по варианту осуществления 30, где слитый белок включает аминокислотную последовательность любой одной из SEQ ID NO: 123 и 134 (PE1, PE2), или аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно любой из SEQ ID NO: 123 и 134 (PE1, PE2).

[2665] Вариант осуществления 64. Набор для использования в клетке, включающей плазмиду для записи клеточных данных по любому одному из вариантов осуществления 30-63.

[2666] Вариант осуществления 65. Набор по варианту осуществления 64, где клетка представляет собой прокариотическую клетку.

[2667] Вариант осуществления 66. Набор по варианту осуществления 64, где клетка представляет собой эукариотическую клетку.

[2668] Вариант осуществления 67. Клетка, включающая плазмиду для записи клеточных данных по любому одному из вариантов осуществления 30-63.

[2669] Вариант осуществления 68. Клетка по варианту осуществления 67, где клетка представляет собой прокариотическую клетку.

[2670] Вариант осуществления 69. Клетка по варианту осуществления 67, где клетка представляет собой эукариотическую клетку.

[2671] Вариант осуществления 70. Клетка по варианту осуществления 69, где эукариотическая клетка представляет собой клетку млекопитающего.

[2672] Вариант осуществления 71. Клетка по варианту осуществления 70, где клетка млекопитающего представляет собой клетку человека.

[2673] Вариант осуществления 72. Способ записи клеточного события с помощью праймерного редактирования, при этом способ включает: (A) внесение в клетку плазмиды для записи клеточных данных по любому из вариантов осуществления 30-63, где слитый белок и/или PEgRNA индуцируется за счет наступления клеточного события и где экспрессия слитого белка и/или PEgRNA приводит к праймерному редактированию участка-мишени редактирования в геноме клетки для внесения детектируемой последовательности, и (B) идентификацию детектируемой последовательности, тем самым обеспечивая идентификацию наступления клеточного события.

[2674] Вариант осуществления 73. Способ по варианту осуществления 72, где этап (A) внесения осуществляется с помощью трансфекции или электропорации.

ЭКВИВАЛЕНТЫ И ОХВАТЫВАЕМЫЙ ОБЪЕМ

[2675] В рамках описания существительные могут обозначать один или более одного, если иное не указано или не следует из контекста. Варианты осуществления или фрагменты описания, которые включают "или" между одним или несколькими элементами группы, являются действительными, если один, более чем один, или все из элементов группы присутствуют, используются, или иным образом связаны с определенным продуктом или процессом, если иное не указано или не следует из контекста. Изобретение включает варианты осуществления, в которых конкретно один элемент группы присутствует, используется, или иным образом связан с определенным продуктом или процессом. Изобретение включает варианты осуществления, в которых более чем один или все из элементов группы присутствуют, используются, или иным образом являются связанными с определенным продуктом или процессом.

[2676] Более того, изобретение охватывает все вариации, комбинации, и пермутации, в рамках которых одно или несколько ограничений, элементов, пунктов, и описательных элементов из одного или нескольких перечисленных пунктов формулы изобретения внесены в другой пункт формулы изобретения. Например, любой пункт формулы изобретения, который зависит от другого пункта формулы изобретения, может быть модифицирован для включения одного или нескольких ограничений, указанных в любом другом пункте формулы изобретения, который зависит от того же первоначального пункта формулы изобретения. Если элементы представлены в виде перечней, например в формате группы Маркуша, каждая подгруппа данных элементов также включается, при этом любой элемент (элементы) могут быть удалены из группы. Следует понимать, что, в целом, при обозначении изобретения или аспектов изобретения в качестве включающих определенные элементы и/или свойства, определенные варианты осуществления изобретения или аспекты изобретения состоят или состоят по существу из подобных элементов и/или свойств. Для упрощения такие варианты осуществления не были явно приведены *in haec verba* здесь. Следует также отметить, что термины "включающий" и "содержащий" являются открытыми и допускают включение дополнительных элементов или этапов. При указании диапазонов их границы включаются. Более того, если иное не указано или не следует из контекста и понимания среднего специалиста в данной области, значения, которые выражены в виде диапазонов, могут соответствовать любому конкретному значению или поддиапазону в рамках указанных диапазонов в различных вариантах

осуществления изобретения до десятой доли единицы нижней границы диапазона, если из контекста явно не следует иное.

[2677] Данная заявка включает ссылки на различные выданные патенты, опубликованные патентные заявки, журнальные статьи, и другие публикации, все из которых приведены здесь в качестве ссылки. В случае несоответствия между какими-либо включенными ссылками и настоящим описанием описание обладает приоритетом. Кроме того, любой отдельный вариант осуществления настоящего изобретения, который соответствует уровню техники, может быть явно исключен из любого одного или нескольких вариантов осуществления. Поскольку такие варианты осуществления рассматриваются в качестве известных среднему специалисту в данной области, они могут быть исключены, даже если исключение явно не указано здесь. Любой конкретный вариант осуществления изобретения может быть исключен из любого варианта осуществления по любой причине, которая связана или не связана с наличием уровня техники.

[2678] Специалистам в данной области будут очевидны или они будут способны определить с помощью стандартных экспериментов многие эквиваленты определенных описанных здесь вариантов осуществления. Объем настоящих описанных здесь вариантов осуществления не ограничивается приведенным выше Описанием, но основан на прилагаемой формуле изобретения. Средним специалистам в данной области будет очевидно, что могут быть реализованы различные изменения и модификации данного изобретения без выхода за рамки духа или объема настоящего изобретения, как определено в рамках приведенной ниже формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Комплекс для праймерного редактирования, включающий:

(i) слитый белок, включающий домен программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и домен, включающий активность РНК-зависимой ДНК-полимеразы; и

(ii) направляющую РНК для праймерного редактирования (PEgRNA).

2. Комплекс по п.1, отличающийся тем, что слитый белок способен осуществлять праймерное редактирование в присутствии направляющей РНК для праймерного редактирования (PEgRNA) для внесения требуемого нуклеотидного изменения в последовательность-мишень.

3. Комплекс по п.11, отличающийся тем, что napDNAbp имеет никазную активность.

4. Комплекс по п.1, отличающийся тем, что napDNAbp представляет собой белок Cas9 или его вариант.

5. Комплекс по п.1, отличающийся тем, что napDNAbp представляет собой имеющий нуклеазную активность Cas9, не имеющий нуклеазной активности Cas9 (dCas9), или никазу Cas9 (nCas9).

6. Комплекс по п.1, отличающийся тем, что napDNAbp представляет собой никазу Cas9 (nCas9).

7. Комплекс по п.1, отличающийся тем, что napDNAbp выбран из группы, состоящей из: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, и Argonaute, и необязательно имеет никазную активность.

8. Комплекс по п.1, отличающийся тем, что домен, включающий активность РНК-зависимой ДНК-полимеразы представляет собой обратную транскриптазу, включающую любую одну из аминокислотных последовательностей SEQ ID NO: 89-

100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

9. Комплекс по п.1, отличающийся тем, что домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу, включающую аминокислотную последовательность, имеющую, по меньшей мере, 80%, 85%, 90%, 95%, 98%, или 99% идентичность последовательности относительно аминокислотной последовательности любой одной из SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700, 701-716, 739-741, и 766.

10. Комплекс по п.1, отличающийся тем, что домен, включающий активность РНК-зависимой ДНК-полимеразы, представляет собой встречающуюся в природе обратную транскриптазу из ретровируса или ретротранспозона.

11. Комплекс по п.1, отличающийся тем, что слитый белок при образовании комплекса с PEgRNA способен связываться с последовательностью-мишенью ДНК.

12. Комплекс по п.1, отличающийся тем, что PEgRNA включает направляющую РНК и, по меньшей мере, одно плечо удлинения нуклеиновой кислоты, включающее матрицу для синтеза ДНК.

13. Комплекс по п.12, отличающийся тем, что плечо удлинения нуклеиновой кислоты расположено на 3′- или 5′-конце направляющей РНК или во внутримолекулярном положении в направляющей РНК, и при этом плечо удлинения нуклеиновой кислоты представляет собой ДНК или РНК.

14. Комплекс по п.12, отличающийся тем, что PEgRNA способна связываться с napDNAbp и направлять napDNAbp к последовательности-мишени ДНК.

15. Комплекс по п.14, отличающийся тем, что последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

16. Комплекс по п.12, отличающийся тем, что направляющая РНК гибридизируется с цепью-мишенью для образования РНК-ДНК гибрида и R-петли.

17. Комплекс по п.12, отличающийся тем, что указанное, по меньшей мере, одно

плечо удлинения нуклеиновой кислоты дополнительно включает участок связывания праймера.

18. Комплекс по п.12, отличающийся тем, что плечо удлинения нуклеиновой кислоты имеет длину, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, по меньшей мере, 15 нуклеотидов, по меньшей мере, 16 нуклеотидов, по меньшей мере, 17 нуклеотидов, по меньшей мере, 18 нуклеотидов, по меньшей мере, 19 нуклеотидов, по меньшей мере, 20 нуклеотидов, по меньшей мере, 21 нуклеотидов, по меньшей мере, 22 нуклеотидов, по меньшей мере, 23 нуклеотидов, по меньшей мере, 24 нуклеотидов, по меньшей мере, 25 нуклеотидов, по меньшей мере, 26 нуклеотидов, по меньшей мере, 27 нуклеотидов, по меньшей мере, 28 нуклеотидов, по меньшей мере, 29 нуклеотидов, по меньшей мере, 30 нуклеотидов, по меньшей мере, 31 нуклеотидов, по меньшей мере, 32 нуклеотидов, по меньшей мере, 33 нуклеотидов, по меньшей мере, 34 нуклеотидов, по меньшей мере, 35 нуклеотидов, по меньшей мере, 36 нуклеотидов, по меньшей мере, 37 нуклеотидов, по меньшей мере, 38 нуклеотидов, по меньшей мере, 39 нуклеотидов, по меньшей мере, 40 нуклеотидов, по меньшей мере, 41 нуклеотидов, по меньшей мере, 42 нуклеотидов, по меньшей мере, 43 нуклеотидов, по меньшей мере, 44 нуклеотидов, по меньшей мере, 45 нуклеотидов, по меньшей мере, 46 нуклеотидов, по меньшей мере, 47 нуклеотидов, по меньшей мере, 48 нуклеотидов, по меньшей мере, 49 нуклеотидов, или, по меньшей мере, 50 нуклеотидов.

19. Комплекс по п.12, отличающийся тем, что матрица для синтеза ДНК имеет длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

20. Комплекс по п.17, отличающийся тем, что участок связывания праймера имеет

длину, по меньшей мере, 3 нуклеотида, по меньшей мере, 4 нуклеотида, по меньшей мере, 5 нуклеотидов, по меньшей мере, 6 нуклеотидов, по меньшей мере, 7 нуклеотидов, по меньшей мере, 8 нуклеотидов, по меньшей мере, 9 нуклеотидов, по меньшей мере, 10 нуклеотидов, по меньшей мере, 11 нуклеотидов, по меньшей мере, 12 нуклеотидов, по меньшей мере, 13 нуклеотидов, по меньшей мере, 14 нуклеотидов, или, по меньшей мере, 15 нуклеотидов.

21. Комплекс по п.12, отличающийся тем, что PEgRNA дополнительно включает, по меньшей мере, одну дополнительную структуру, выбранную из группы, состоящей из линкера, структуры типа "стебель-петля", шпильки, опорной петли, аптамера, или домена рекрутирования ДНК-белка.

22. Комплекс по п.12, отличающийся тем, что матрица для синтеза ДНК кодирует одноцепочечный ДНК-флэп, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный ДНК-флэп включает требуемое нуклеотидное изменение.

23. Комплекс по п.22, отличающийся тем, что одноцепочечный ДНК-флэп вытесняет эндогенную одноцепочечную ДНК, имеющую 5′-конец, в последовательности-мишени ДНК, которая была никирована, и при этом эндогенная одноцепочечная ДНК является непосредственно смежной и расположенной ниже участка разрыва.

24. Комплекс по п.23, отличающийся тем, что эндогенная одноцепочечная ДНК, имеющая свободный 5′-конец, подвергается эксцизии клеткой.

25. Комплекс по п.23, отличающийся тем, что клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование требуемого продукта.

26. Комплекс по п.12, отличающийся тем, что PEgRNA включает нуклеотидную последовательность SEQ ID NO: 18-36 или нуклеотидную последовательность, имеющую, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% идентичность последовательности относительно любой одной из SEQ ID NO: 101-104, 181-183, 223-

244, 277, 325-334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 499-505, 735-761, 776-777.

27. Комплекс по п.12, отличающийся тем, что матрица для синтеза ДНК включает нуклеотидную последовательность, которая, по меньшей мере, на 80%, или 85%, или 90%, или 95%, или 99% идентична эндогенной ДНК-мишени.

28. Комплекс по п.17, отличающийся тем, что участок связывания праймера гибридизируется со свободным 3'-концом разрезанной ДНК.

29. Комплекс по п.21, отличающийся тем, что указанная, по меньшей мере, одна дополнительная структура расположена на 3'- или 5'-конце PEgRNA.

30. Комплекс по п. 29, отличающийся тем, что линкер включает нуклеотидную последовательность, выбранную из группы, состоящей из SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

31. Комплекс по п.29, отличающийся тем, что структура типа "стебель-петля" включает нуклеотидную последовательность, выбранную из описанных здесь структур типа "стебель-петля".

32. Комплекс по п.29, отличающийся тем, что шпилька включает нуклеотидную последовательность, выбранную из описанных здесь шпилек.

33. Комплекс по п.29, отличающийся тем, что опорная петля включает нуклеотидную последовательность, выбранную из описанных здесь опорных петель.

34. Комплекс по п.29, отличающийся тем, что аптамер включает нуклеотидную последовательность, выбранную из описанных здесь аптамеров.

35. Комплекс по п.29, отличающийся тем, что домен рекрутирования РНК-белка включает нуклеотидную последовательность, выбранную из описанных здесь доменов рекрутирования РНК-белка.

36. Комплекс по п.1, отличающийся тем, что последовательность-мишень ДНК включает цепь-мишень и комплементарную не являющуюся мишенью цепь.

37. Комплекс по п.36, отличающийся тем, что R-петля включает (i) РНК-ДНК гибрид, включающий PEgRNA и цепь-мишень, и (ii) комплементарную не являющуюся мишенью цепь.

38. Комплекс по п.37, отличающийся тем, что мишень или комплементарная не являющаяся мишенью цепь никируется для образования праймерной последовательности, имеющей свободный 3´-конец.

39. Комплекс по п.38, отличающийся тем, что участок разрыва расположен выше последовательности PAM на цепи-мишени.

40. Комплекс по п.38, отличающийся тем, что участок разрыва расположен выше последовательности PAM на не являющейся мишенью цепи.

41. Комплекс по п.38, отличающийся тем, что участок разрыва расположен -1, -2, -3, -4, -5, -6, -7, -8, или -9 относительно 5'-конца последовательности PAM.

42. Комплекс по п.22, отличающийся тем, что одноцепочечный ДНК-флэп гибридизируется с последовательностью эндогенной ДНК, являющейся смежной относительно участка разрыва, тем самым обеспечивая внесение требуемого нуклеотидного изменения в цепь-мишень.

43. Комплекс по п.22, отличающийся тем, что одноцепочечный ДНК-флэп замещает последовательность эндогенной ДНК, являющуюся смежной относительно участка разрыва и имеющую свободный 5'-конец.

44. Комплекс по п.22, отличающийся тем, что последовательность эндогенной ДНК, имеющая 5´-конец, подвергается эксцизии клеткой.

45. Комплекс по п.44, отличающийся тем, что последовательность эндогенной ДНК, имеющая 5'-конец, подвергается эксцизии флэп-эндонуклеазой.

46. Комплекс по п.43, отличающийся тем, что клеточная репарация одноцепочечного ДНК-флэпа приводит к внесению требуемого нуклеотидного изменения в не являющуюся мишенью цепь, тем самым обеспечивая образование требуемого продукта.

47. Комплекс по п.46, отличающийся тем, что требуемое нуклеотидное изменение вносится в окне редактирования, которое расположено между приблизительно от -4 до +10 последовательности PAM, или между приблизительно от -10 до +20 последовательности PAM, или между приблизительно от -20 до +40 последовательности PAM, или между приблизительно от -30 до +100 последовательности PAM.

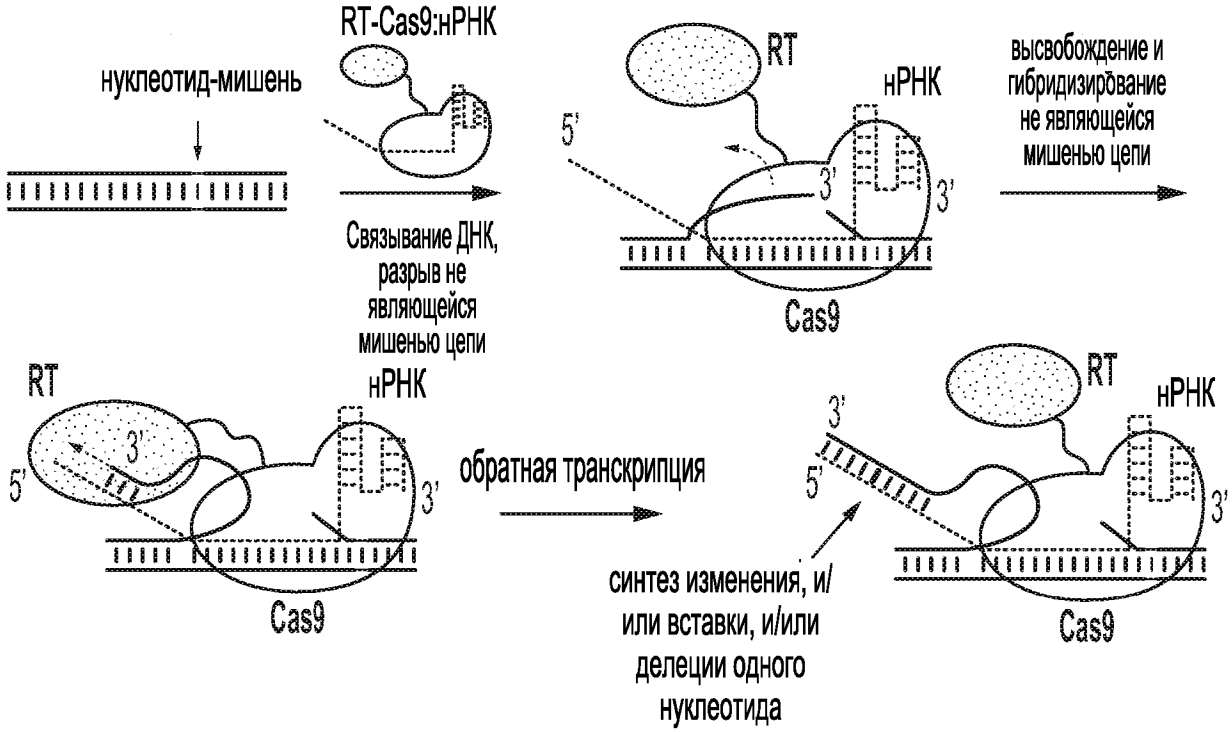
48. Комплекс по п.47, отличающийся тем, что требуемое нуклеотидное изменение вносится, по меньшей мере, на 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, или 100 нуклеотидов ниже участка разрыва.

49. Комплекс по любому одному из предыдущих пунктов, отличающийся тем, что слитый белок включает структуру NH_2-[napDNAbp]-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-COOH или NH_2-[домен, включающий активность РНК-зависимой ДНК-полимеразы]-[napDNAbp]-COOH, при этом каждый экземпляр "]-[" указывает на наличие необязательной последовательности линкера.

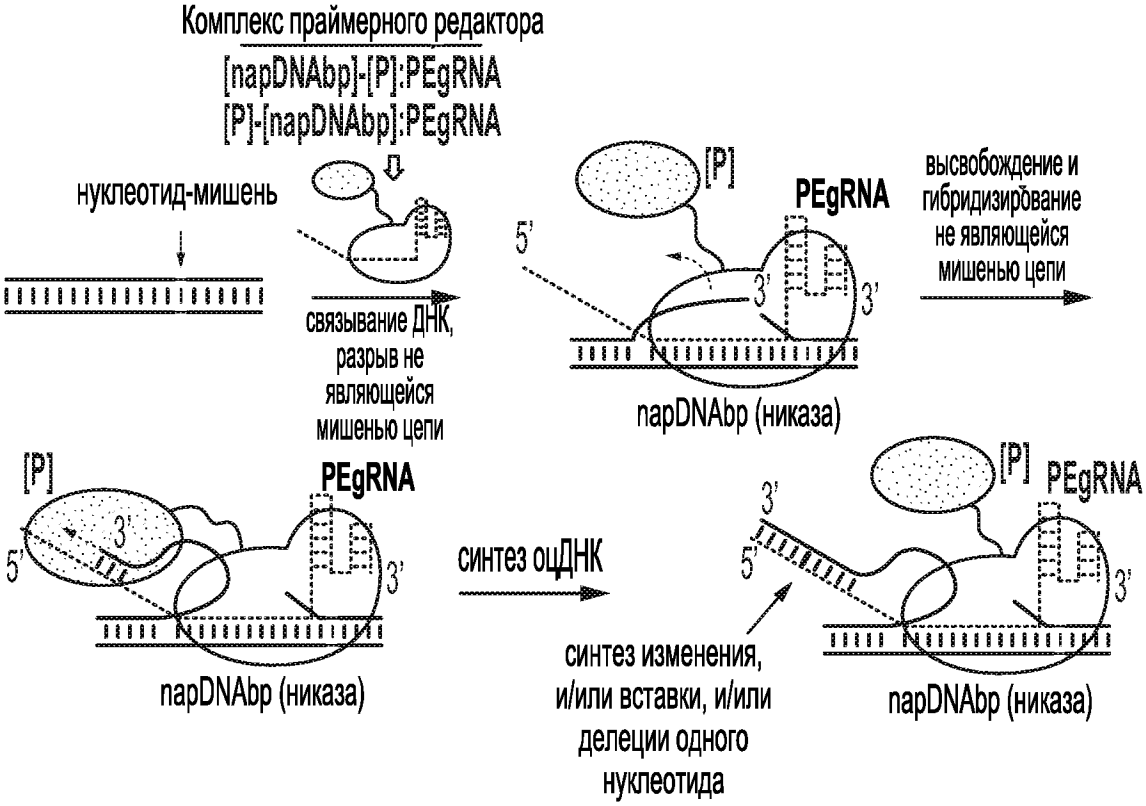
50. Комплекс по п.49, отличающийся тем, что последовательность линкера включает аминокислотную последовательность SEQ ID NO: 127, 165-176, 446, 453, и 767-769.

51. Комплекс по п.1, отличающийся тем, что слитый белок дополнительное включает линкер, который соединяет napDNAbp и домен, включающий активность РНК-зависимой ДНК-полимеразы.

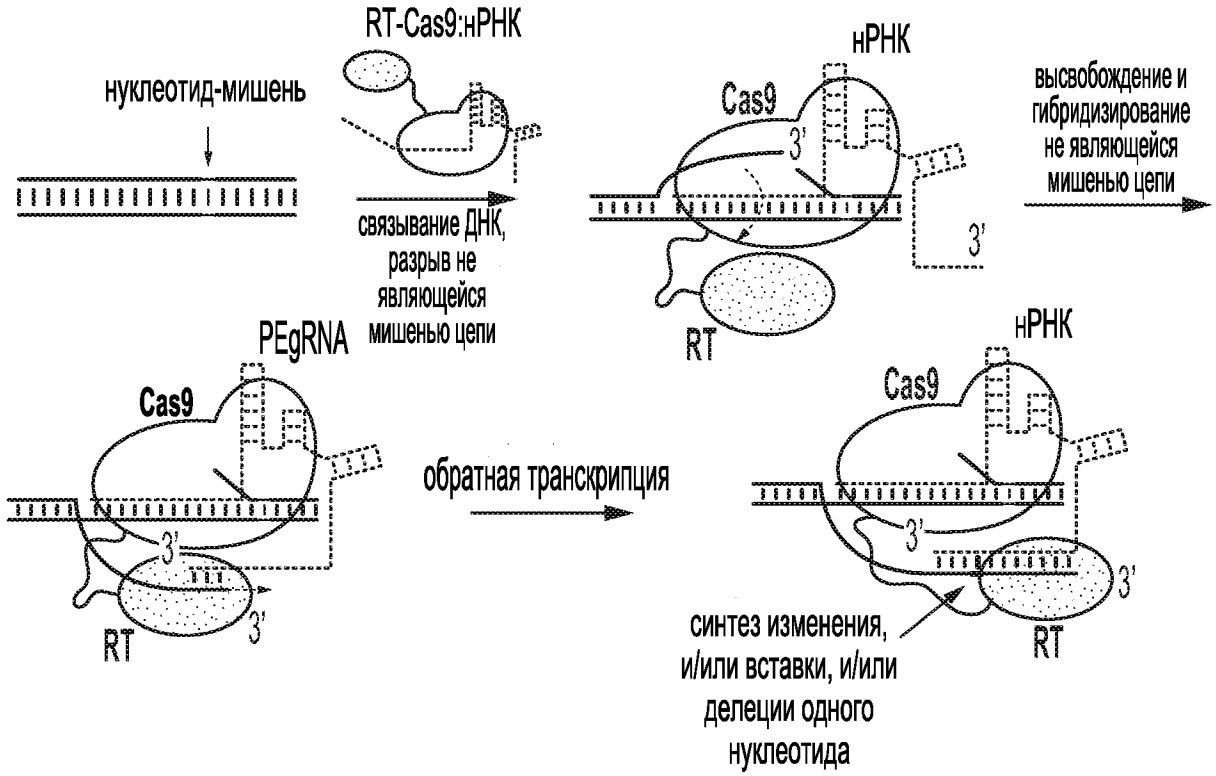
52. Комплекс по п.51, отличающийся тем, что последовательность линкера включает аминокислотную последовательность SEQ ID NO: 3887 (1x SGGS), 3888 (2x SGGS), 3889 (3x SGGS), 3890 (1x XTEN), 3891 (1x EAAAK), 3892 (2x EAAAK), и 3893 (3x EAAAK).


RT-Cas9:нРНК
нуклеотид-мишень
Связывание ДНК, разрыв не являющейся мишенью цепи
RT
нРНК
5'
3'
Cas9
высвобождение и гибридизирбвание не являющейся мишенью цепи
обратная транскрипция
синтез изменения, и/или вставки, и/или делеции одного нуклеотида

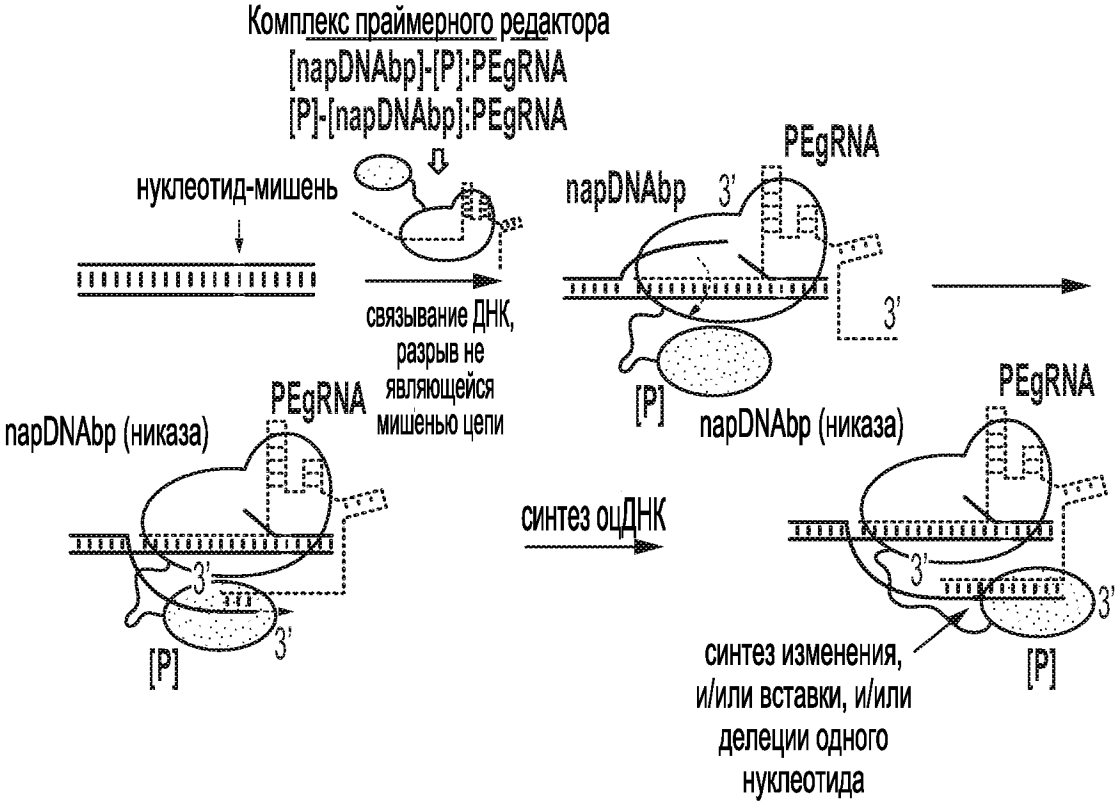
ФИГ.1A


Комплекс праймерного редактора
[napDNAbp]-[P]:PEgRNA
[P]-[napDNAbp]:PEgRNA
нуклеотид-мишень
связывание ДНК, разрыв не являющейся мишенью цепи
[P]
PEgRNA
napDNAbp (никаза)
высвобождение и гибридизирбвание не являющейся мишенью цепи
синтез оцДНК
синтез изменения, и/или вставки, и/или делеции одного нуклеотида

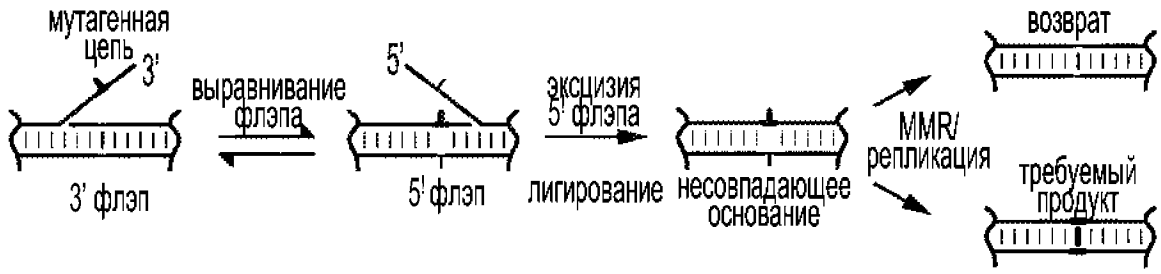
ФИГ.1B


RT-Cas9:нРНК
нуклеотид-мишень
связывание ДНК, разрыв не являющейся мишенью цепи
нРНК
Cas9
3'
3'
RT
высвобождение и гибридизирование не являющейся мишенью цепи
PEgRNA
Cas9
3'
3'
RT
обратная транскрипция
нРНК
Cas9
3'
3'
RT
синтез изменения, и/или вставки, и/или делеции одного нуклеотида

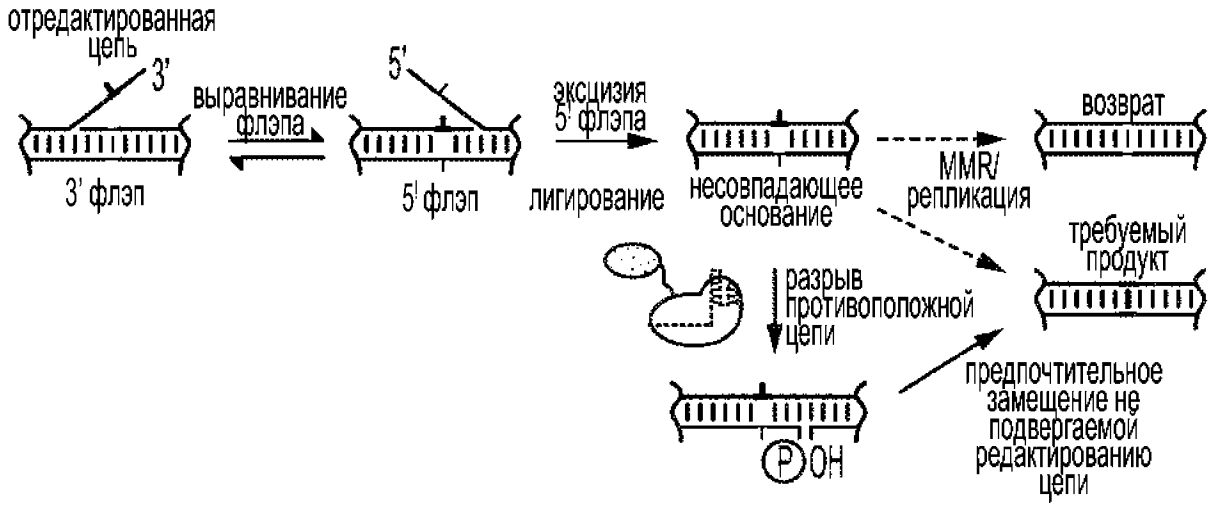
ФИГ.1C


Комплекс праймерного редактора
[napDNAbp]-[P]:PEgRNA
[P]-[napDNAbp]:PEgRNA
нуклеотид-мишень
связывание ДНК, разрыв не являющейся мишенью цепи
PEgRNA
napDNAbp
3'
3'
[P]
napDNAbp (никаза)
PEgRNA
napDNAbp (никаза)
3'
3'
[P]
синтез оцДНК
PEgRNA
3'
3'
[P]
синтез изменения, и/или вставки, и/или делеции одного нуклеотида

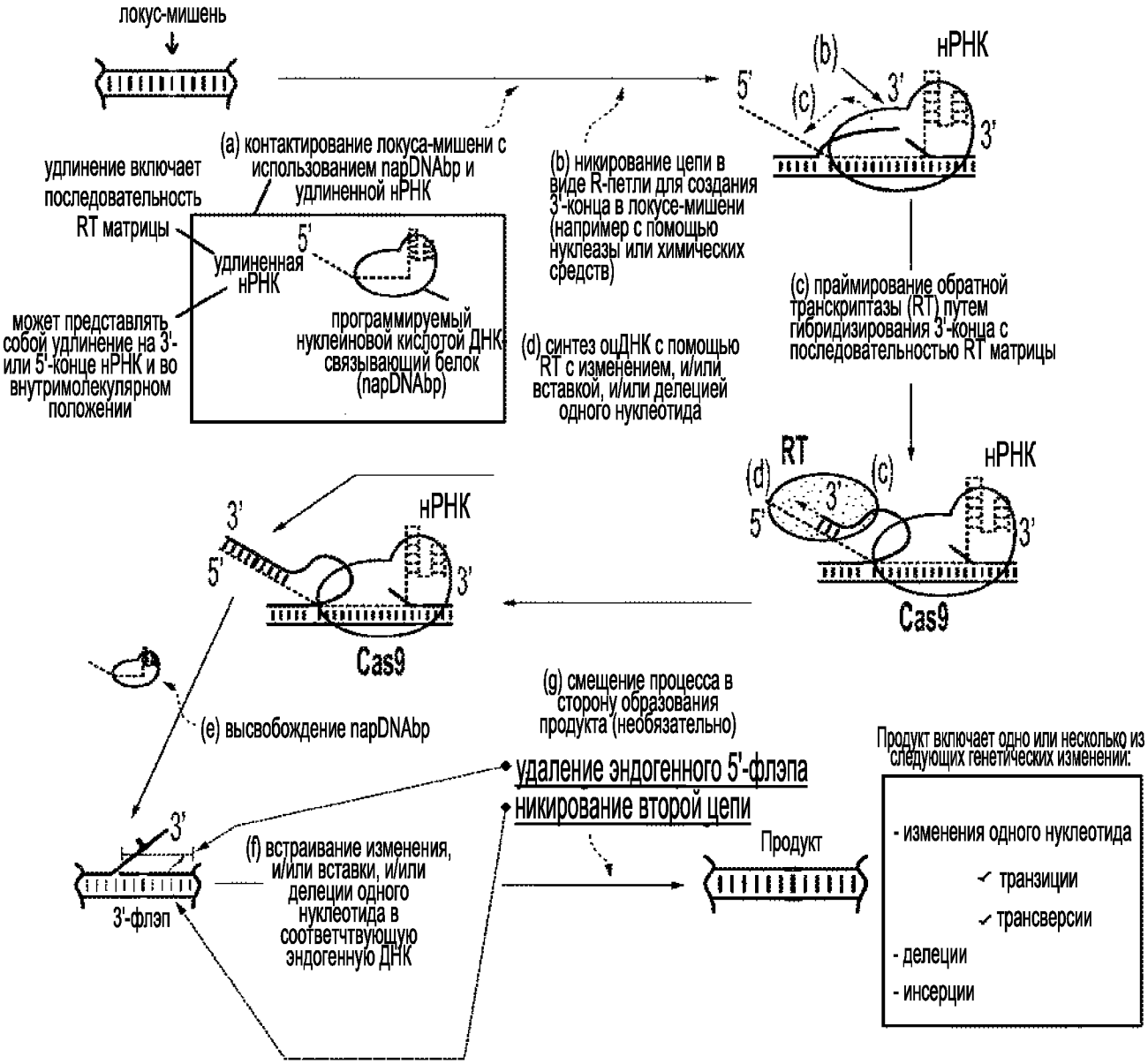
ФИГ.1D


мутагенная цепь
3'
выравнивание флэпа
5'
эксцизия 5' флэпа
возврат
MMR/ репликация
требуемый продукт
3' флэп
5' флэп
лигирование
несовпадающее основание

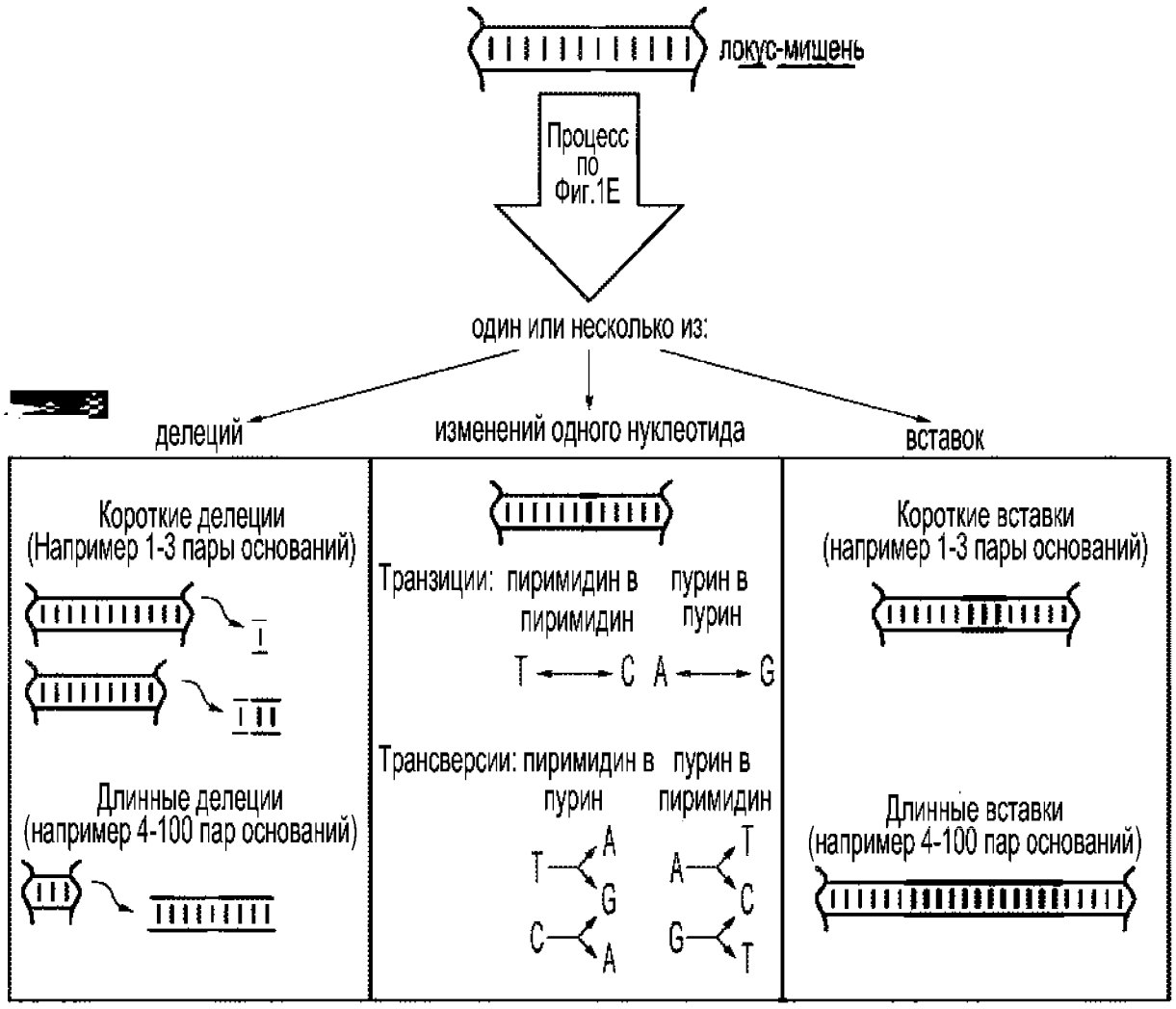
ФИГ.1E


отредактированная цепь
3'
выравнивание флэпа
5'
эксцизия 5' флэпа
возврат
MMR/ репликация
требуемый продукт
3' флэп
5' флэп
лигирование
несовпадающее основание
разрыв противоположной цепи
предпочтительное замещение не подвергаемой редактированию цепи
P
OH

ФИГ.1F


локус-мишень
(a) контактирование локуса-мишени с использованием napDNAbp и удлиненной нРНК
удлинение включает последовательность RT матрицы
может представлять собой удлинение на 3'- или 5'-конце нРНК и во внутримолекулярном положении
удлиненная нРНК
программируемый нуклеиновой кислотой ДНК связывающий белок (napDNAbp)
(b) никирование цепи в виде R-петли для создания 3'-конца в локусе-мишени (например с помощью нуклеазы или химических средств)
нРНК
(c) праймирование обратной транскриптазы (RT) путем гибридизирования 3'-конца с последовательностью RT матрицы
(d) синтез оцДНК с помощью RT с изменением, и/или вставкой, и/или делецией одного нуклеотида
RT
Cas9
(e) высвобождение napDNAbp
(f) встраивание изменения, и/или вставки, и/или делеции одного нуклеотида в соответствующую эндогенную ДНК
3'-флэп
(g) смещение процесса в сторону образования продукта (необязательно)
удаление эндогенного 5'-флэпа
никирование второй цепи
Продукт
Продукт включает одно или несколько из следующих генетических изменений:
- изменения одного нуклеотида
✓ транзиции
✓ трансверсии
- делеции
- инсерции

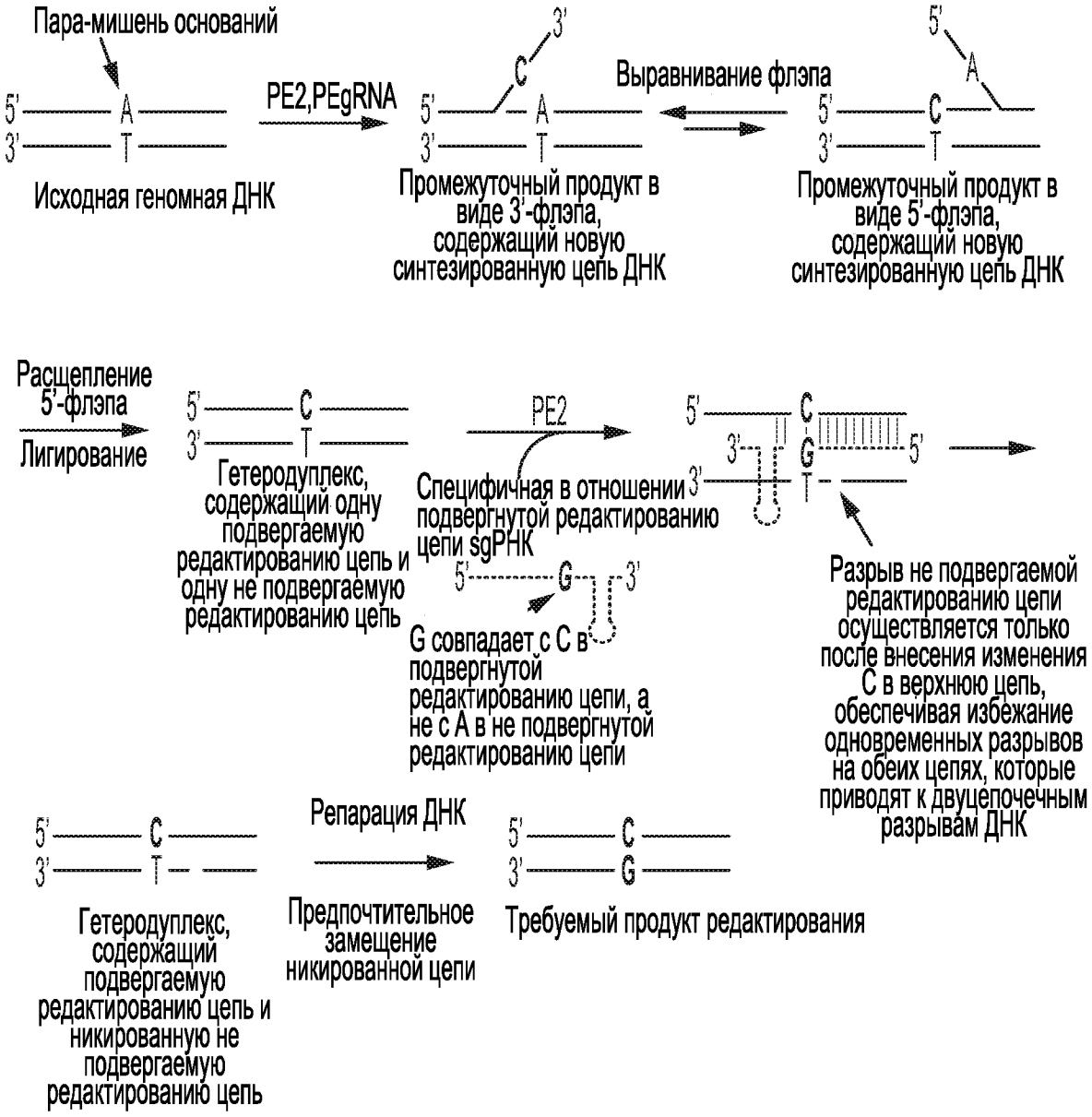
ФИГ.1G


локус-мишень
Процесс по Фиг.1E
один или несколько из:
делеций
изменений одного нуклеотида
вставок
Короткие делеции
(Например 1-3 пары оснований)
Длинные делеции
(например 4-100 пар оснований)
Транзиции: пиримидин в пиримидин
пурин в пурин
T ⟷ C A ⟷ G
Трансверсии: пиримидин в пурин
пурин в пиримидин
Короткие вставки
(например 1-3 пары оснований)
Длинные вставки
(например 4-100 пар оснований)

ФИГ.1H

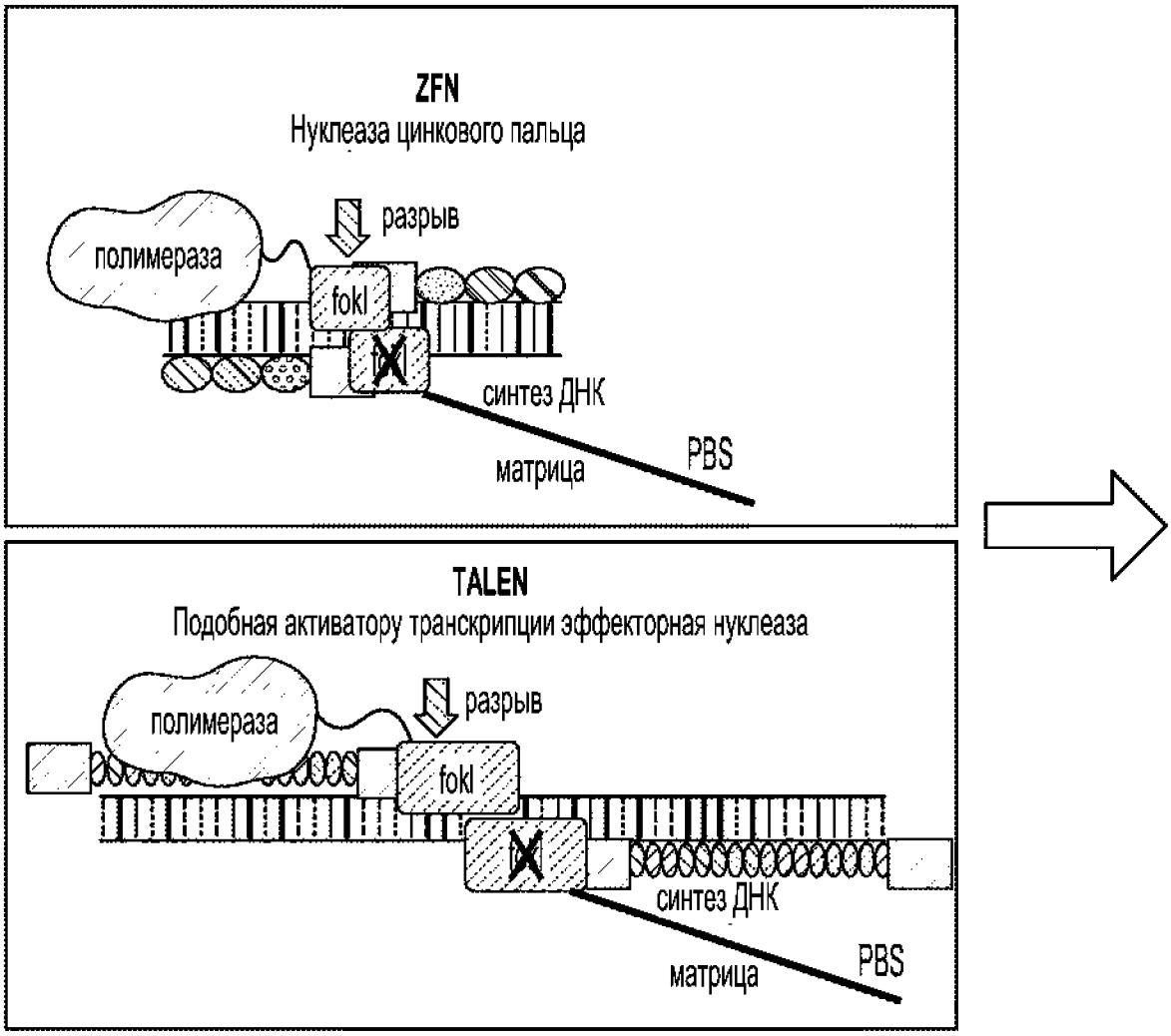
Временно́е никирование в рамках PE3b

Система редактирования PE3b: специфическое в плане редактирование никирование комплементарной (не подвергнутой редактированию цепи)



ФИГ.1I

Праймерное редактирование с использованием любой программируемой нуклеазы

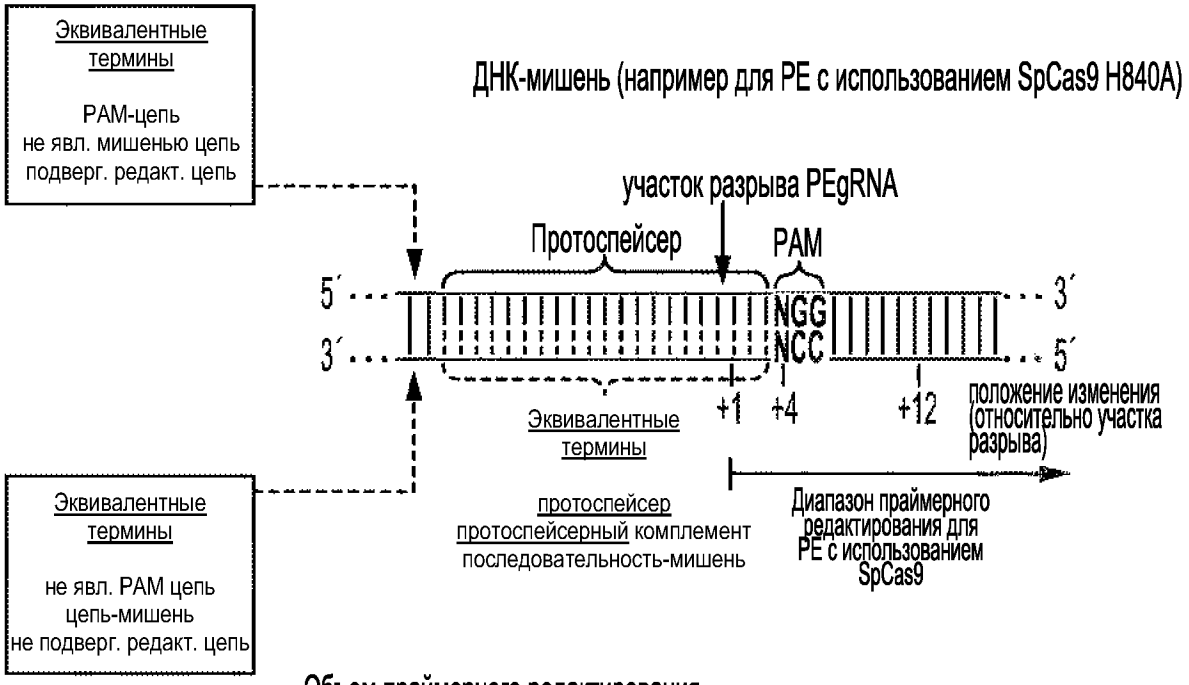


ФИГ.1J

Зависимый от полимеразы синтез замещающей цепи, кодируемой полимеразой с использованием матрицы для синтеза ДНК, которая содержит требуемое изменение. Предполагается, что механизм работает аналогично праймерному редактированию. Дополнительные факторы или эффекторы могут быть добавлены in trans или в виде продуктов слияния для обеспечения реакции в целом (например (a) хеликаза для раскручивания ДКН в участке разрыва для получения цепи с разрывом с 3'-концом, доступным в качестве праймера, (b) FEN1 для осуществления удаления эндогенной цепи на цепи с разрывом для направления реакции к замещению эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:нРНК для создания второго участка разрыва, который может обеспечивать встраивание синтезированной коррекции за счет предпочтительной клеточной репарации не подвергаемой редактированию цепи)

ФИГ.1J

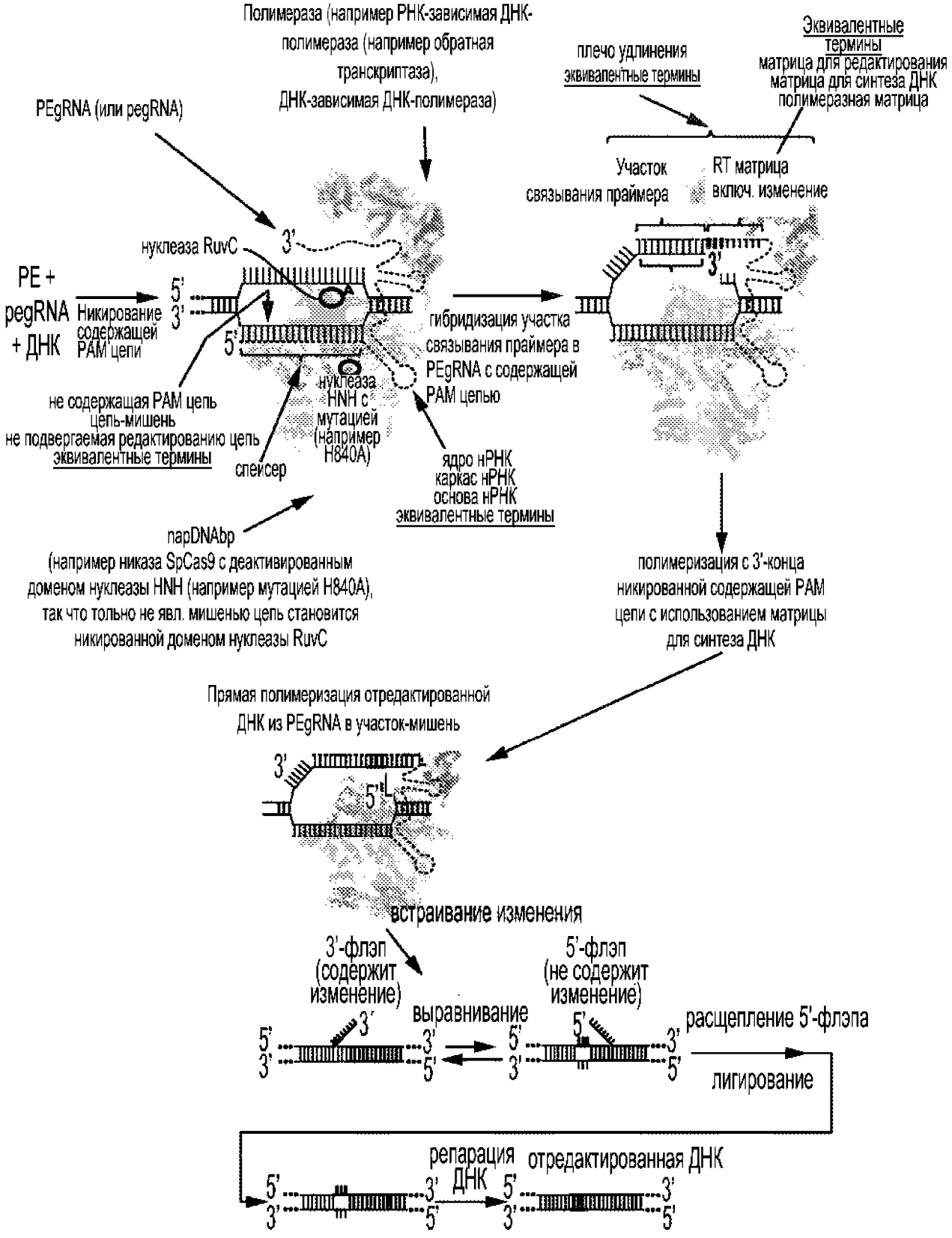
ПРОДОЛЖЕНИЕ



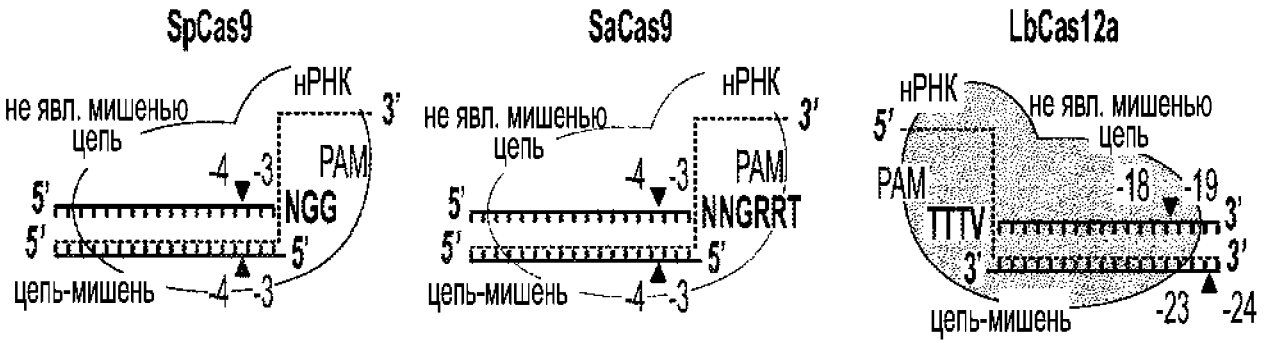
Объем праймерного редактирования

Все 4 переходные точечные мутации
Все 8 трансверсивных точечных мутаций
Вставки (от 1 bp до >44 bp)
Делеции (от 1 bp до >80 bp)
Комбинации указанных выше

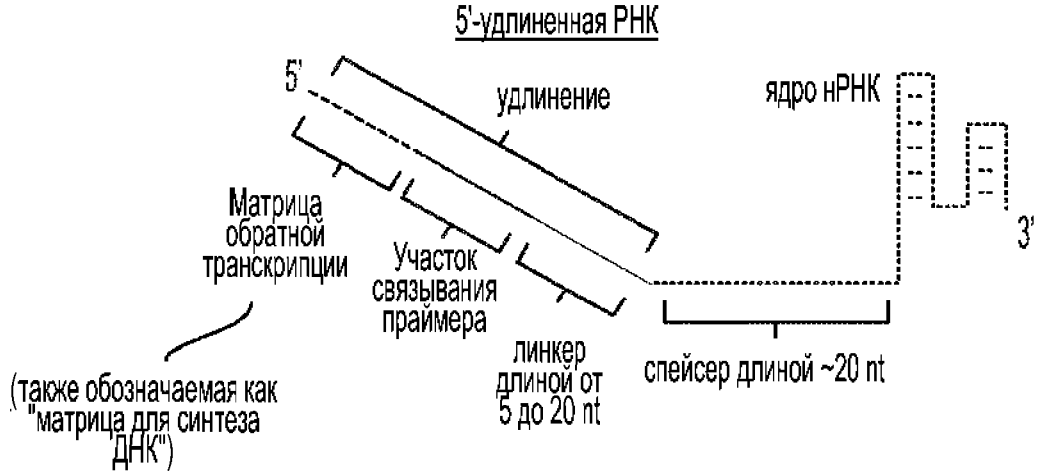
ФИГ.1K


Полимераза (например РНК-зависимая ДНК-полимераза (например обратная транскриптаза), ДНК-зависимая ДНК-полимераза)
плечо удлинения
эквивалентные термины
Эквивалентные термины
матрица для редактирования
матрица для синтеза ДНК
полимеразная матрица
PEgRNA (или pegRNA)
Участок связывания праймера
RT матрица включ. изменение
нуклеаза RuvC
PE + pegRNA + ДНК
Никирование содержащей PAM цепи
гибридизация участка связывания праймера в PEgRNA с содержащей PAM цепью
не содержащая PAM цепь
цепь-мишень
не подвергаемая редактированию цепь
эквивалентные термины
спейсер
нуклеаза HNH с мутацией (например H840A)
ядро нРНК
каркас нРНК
основа нРНК
эквивалентные термины
napDNAbp
(например никаза SpCas9 с деактивированным доменом нуклеазы HNH (например мутацией H840A), так что только не явл. мишенью цепь становится никированной доменом нуклеазы RuvC
полимеризация с 3'-конца никированной содержащей PAM цепи с использованием матрицы для синтеза ДНК
Прямая полимеризация отредактированной ДНК из PEgRNA в участок-мишень
встраивание изменения
3'-флэп (содержит изменение)
5'-флэп (не содержит изменение)
выравнивание
расщепление 5'-флэпа
лигирование
репарация ДНК
отредактированная ДНК

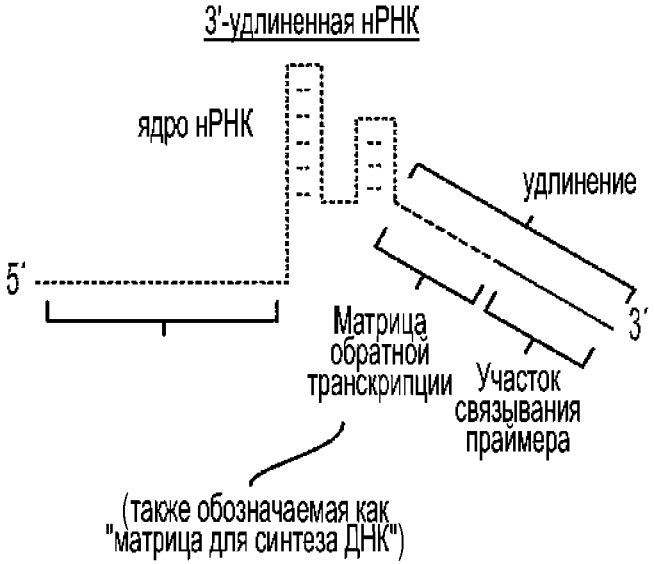
ФИГ.1L


SpCas9
не явл. мишенью цепь
нРНК
3'
PAM
NGG
5'
5'
-4
-3
цепь-мишень
SaCas9
не явл. мишенью цепь
нРНК
3'
PAM
NNGRRT
цепь-мишень
LbCas12a
нРНК
не явл. мишенью цепь
5'
PAM
TTTV
-18
-19
3'
цепь-мишень
-23
-24

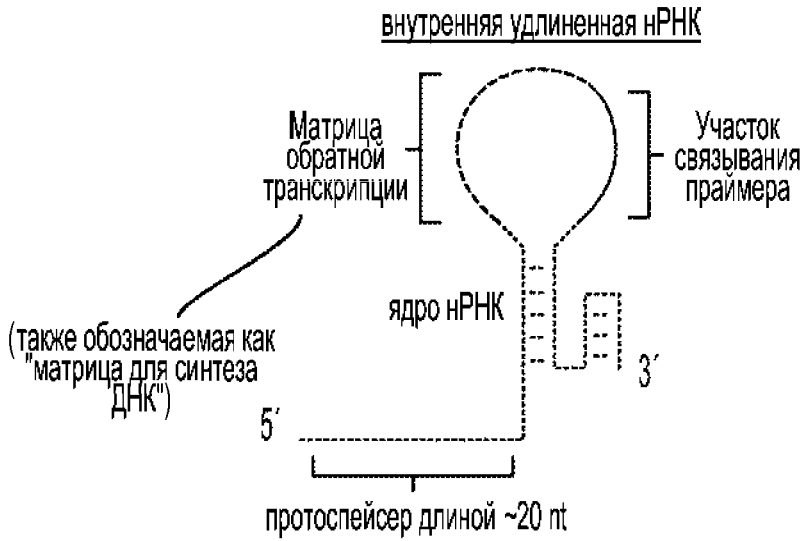
ФИГ.2


5'-удлиненная РНК
5'
удлинение
ядро нРНК
3'
Матрица обратной транскрипции
Участок связывания праймера
линкер длиной от 5 до 20 nt
спейсер длиной ~20 nt
(также обозначаемая как "матрица для синтеза ДНК")

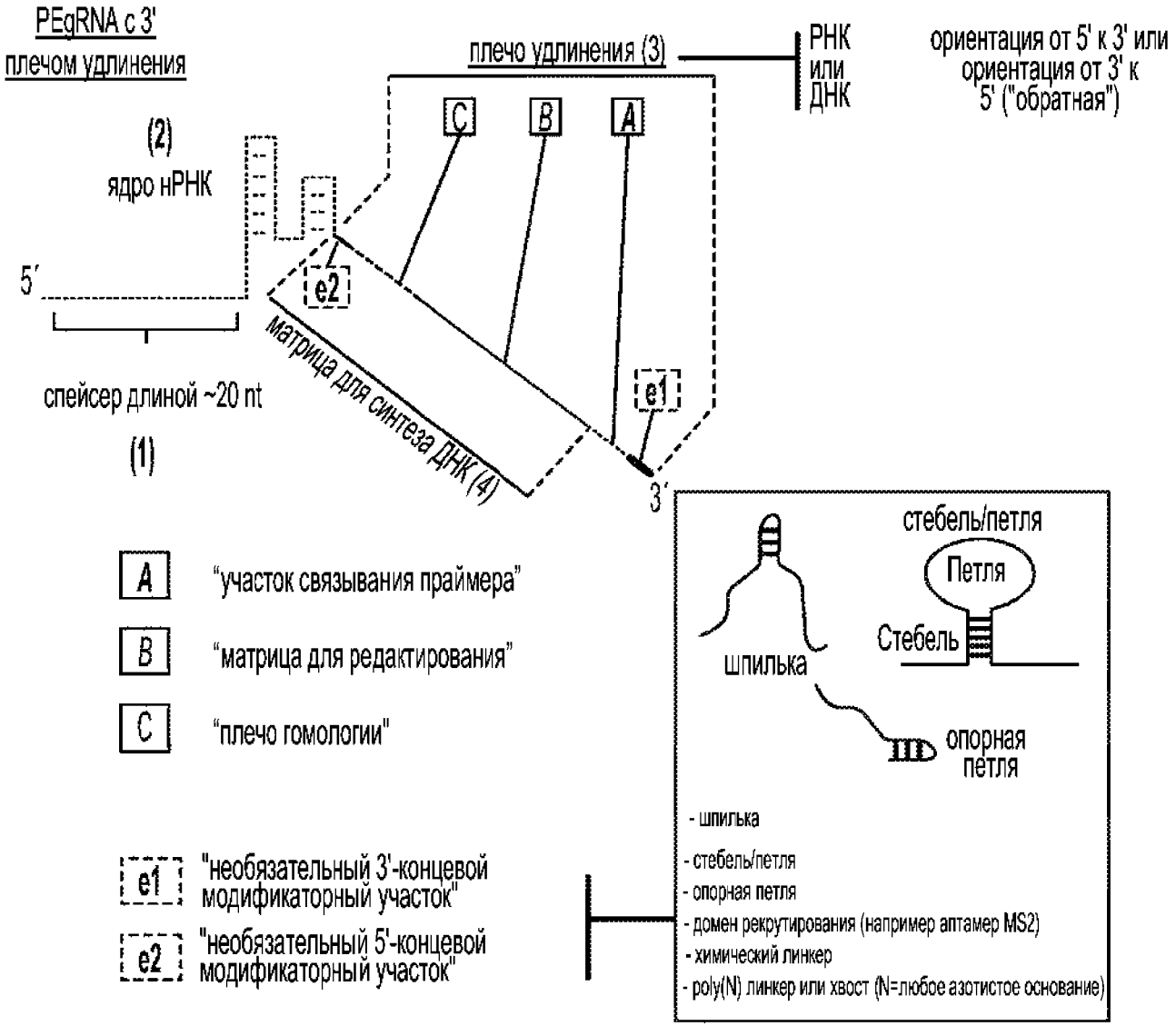
ФИГ.3А


3'-удлиненная нРНК
ядро нРНК
удлинение
5´
3´
Матрица обратной транскрипции
Участок связывания праймера
(также обозначаемая как "матрица для синтеза ДНК")

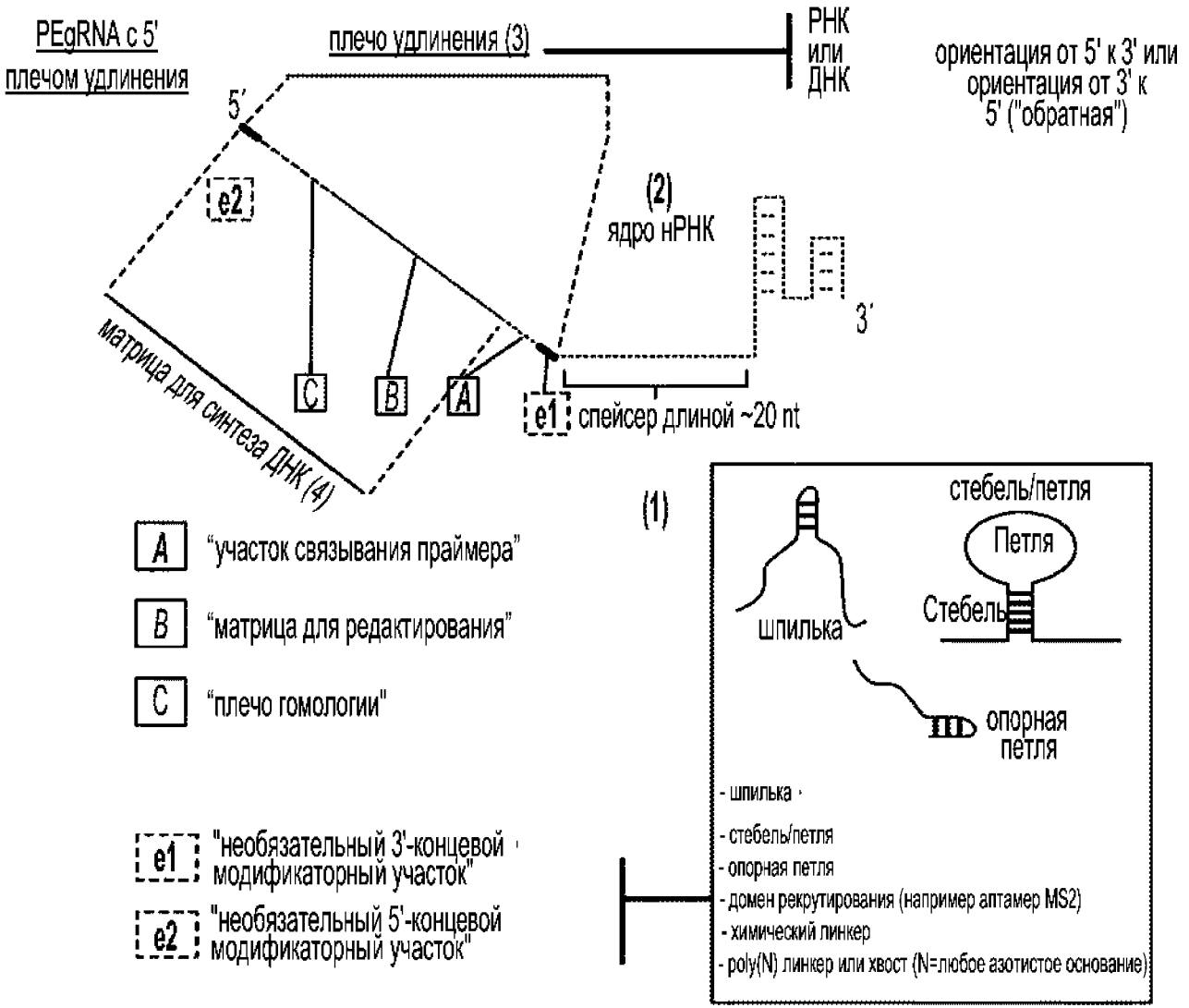
ФИГ.В


внутренняя удлиненная нРНК
Матрица обратной транскрипции
Участок связывания праймера
(также обозначаемая как "матрица для синтеза ДНК")
ядро нРНК
3´
5´
протоспейсер длиной ~20 nt

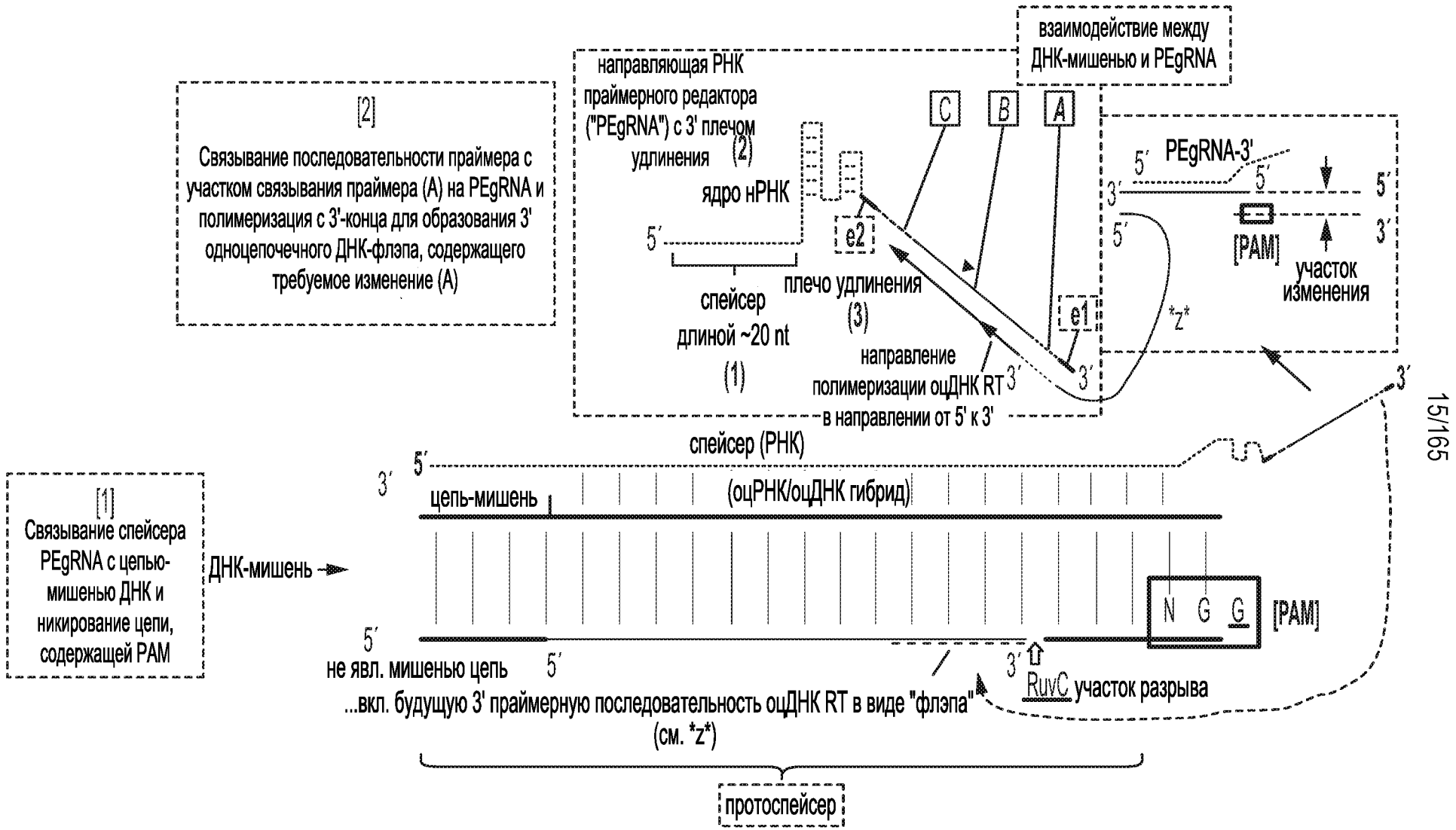
ФИГ.3С


PEgRNA с 3'
плечом удлинения
(2)
ядро нРНК
5'
спейсер длиной ~20 nt
(1)
плечо удлинения (3)
РНК
или
ДНК
ориентация от 5' к 3' или
ориентация от 3' к
5' ("обратная")
C
B
A
e2
e1
матрица для синтеза ДНК (4)
3'
A
"участок связывания праймера"
B
"матрица для редактирования"
C
"плечо гомологии"
e1
"необязательный 3'-концевой
модификаторный участок"
e2
"необязательный 5'-концевой
модификаторный участок"
стебель/петля
Петля
Стебель
шпилька
опорная
петля
- шпилька
- стебель/петля
- опорная петля
- домен рекрутирования (например аптамер MS2)
- химический линкер
- poly(N) линкер или хвост (N=любое азотистое основание)

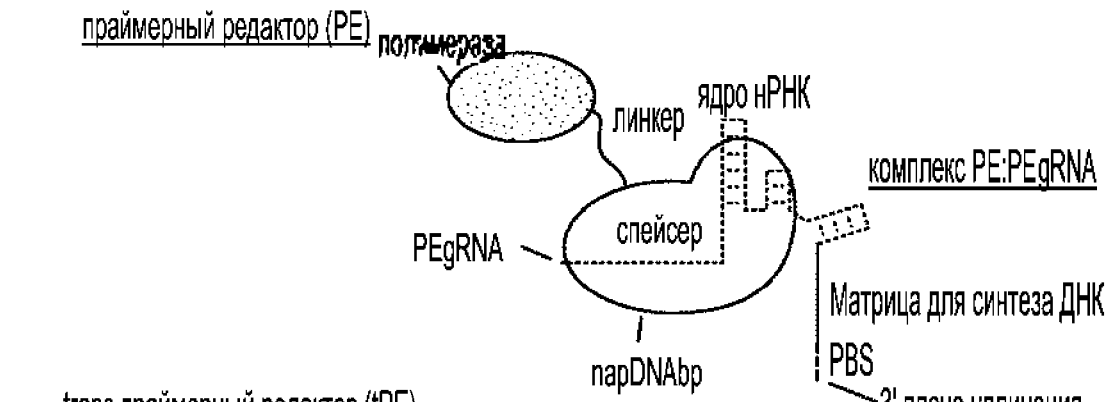
ФИГ.3D

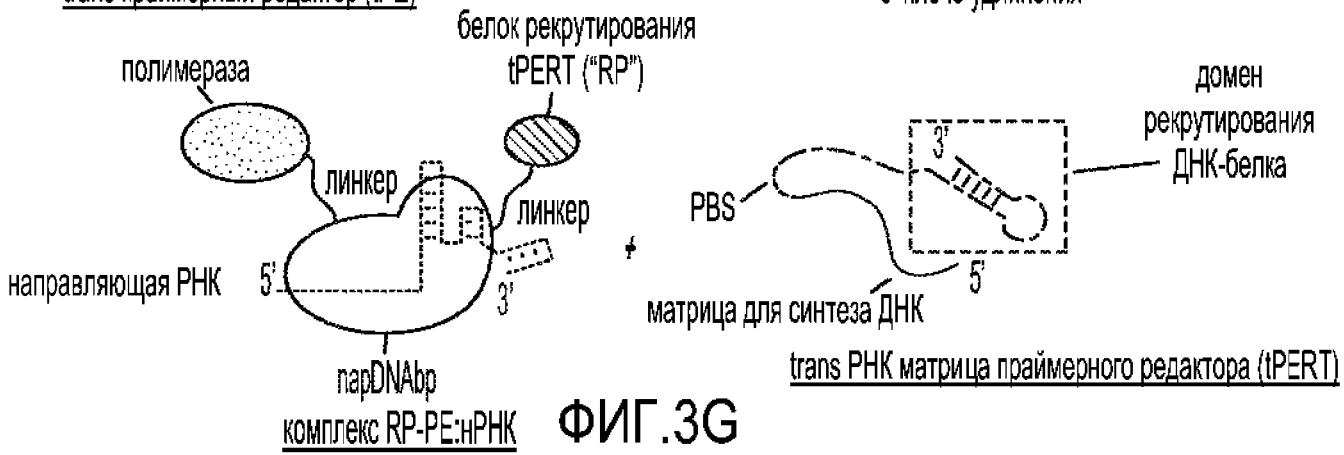

PEgRNA с 5' плечом удлинения
плечо удлинения (3)
РНК или ДНК
ориентация от 5' к 3' или ориентация от 3' к 5' ("обратная")
5'
e2
(2) ядро нРНК
3'
матрица для синтеза ДНК (4)
C
B
A
e1
спейсер длиной ~20 nt
A "участок связывания праймера"
B "матрица для редактирования"
C "плечо гомологии"
e1 "необязательный 3'-концевой модификаторный участок"
e2 "необязательный 5'-концевой модификаторный участок"
(1)
стебель/петля
Петля
Стебель
шпилька
опорная петля
- шпилька
- стебель/петля
- опорная петля
- домен рекрутирования (например аптамер MS2)
- химический линкер
- poly(N) линкер или хвост (N=любое азотистое основание)

ФИГ.3E

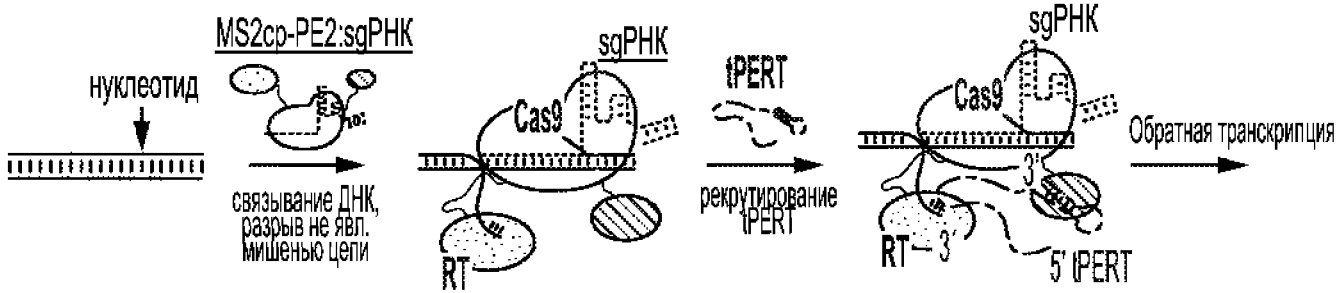

[2]
Связывание последовательности праймера с участком связывания праймера (A) на PEgRNA и полимеризация с 3'-конца для образования 3' одноцепочечного ДНК-флэпа, содержащего требуемое изменение (A)
направляющая РНК праймерного редактора ("PEgRNA") с 3' плечом удлинения
(2)
ядро нРНК
взаимодействие между ДНК-мишенью и PEgRNA
C
B
A
PEgRNA-3'
5'
3'
[PAM]
участок изменения
e2
e1
z
спейсер длиной ~20 nt
(1)
плечо удлинения
(3)
направление полимеризации оцДНК RT
в направлении от 5' к 3'
спейсер (РНК)
[1]
Связывание спейсера PEgRNA с цепью-мишенью ДНК и никирование цепи, содержащей PAM
ДНК-мишень
цепь-мишень
(оцРНК/оцДНК гибрид)
N G G
[PAM]
не явл. мишенью цепь
RuvC участок разрыва
...вкл. будущую 3' праймерную последовательность оцДНК RT в виде "флэпа"
(см. *z*)
протоспейсер

ФИГ.3F

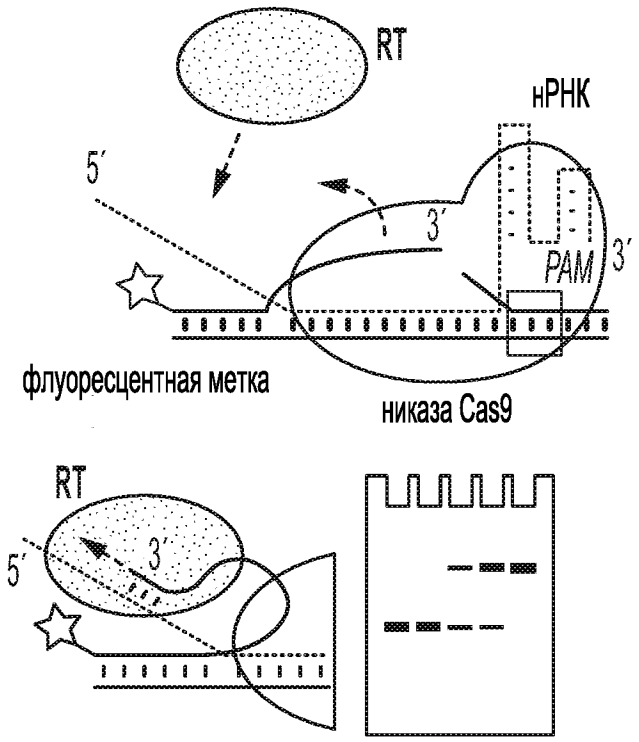

праймерный редактор (PE)
полимераза
линкер
ядро нРНК
комплекс PE:PEgRNA
спейсер
PEgRNA
Матрица для синтеза ДНК
PBS
napDNAbp
3' плечо удлинения


trans праймерный редактор (tPE)
белок рекрутирования
tPERT ("RP")
полимераза
линкер
линкер
направляющая РНК
5'
3'
napDNAbp
комплекс RP-PE:нРНК
домен
рекрутирования
ДНК-белка
PBS
3'
5'
матрица для синтеза ДНК
trans РНК матрица праймерного редактора (tPERT)

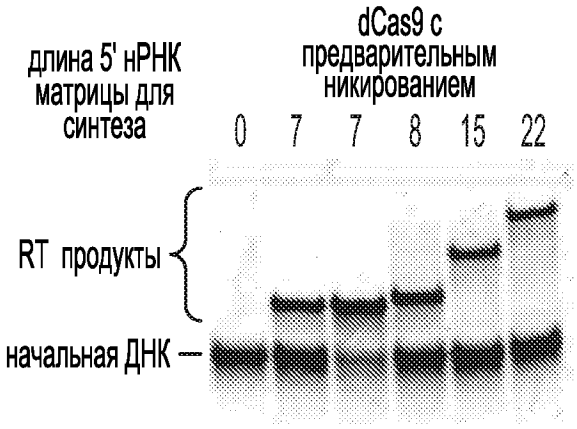
ФИГ.3G


MS2cp-PE2:sgРНК
нуклеотид
связывание ДНК,
разрыв не явл.
мишенью цепи
sgРНК
Cas9
RT
tPERT
рекрутирование
tPERT
sgРНК
Cas9
3'
RT
3'
5' tPERT
Обратная транскрипция

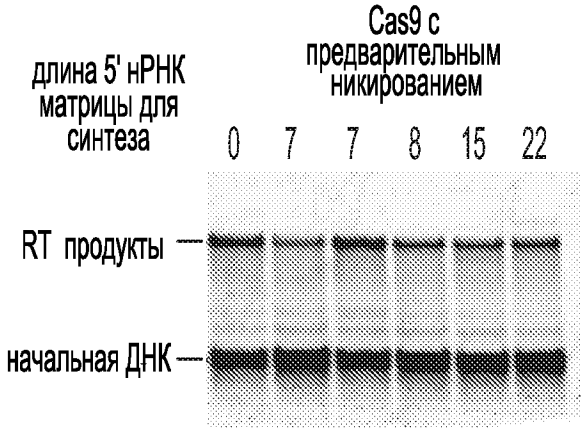
ФИГ.3H



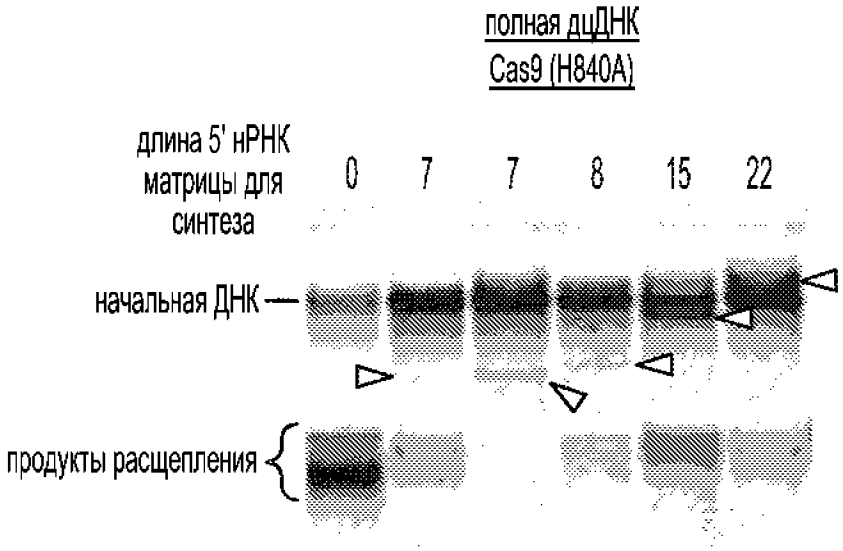
ФИГ.4А



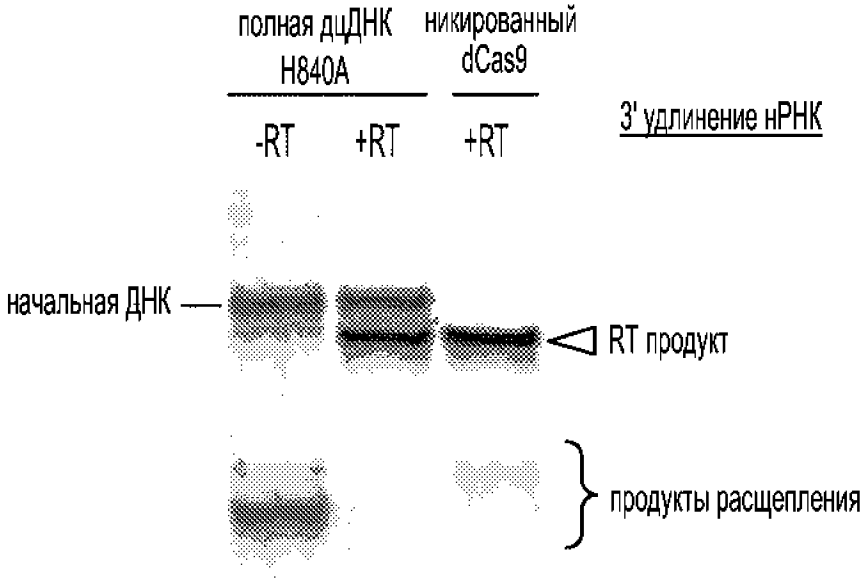
ФИГ.4В



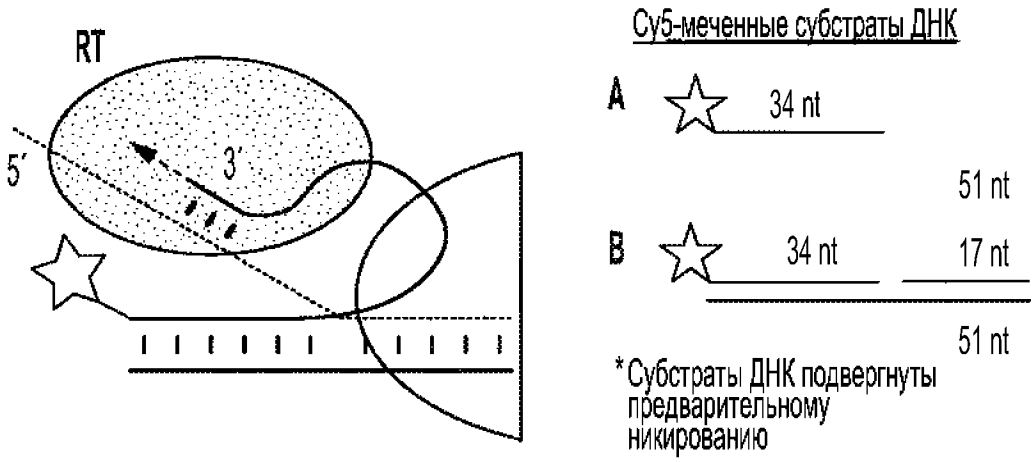
ФИГ.4С

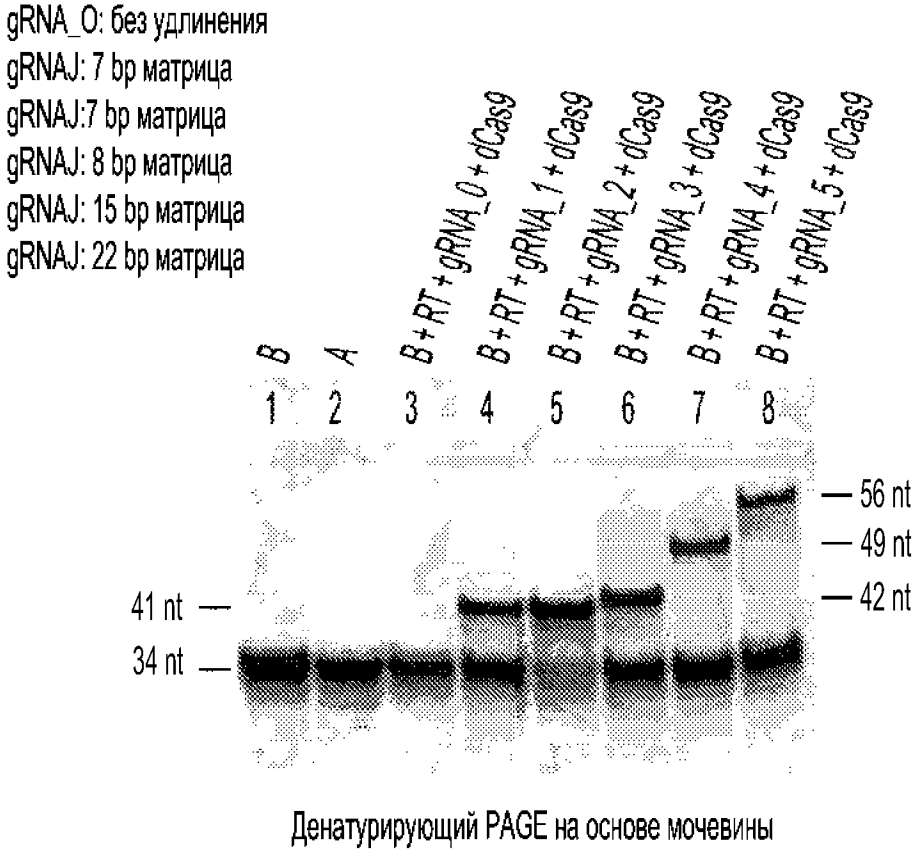

полная дцДНК
Cas9 (H840A)
длина 5' нРНК матрицы для синтеза
0
7
7
8
15
22
начальная ДНК
продукты расщепления

ФИГ.4D

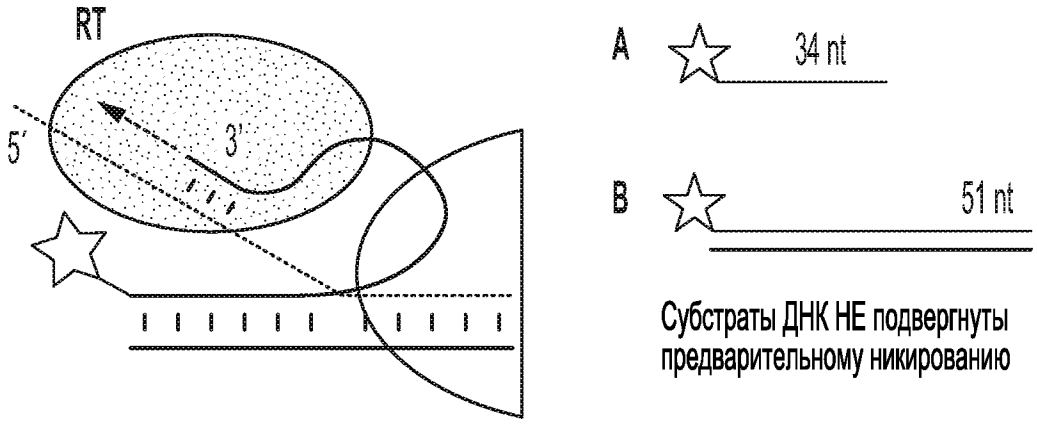

полная дцДНК H840A
никированный dCas9
-RT
+RT
+RT
3' удлинение нРНК
начальная ДНК
RT продукт
продукты расщепления

ФИГ.4E

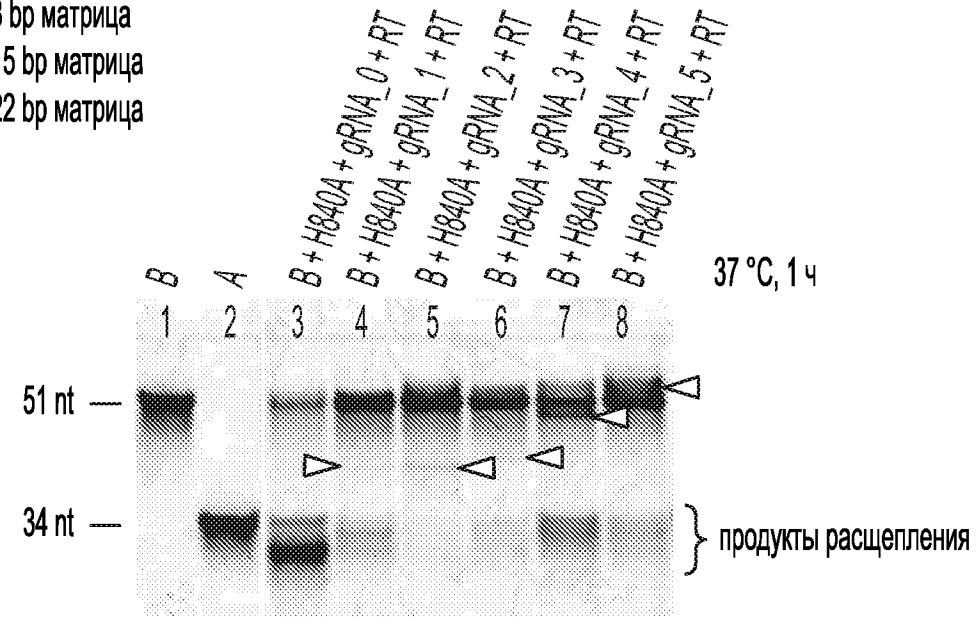

RT
5'
3'
Cy5-меченные субстраты ДНК
A
34 nt
51 nt
B
34 nt
17 nt
51 nt
*Субстраты ДНК подвергнуты предварительному никированию


gRNA_O: без удлинения
gRNAJ: 7 bp матрица
gRNAJ:7 bp матрица
gRNAJ: 8 bp матрица
gRNAJ: 15 bp матрица
gRNAJ: 22 bp матрица
B
A
B+RT+gRNA_0+dCas9
B+RT+gRNA_1+dCas9
B+RT+gRNA_2+dCas9
B+RT+gRNA_3+dCas9
B+RT+gRNA_4+dCas9
B+RT+gRNA_5+dCas9
1
2
3
4
5
6
7
8
56 nt
49 nt
42 nt
41 nt
34 nt
Денатурирующий PAGE на основе мочевины

ФИГ.5

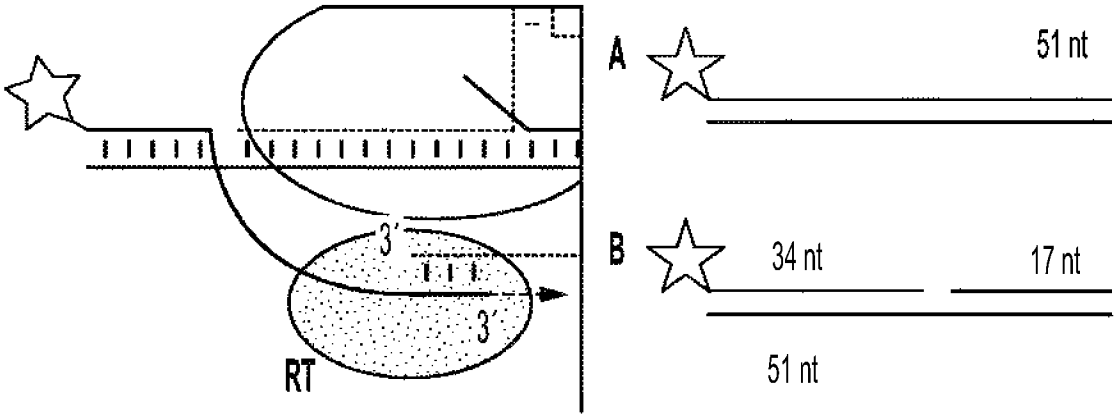


gRNA_0: без удлинения
gRNAJ: 7 bp матрица
gRNA_2:7 bp матрица
gRNAJ: 8 bp матрица
gRNAJ: 15 bp матрица
gRNAJ: 22 bp матрица

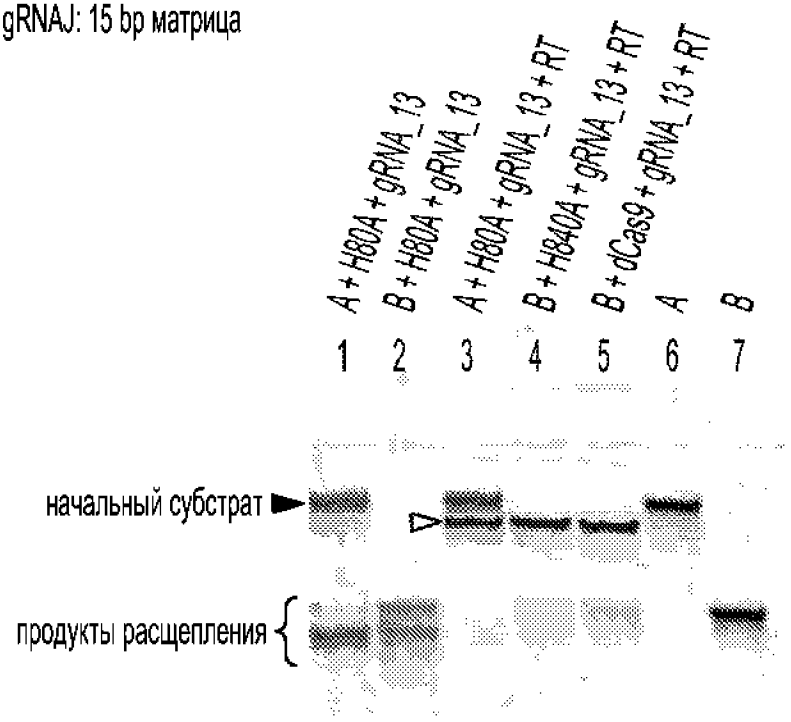


15% денатурирующий PAGE на основе мочевины

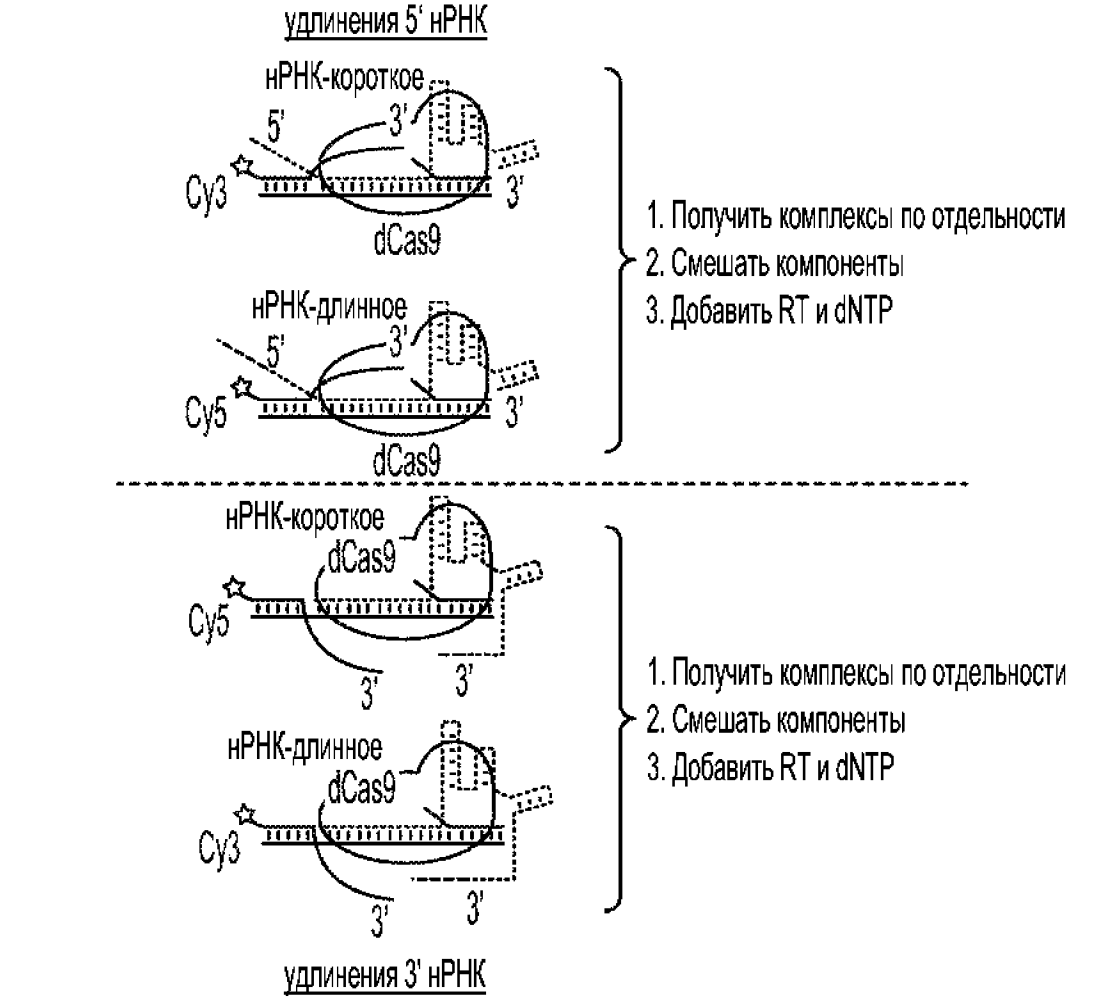
ФИГ.6



gRNAJ: без удлинения
gRNAJ: 15 bp матрица

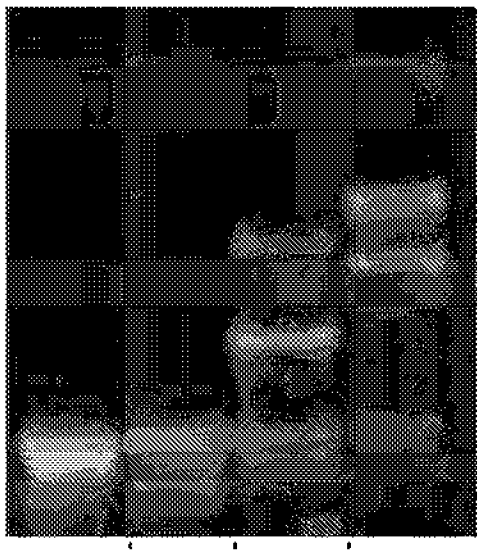


ФИГ.7



$$\frac{cis_{Cy3}/trans_{Cy3}}{trans_{Cy5}/cis_{Cy5}}$$

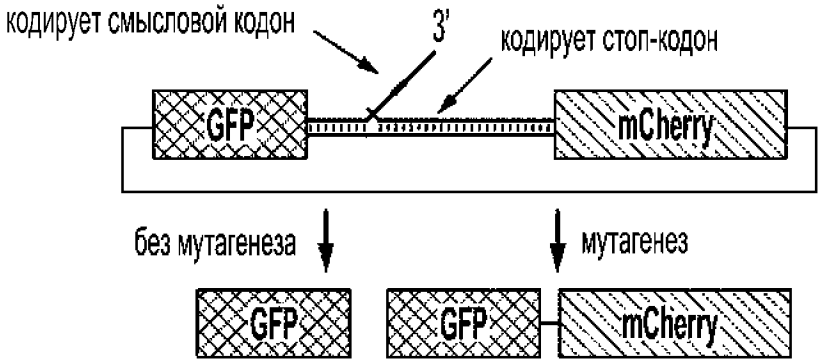
Соотношения продуктов: 7.4 4.7



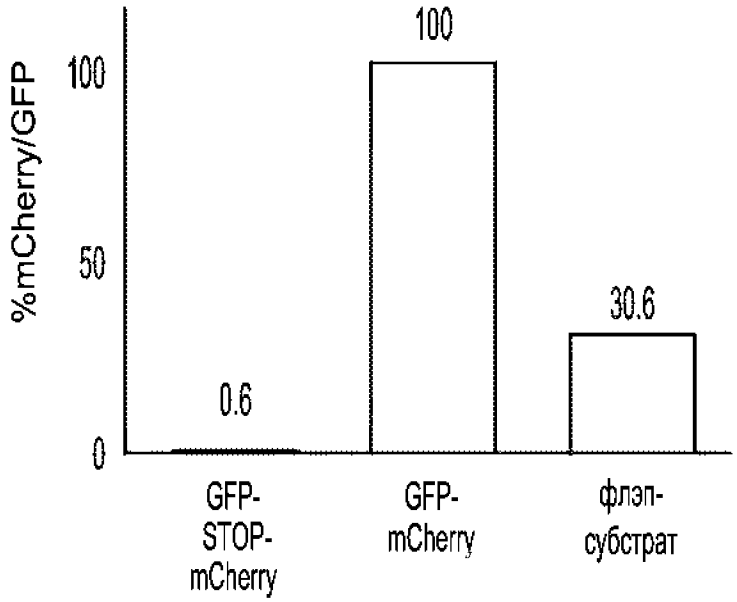
продукты RT

Cy3 | Cy5 | 5'-удл. | 3'-удл.

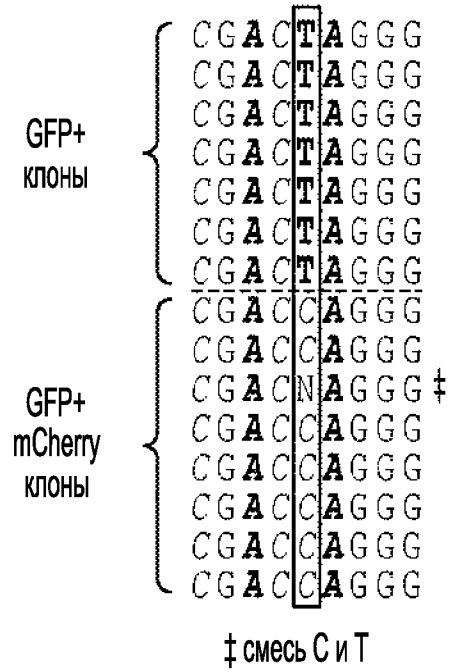
ФИГ.8


кодирует смысловой кодон
3'
кодирует стоп-кодон
GFP
mCherry
без мутагенеза
мутагенез
GFP
GFP
mCherry

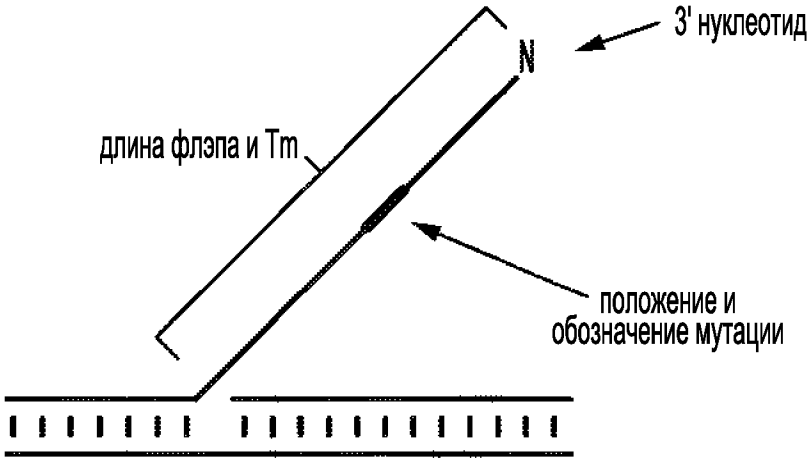
ФИГ.9А


%mCherry/GFP
100
50
0
100
0.6
30.6
GFP-STOP-mCherry
GFP-mCherry
флэп-субстрат

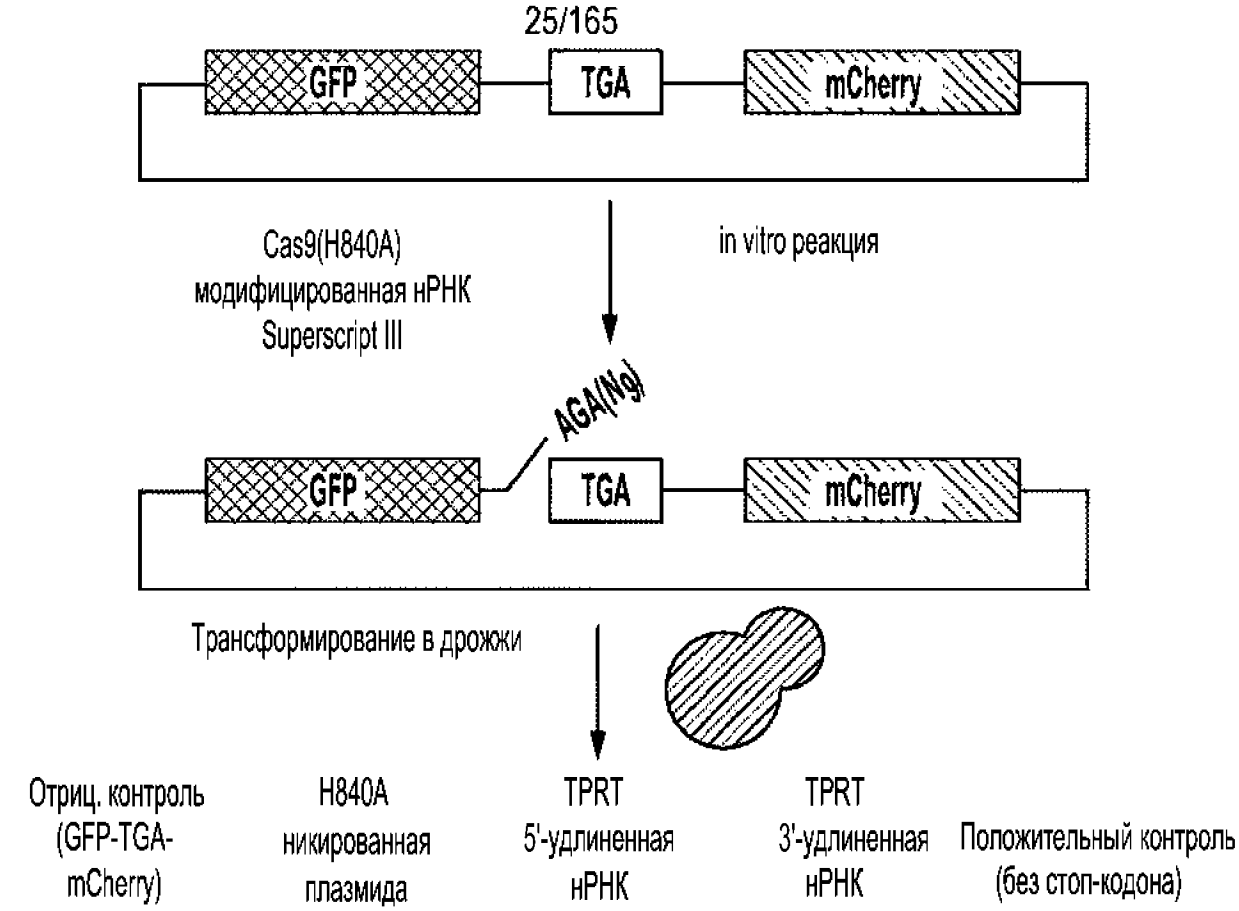
ФИГ.9В


GFP+
клоны
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
GFP+
mCherry
клоны
CGACCAGGG
CGACCAGGG
CGACNAGGG ‡
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
‡ смесь C и T

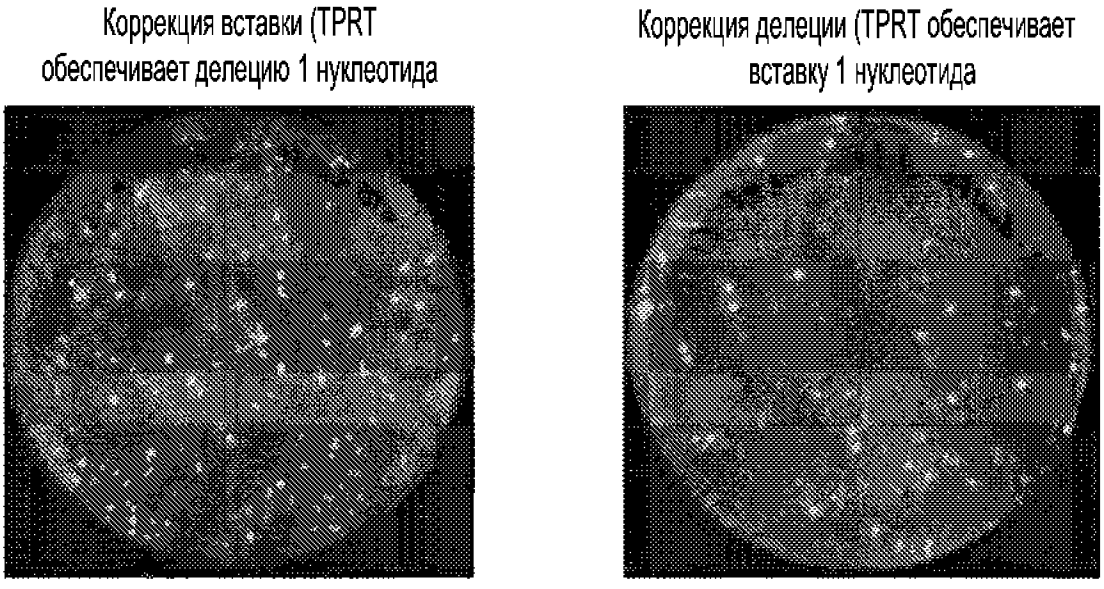
ФИГ.9C


3' нуклеотид
N
длина флэпа и Tm
положение и
обозначение мутации

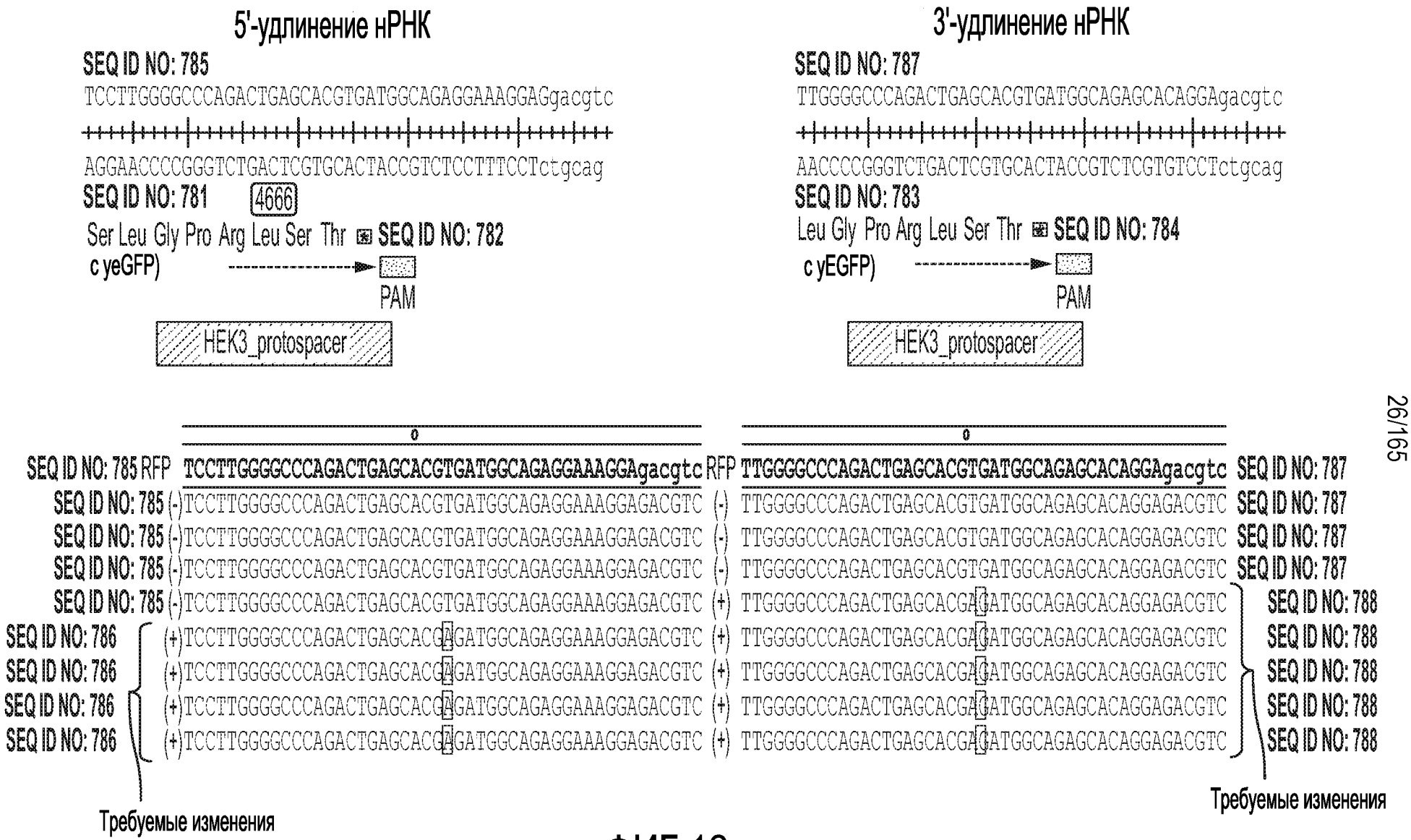
ФИГ.9D


GFP
TGA
mCherry
Cas9(H840A)
модифицированная нРНК
Superscript III
in vitro реакция
AGA(Ng)
GFP
TGA
mCherry
Трансформирование в дрожжи
Отриц. контроль (GFP-TGA-mCherry)
H840A никированная плазмида
TPRT 5'-удлиненная нРНК
TPRT 3'-удлиненная нРНК
Положительный контроль (без стоп-кодона)

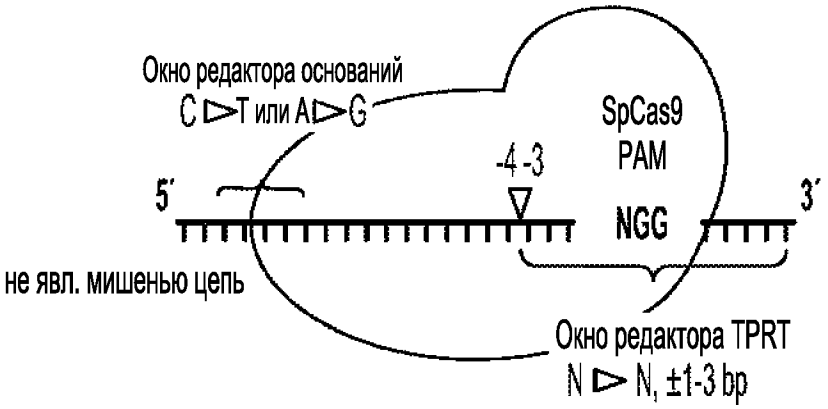
ФИГ.10


Коррекция вставки (TPRT обеспечивает делецию 1 нуклеотида
Коррекция делеции (TPRT обеспечивает вставку 1 нуклеотида

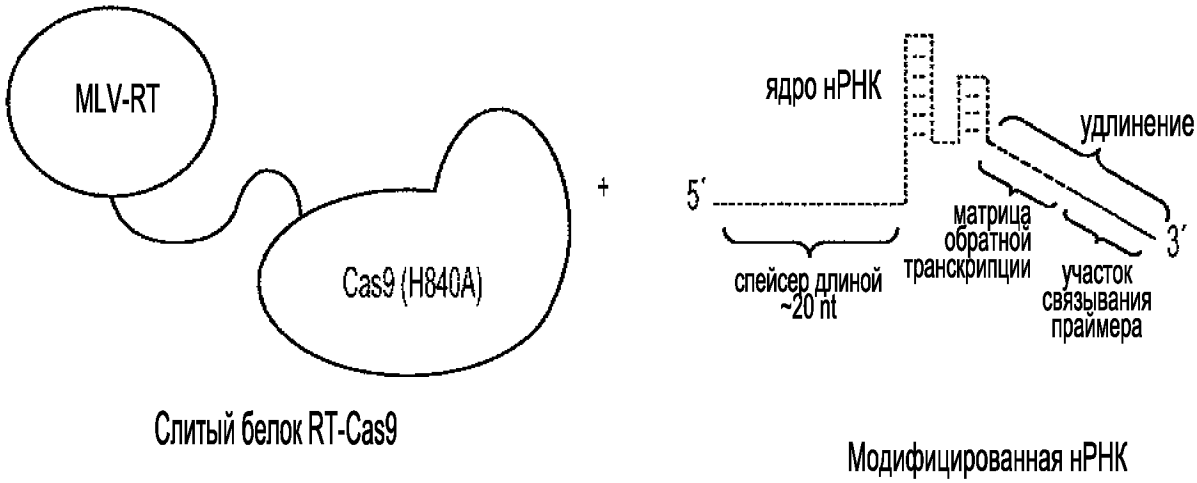
ФИГ.11


5'-удлинение нРНК
SEQ ID NO: 785
TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGgacgtc
AGGAACCCCGGGTCTGACTCGTGCACTACCGTCTCCTTTCCTctgcag
SEQ ID NO: 781
4666
Ser Leu Gly Pro Arg Leu Ser Thr
SEQ ID NO: 782
c yeGFP)
PAM
HEK3_protospacer
3'-удлинение нРНК
SEQ ID NO: 787
TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAgacgtc
AACCCCGGGTCTGACTCGTGCACTACCGTCTCGTGTCCTctgcag
SEQ ID NO: 783
Leu Gly Pro Arg Leu Ser Thr
SEQ ID NO: 784
c yEGFP)
PAM
HEK3_protospacer
0
0
SEQ ID NO: 785 RFP TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAgacgtc RFP TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAgacgtc SEQ ID NO: 787
SEQ ID NO: 785 (-) TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGACGTC (-) TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 787
SEQ ID NO: 785 (-) TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGACGTC (-) TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 787
SEQ ID NO: 785 (-) TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGACGTC (-) TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 787
SEQ ID NO: 785 (-) TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGACGTC (+) TTGGGGCCCAGACTGAGCACGAGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 788
SEQ ID NO: 786 (+) TCCTTGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAGACGTC (+) TTGGGGCCCAGACTGAGCACGAGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 788
SEQ ID NO: 786 (+) TCCTTGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAGACGTC (+) TTGGGGCCCAGACTGAGCACGAGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 788
SEQ ID NO: 786 (+) TCCTTGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAGACGTC (+) TTGGGGCCCAGACTGAGCACGAGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 788
SEQ ID NO: 786 (+) TCCTTGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAGACGTC (+) TTGGGGCCCAGACTGAGCACGAGATGGCAGAGCACAGGAGACGTC SEQ ID NO: 788
Требуемые изменения
Требуемые изменения

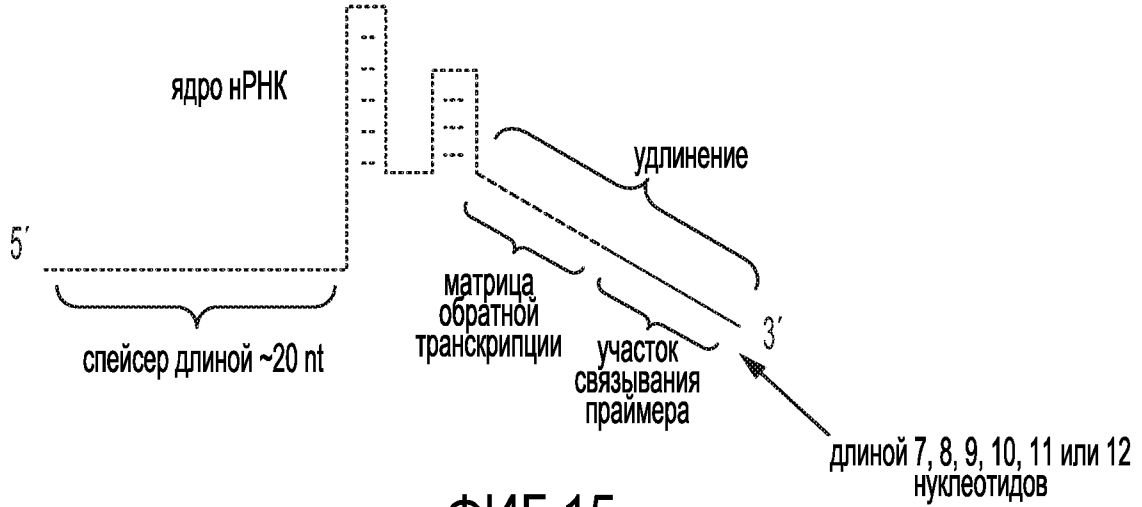
ФИГ.12


Окно редактора оснований
C ▷T или A ▷G
SpCas9
PAM
-4 -3
5′
3′
NGG
не явл. мишенью цепь
Окно редактора TPRT
N ▷ N, ±1-3 bp

ФИГ.13


MLV-RT
+
Cas9 (H840A)
Слитый белок RT-Cas9
ядро нРНК
удлинение
5′
3′
матрица обратной транскрипции
спейсер длиной ~20 nt
участок связывания праймера
Модифицированная нРНК

ФИГ.14



ФИГ.15

локус HEK3

SEQ ID NO: 789

5′ ATGTGGGCTGCCTAGAAAGGCATGGATGAGAGAAGCCTGGAGACAGGGATCCCAGGGAAACGCCCATGCAATTAGTCTATTTCTGCTGCAAGTAAGCATG

3′ TACACCCGACGGATCTTTCCGTACCTACTCTCTTCGGACCTCTGTCCCTAGGGTCCCTTTGCGGGTACGTTAATCAGATAAAGACGACGTTCATTCGTAC

SEQ ID NO: 790

spacer_3

CATTTGTAGGCTTGATGCTTTTTTTCTGCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAGCCC

GTAAACATCCGAACTACGAAAAAAAGACGAAGAGGTCGGGACCGGACCCAGTTAGGAACCCCGGGTCTGACTCGTGCACTACCGTCTCCTTTCCTTCGGG

spacer_3 B

spacer_2 A

мишень редактирования

PAM

TGCTTCCTCCAGAGGGCGTCGCAGGACAGCTTTTCCTAGACAGGGGCTAGTATGTGCAGCTCCTGCACCGGGATACTGGTTGACAAGTTTGGCTGGG 3′

ACGAAGGAGGTCTCCCGCAGCGTCCTGTCGAAAAGGATCTGTCCCCGATCATACACGTCGAGGACGTGGCCCTATGACCAACTGTTCAAACCGACCC 5′

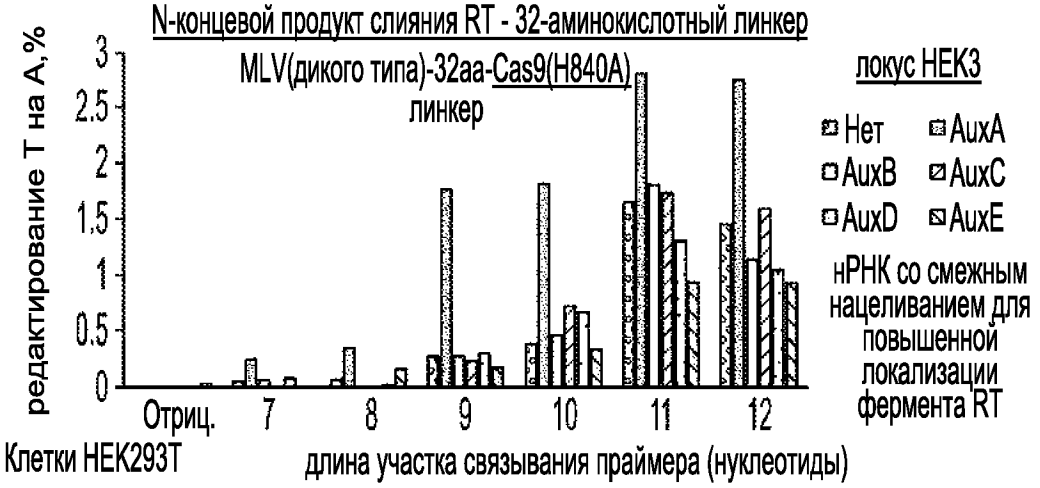
297

spacer_4_rev C

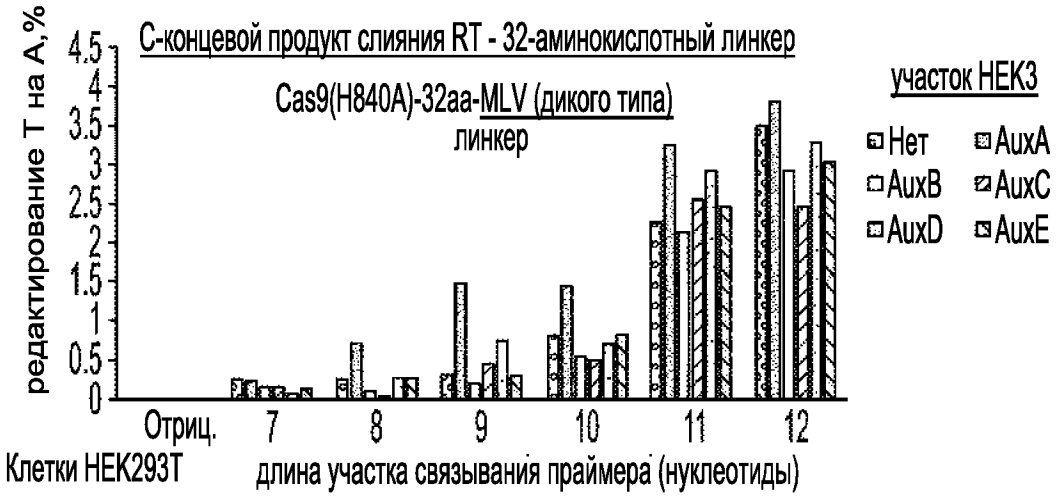
spacer_5_rev D

spacer_6 E

ФИГ.16



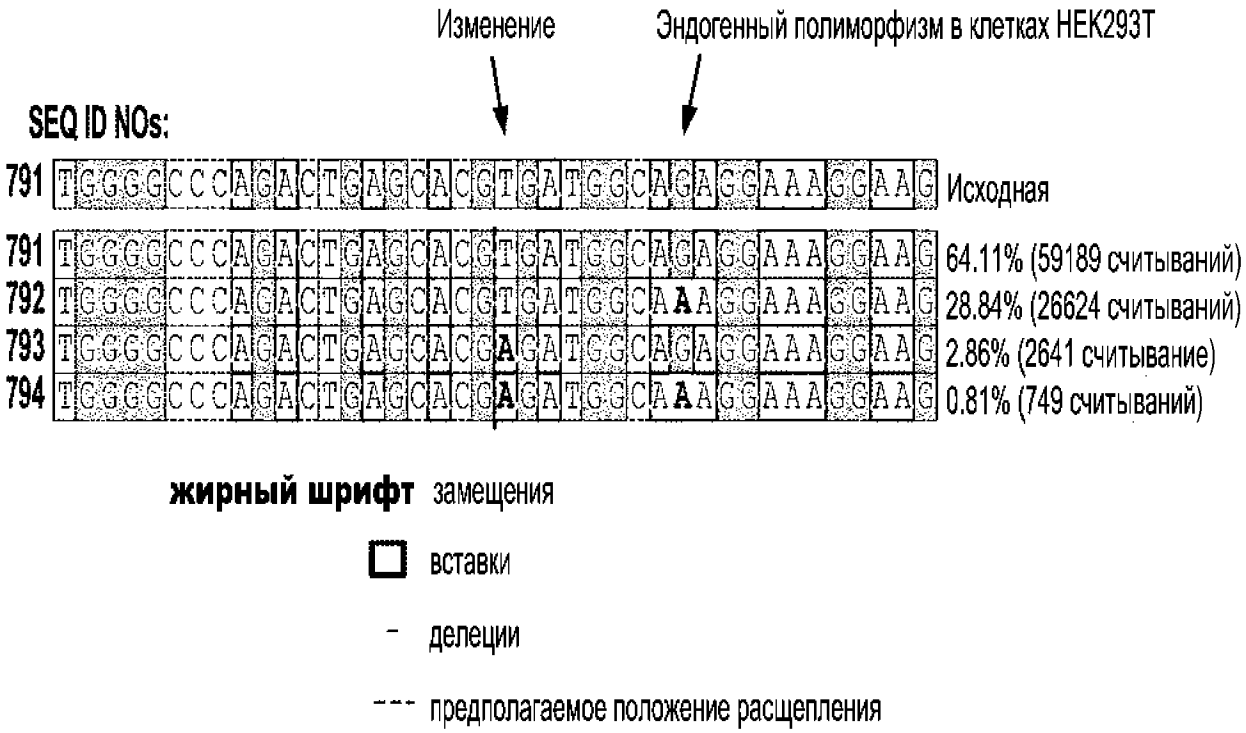
ФИГ.17A



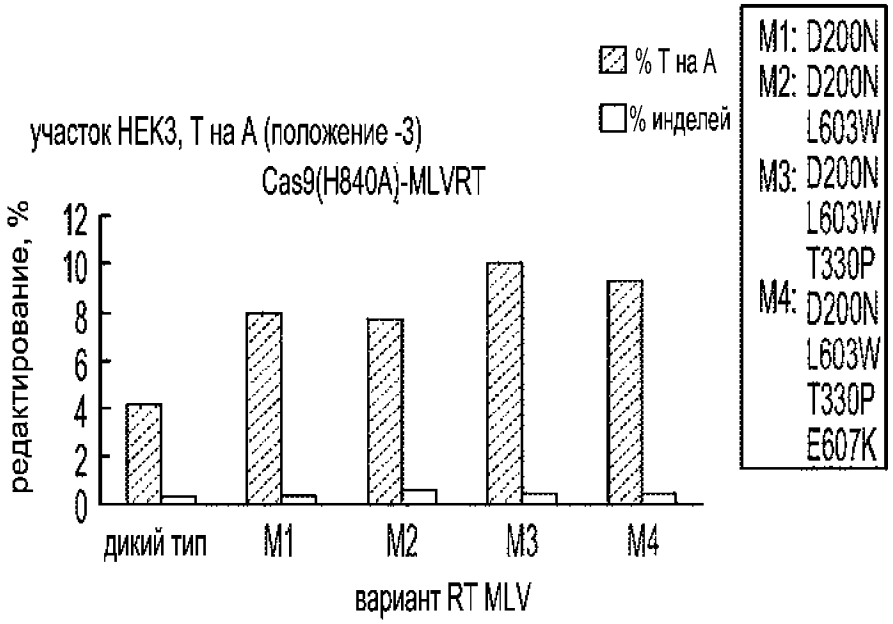
ФИГ.17B



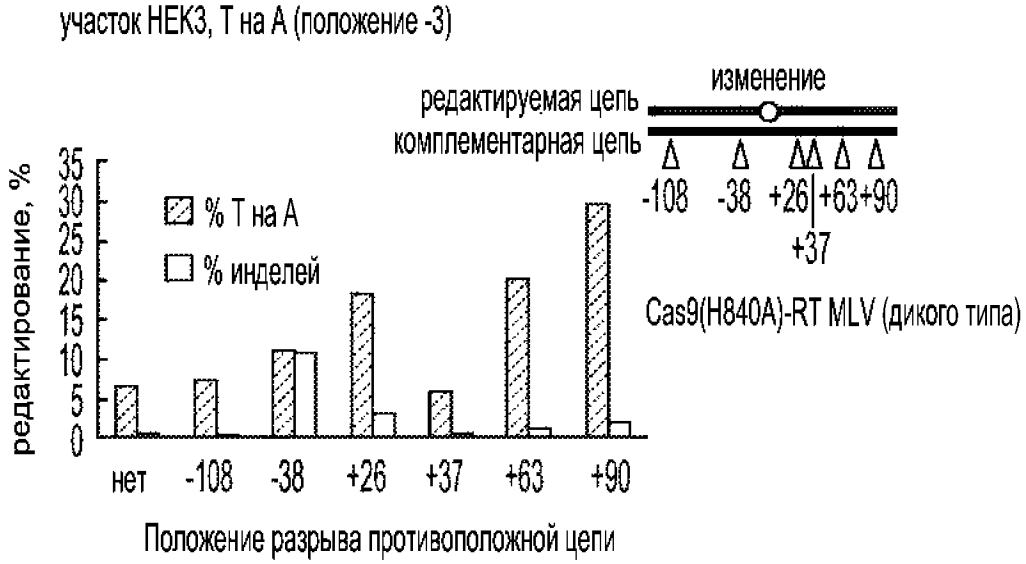
ФИГ.17C


Изменение
Эндогенный полиморфизм в клетках HEK293T
SEQ ID NOs:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG Исходная
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 64.11% (59189 считываний)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 28.84% (26624 считываний)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 2.86% (2641 считывание)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 0.81% (749 считываний)
жирный шрифт замещения
вставки
делеции
предполагаемое положение расщепления

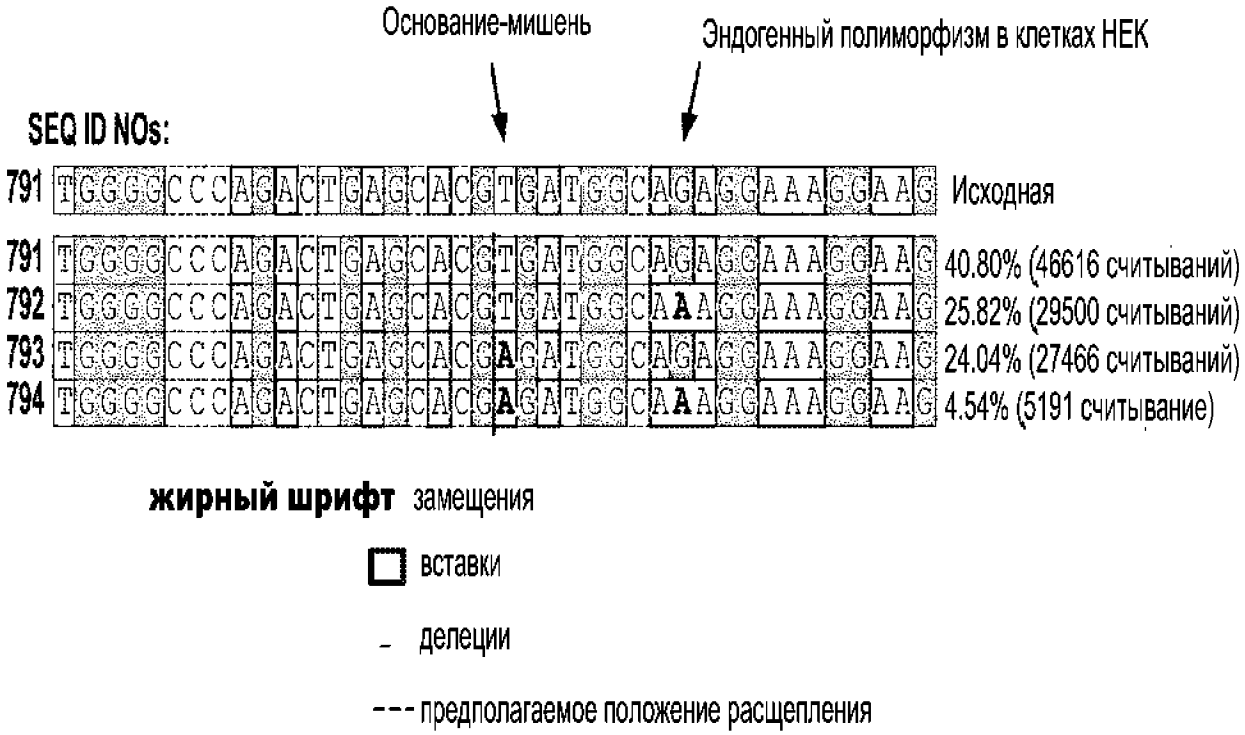
ФИГ.18


% T на A
% инделей
M1: D200N
M2: D200N
L603W
M3: D200N
L603W
T330P
M4: D200N
L603W
T330P
E607K
участок HEK3, T на A (положение -3)
Cas9(H840A)-MLVRT
редактирование, %
0
2
4
6
8
10
12
дикий тип
M1
M2
M3
M4
вариант RT MLV

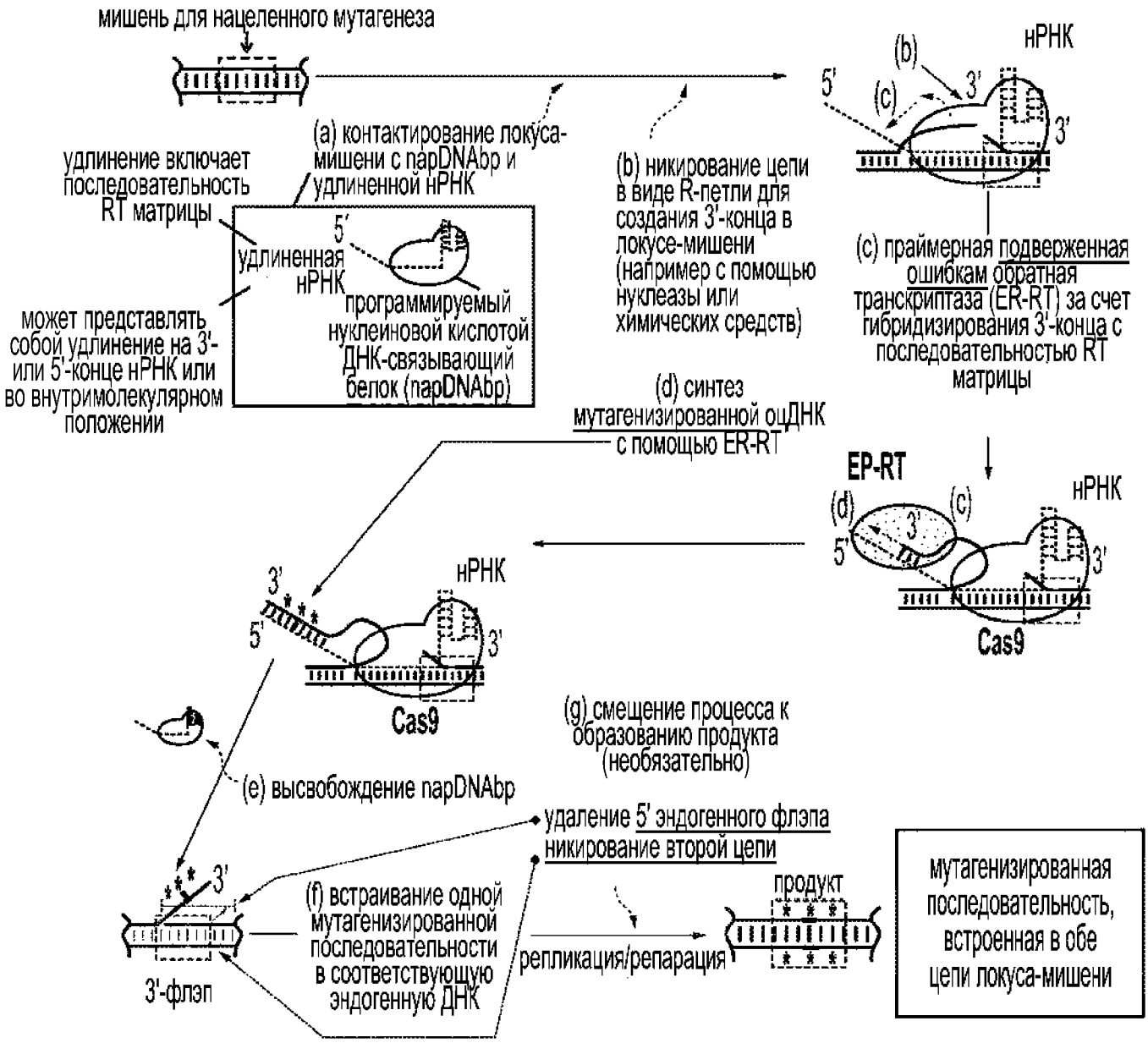
ФИГ.19


участок HEK3, T на A (положение -3)
редактируемая цепь
комплементарная цепь
изменение
Δ Δ ΔΔ Δ Δ
-108 -38 +26 +63 +90
+37
Cas9(H840A)-RT MLV (дикого типа)
% T на A
% инделей
редактирование, %
0
5
10
15
20
25
30
35
нет
-108
-38
+26
+37
+63
+90
Положение разрыва противоположной цепи

ФИГ.20


Основание-мишень
Эндогенный полиморфизм в клетках HEK
SEQ ID NOs:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG Исходная
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 40.80% (46616 считываний)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 25.82% (29500 считываний)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 24.04% (27466 считываний)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 4.54% (5191 считывание)
жирный шрифт замещения
вставки
делеции
предполагаемое положение расщепления

ФИГ.21

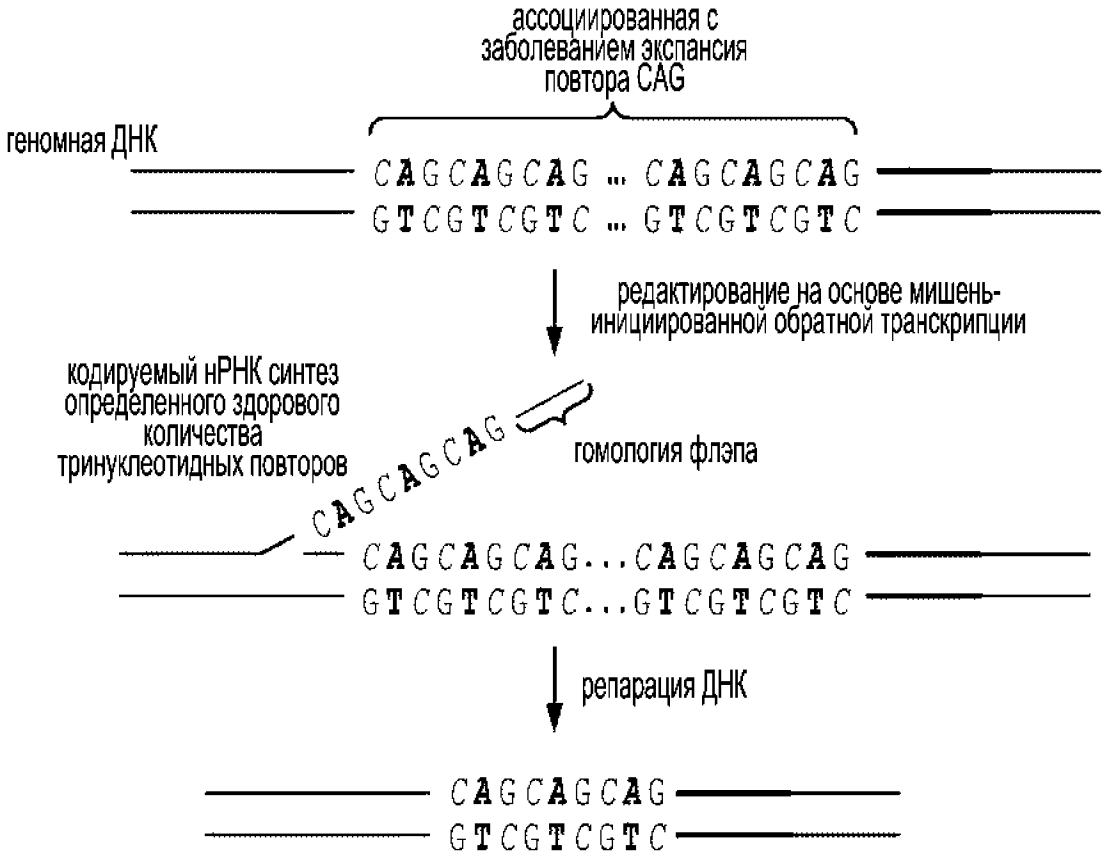

мишень для нацеленного мутагенеза
(a) контактирование локуса-мишени с napDNAbp и удлиненной нРНК
удлинение включает последовательность RT матрицы
удлиненная нРНК
5'
программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp)
может представлять собой удлинение на 3'- или 5'-конце нРНК или во внутримолекулярном положении
(b) никирование цепи в виде R-петли для создания 3'-конца в локусе-мишени (например с помощью нуклеазы или химических средств)
нРНК
(b)
(c)
5'
3'
3'
(c) праймерная подверженная ошибкам обратная транскриптаза (ER-RT) за счет гибридизирования 3'-конца с последовательностью RT матрицы
(d) синтез мутагенизированной оцДНК с помощью ER-RT
EP-RT
нРНК
(d)
(c)
5'
3'
3'
Cas9
3'
5'
нРНК
3'
Cas9
(g) смещение процесса к образованию продукта (необязательно)
(e) высвобождение napDNAbp
удаление 5' эндогенного флэпа
никирование второй цепи
3'
(f) встраивание одной мутагенизированной последовательности в соответствующую эндогенную ДНК
продукт
репликация/репарация
3'-флэп
мутагенизированная последовательность, встроенная в обе цепи локуса-мишени

ФИГ.22

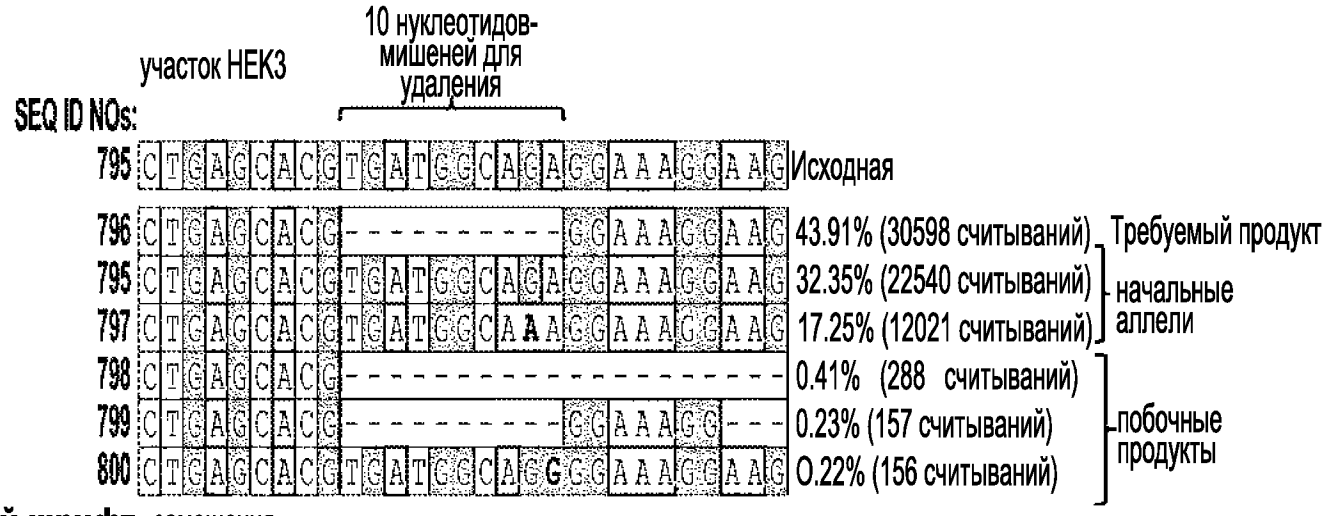
Конструирование нРНК для сокращения последовательностей тринуклеотидных повторов



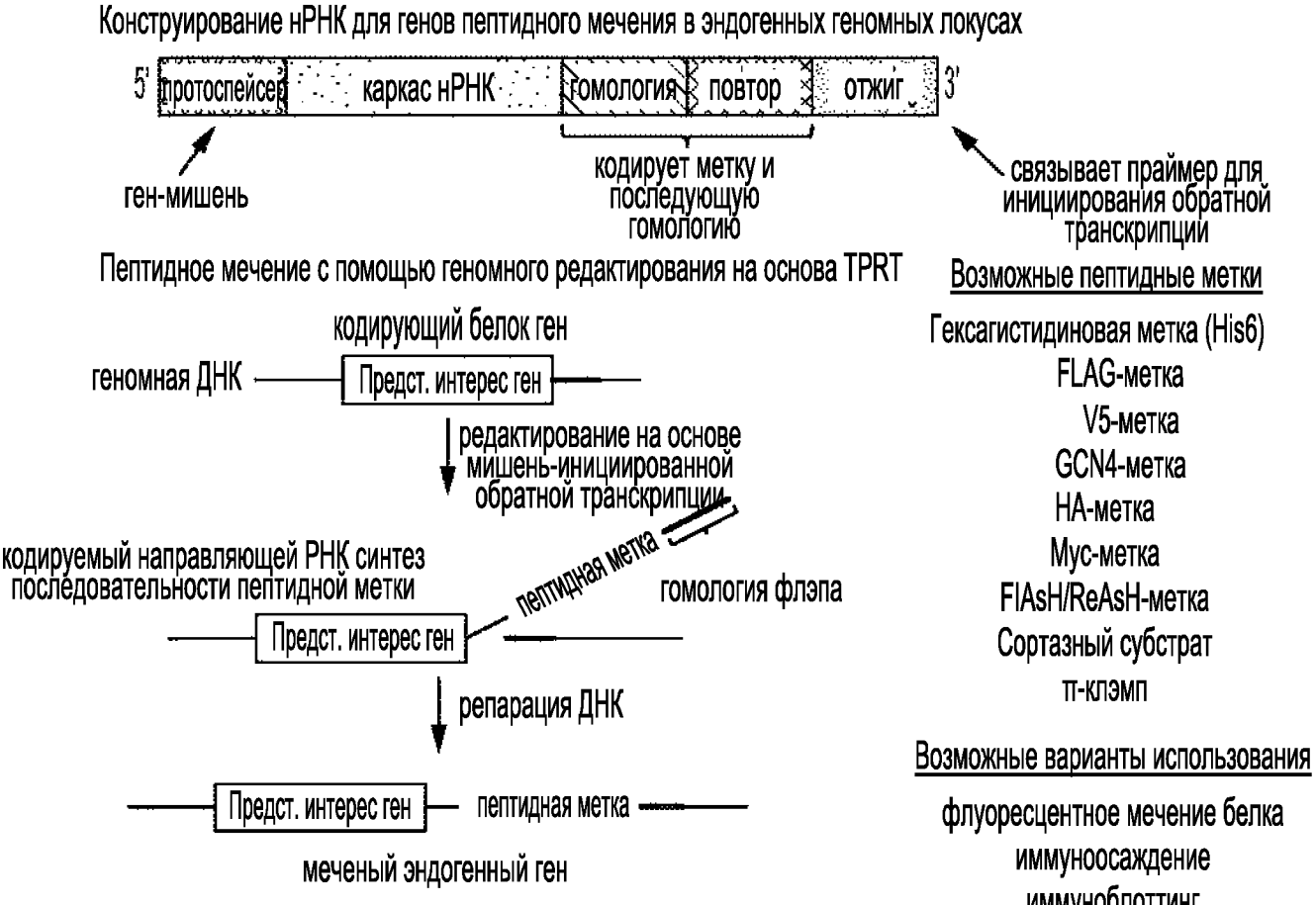
Сокращение тринуклеотидных повторов с помощью геномного редактирования на основе TPRT



ФИГ.23

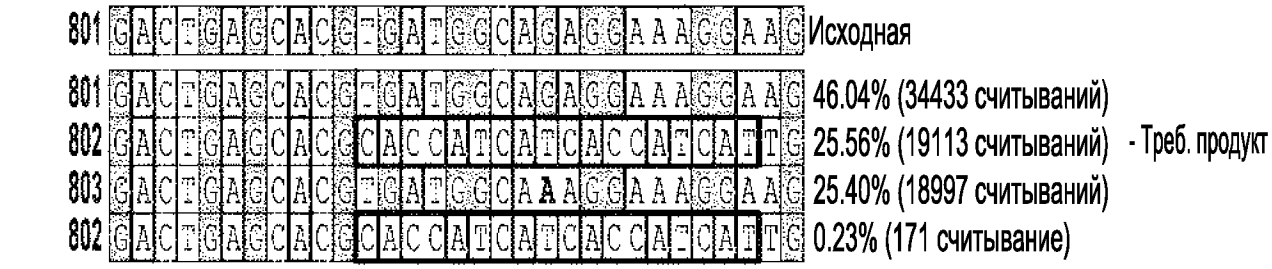

участок HEK3
10 нуклеотидов-мишеней для удаления
SEQ ID NOs:
795 CTGAGCACGTGATGGCAGAGGAAAGGAAG Исходная
796 CTGAGCACG----------GGAAAGGAAG 43.91% (30598 считываний) Требуемый продукт
795 CTGAGCACGTGATGGCAGAGGAAAGGAAG 32.35% (22540 считываний) начальные аллели
797 CTGAGCACGTGATGGCAAAAGGAAAGGAAG 17.25% (12021 считываний) начальные аллели
798 CTGAGCACG---------- 0.41% (288 считываний) побочные продукты
799 CTGAGCACG----------GGAAAGG--- 0.23% (157 считываний) побочные продукты
800 CTGAGCACGTGATGGCAGGGGAAAGGAAG 0.22% (156 считываний) побочные продукты
жирный шрифт замещения
вставки
- делеции
--- предполагаемое положение расщепления

ФИГ.24


Конструирование нРНК для генов пептидного мечения в эндогенных геномных локусах
5' протоспейсер каркас нРНК гомология повтор отжиг 3'
ген-мишень
кодирует метку и последующую гомологию
связывает праймер для инициирования обратной транскрипции
Пептидное мечение с помощью геномного редактирования на основа TPRT
кодирующий белок ген
геномная ДНК
Предст. интерес ген
редактирование на основе мишень-инициированной обратной транскрипции
кодируемый направляющей РНК синтез последовательности пептидной метки
пептидная метка
гомология флэпа
Предст. интерес ген
репарация ДНК
Предст. интерес ген
пептидная метка
меченый эндогенный ген
Возможные пептидные метки
Гексагистидиновая метка (His6)
FLAG-метка
V5-метка
GCN4-метка
HA-метка
Myc-метка
FlAsH/ReAsH-метка
Сортазный субстрат
тт-клэмп
Возможные варианты использования
флуоресцентное мечение белка
иммуноосаждение
иммуноблоттинг
иммуногистохимия
рекрутирование белка
индуцируемые белковые дегроны
скрининг в рамках всего генома

ФИГ.25

Участок HEK3, Вставка His-метки



жирный шрифт замещения

вставки

\- делеции

--- предполагаемое положение расщепления

Участок HEK3, Вставка FLAG-метки



жирный шрифт замещения

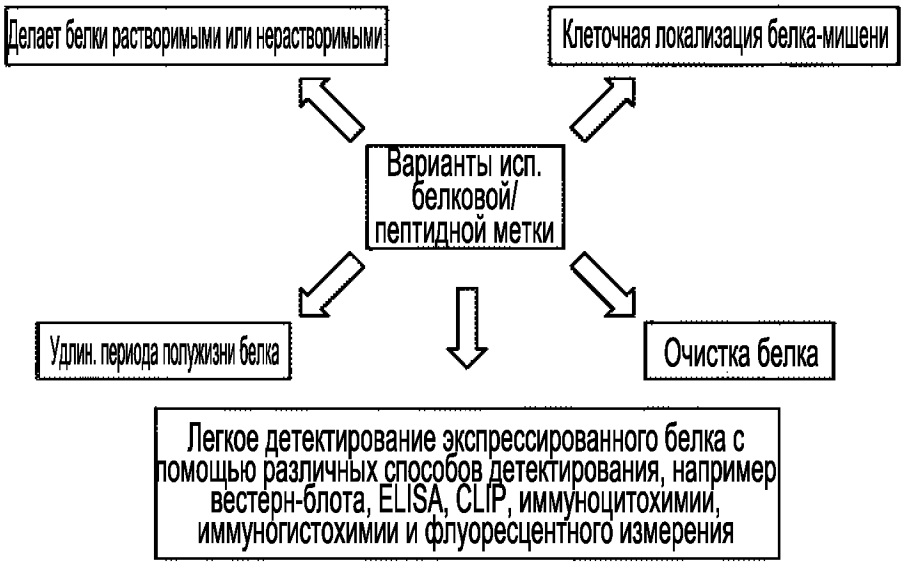
вставки

\- делеции

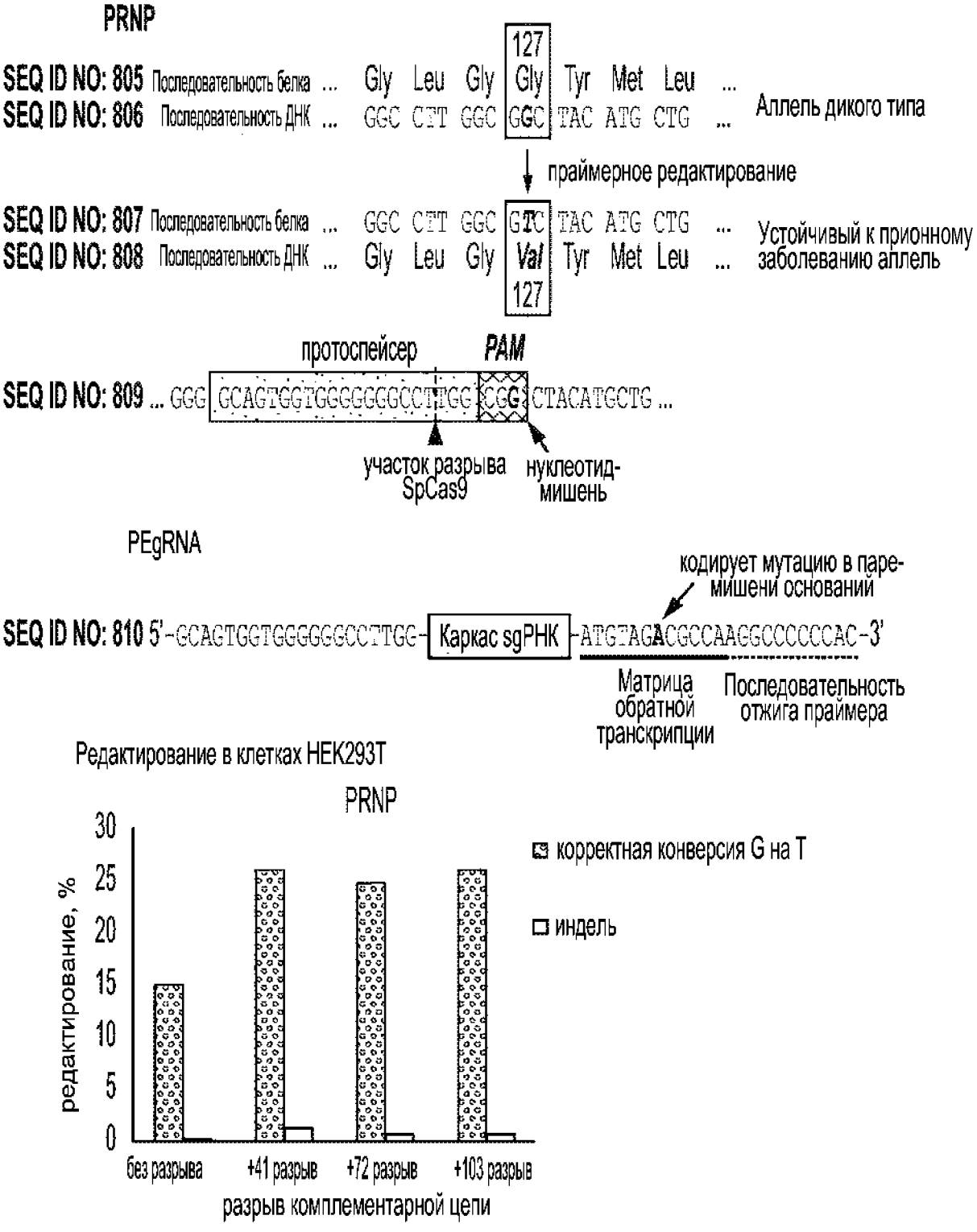
--- предполагаемое положение расщепления

след. обратить внимание, что вставка длиной 24 nt не показана

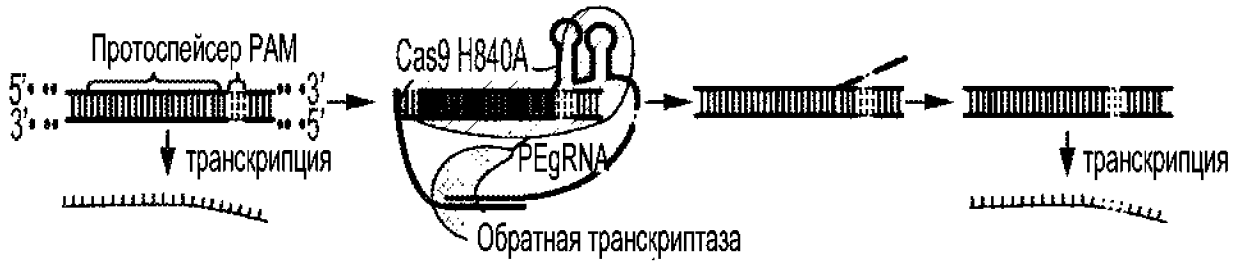
ФИГ.26A



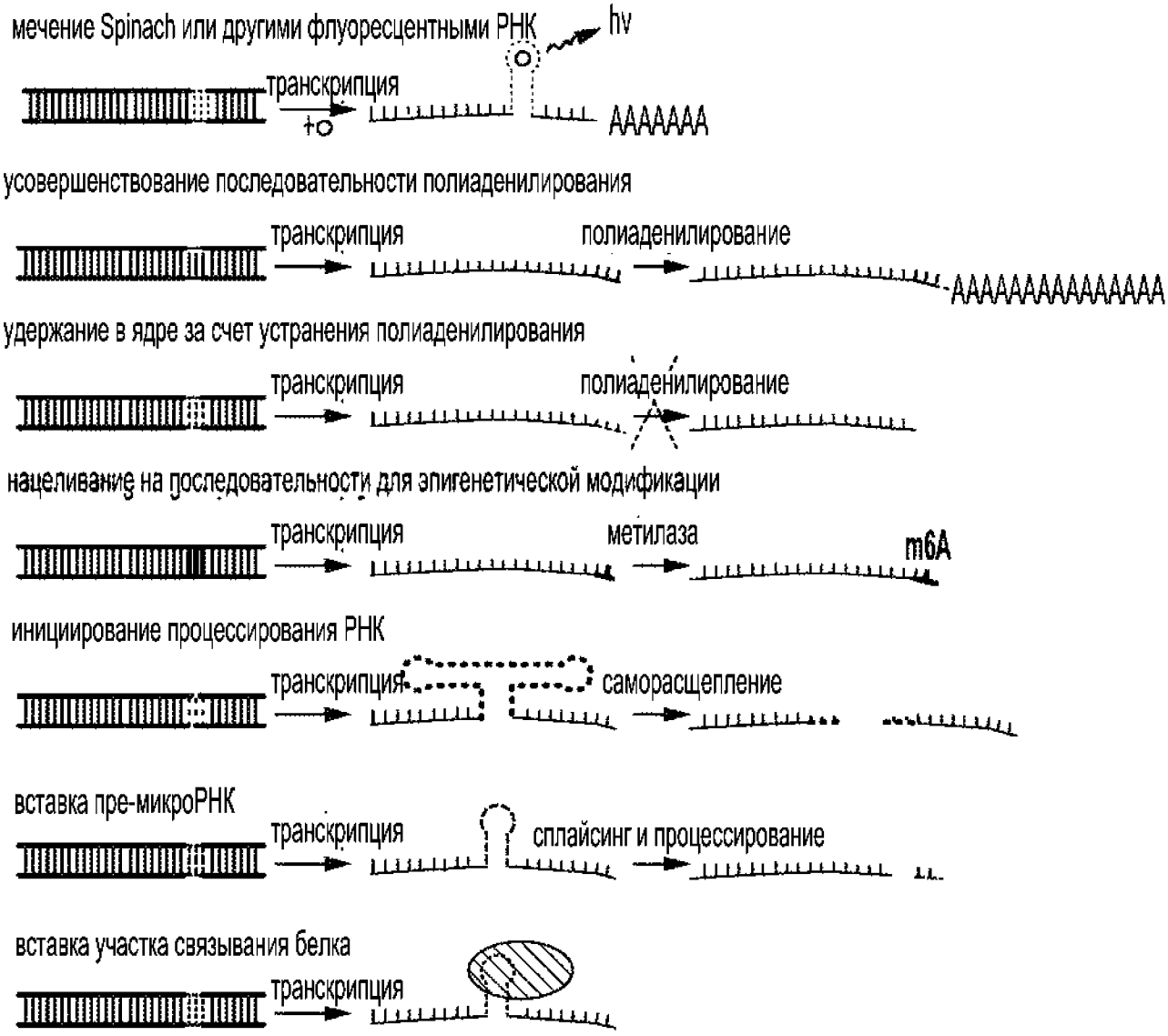
ФИГ.26B


PRNP
SEQ ID NO: 805 Последовательность белка ... Gly Leu Gly 127 Gly Tyr Met Leu ...
SEQ ID NO: 806 Последовательность ДНК ... GGC CTT GGC GGC TAC ATG CTG ...
Аллель дикого типа
праймерное редактирование
SEQ ID NO: 807 Последовательность белка ... GGC CTT GGC GTC TAC ATG CTG ...
SEQ ID NO: 808 Последовательность ДНК ... Gly Leu Gly Val Tyr Met Leu ...
127
Устойчивый к прионному заболеванию аллель
протоспейсер
PAM
SEQ ID NO: 809 ... GGG GCAGTGGTGGGGGGGCCTTGG CGG CTACATGCTG ...
участок разрыва SpCas9
нуклеотид-мишень
PEgRNA
кодирует мутацию в паре-мишени оснований
SEQ ID NO: 810 5'-GCAGTGGTGGGGGGGCCTTGG- Каркас sgPHK -ATGTAGACGCCAAGGCCCCCCAC-3'
Матрица обратной транскрипции
Последовательность отжига праймера
Редактирование в клетках HEK293T
PRNP
редактирование, %
30
25
20
15
10
5
0
корректная конверсия G на T
индель
без разрыва
+41 разрыв
+72 разрыв
+103 разрыв
разрыв комплементарной цепи

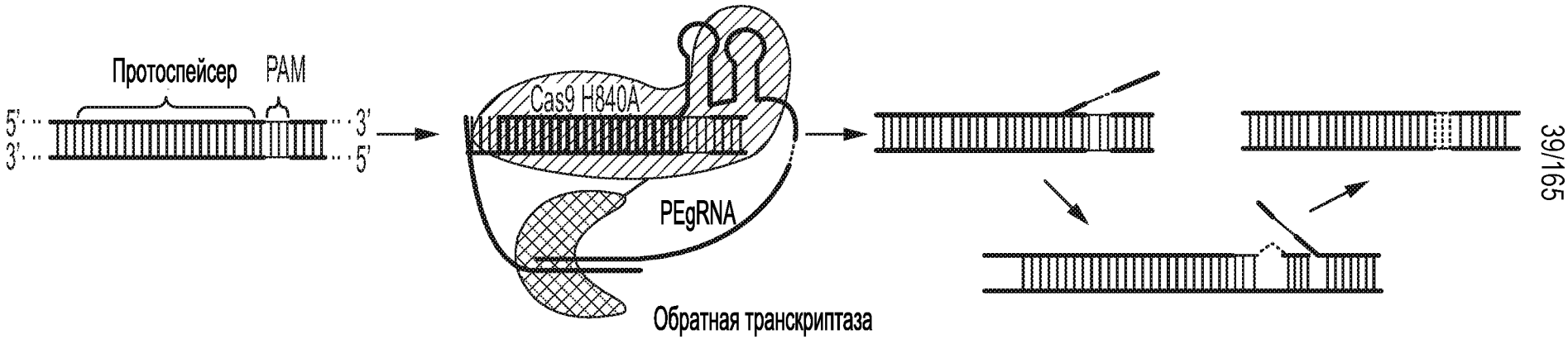
ФИГ.27



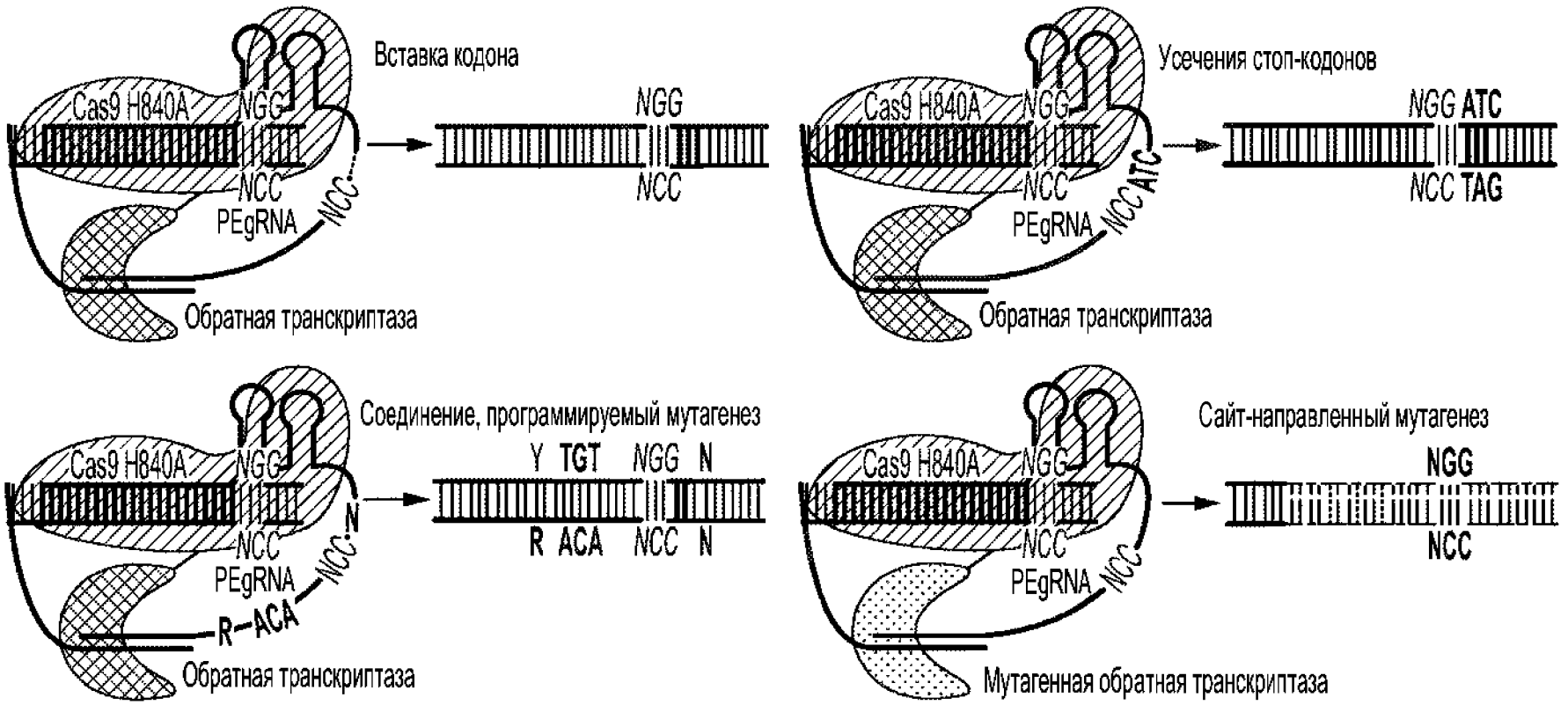
ФИГ.28A



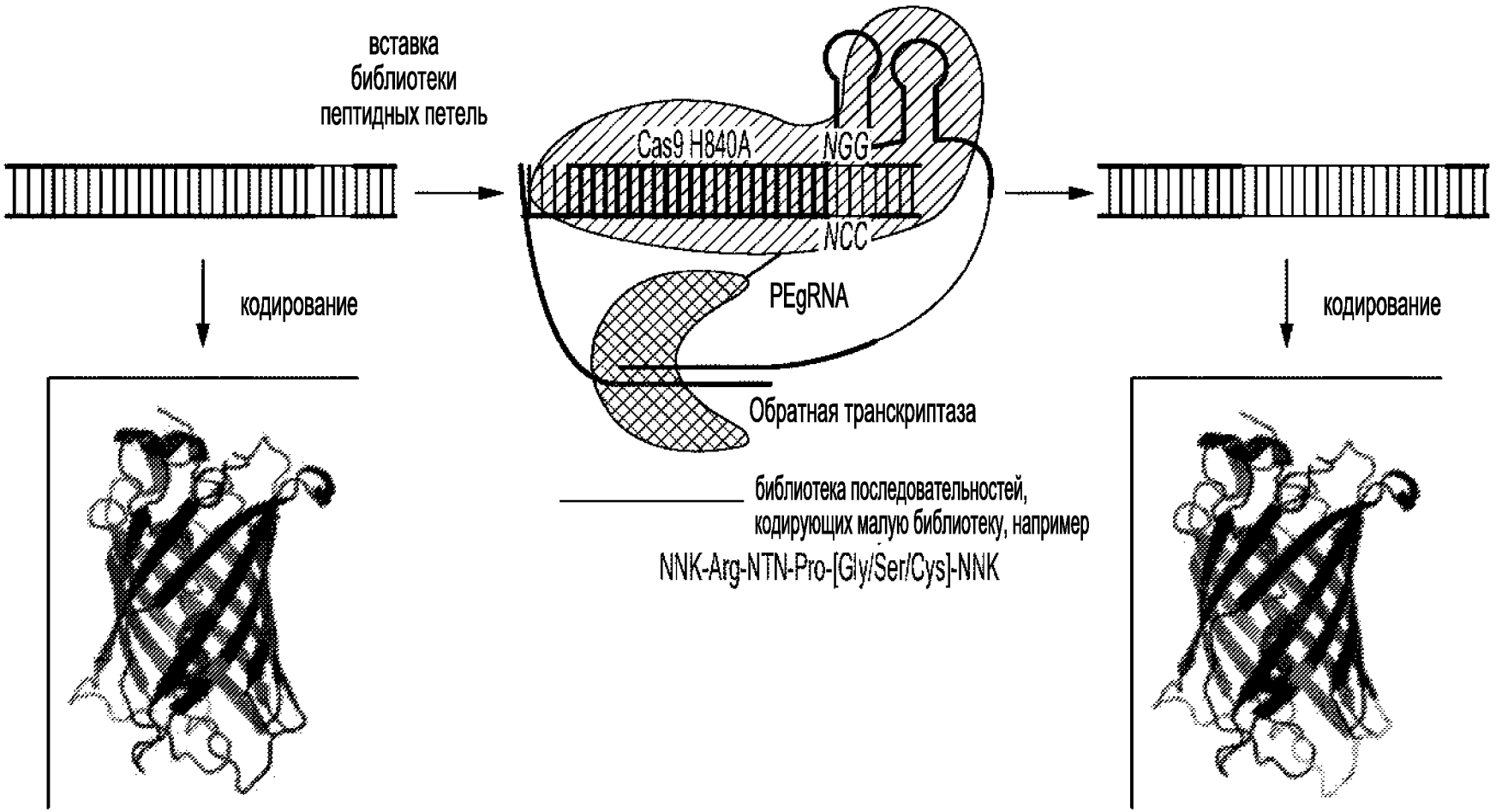
ФИГ.28B


Протоспейсер
PAM
5'
3'
3'
5'
Cas9 H840A
PEgRNA
Обратная транскриптаза

ФИГ.29А


Вставка кодона
Cas9 H840A
NGG
NCC
PEgRNA
NCC
Обратная транскриптаза
NGG
NCC
Усечения стоп-кодонов
Cas9 H840A
NGG
NCC
PEgRNA
NCC ATC
Обратная транскриптаза
NGG ATC
NCC TAG
Соединение, программируемый мутагенез
Cas9 H840A
NGG
NCC
PEgRNA
N
NCC
R–ACA
Обратная транскриптаза
Y TGT NGG N
R ACA NCC N
Сайт-направленный мутагенез
Cas9 H840A
NGG
NCC
PEgRNA
NCC
Мутагенная обратная транскриптаза
NGG
NCC

ФИГ.29В

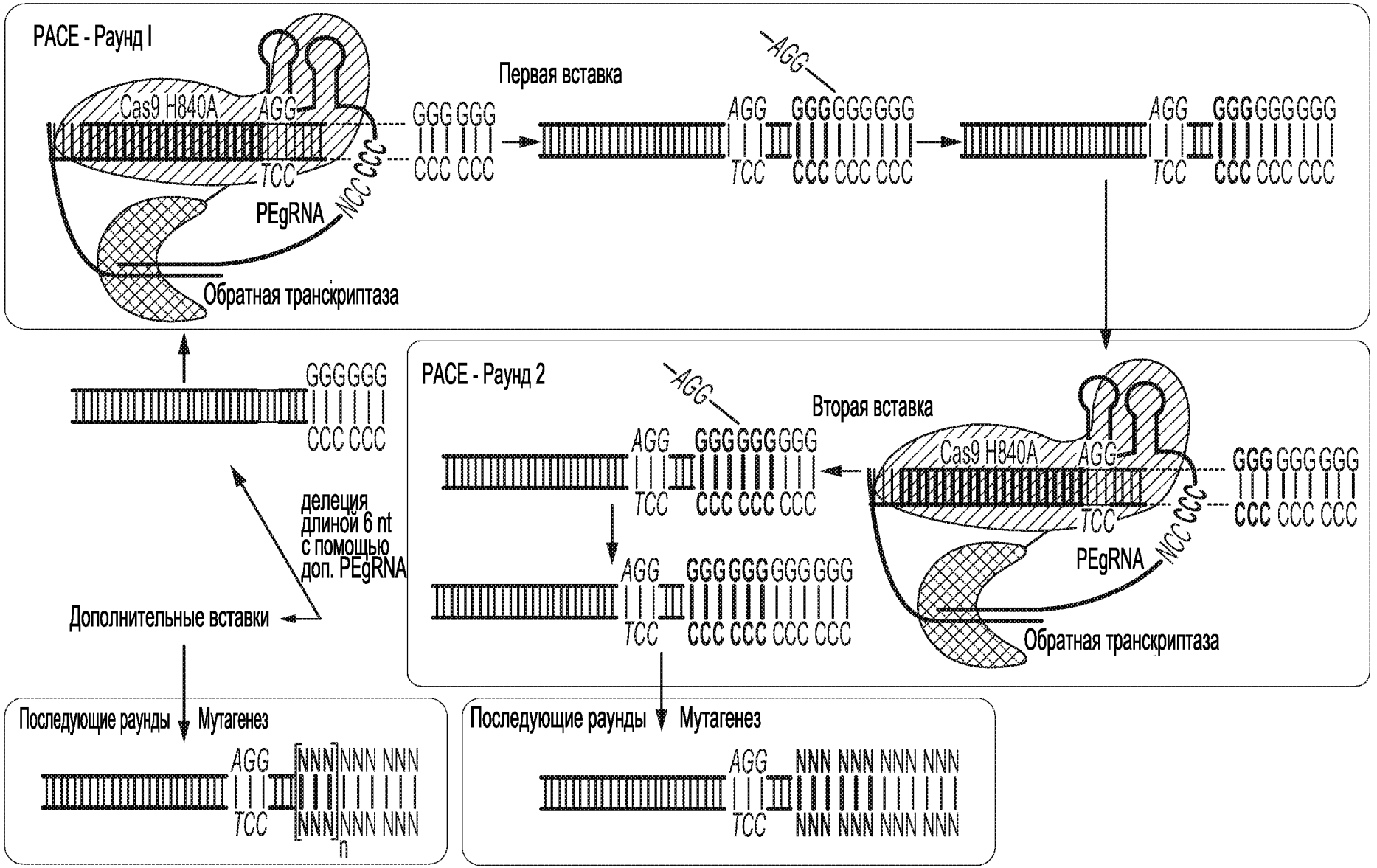

вставка
библиотеки
пептидных петель
Cas9 H840A
NGG
NCC
PEgRNA
Обратная транскриптаза
кодирование
кодирование
библиотека последовательностей,
кодирующих малую библиотеку, например
NNK-Arg-NTN-Pro-[Gly/Ser/Cys]-NNK

ФИГ.28С

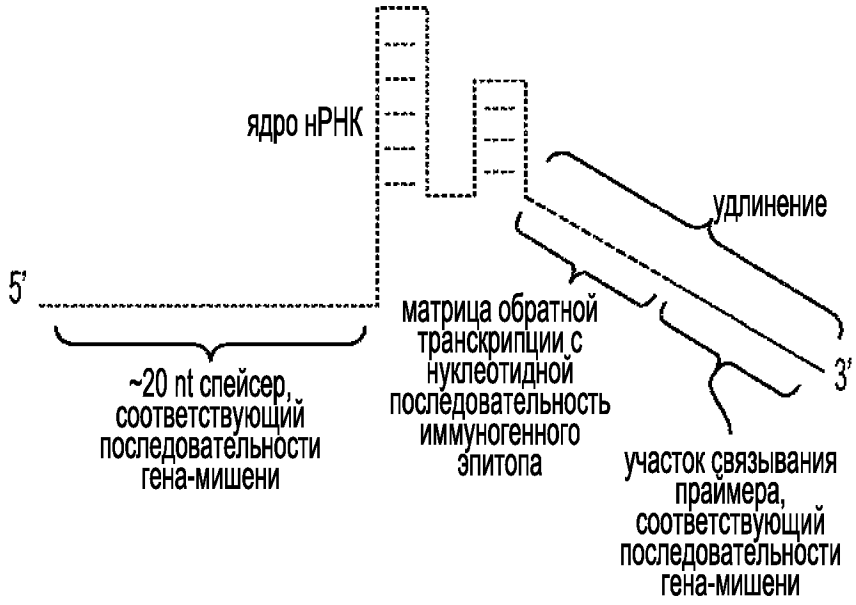
Программируемые делеции и усечения кодонов

SEQ ID NO: 811	**GFP**	NFKIRHNIEDGSVQLADHYQQNTPI
SEQ ID NO: 812	del5	NFKI-HNIEDGSVQLADHYQQNTPI
SEQ ID NO: 813	del6	NFKIR-NIEDGSVQLADHYQQNTPI
SEQ ID NO: 814	del8	NFKIRHN-EDGSVQLADHYQQNTPI
SEQ ID NO: 815	del 8-13	NFKIRHN------QLADHYQQNTPI
SEQ ID NO: 816	del 1-5	-----HNIEDGSVQLADHYQQNTPI
SEQ ID NO: 817	del 15-25	NFKIRHNIEDGSVQ----------

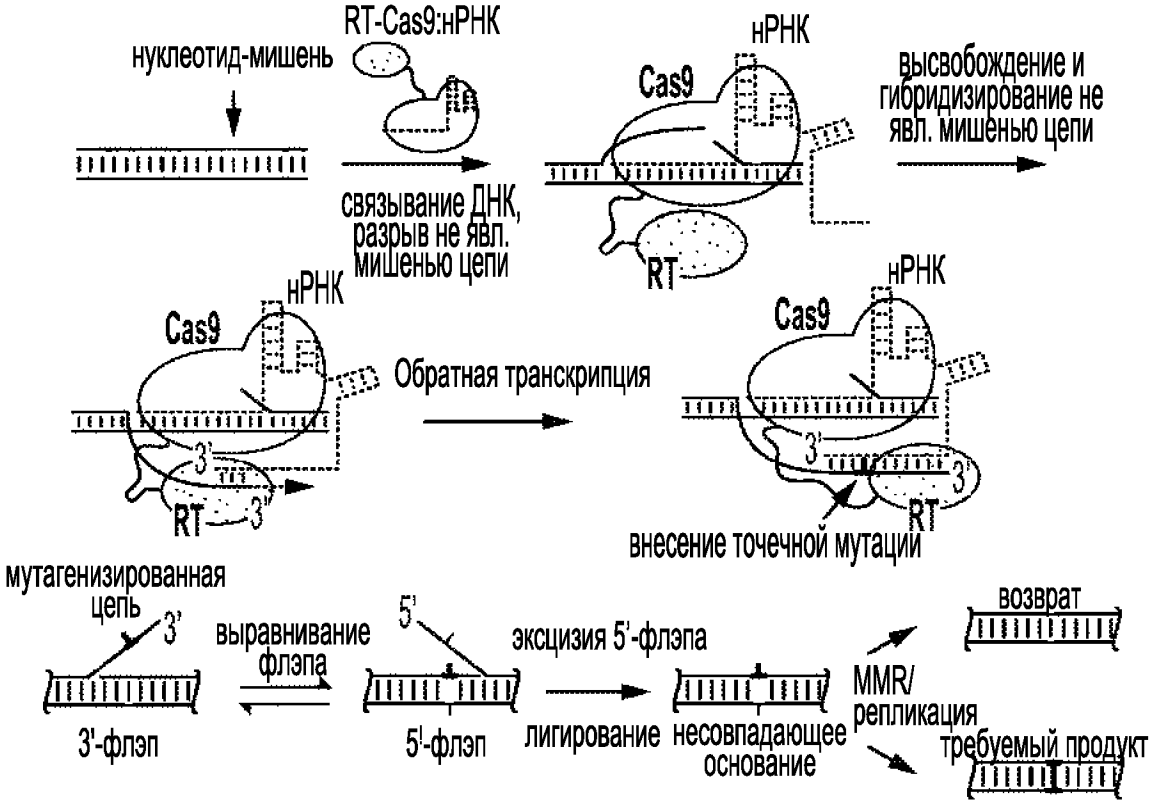
ФИГ.29D


PACE - Раунд I
Cas9 H840A
AGG
TCC
CCC
NCC
PEgRNA
Обратная транскриптаза
GGG GGG
CCC CCC
Первая вставка
AGG
GGG GGG GGG
TCC
CCC CCC CCC
PACE - Раунд 2
Вторая вставка
делеция
длиной 6 nt
с помощью
доп. PEgRNA
Дополнительные вставки
Последующие раунды
Мутагенез
NNN NNN NNN NNN
n

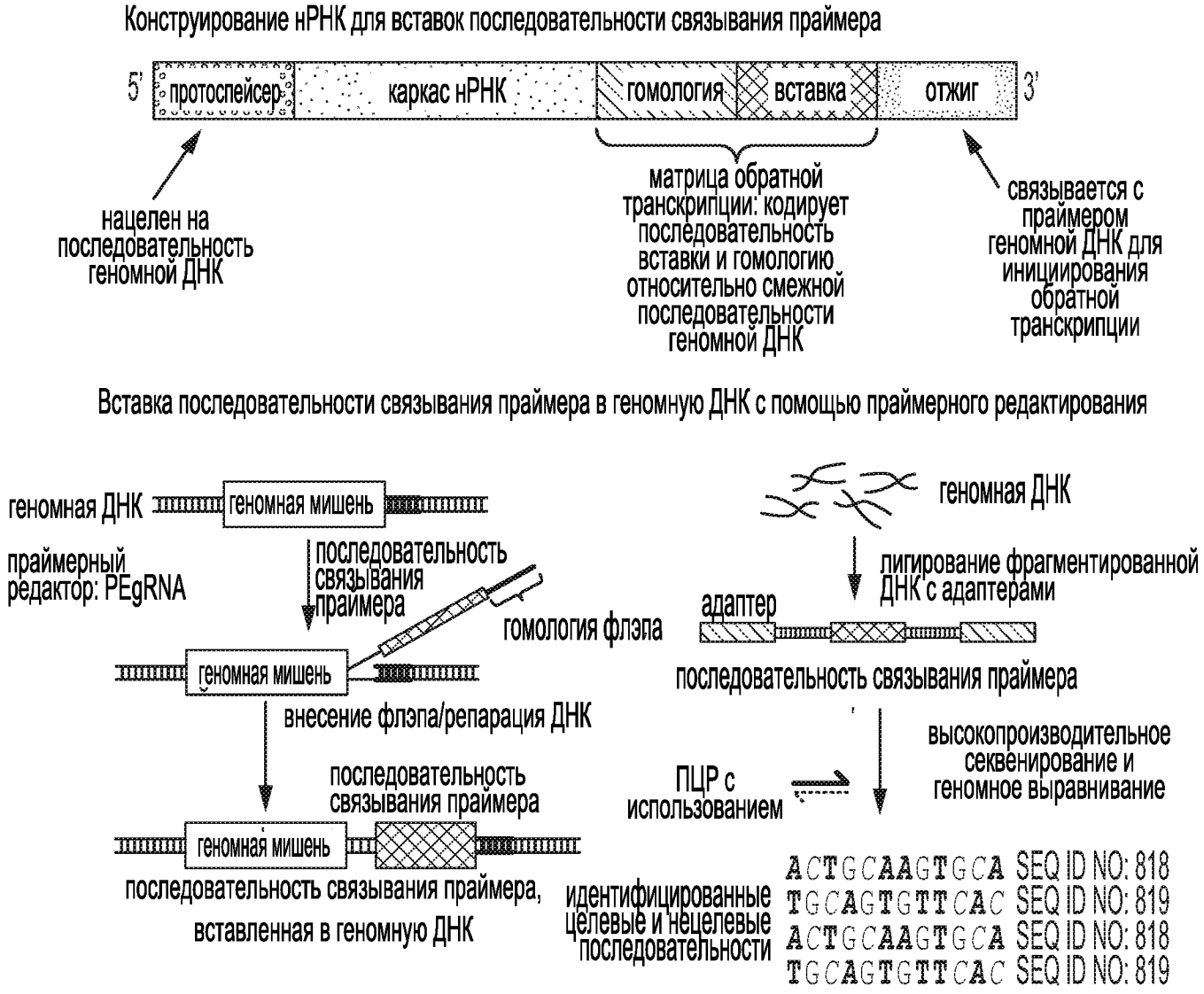
ФИГ.30


ядро нРНК
удлинение
5'
матрица обратной транскрипции с нуклеотидной последовательность иммуногенного эпитопа
3'
~20 nt спейсер, соответствующий последовательности гена-мишени
участок связывания праймера, соответствующий последовательности гена-мишени

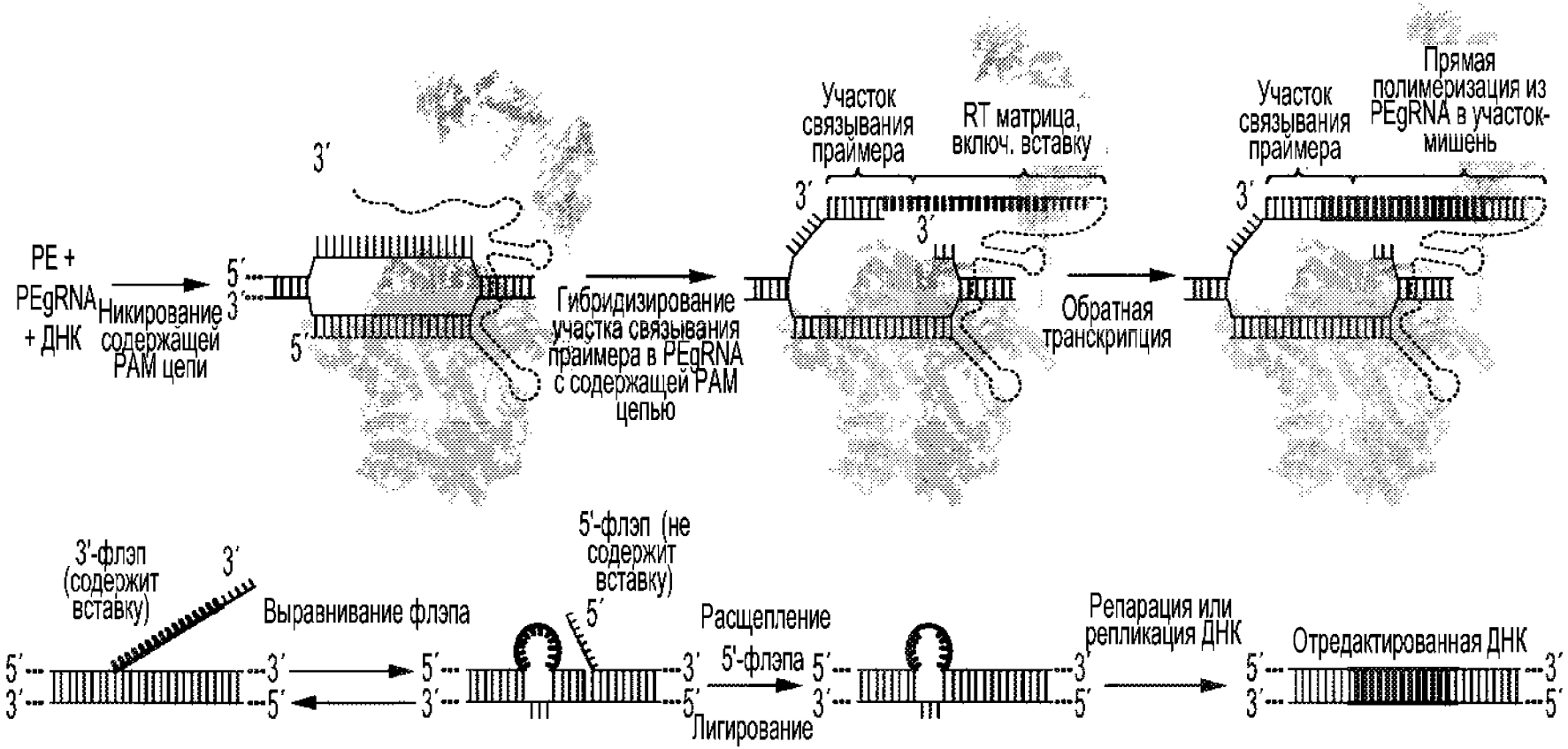
ФИГ.31


RT-Cas9:нРНК
нуклеотид-мишень
нРНК
Cas9
высвобождение и гибридизирование не явл. мишенью цепи
связывание ДНК, разрыв не явл. мишенью цепи
RT
нРНК
Cas9
Обратная транскрипция
нРНК
Cas9
3'
3''
RT
3'
3'
RT
внесение точечной мутации
мутагенизированная цепь
3'
выравнивание флэпа
5'
эксцизия 5'-флэпа
возврат
MMR/ репликация
3'-флэп
5'-флэп
лигирование
несовпадающее основание
требуемый продукт

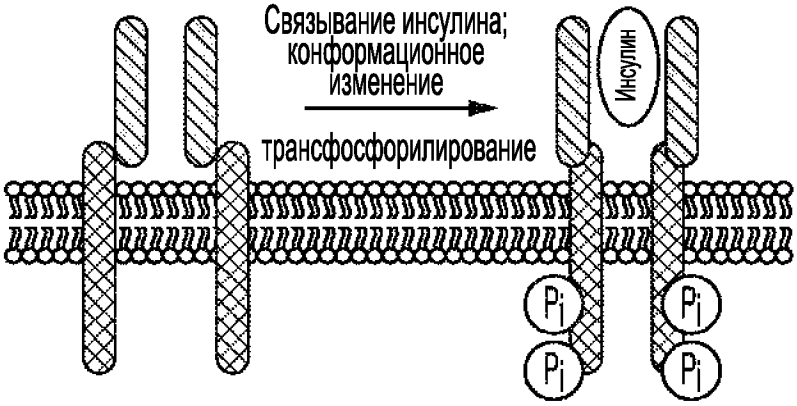
ФИГ.32


Конструирование нРНК для вставок последовательности связывания праймера
5'
протоспейсер
каркас нРНК
гомология
вставка
отжиг
3'
нацелен на последовательность геномной ДНК
матрица обратной транскрипции: кодирует последовательность вставки и гомологию относительно смежной последовательности геномной ДНК
связывается с праймером геномной ДНК для инициирования обратной транскрипции
Вставка последовательности связывания праймера в геномную ДНК с помощью праймерного редактирования
геномная ДНК
геномная мишень
праймерный редактор: PEgRNA
последовательность связывания праймера
гомология флэпа
геномная мишень
внесение флэпа/репарация ДНК
последовательность связывания праймера
геномная мишень
последовательность связывания праймера, вставленная в геномную ДНК
геномная ДНК
лигирование фрагментированной ДНК с адаптерами
адаптер
последовательность связывания праймера
ПЦР с использованием
высокопроизводительное секвенирование и геномное выравнивание
идентифицированные целевые и нецелевые последовательности
ACTGCAAGTGCA SEQ ID NO: 818
TGCAGTGTTCAC SEQ ID NO: 819
ACTGCAAGTGCA SEQ ID NO: 818
TGCAGTGTTCAC SEQ ID NO: 819

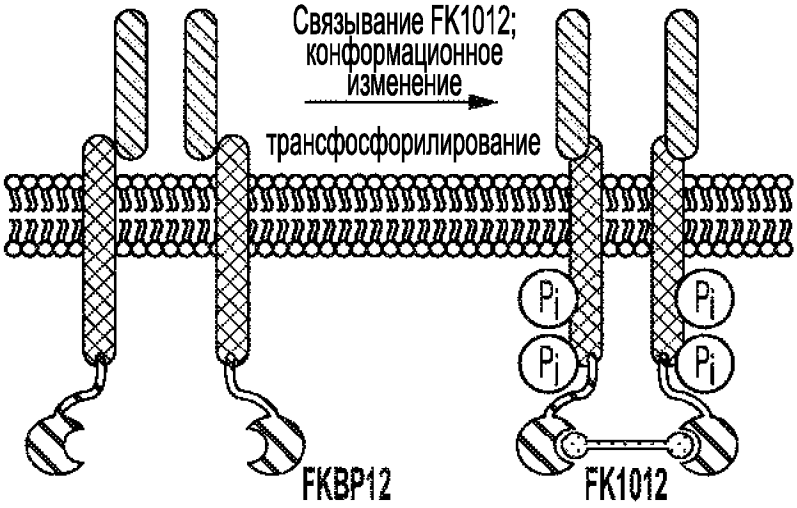
ФИГ.33


PE +
PEgRNA
+ ДНК
Никирование содержащей PAM цепи
Гибридизирование участка связывания праймера в PEgRNA с содержащей PAM цепью
Участок связывания праймера
RT матрица, включ. вставку
Обратная транскрипция
Участок связывания праймера
Прямая полимеризация из PEgRNA в участок-мишень
3'-флэп (содержит вставку)
Выравнивание флэпа
5'-флэп (не содержит вставку)
Расщепление 5'-флэпа
Лигирование
Репарация или репликация ДНК
Отредактированная ДНК

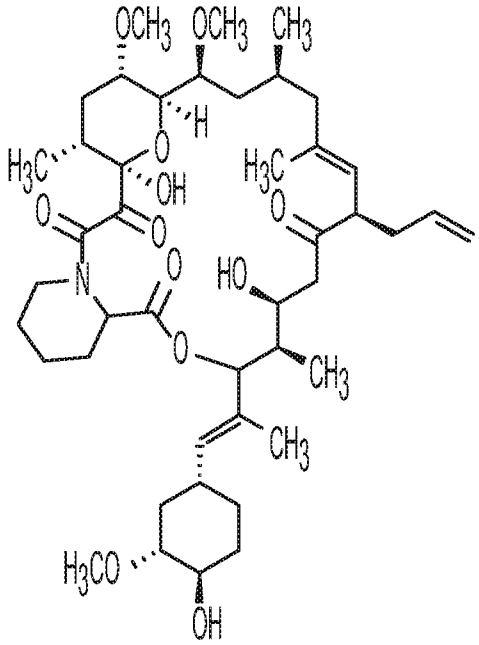
ФИГ.34



ФИГ.35A

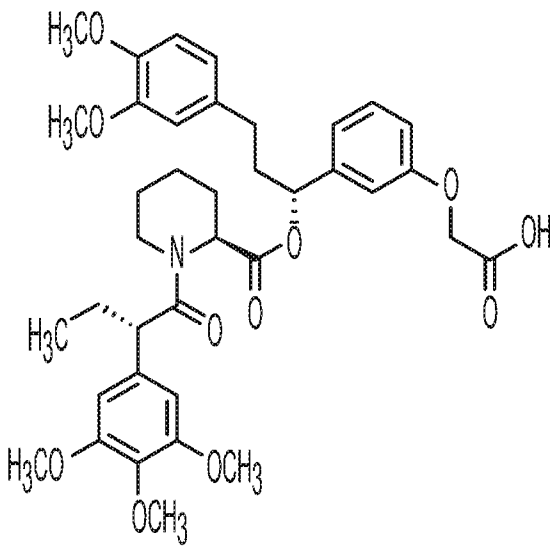


ФИГ.35B



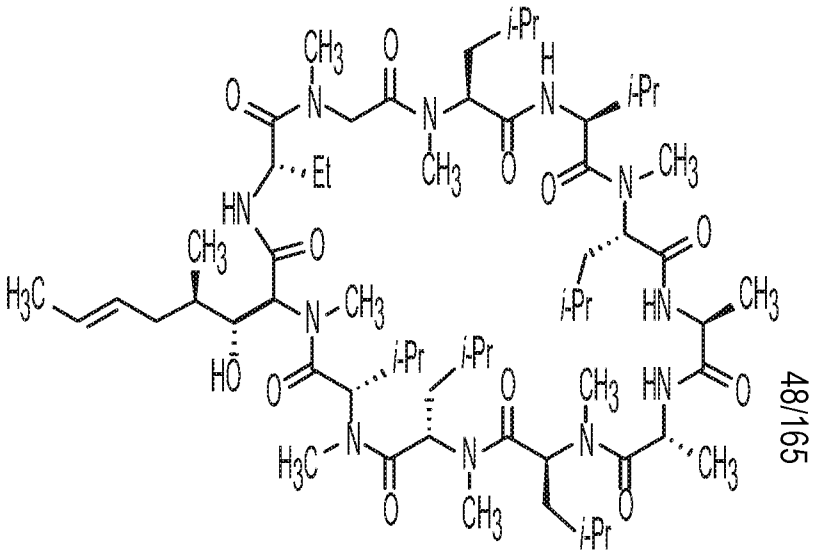
1

FK506
мишень: FKBP12



2

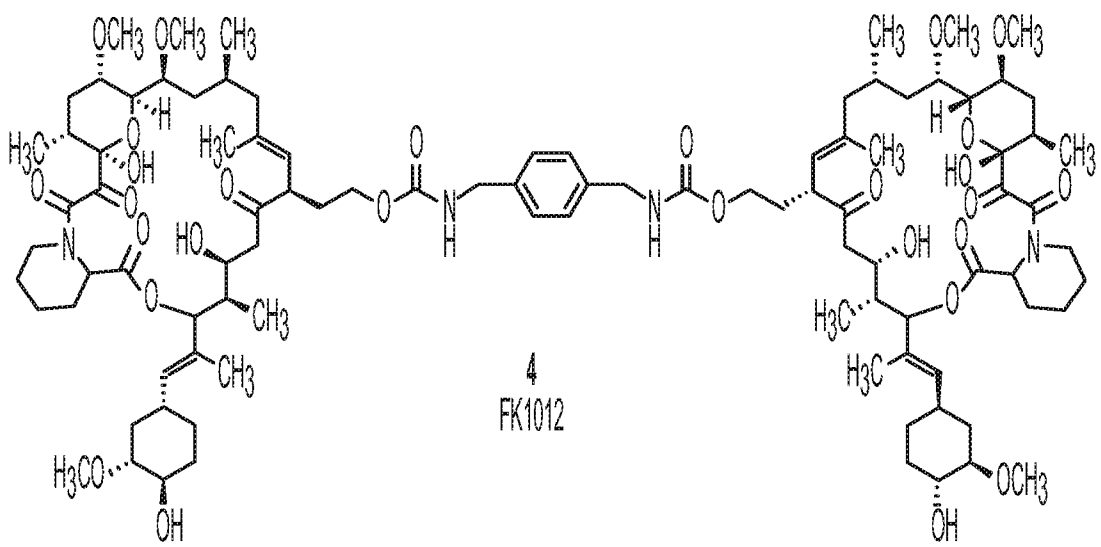
имеющий "выступы"
миметик FK506
мишень: FKBP12 F36V



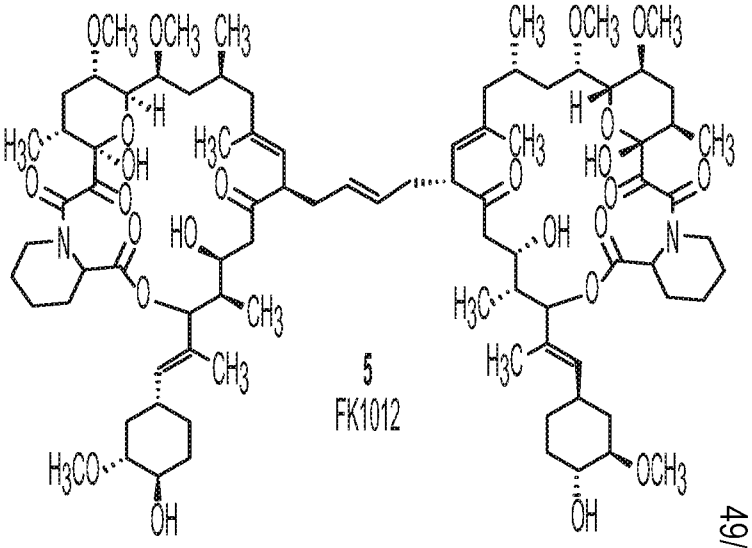
3

циклоспорин А
мишень: циклофилин

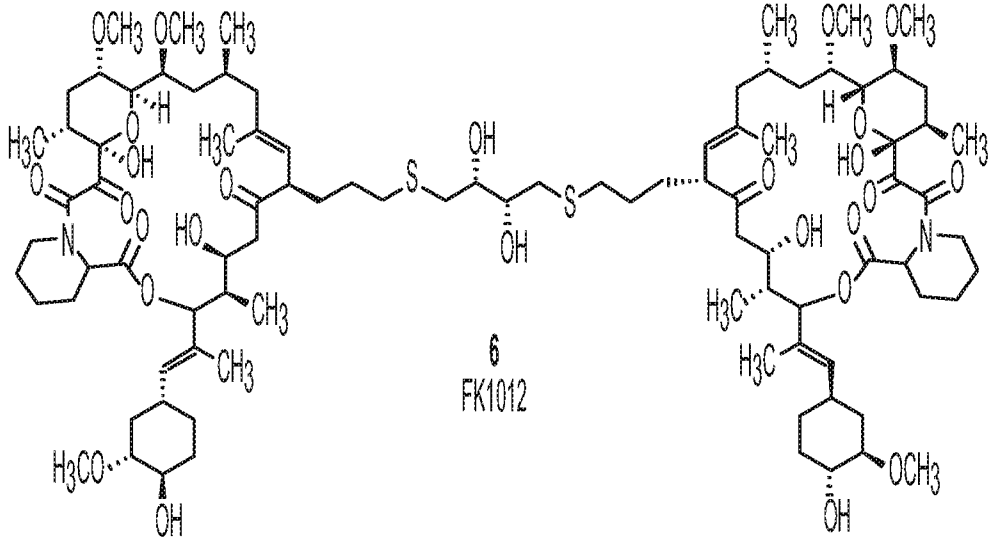
ФИГ.36



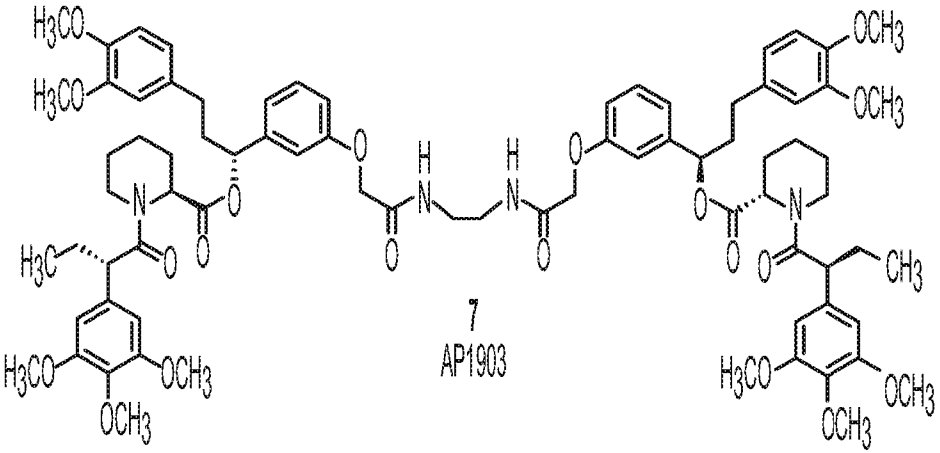
4
FK1012



5
FK1012

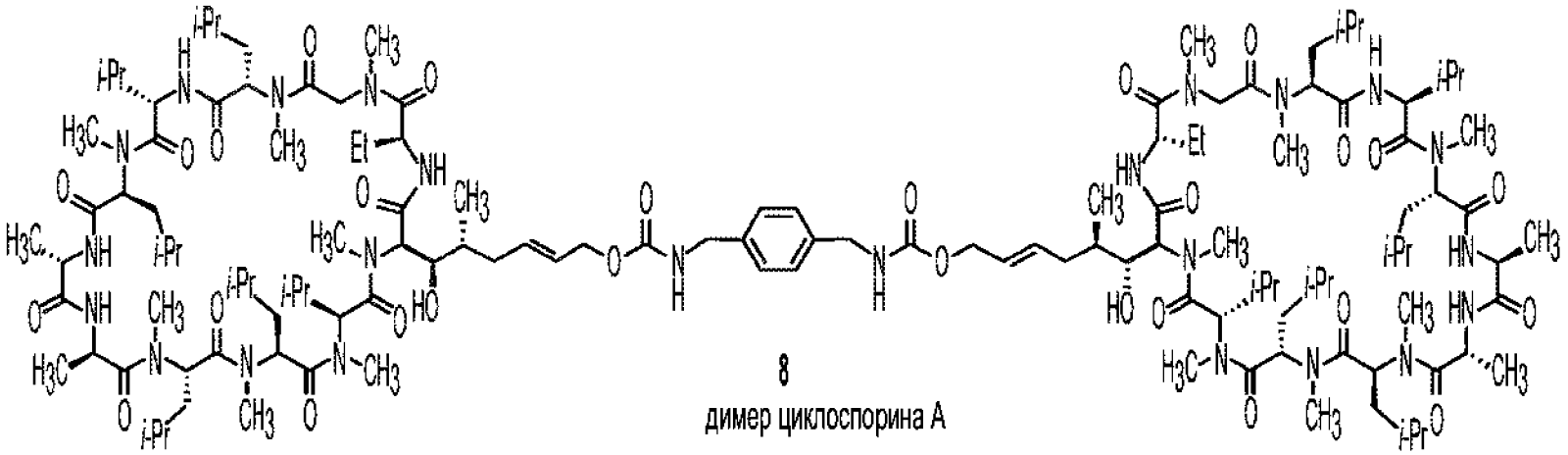


6
FK1012

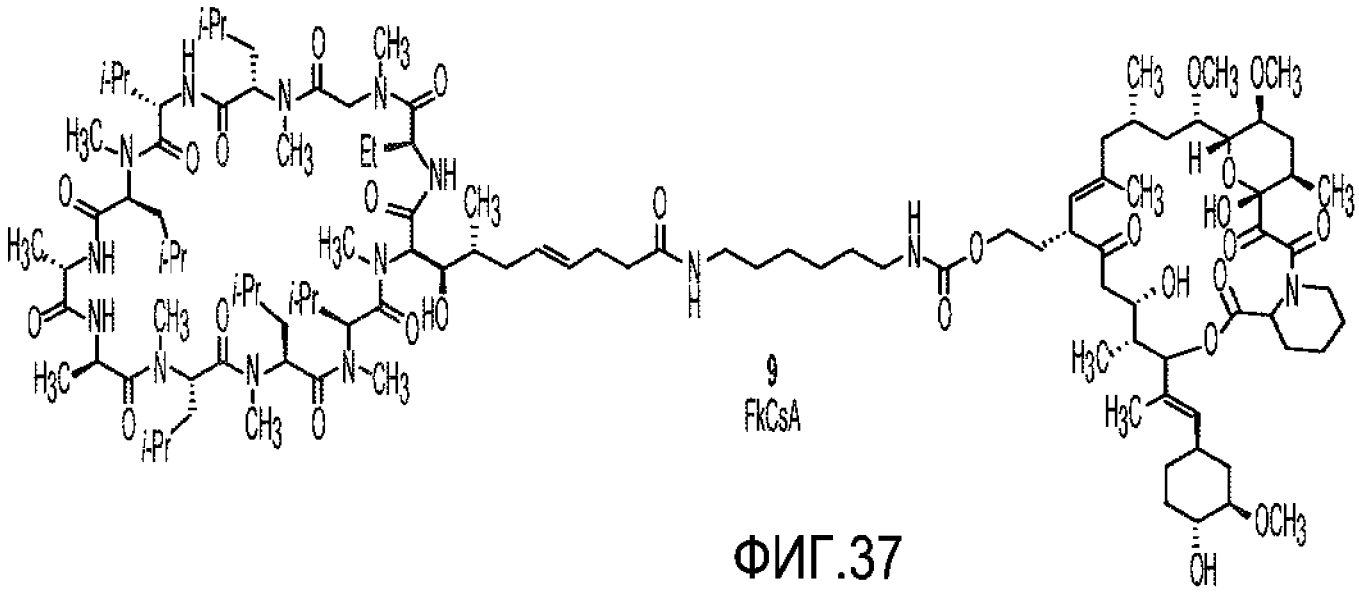


7
AP1903

ФИГ.37



8
димер циклоспорина А



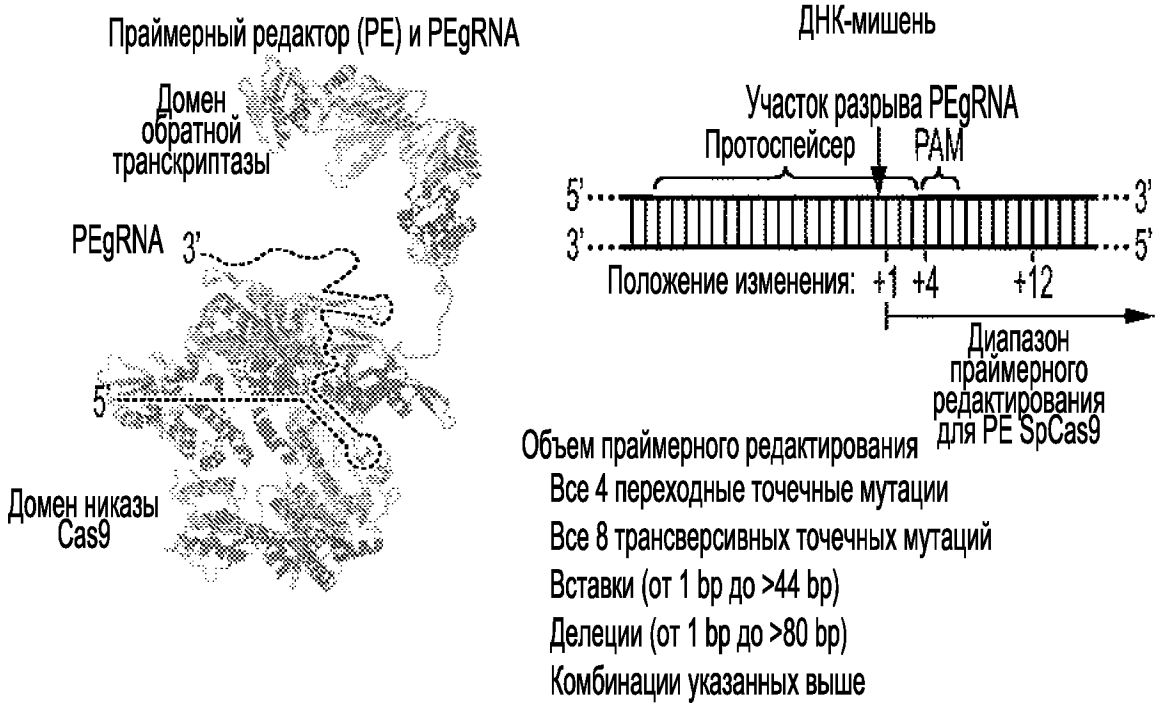
9
FkCsA

ФИГ.37

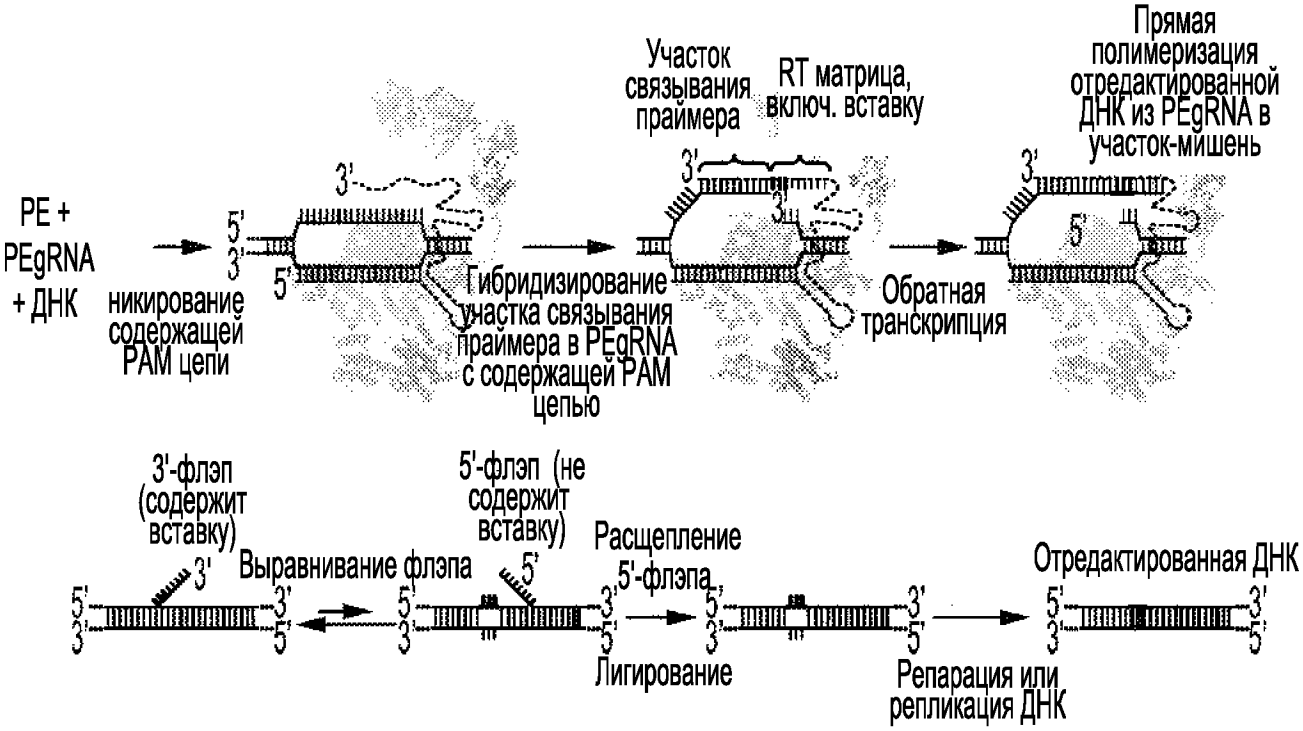
ПРОДОЛЖЕНИНЕ



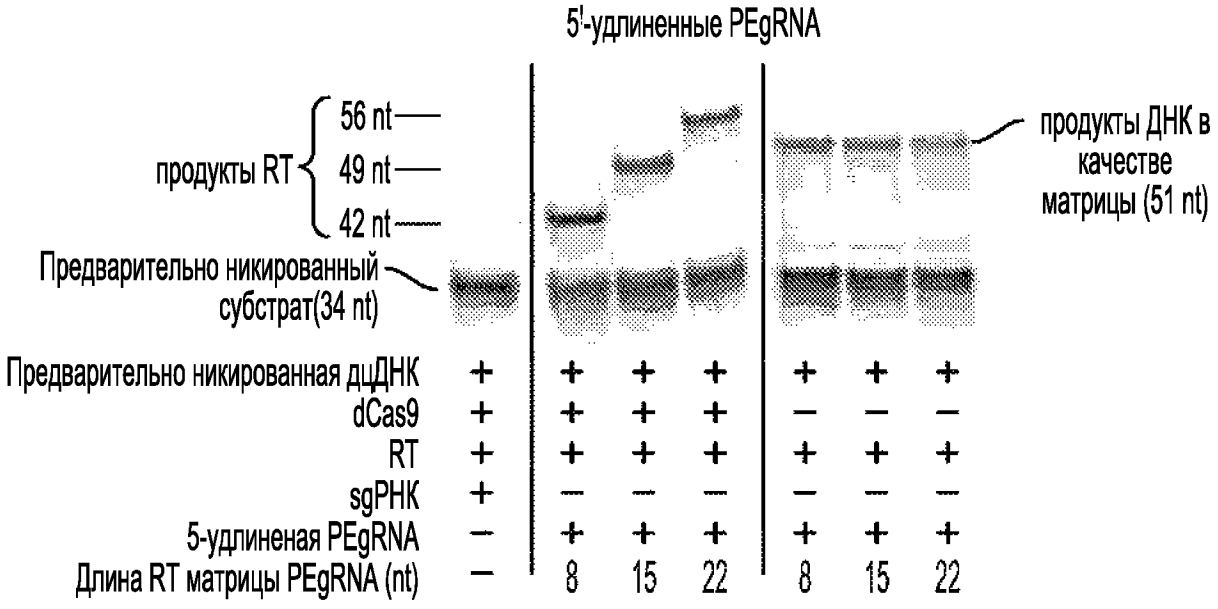
ФИГ.38A



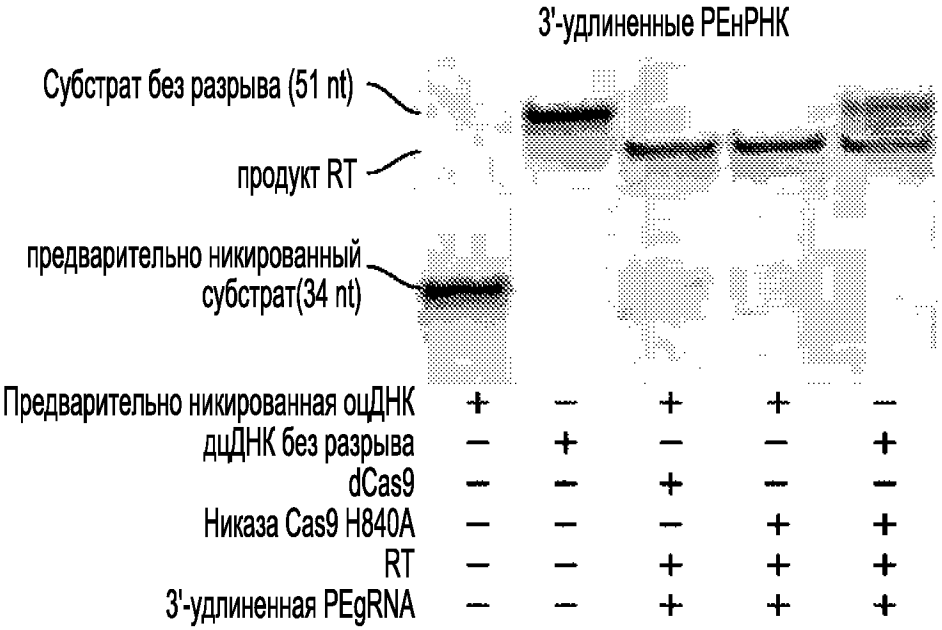
ФИГ.38B



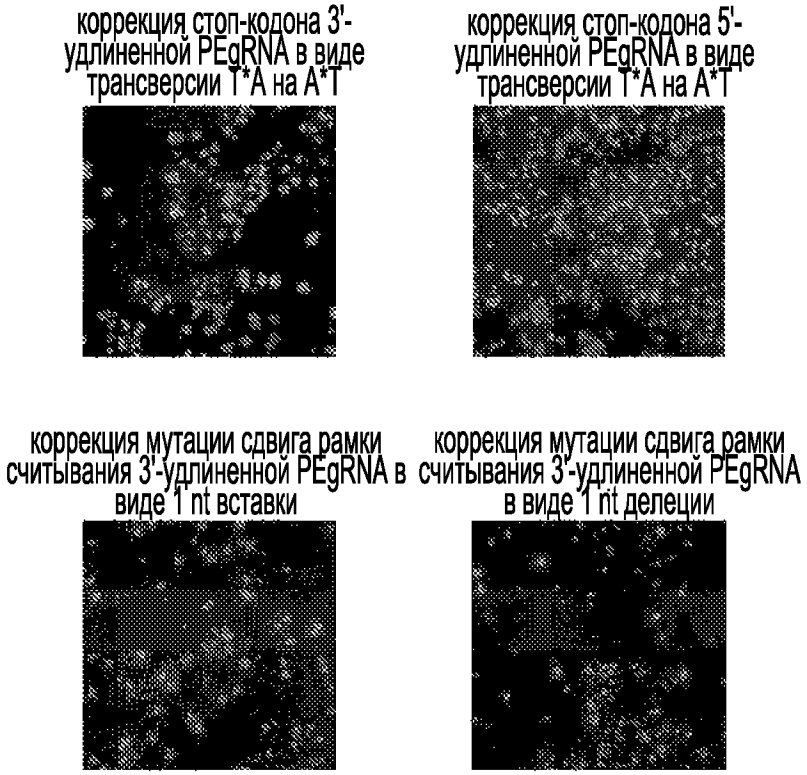
ФИГ.38C



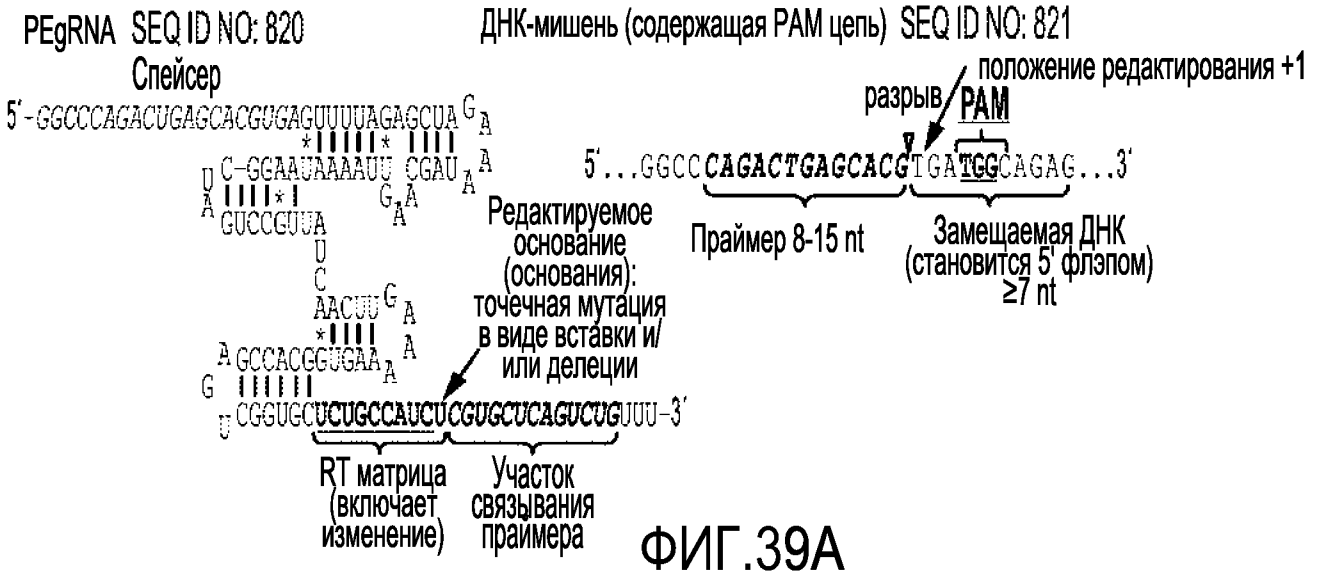
ФИГ.38D



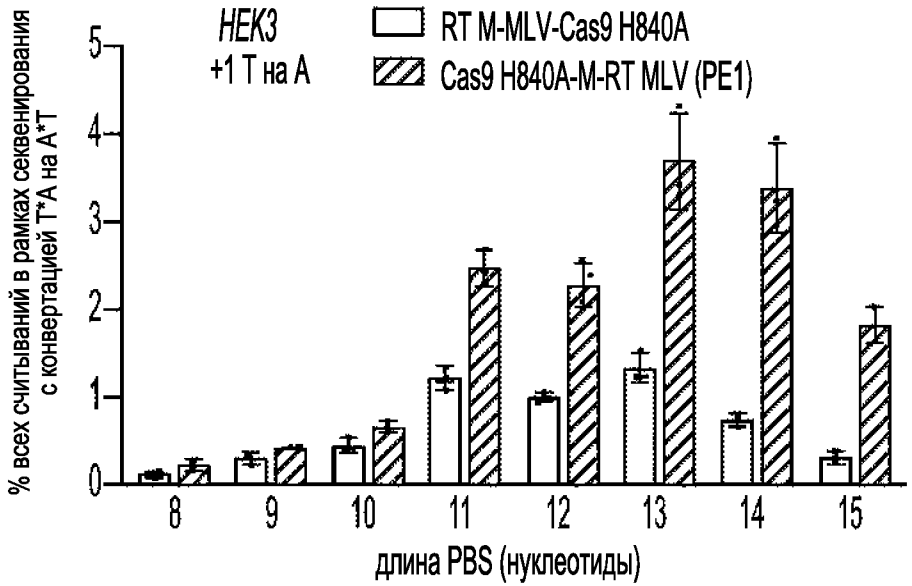
ФИГ.38E



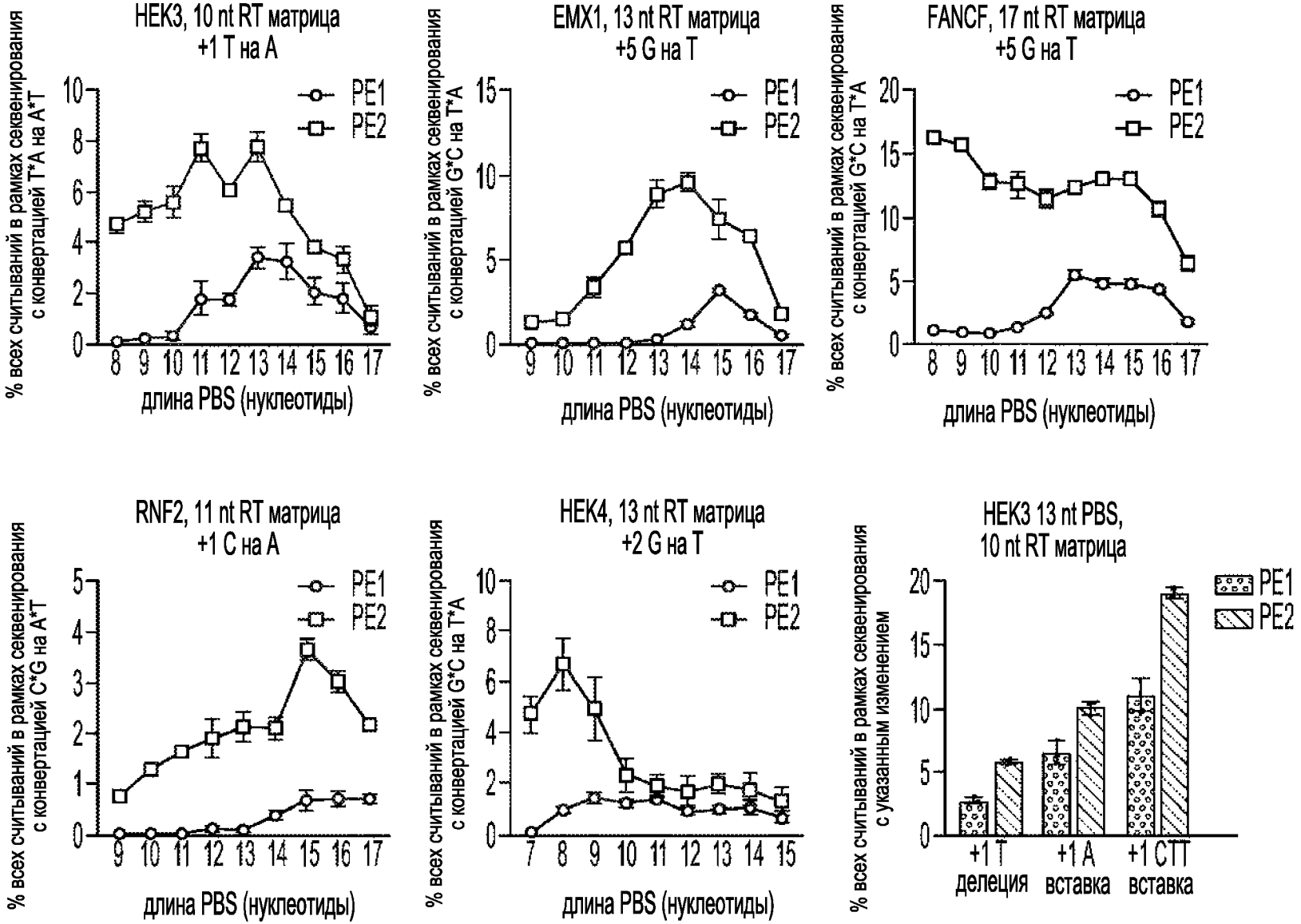
ФИГ.38F



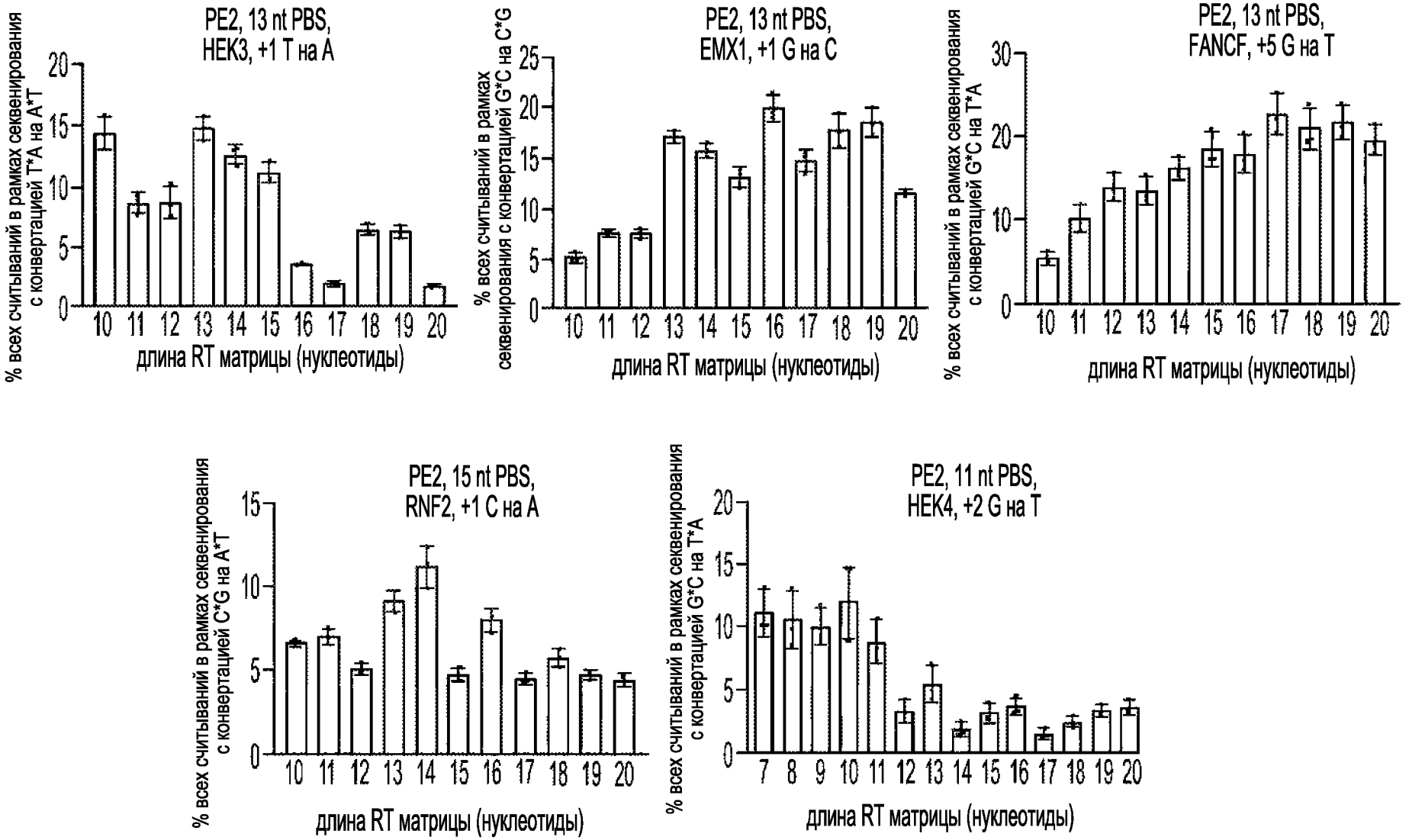
ФИГ.39A



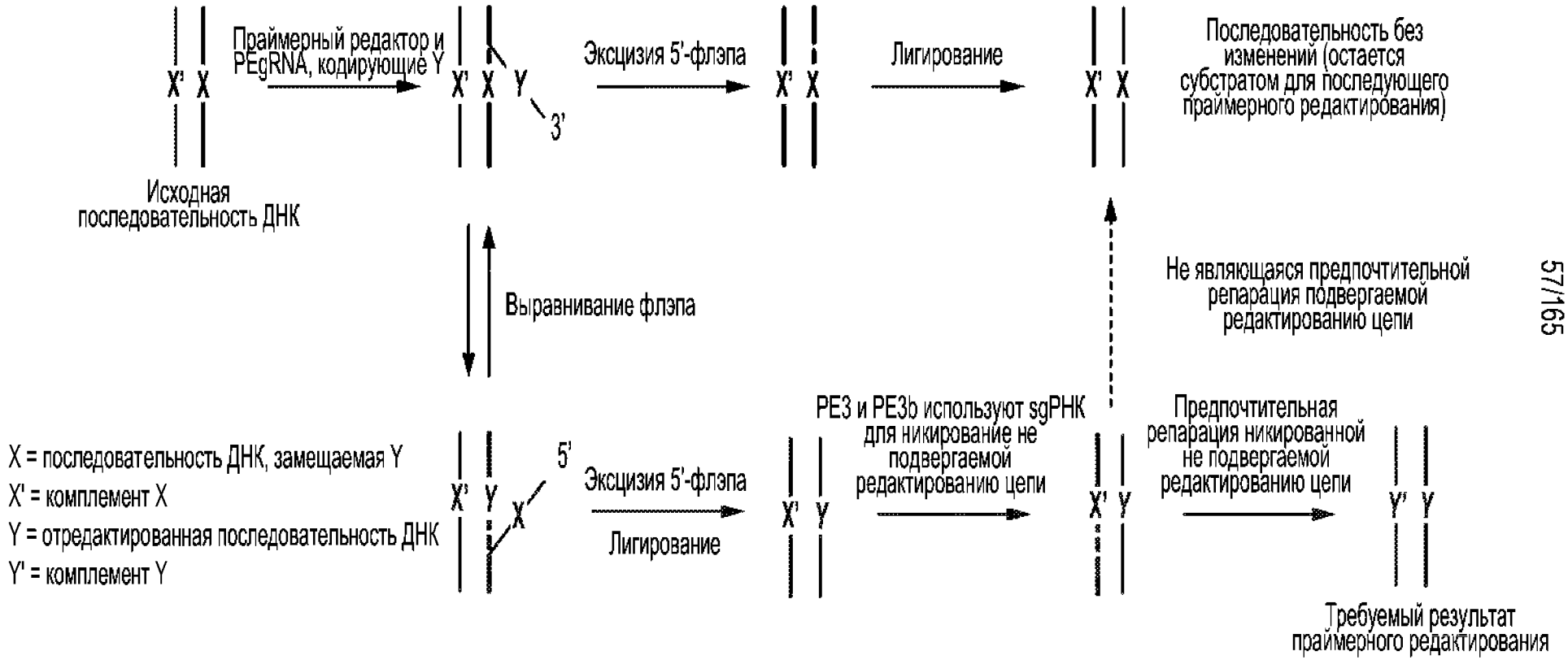
ФИГ.39B


HEK3, 10 nt RT матрица
+1 T на A
% всех считываний в рамках секвенирования с конвертацией T*A на A*T
PE1
PE2
длина PBS (нуклеотиды)
EMX1, 13 nt RT матрица
+5 G на T
% всех считываний в рамках секвенирования с конвертацией G*C на T*A
FANCF, 17 nt RT матрица
+5 G на T
RNF2, 11 nt RT матрица
+1 C на A
% всех считываний в рамках секвенирования с конвертацией C*G на A*T
HEK4, 13 nt RT матрица
+2 G на T
HEK3 13 nt PBS,
10 nt RT матрица
% всех считываний в рамках секвенирования с указанным изменением
+1 T
делеция
+1 A
вставка
+1 CTT
вставка

ФИГ.39C

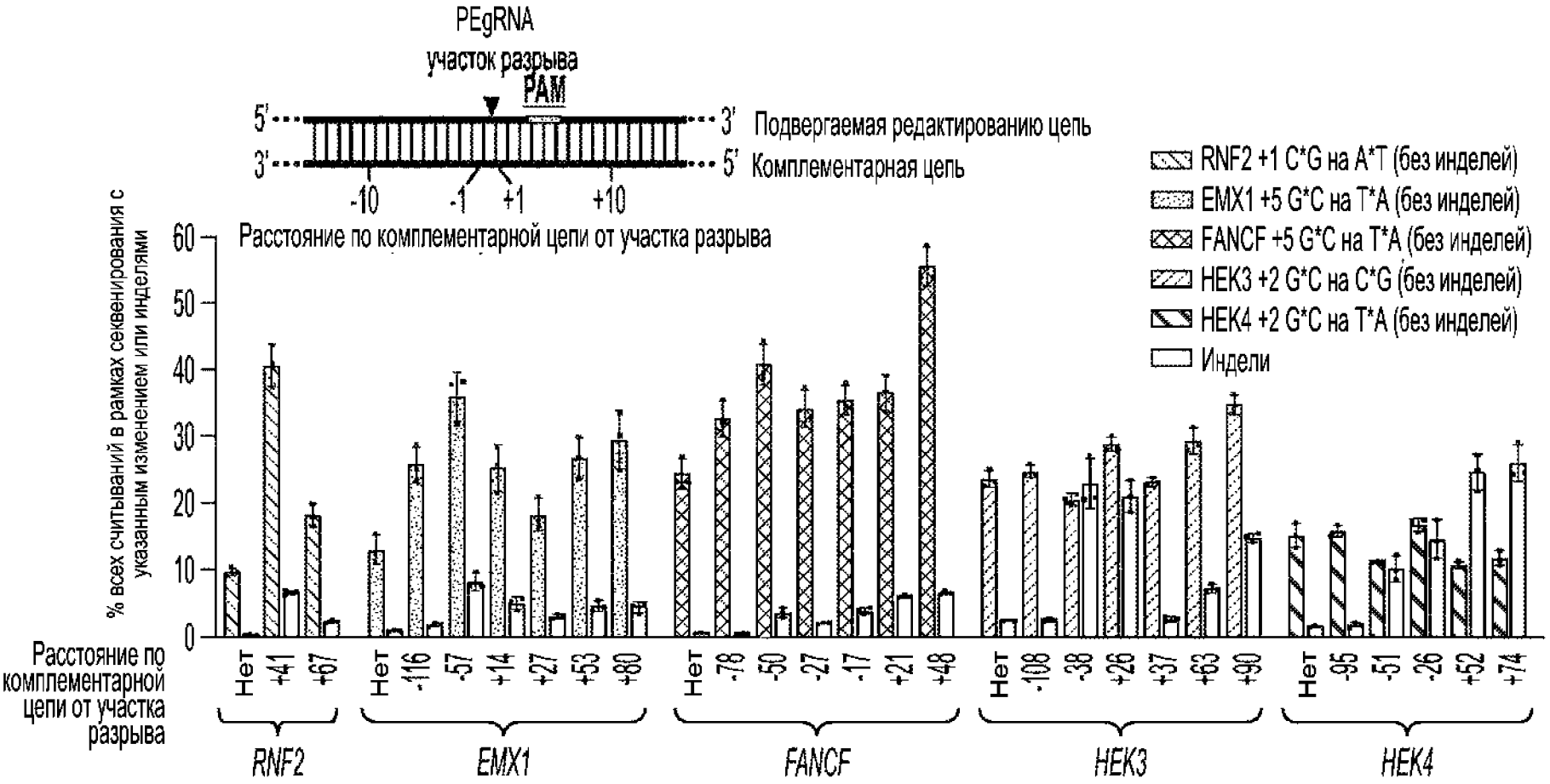

PE2, 13 nt PBS,
HEK3, +1 T на A
% всех считываний в рамках секвенирования с конвертацией T*A на A*T
длина RT матрицы (нуклеотиды)
PE2, 13 nt PBS,
EMX1, +1 G на C
% всех считываний в рамках секвенирования с конвертацией G*C на C*G
длина RT матрицы (нуклеотиды)
PE2, 13 nt PBS,
FANCF, +5 G на T
% всех считываний в рамках секвенирования с конвертацией G*C на T*A
длина RT матрицы (нуклеотиды)
PE2, 15 nt PBS,
RNF2, +1 C на A
% всех считываний в рамках секвенирования с конвертацией C*G на A*T
длина RT матрицы (нуклеотиды)
PE2, 11 nt PBS,
HEK4, +2 G на T
% всех считываний в рамках секвенирования с конвертацией G*C на T*A
длина RT матрицы (нуклеотиды)

ФИГ.39D



57/165

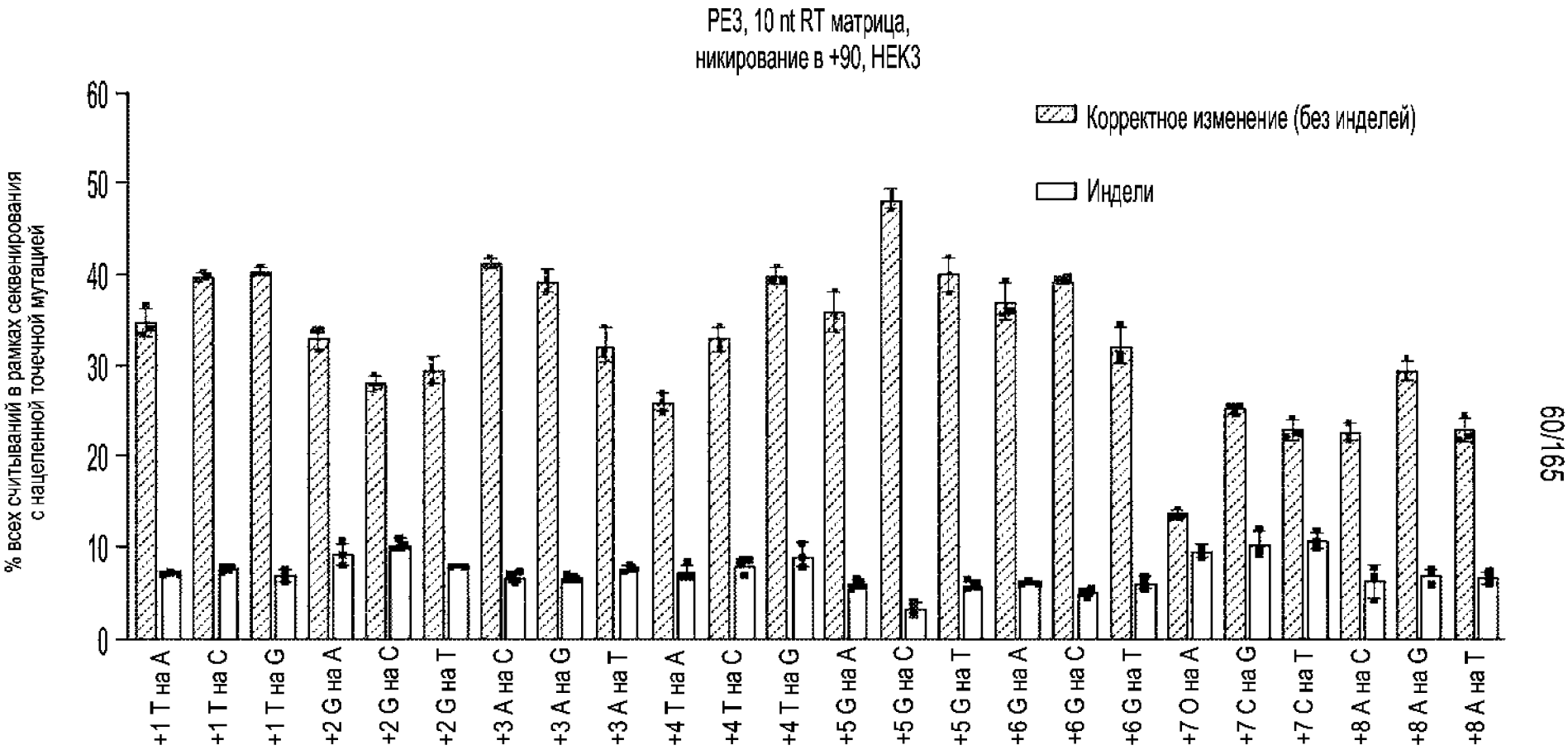
ФИГ.40A


PEgRNA
участок разрыва
PAM
5'
3'
3'
5'
Подвергаемая редактированию цепь
Комплементарная цепь
-10
-1
+1
+10
Расстояние по комплементарной цепи от участка разрыва
% всех считываний в рамках секвенирования с указанным изменением или инделями
RNF2 +1 C*G на A*T (без инделей)
EMX1 +5 G*C на T*A (без инделей)
FANCF +5 G*C на T*A (без инделей)
HEK3 +2 G*C на C*G (без инделей)
HEK4 +2 G*C на T*A (без инделей)
Индели
Расстояние по комплементарной цепи от участка разрыва
Нет +41 +67
Нет -116 -57 +14 +27 +53 +80
Нет -78 -50 -27 -17 +21 +48
Нет -108 -38 +26 +37 +63 +90
Нет -95 -51 -26 +52 +74
RNF2
EMX1
FANCF
HEK3
HEK4

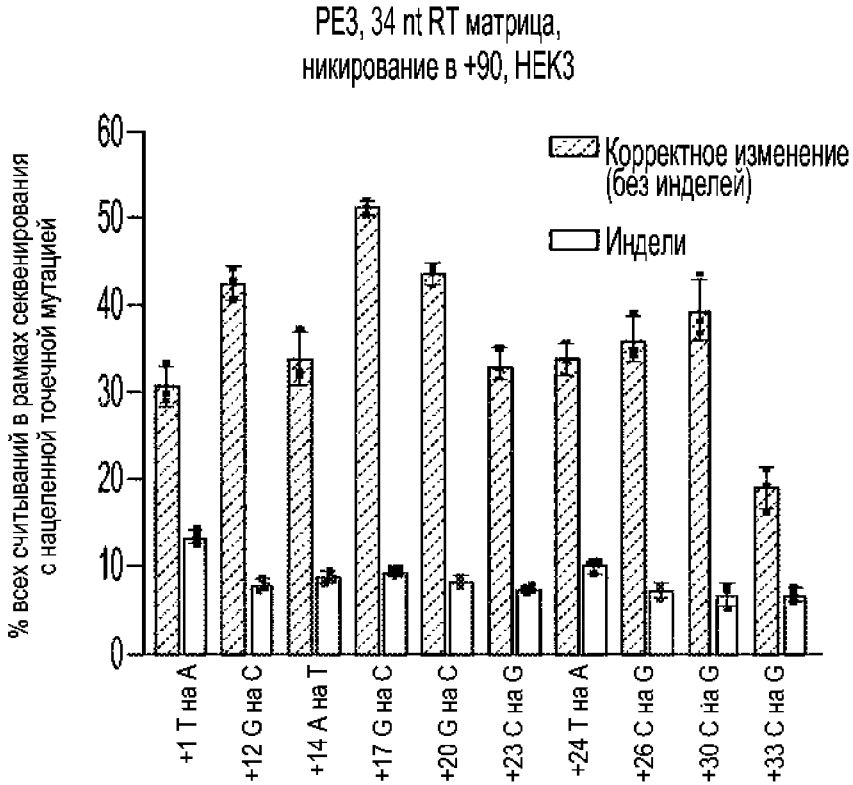
ФИГ.40В

RNF2
% всех считываний в рамках секвенирования с указанным изменением или инделями
60
40
20
0
Корректное изменение (без инделей)
Индели
PE2 +4 A на C
PE3 +4 A на C +41 разрыв
PE3b +4 A на C +5 разрыв
PE2 +4 A на G
PE3 +4 A на G +41 разрыв
PE3b +4 A на C +5 разрыв
FANCF
% всех считываний в рамках секвенирования с указанным изменением или инделями
60
50
40
30
20
10
0
Корректное изменение (без инделей)
Индели
PE2 +5 G на T
PE3 +5 G на T +48 разрыв
PE3b +5 G на T +7 разрыв
PE2 +7 A на C
PE3 +7 A на C +48 разрыв
PE3b +7 A на C +7 разрыв
HEK3
% всех считываний в рамках секвенирования с указанным изменением или инделями
60
40
20
0
Корректное изменение (без инделей)
Индели
PE2 +1 T на A
PE3 +1 T на A +90 разрыв
PE3b +1 T на A -8 разрыв

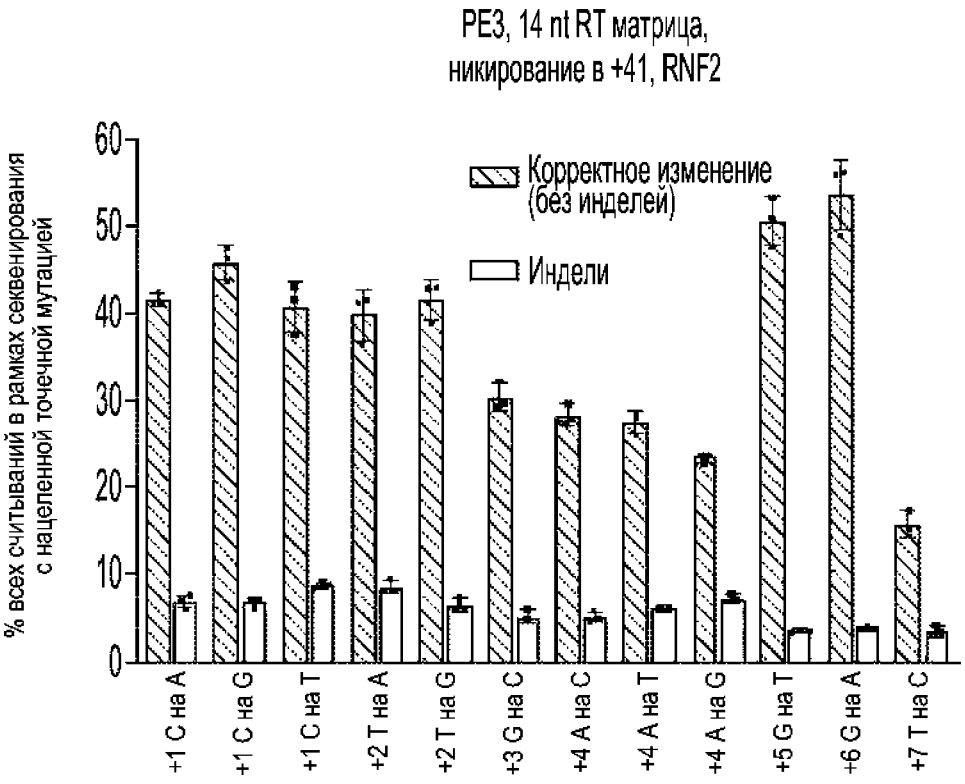
ФИГ.40C


PE3, 10 nt RT матрица,
никирование в +90, HEK3
Корректное изменение (без инделей)
Индели
% всех считываний в рамках секвенирования с нацеленной точечной мутацией
0
10
20
30
40
50
60
+1 T на A
+1 T на C
+1 T на G
+2 G на A
+2 G на C
+2 G на T
+3 A на C
+3 A на G
+3 A на T
+4 T на A
+4 T на C
+4 T на G
+5 G на A
+5 G на C
+5 G на T
+6 G на A
+6 G на C
+6 G на T
+7 O на A
+7 C на G
+7 C на T
+8 A на C
+8 A на G
+8 A на T

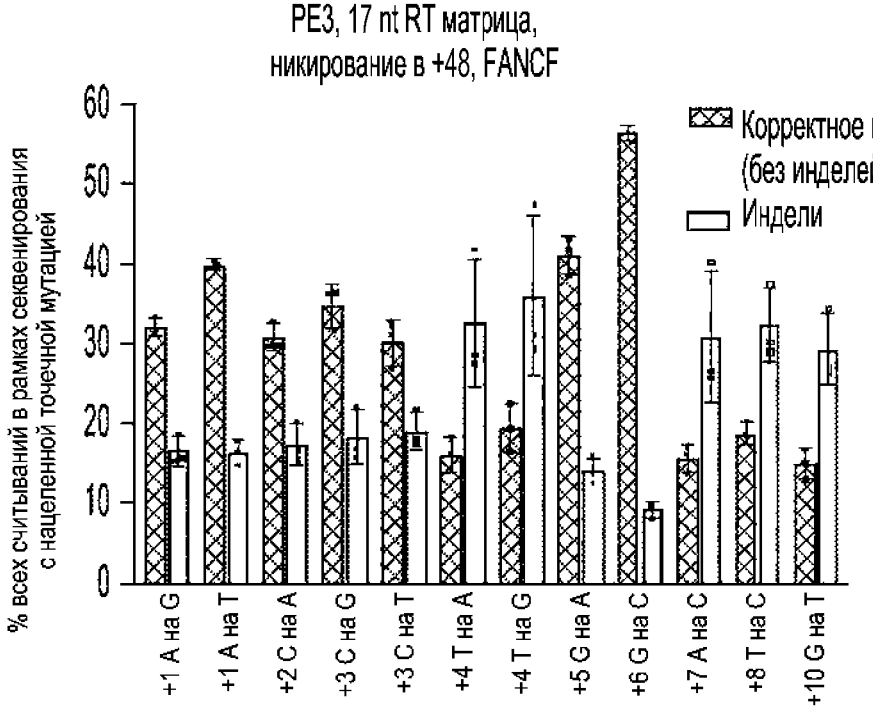
ФИГ.41A



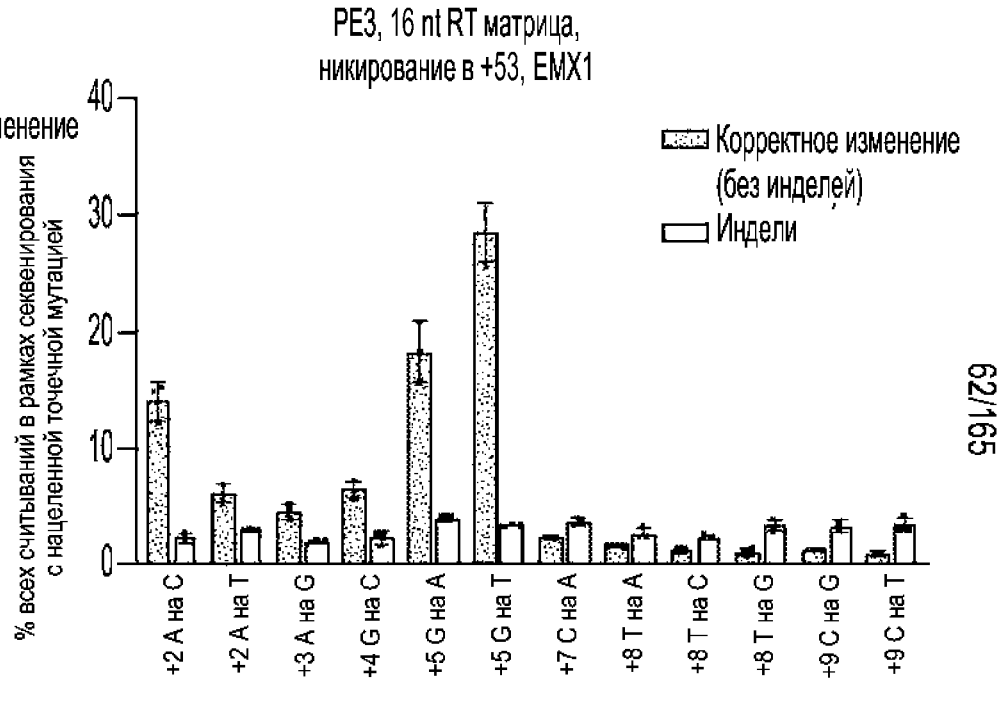
ФИГ.41B



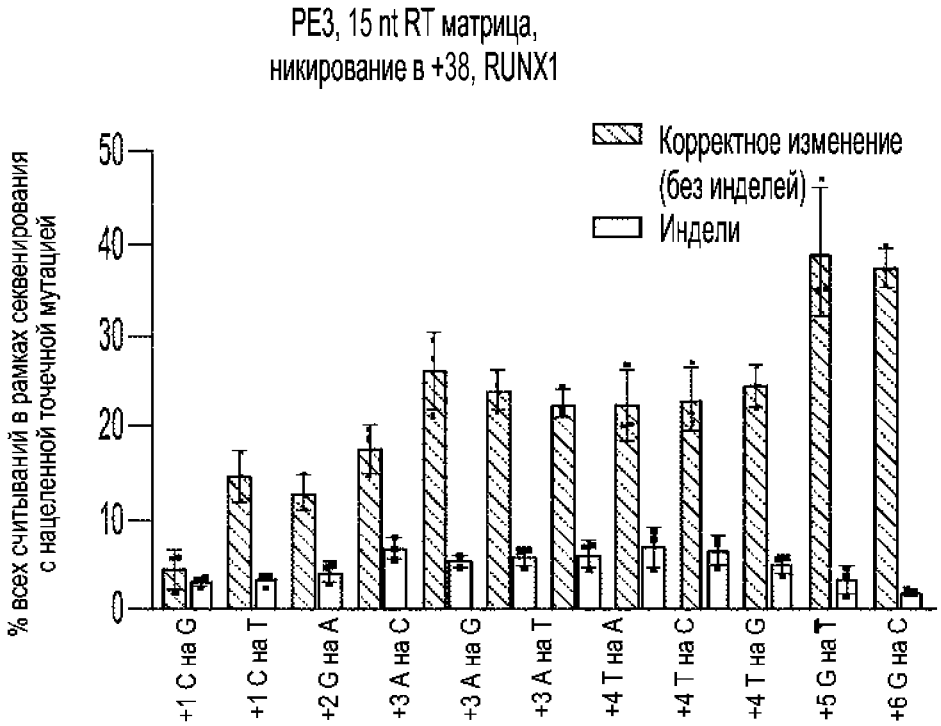
ФИГ.41C



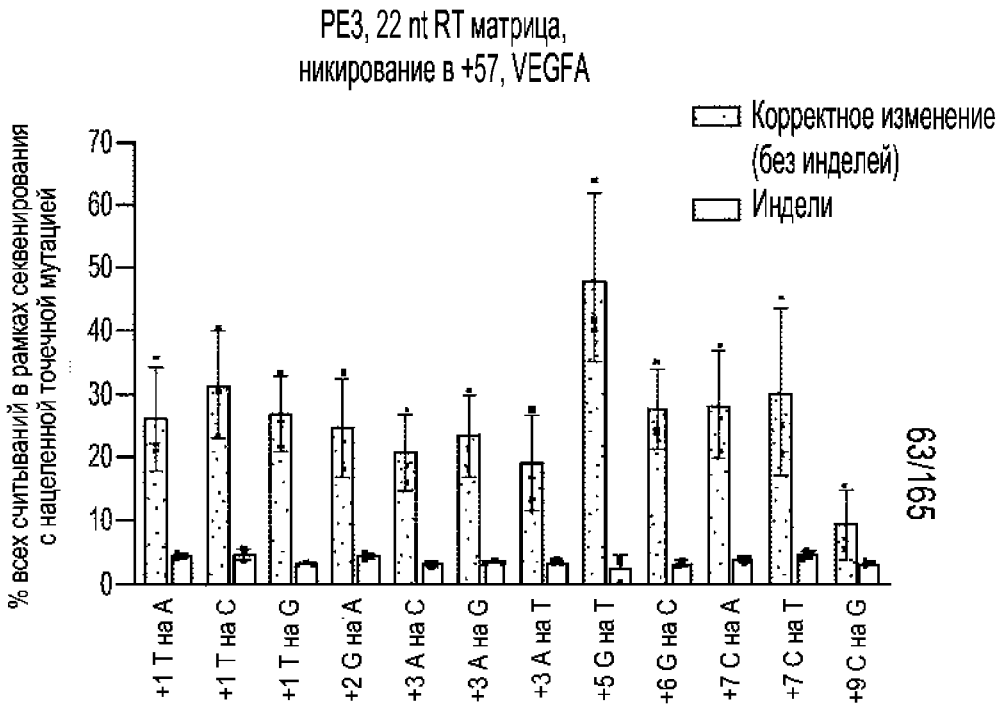
ФИГ.41D



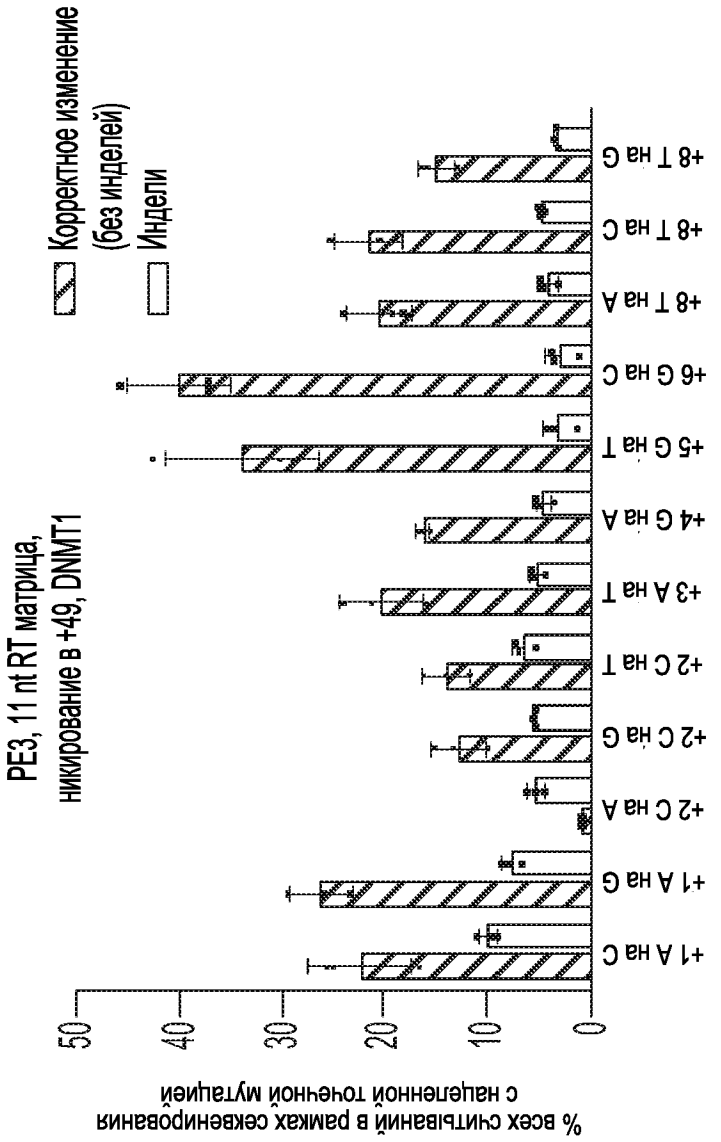
ФИГ.41C



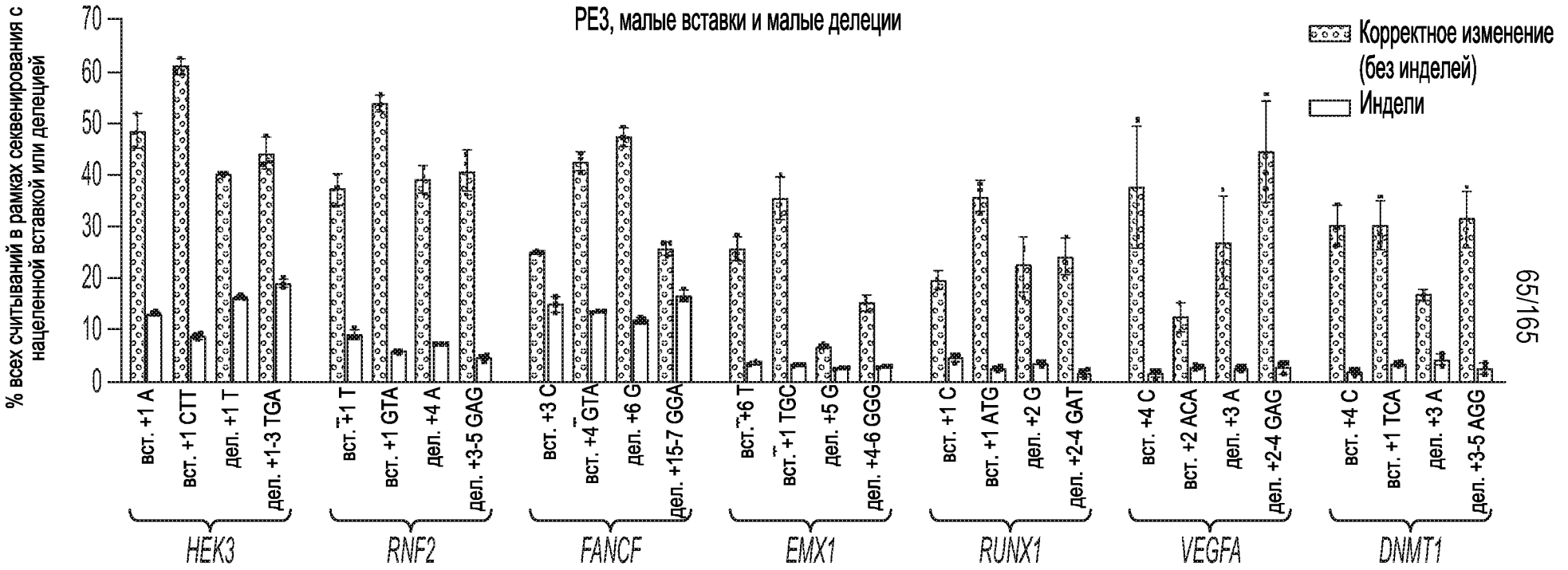
ФИГ.41F



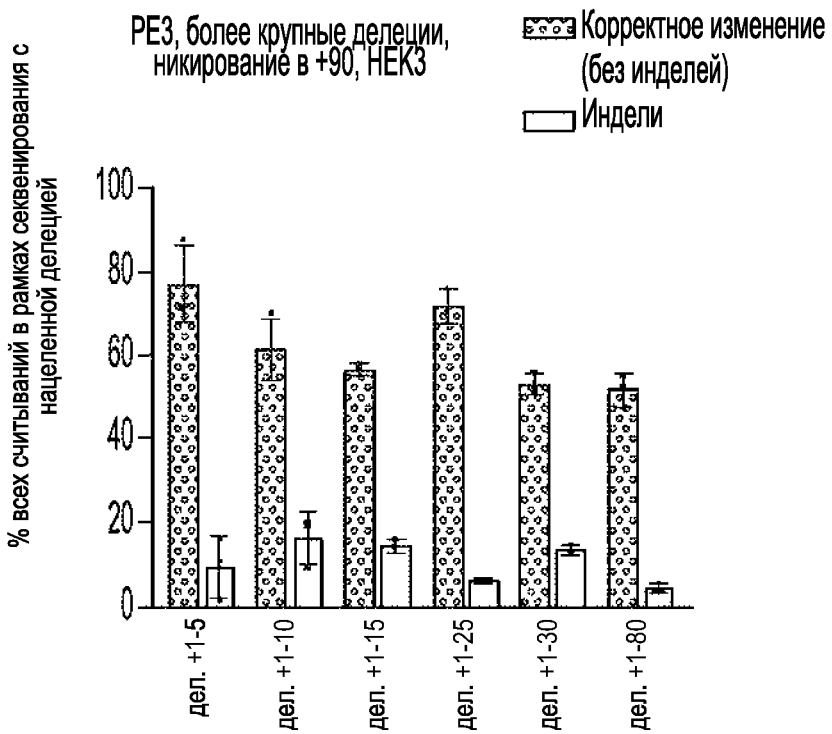
ФИГ.41G


PE3, 11 nt RT матрица,
никирование в +49, DNMT1
Корректное изменение
(без инделей)
Индели
% всех считываний в рамках секвенирования
с нацеленной точечной мутацией
0
10
20
30
40
50
+1 A на C
+1 A на G
+2 C на A
+2 C на G
+2 C на T
+3 A на T
+4 G на A
+5 G на T
+6 G на C
+8 T на A
+8 T на C
+8 T на G

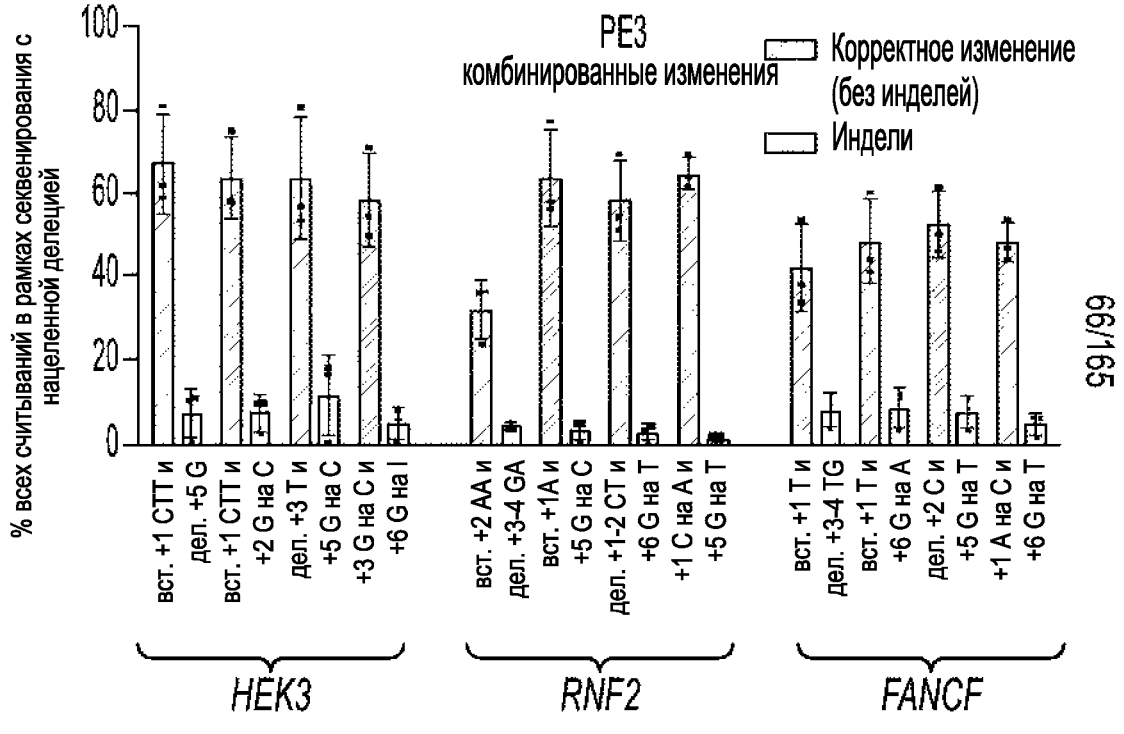
ФИГ.41H


PE3, малые вставки и малые делеции
% всех считываний в рамках секвенирования с нацеленной вставкой или делецией
Корректное изменение (без инделей)
Индели
0
10
20
30
40
50
60
70
вст. +1 A
вст. +1 CTT
дел. +1 T
дел. +1-3 TGA
вст. +1 T
вст. +1 GTA
дел. +4 A
дел. +3-5 GAG
вст. +3 C
вст. +4 GTA
дел. +6 G
дел. +15-7 GGA
вст. +6 T
вст. +1 TGC
дел. +5 G
дел. +4-6 GGG
вст. +1 C
вст. +1 ATG
дел. +2 G
дел. +2-4 GAT
вст. +4 C
вст. +2 ACA
дел. +3 A
дел. +2-4 GAG
вст. +4 C
вст. +1 TCA
дел. +3 A
дел. +3-5 AGG
HEK3
RNF2
FANCF
EMX1
RUNX1
VEGFA
DNMT1

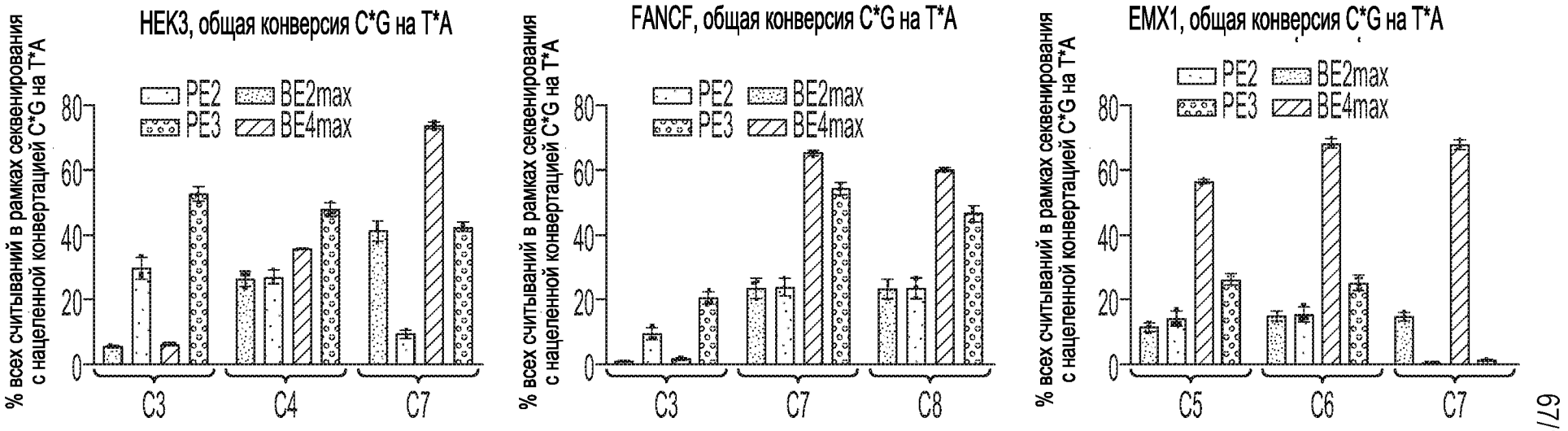
ФИГ.41I



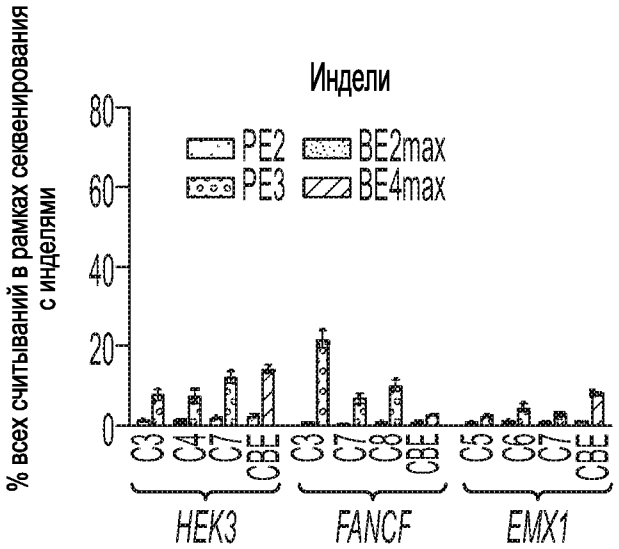
ФИГ.41J



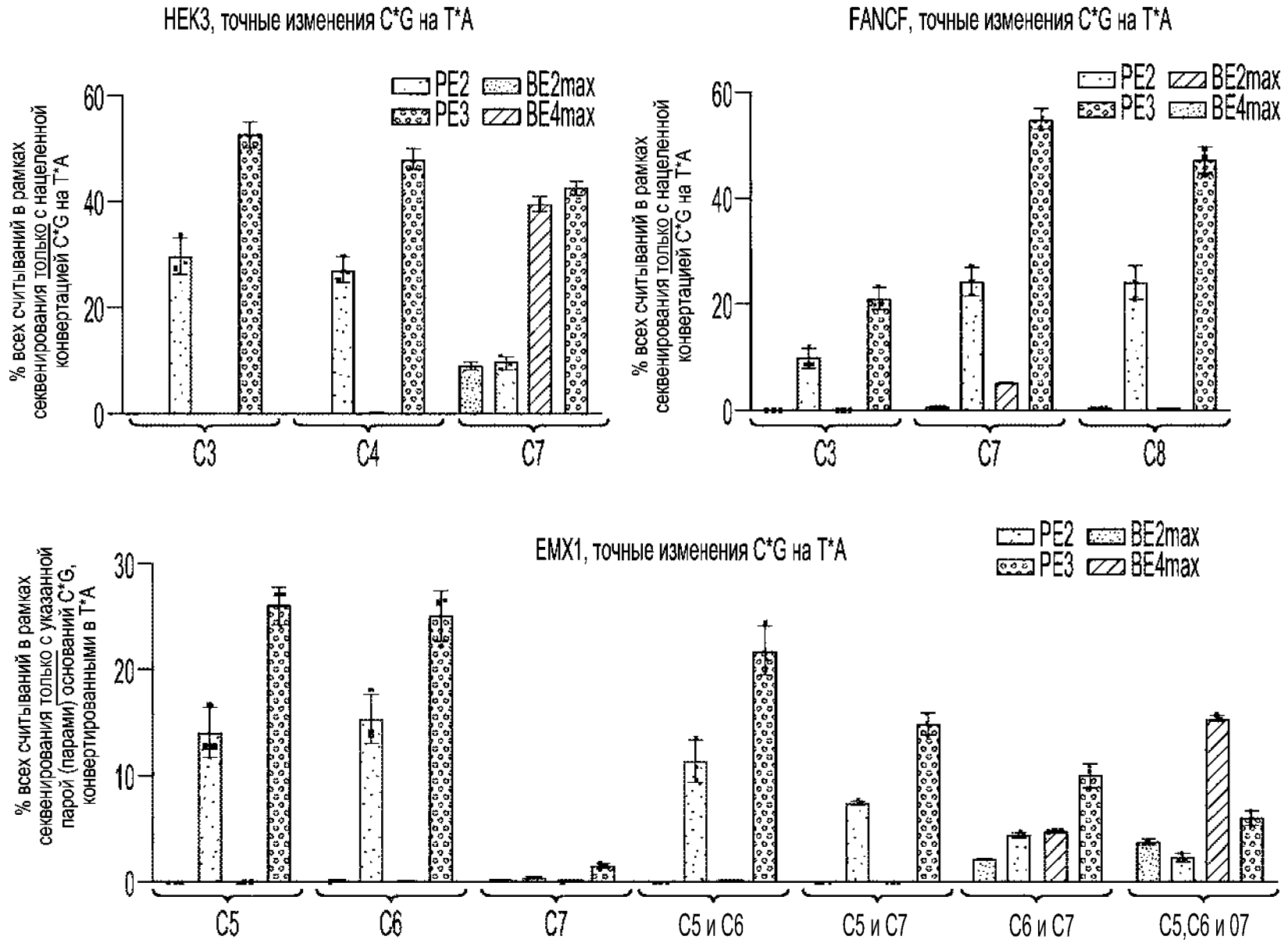
ФИГ.41K



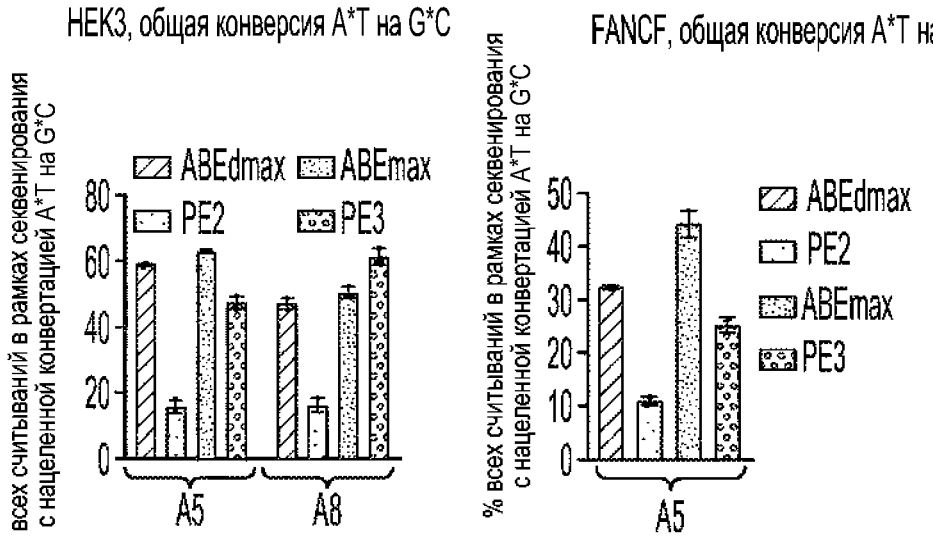
ФИГ.42A



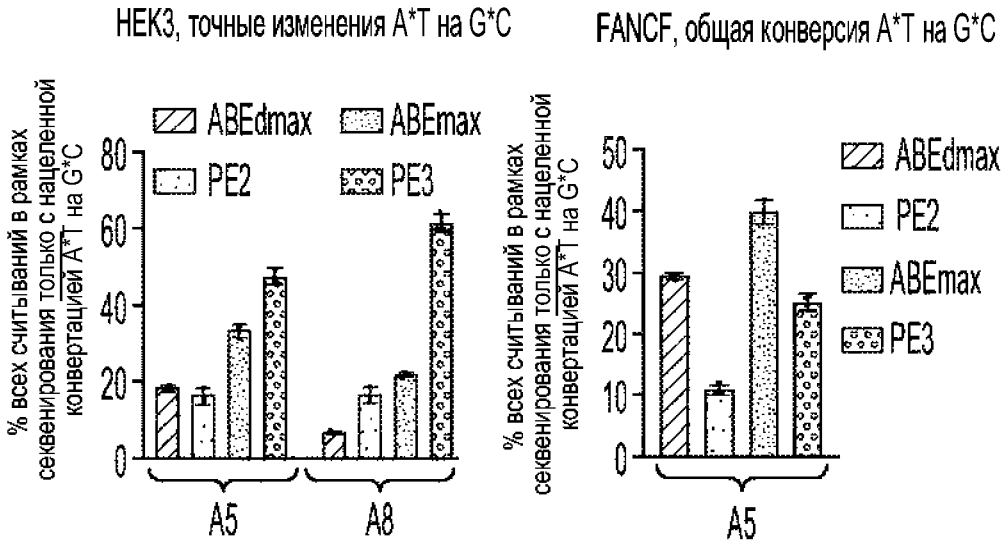
ФИГ.42B


HEK3, точные изменения C*G на T*A
% всех считываний в рамках секвенирования только с нацеленной конвертацией C*G на T*A
PE2
PE3
BE2max
BE4max
60
40
20
0
C3
C4
C7
FANCF, точные изменения C*G на T*A
% всех считываний в рамках секвенирования только с нацеленной конвертацией C*G на T*A
PE2
PE3
BE2max
BE4max
60
40
20
0
C3
C7
C8
EMX1, точные изменения C*G на T*A
% всех считываний в рамках секвенирования только с указанной парой (парами) оснований C*G, конвертированными в T*A
PE2
PE3
BE2max
BE4max
30
20
10
0
C5
C6
C7
C5 и C6
C5 и C7
C6 и C7
C5,C6 и 07

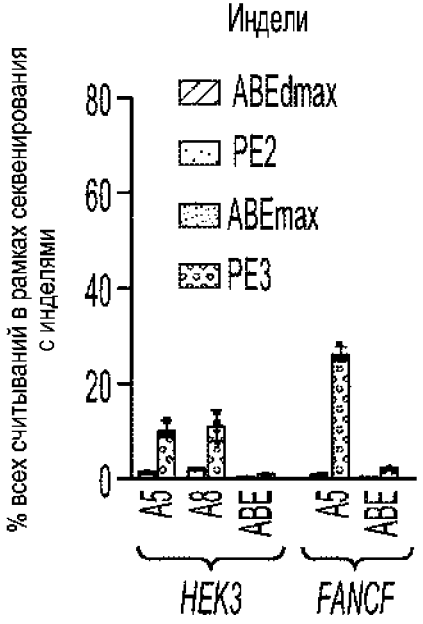
ФИГ.42C



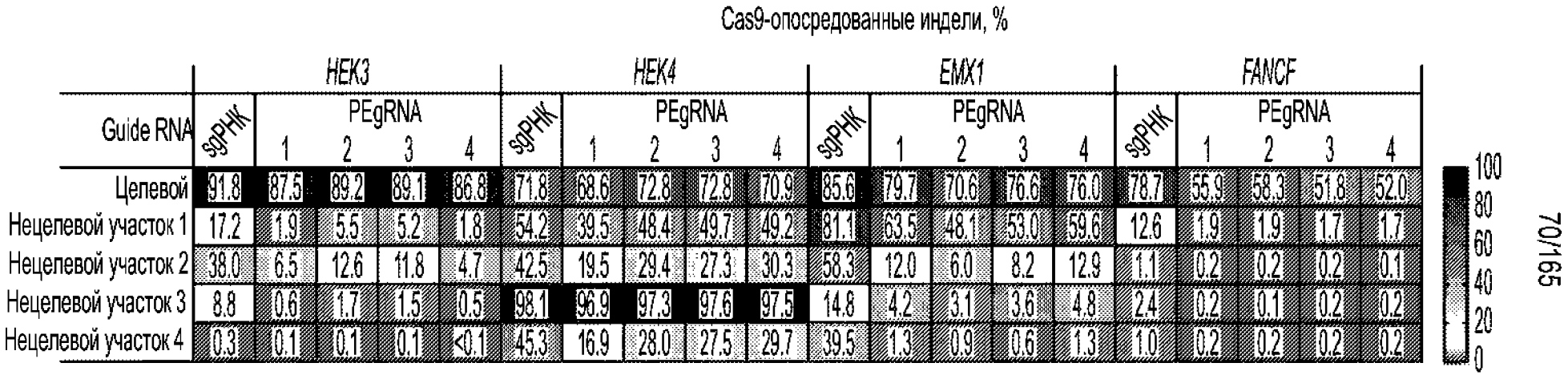
ФИГ.42D



ФИГ.42E



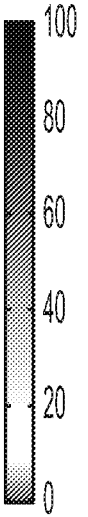
ФИГ.42F


Cas9-опосредованные индели, %
HEK3
HEK4
EMX1
FANCF
Guide RNA
sgРНК
PEgRNA
1
2
3
4
Целевой
Нецелевой участок 1
Нецелевой участок 2
Нецелевой участок 3
Нецелевой участок 4
91.8 87.5 89.2 89.1 86.8 71.8 68.6 72.8 72.8 70.9 85.6 79.7 70.6 76.6 76.0 78.7 55.9 58.3 51.8 52.0
17.2 1.9 5.5 5.2 1.8 54.2 39.5 48.4 49.7 49.2 81.1 63.5 48.1 53.0 59.6 12.6 1.9 1.9 1.7 1.7
38.0 6.5 12.6 11.8 4.7 42.5 19.5 29.4 27.3 30.3 58.3 12.0 6.0 8.2 12.9 1.1 0.2 0.2 0.2 0.1
8.8 0.6 1.7 1.5 0.5 98.1 96.9 97.3 97.6 97.5 14.8 4.2 3.1 3.6 4.8 2.4 0.2 0.1 0.2 0.2
0.3 0.1 0.1 0.1 <0.1 45.3 16.9 28.0 27.5 29.7 39.5 1.3 0.9 0.6 1.3 1.0 0.2 0.2 0.2 0.2
100
80
60
40
20
0

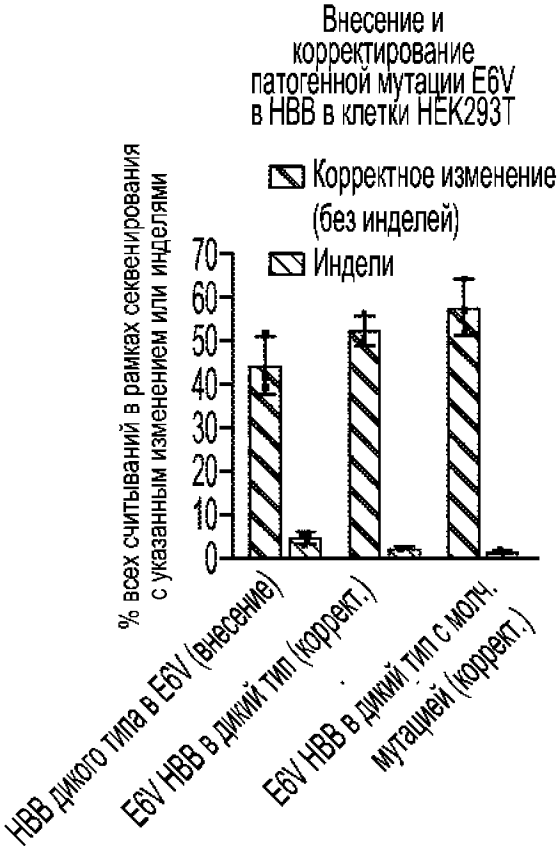
ФИГ.42G

PE-опосредованные изменения (%) и индели (%) (в скобках)

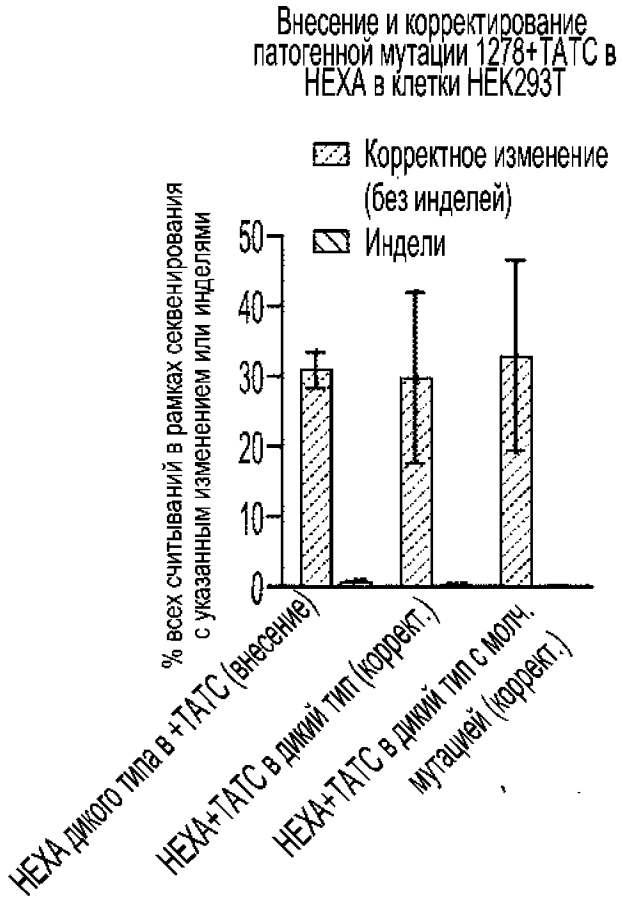
Направляющая РНК	*HEK3* (PE3) PEgRNA 1	2	3	4	*HEK4* (PE2) PEgRNA 1	2	3	4	*EMX1* (PE3) PEgRNA 1	2	3	4	*FANCF* (PE3) PEgRNA 1	2	3	4
Целевой	44.2 (11.9)	61.2 (8.8)	40.4 (16.5)	48.4 (3.3)	18.2 (0.9)	14.4 (1.8)	9.8 (2.0)	7.9 (2.2)	28.6 (3.5)	14.1 (2.4)	35.7 (3.3)	15.4 (2.9)	56.9 (9.3)	32.4 (16.7)	42.8 (13.6)	47.6 (12.0)
Нецелевой участок 1	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	0.4 (<0.1)	<0.1 (<0.1)	0.4 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	0.7 (0.1)	<0.1 (0.1)	<0.1 (0.1)
Нецелевой участок 2	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (<0.1)
Нецелевой участок 3	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	0.2 (0.5)	6.8 (1.9)	19.2 (0.5)	7.9 (3.5)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Нецелевой участок 4	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)



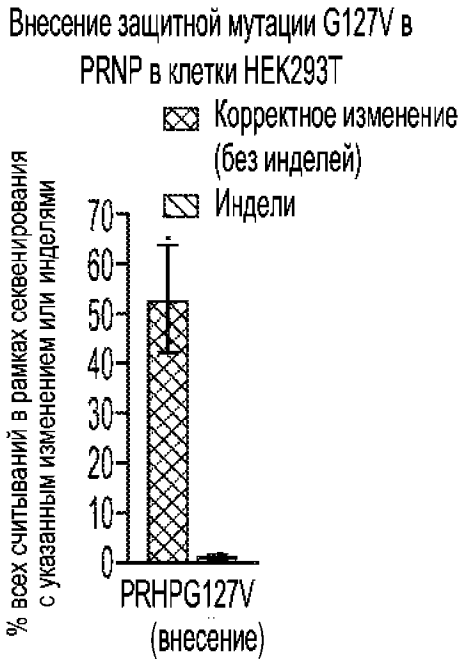
ФИГ.42H



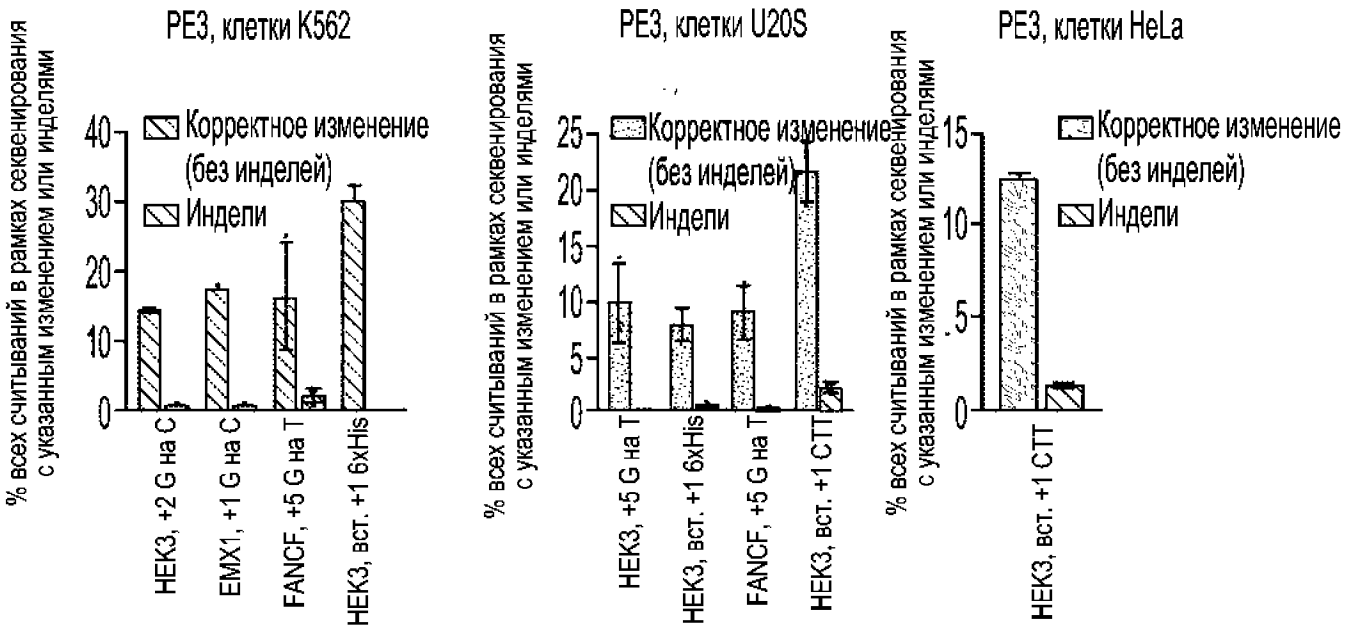
ФИГ.43A



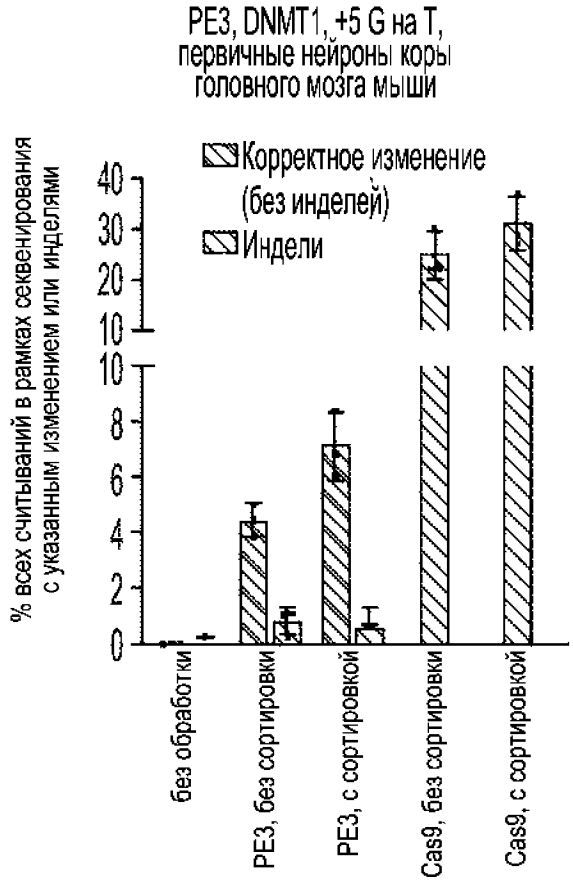
ФИГ.43B



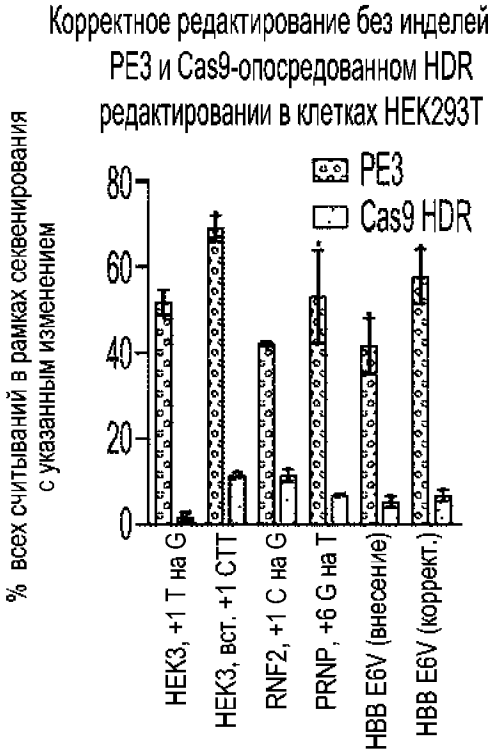
ФИГ.43C



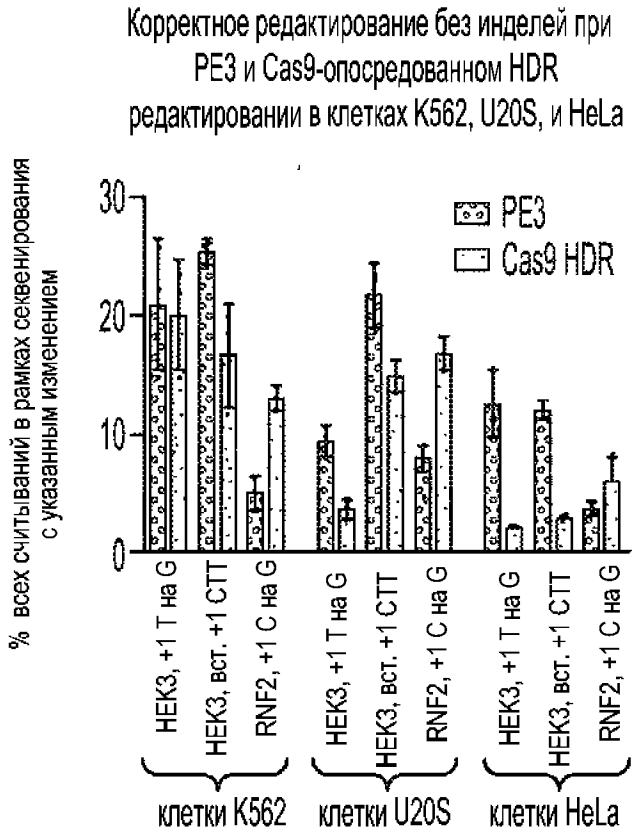
ФИГ.43D



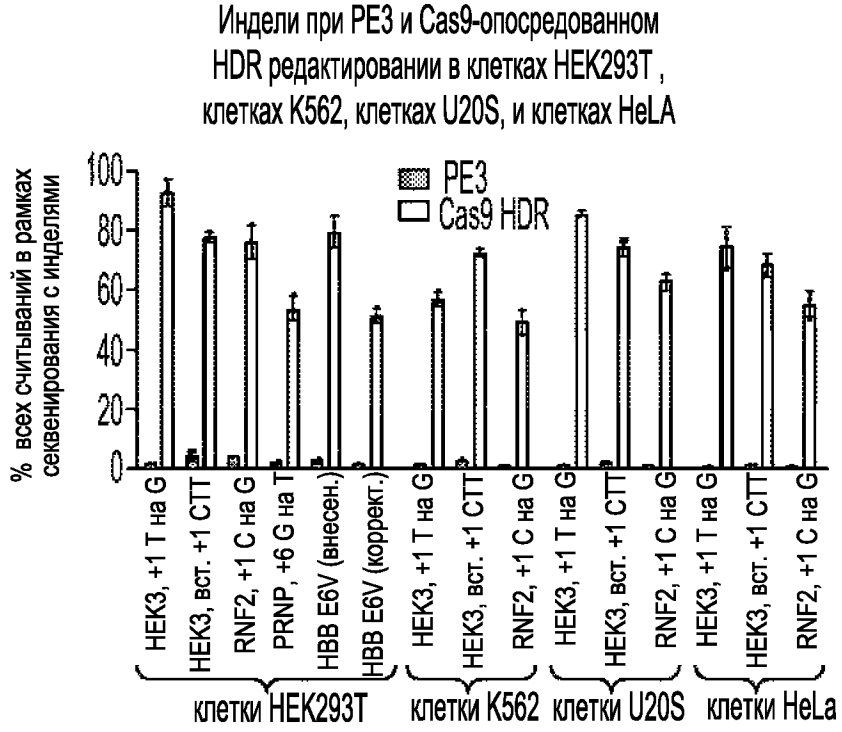
ФИГ.43E



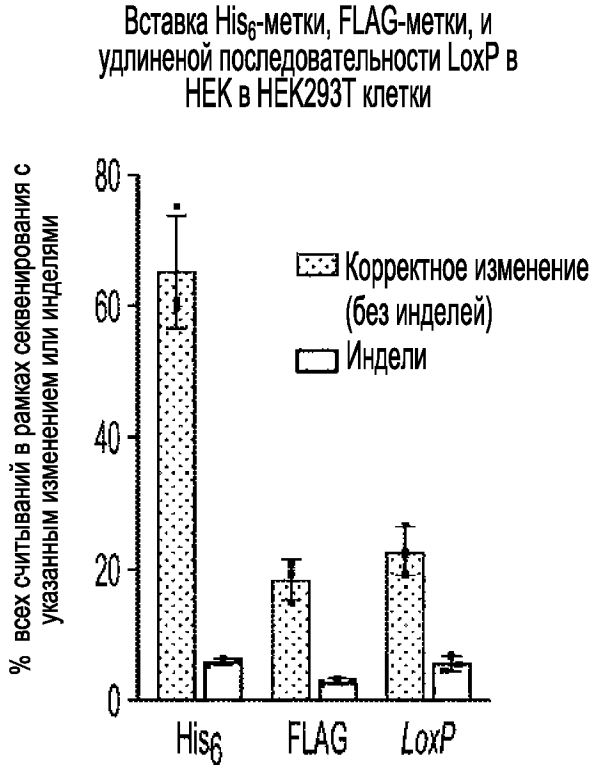
ФИГ.43F



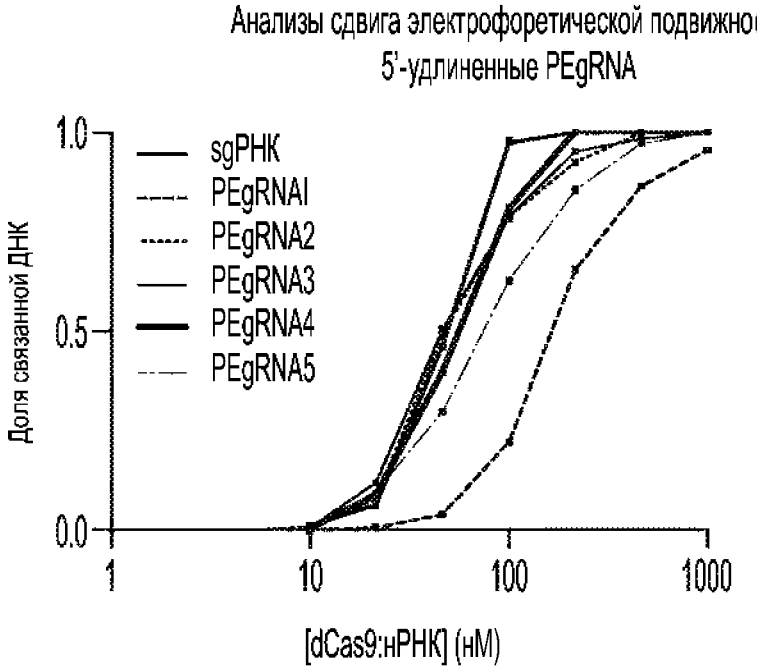
ФИГ.43G



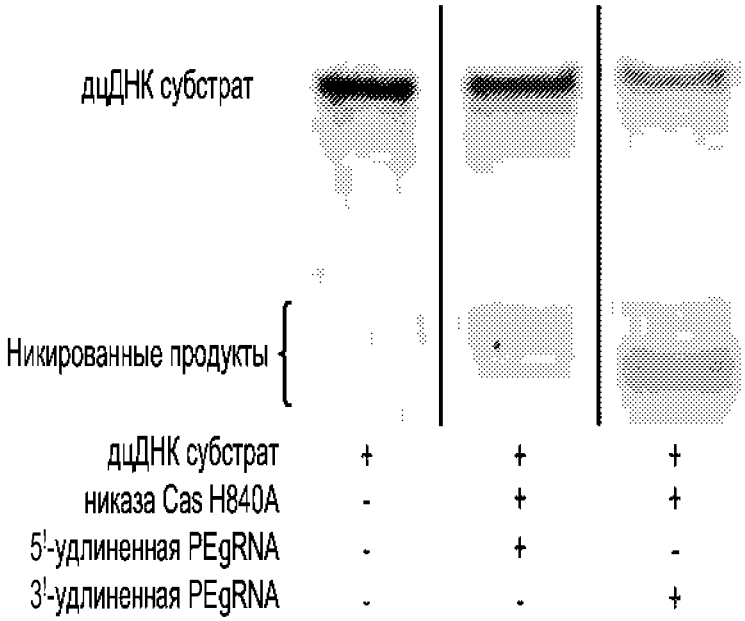
ФИГ.43H



ФИГ.43I

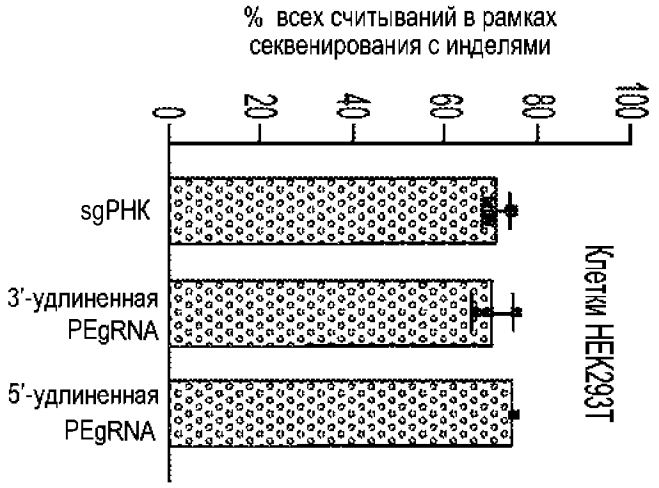


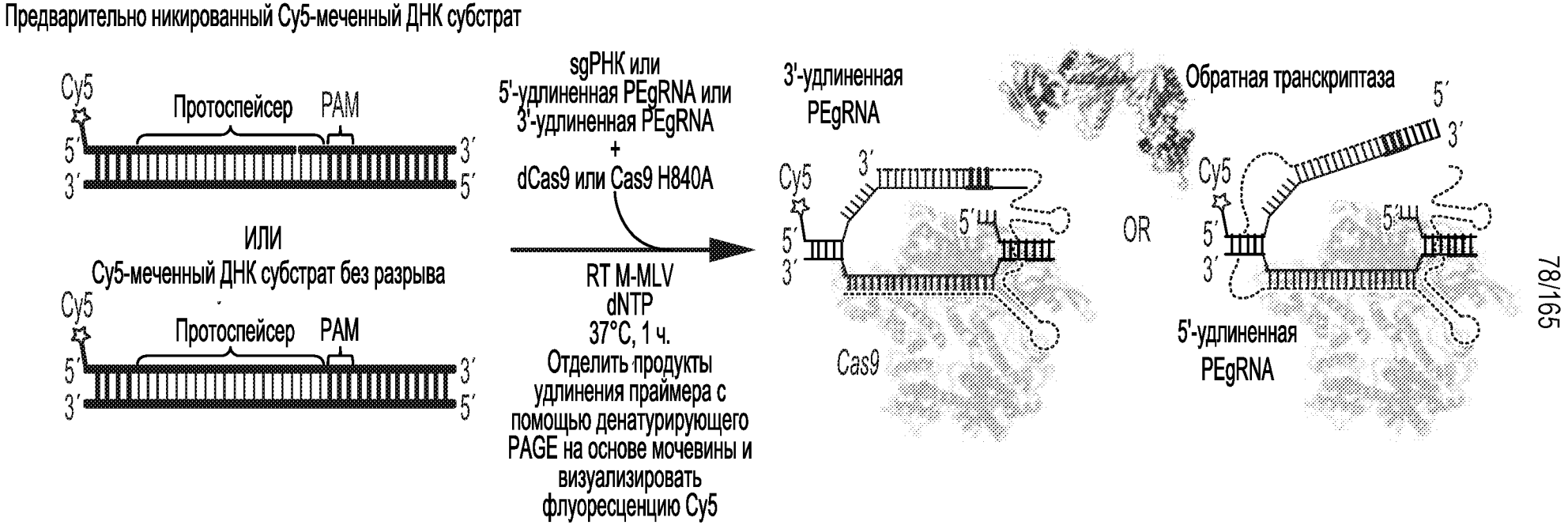
ФИГ.44A



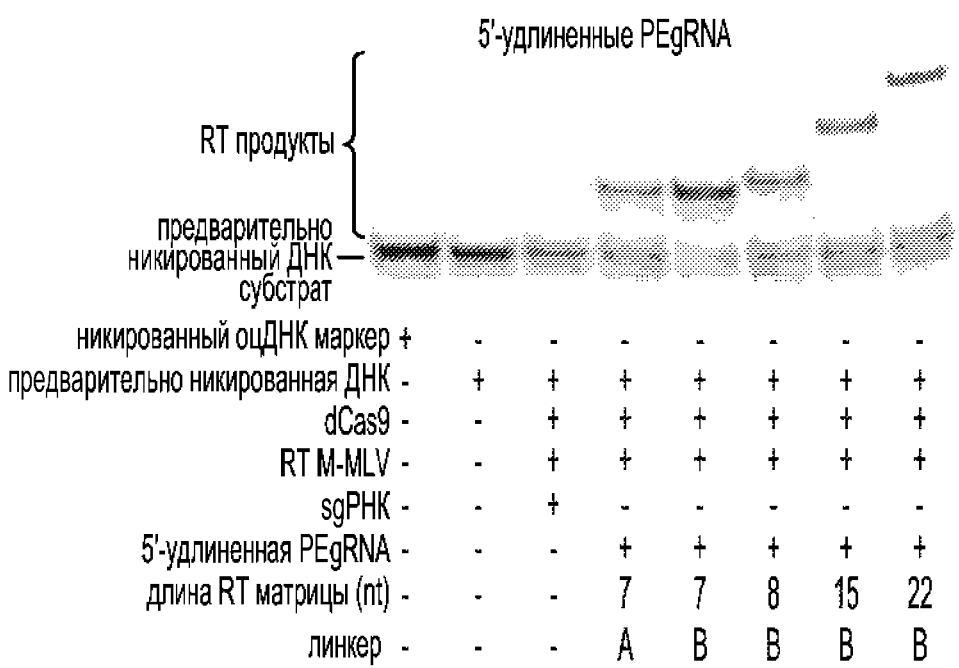
ФИГ.44B

ФИГ.44С

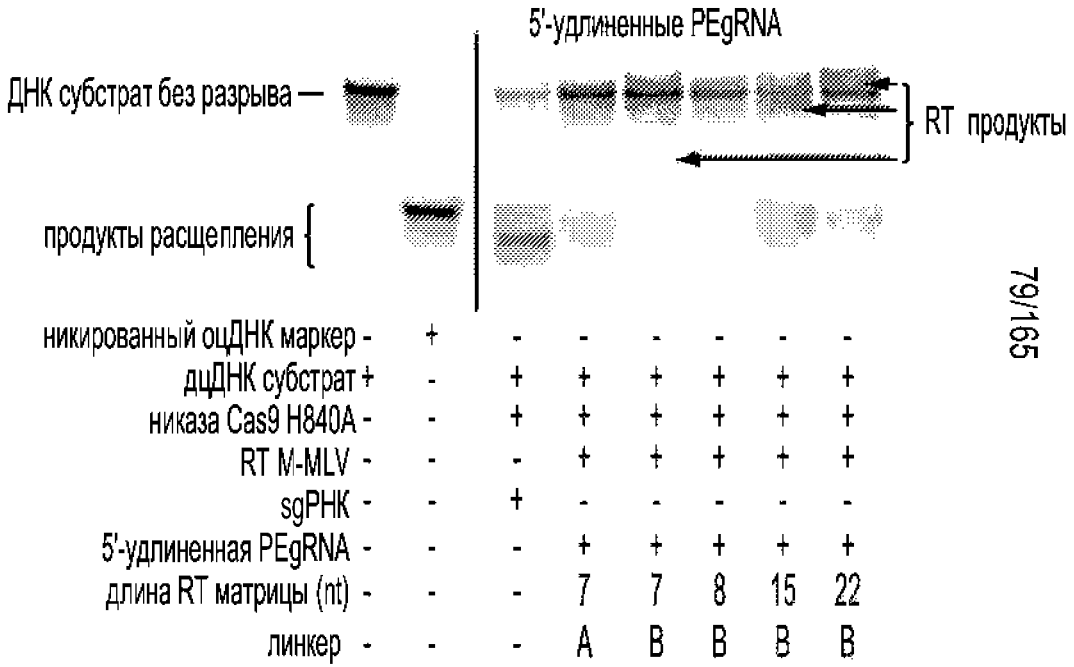

% всех считываний в рамках секвенирования с инделями
0
20
40
60
80
100
sgPHK
3'-удлиненная PEgRNA
5'-удлиненная PEgRNA
Клетки HEK293T


Предварительно никированный Cy5-меченный ДНК субстрат
Cy5
Протоспейсер
PAM
5′
3′
3′
5′
ИЛИ
Cy5-меченный ДНК субстрат без разрыва
Cy5
Протоспейсер
PAM
5′
3′
3′
5′
sgРНК или
5'-удлиненная PEgRNA или
3'-удлиненная PEgRNA
+
dCas9 или Cas9 H840A
RT M-MLV
dNTP
37°C, 1 ч.
Отделить продукты удлинения праймера с помощью денатурирующего PAGE на основе мочевины и визуализировать флуоресценцию Cy5
3'-удлиненная PEgRNA
Cy5
3′
5′
3′
5′
Cas9
OR
Обратная транскриптаза
5′
3′
Cy5
5′
3′
5′
3′
5'-удлиненная PEgRNA

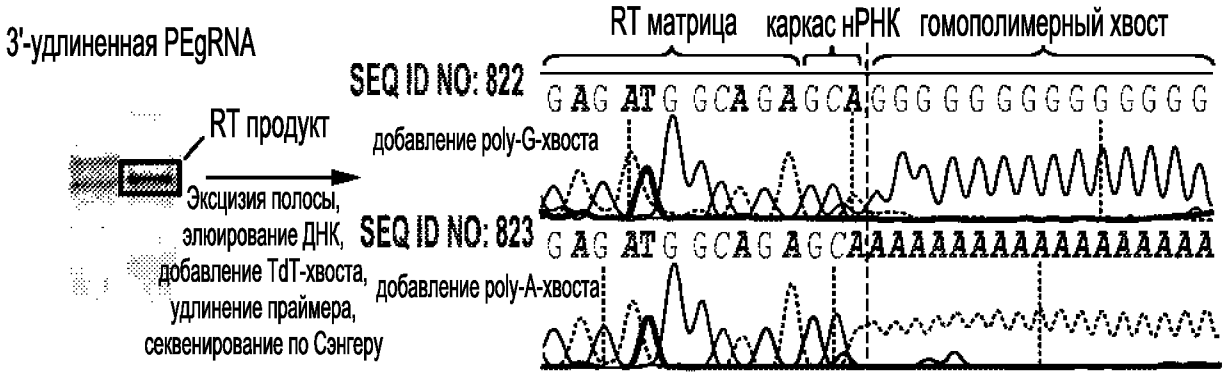
ФИГ.44D



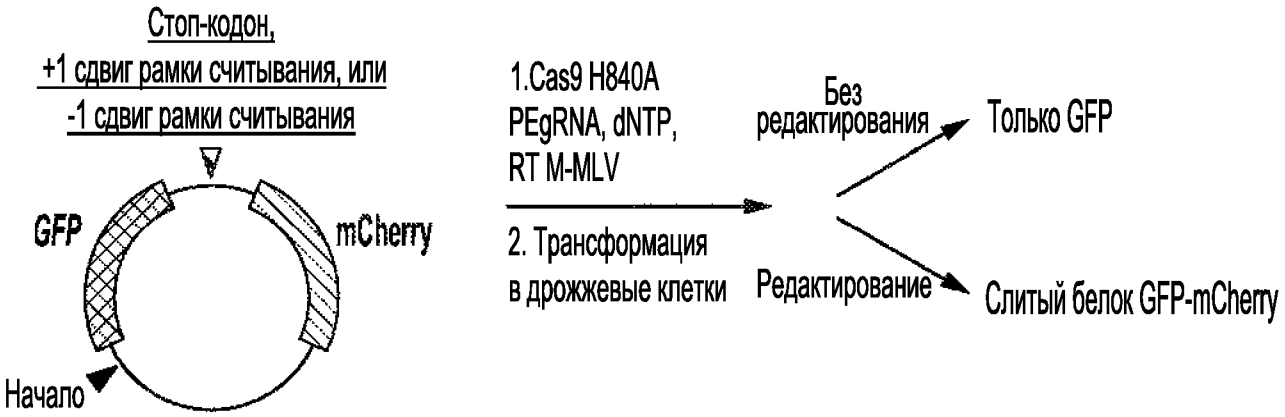
ФИГ.44E



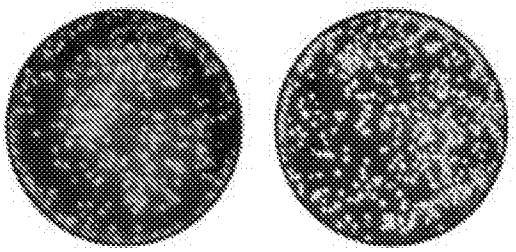
ФИГ.44F



ФИГ.44G

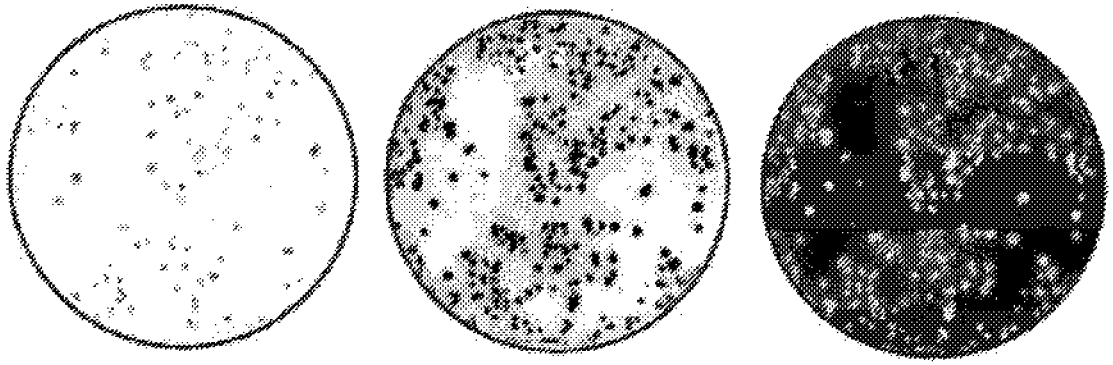


ФИГ.45A


Не подвергаемый редактированию GFP-STOP-mCherry отрицательный контроль
Предварительно отредактированный GFP-Cherry отрицательный контроль

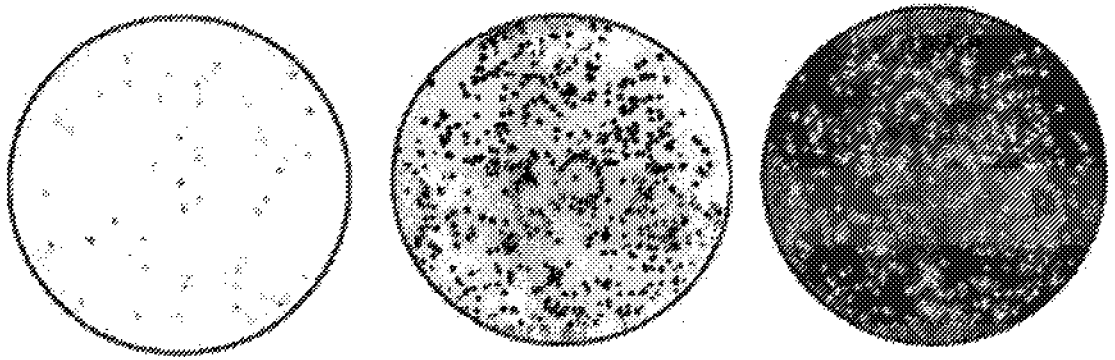
ФИГ.45В

3'-удлиненная PEgRNA, корректирование стоп-кодона (T*A на A*T)



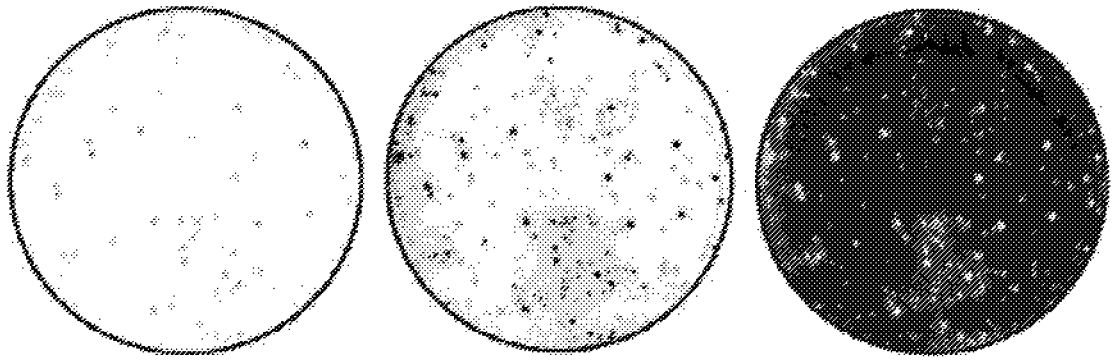
ФИГ.45C

5'-удлиненная PEgRNA, корректирование стоп-кодона (T*A на A*T)



ФИГ.45D

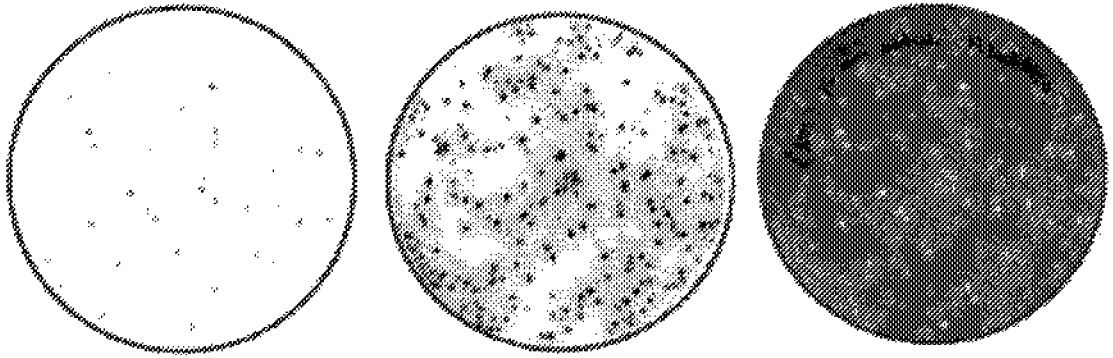
3'-удлиненная PEgRNA, корректирование +1 сдвига рамки считывания (1 bp делеция)



mCherry GFP совмещение

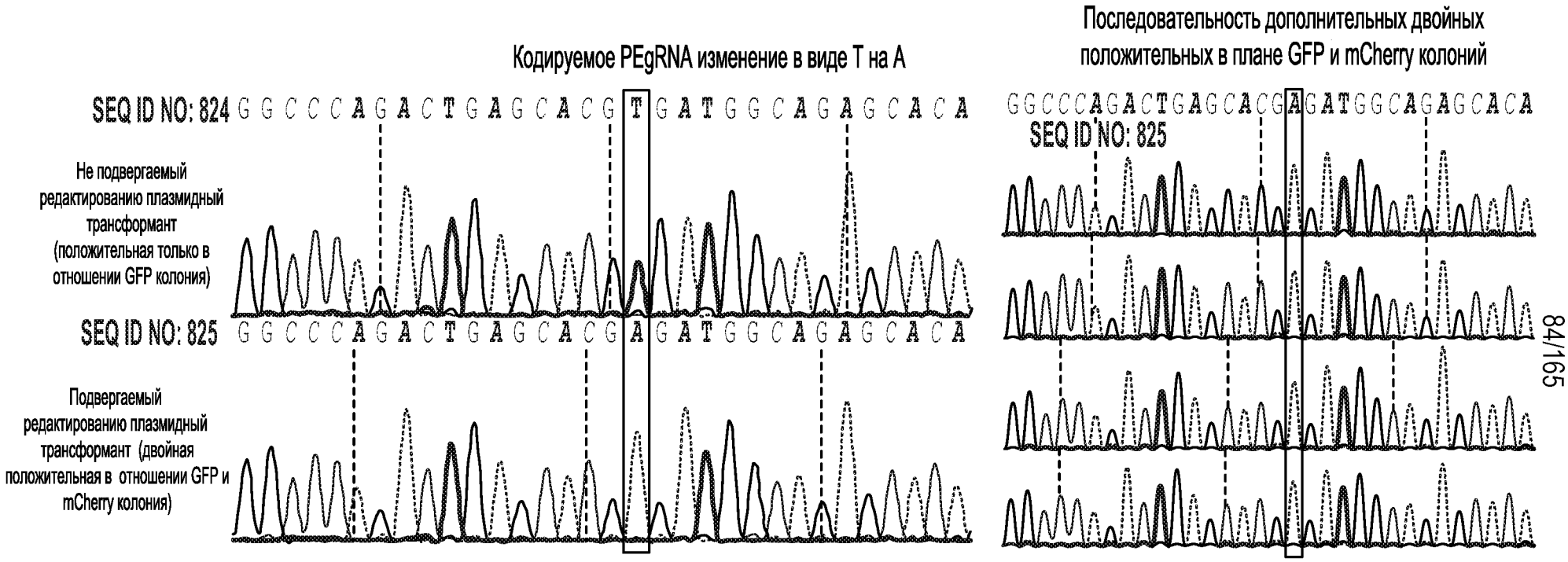
ФИГ.45E

3'-удлиненная PEgRNA, корректирование -1 сдвига рамки считывания (1 bp вставка)

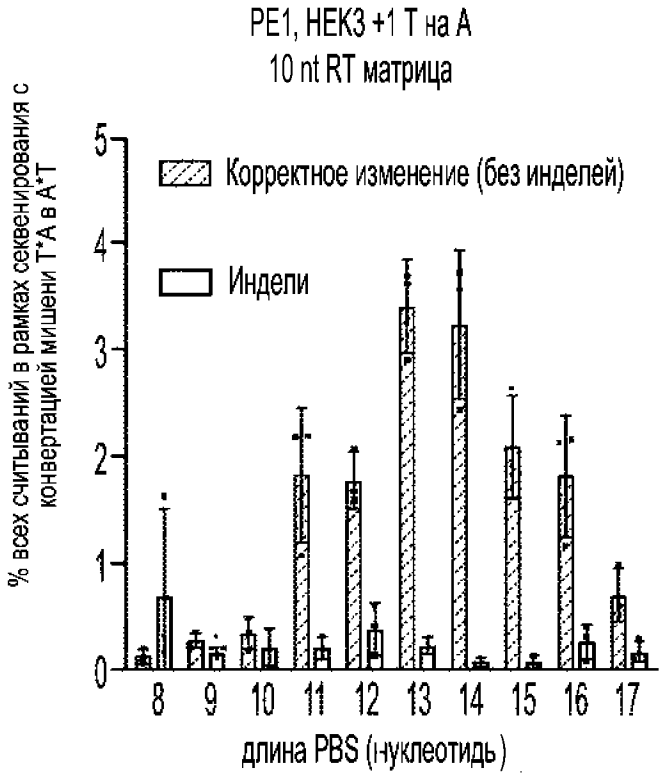


mCherry GFP совмещение

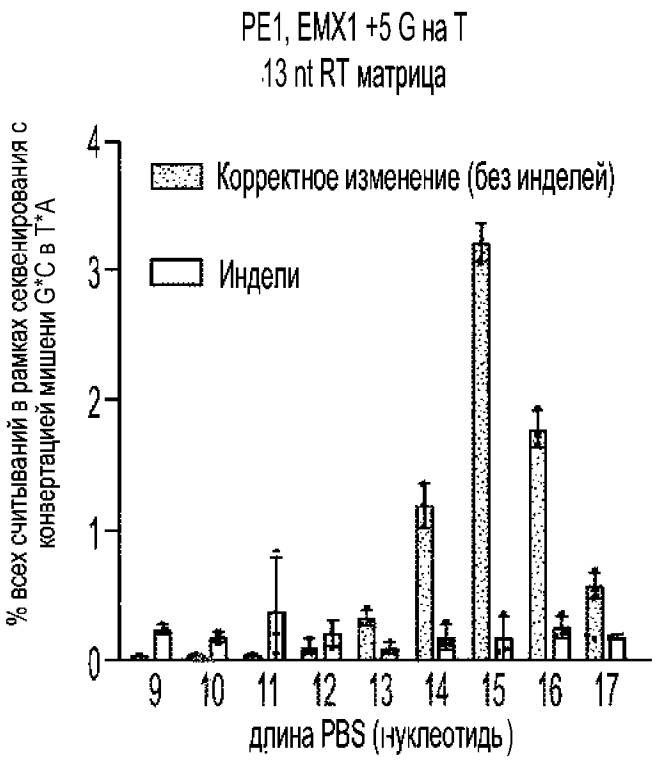
ФИГ.45F


Кодируемое PEgRNA изменение в виде T на A
Последовательность дополнительных двойных положительных в плане GFP и mCherry колоний
SEQ ID NO: 824
G G C C C A G A C T G A G C A C G T G A T G G C A G A G C A C A
Не подвергаемый редактированию плазмидный трансформант (положительная только в отношении GFP колония)
SEQ ID NO: 825
G G C C C A G A C T G A G C A C G A G A T G G C A G A G C A C A
Подвергаемый редактированию плазмидный трансформант (двойная положительная в отношении GFP и mCherry колония)
G G C C C A G A C T G A G C A C G A G A T G G C A G A G C A C A
SEQ ID NO: 825

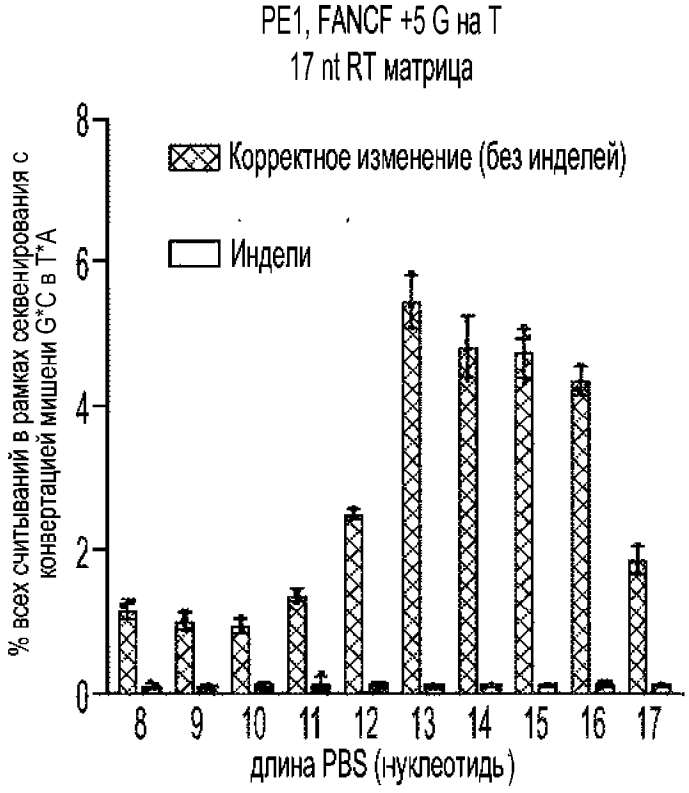
ФИГ.45G


PE1, HEK3 +1 T на A
10 nt RT матрица
% всех считываний в рамках секвенирования с конвертацией мишени T*A в A*T
Корректное изменение (без инделей)
Индели
0
1
2
3
4
5
8
9
10
11
12
13
14
15
16
17
длина PBS (нуклеотидь)

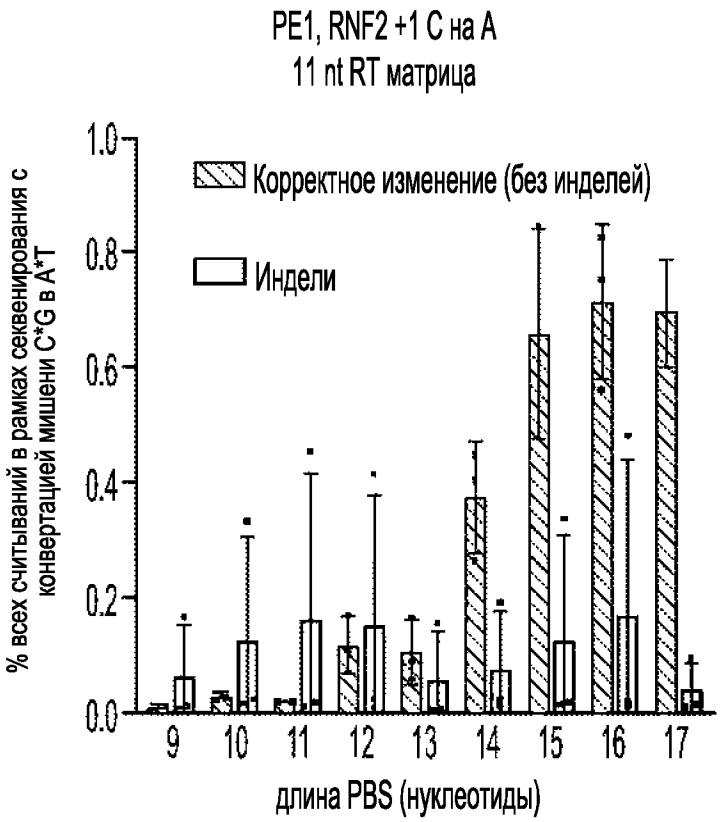
ФИГ.46A



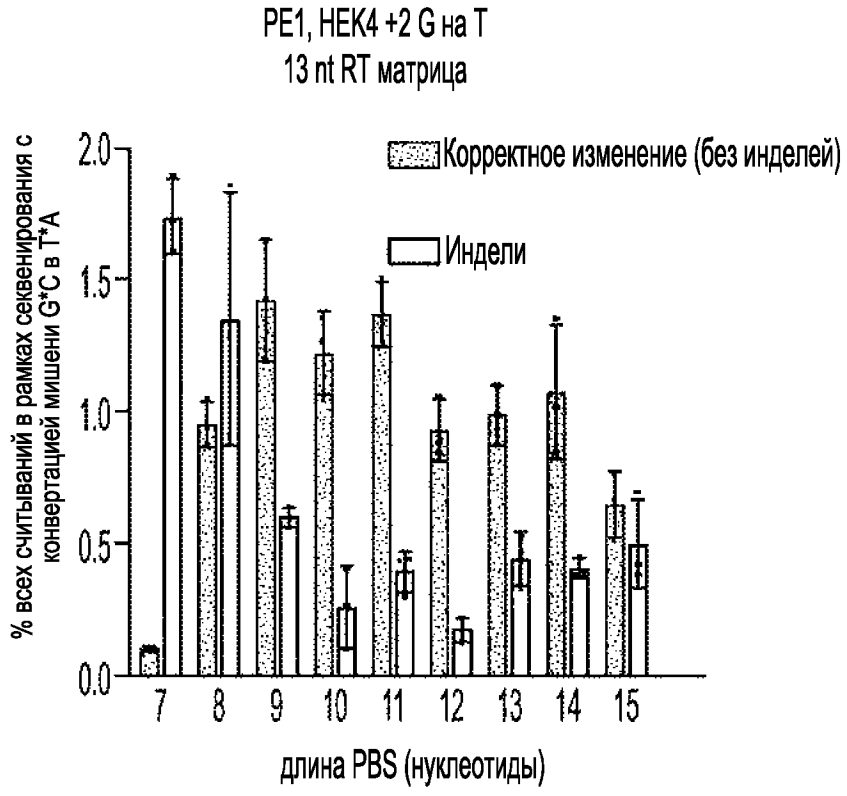
ФИГ.46B



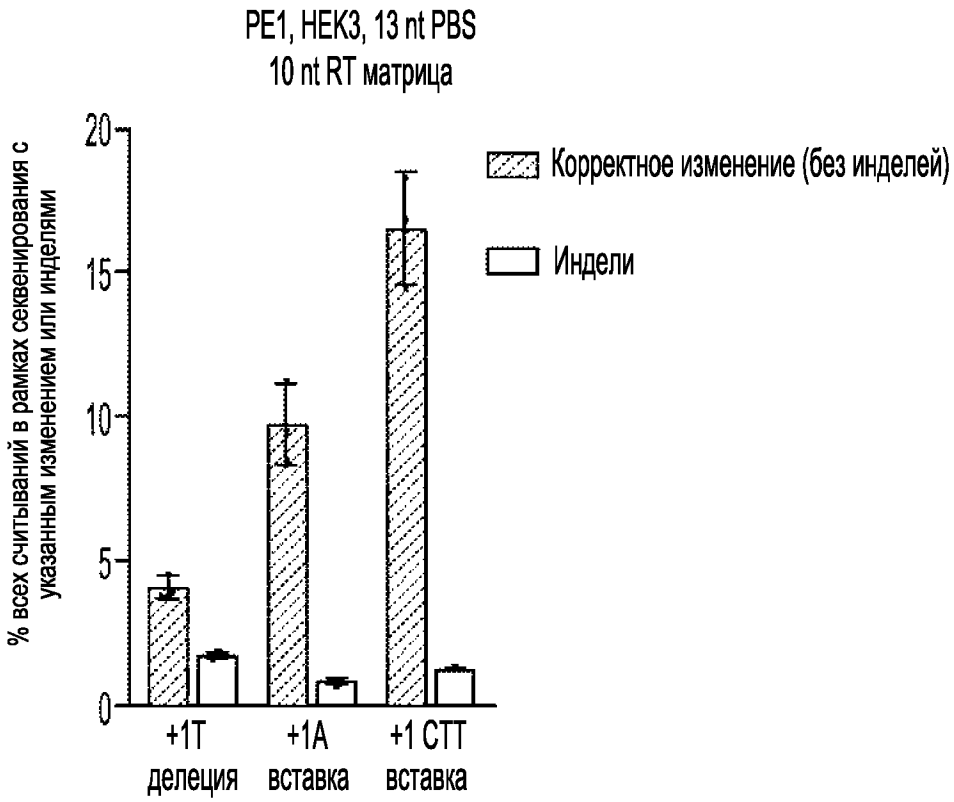
ФИГ.46C



ФИГ.46D



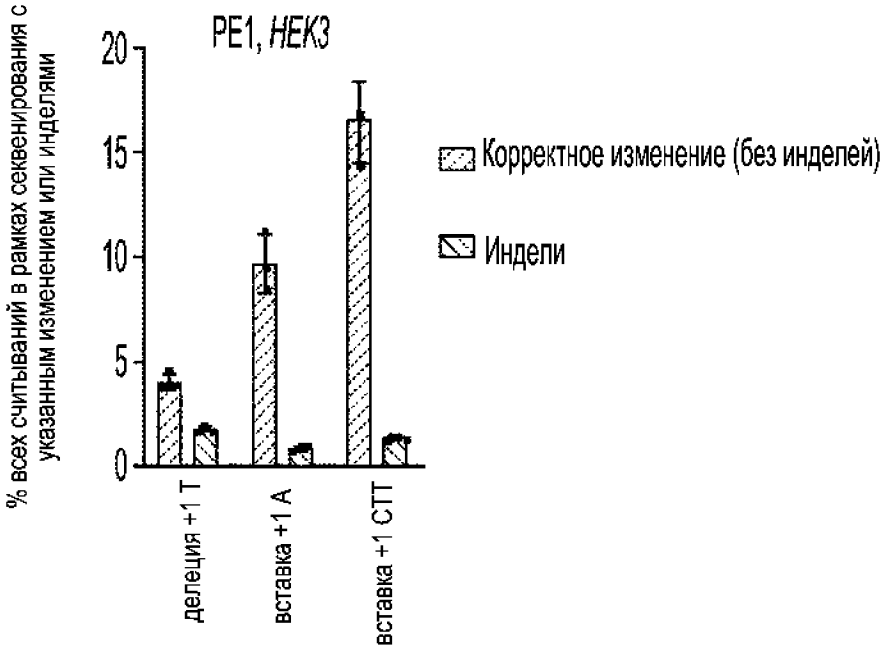
ФИГ.46E



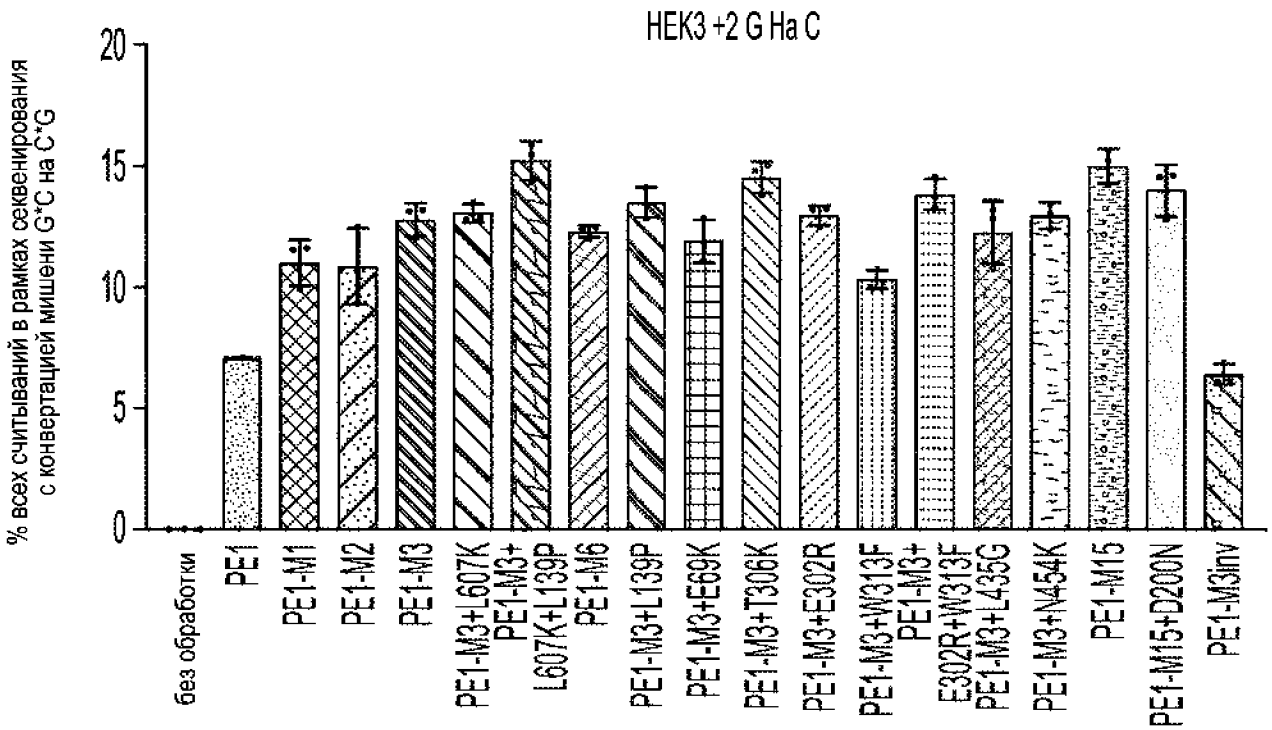
ФИГ.46F

PE1	Cas9 H840A-RT M-MIV (дикий типа)
PE1-M1	Cas9 H840A-RT M-MLV (D200N)
PE1-M2	Cas9 H840A-RT M-MLV (D200N+L603W)
PE1 -M3	Cas9 H840A-RT M-MIV (D200N+L603W+T330P)
PE1-M6	Cas9 H840A-RT M-MLV (D200N+L603W+T330P+D524G+E562Q+D583N)
PE1 -M15	Cas9 H840A-RT M-MLV (P51L+S67R+E67K+T197A+H204R+E302K+F309N+W1331F+T330P+L345G+ N454K+D524G+D583N+H594Q+D653N)
PE1-M3inv	RT M-MLV (D200N+L603W+T330P)-Cas9 H840A
PE2	Cas9 H840A-M-RT MLV (D200N+L603W+T330P+T306K+VV313F)

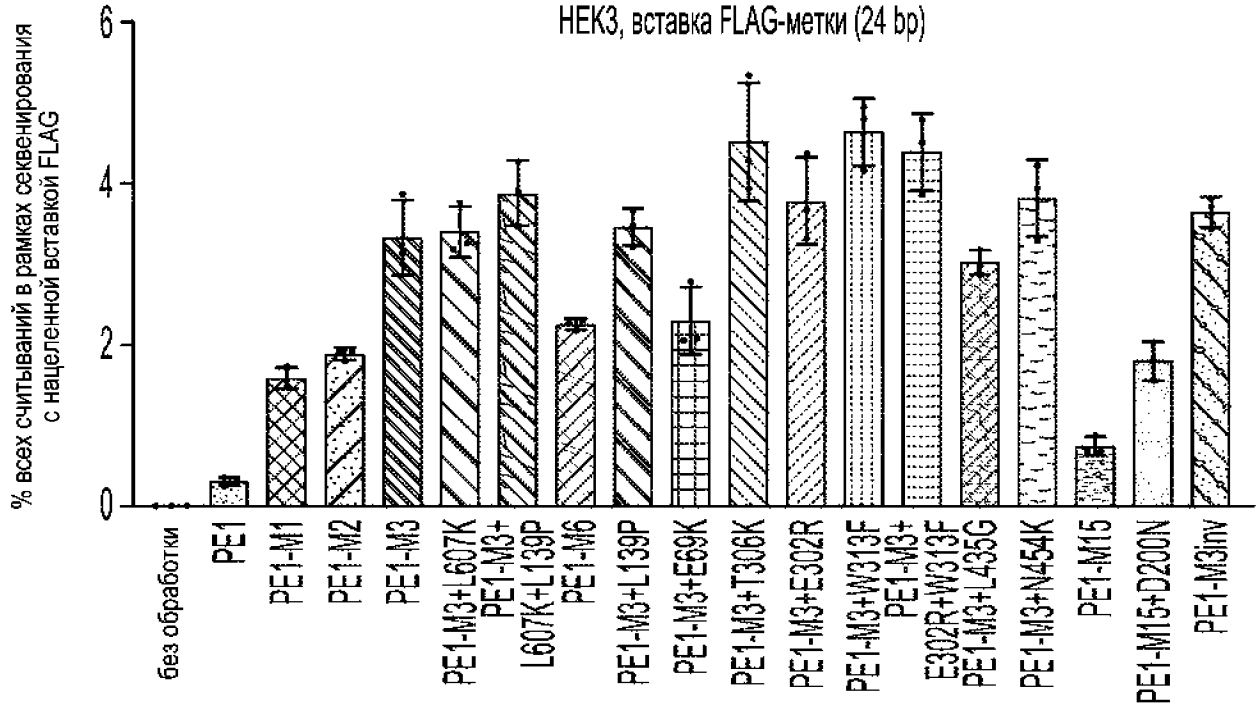
ФИГ.47A



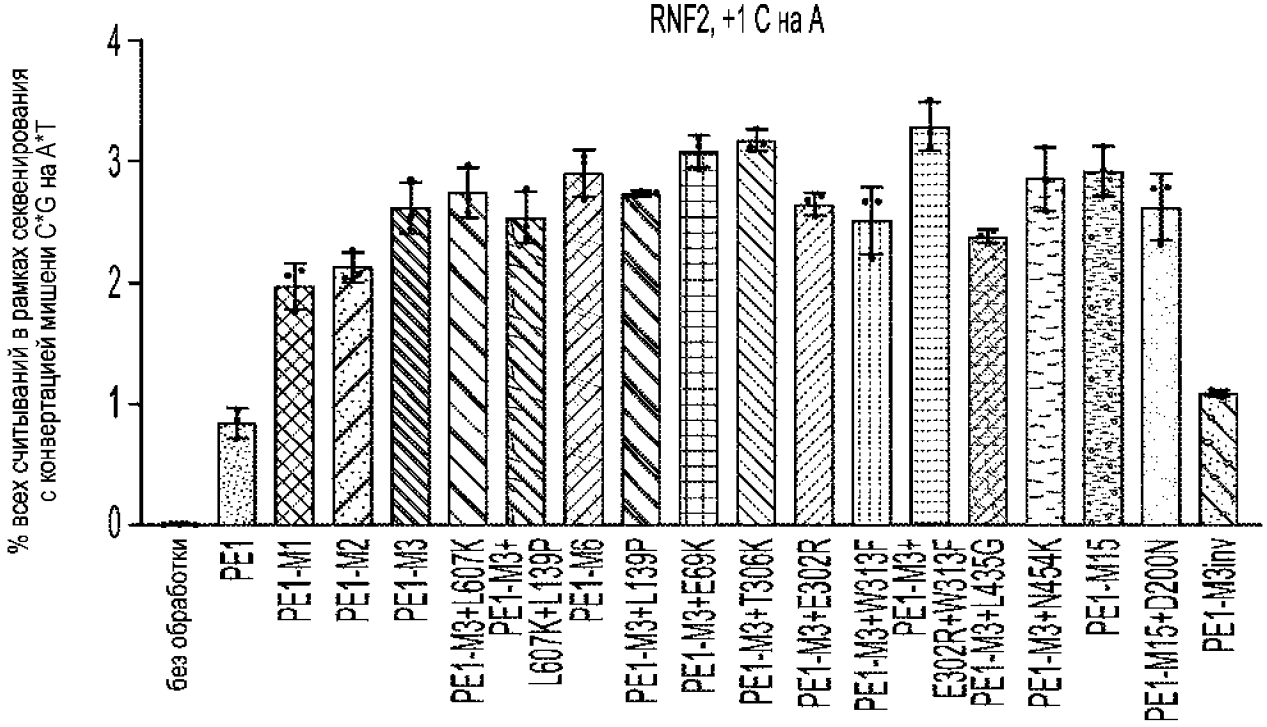
ФИГ.47B



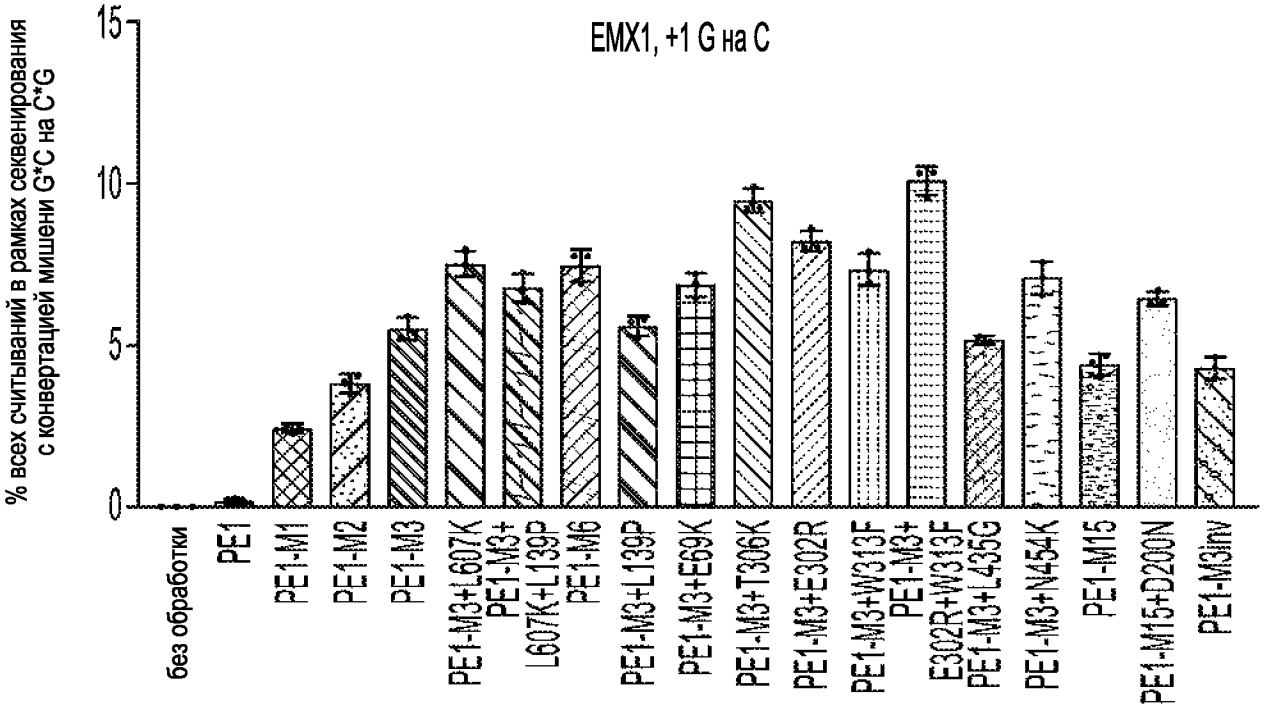
ФИГ.47C



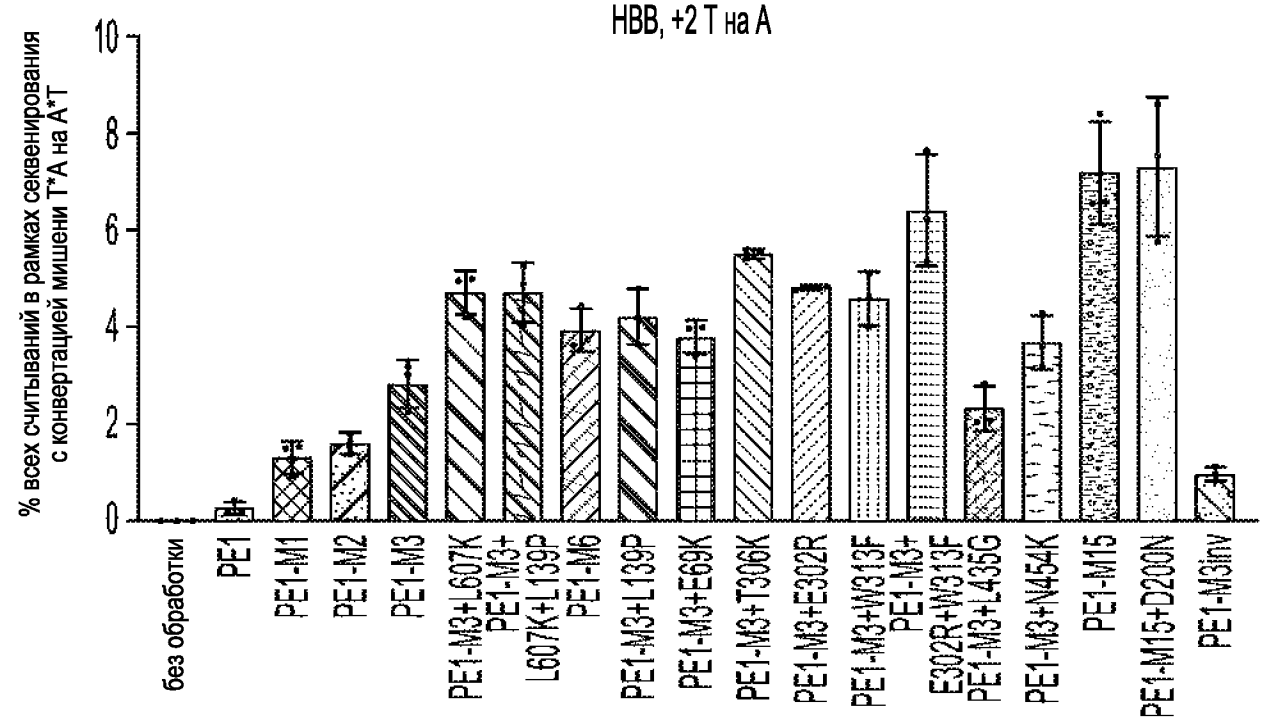
ФИГ.47D



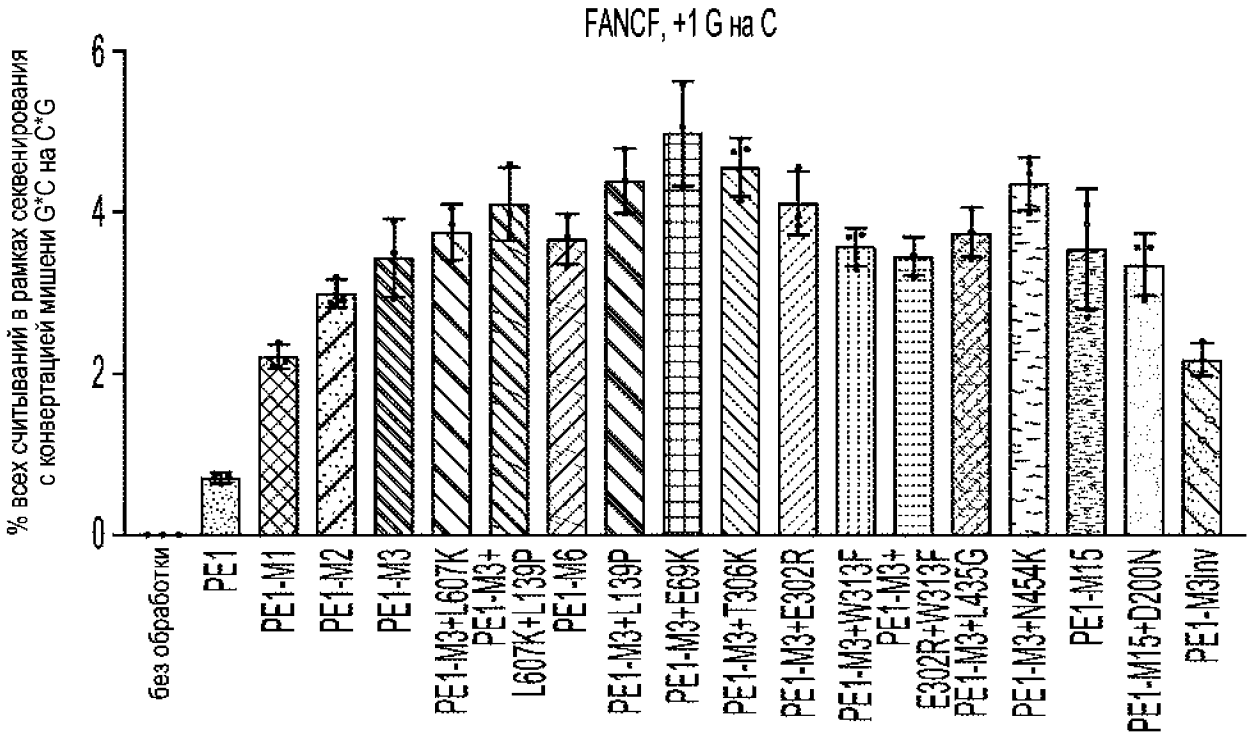
ФИГ.47E



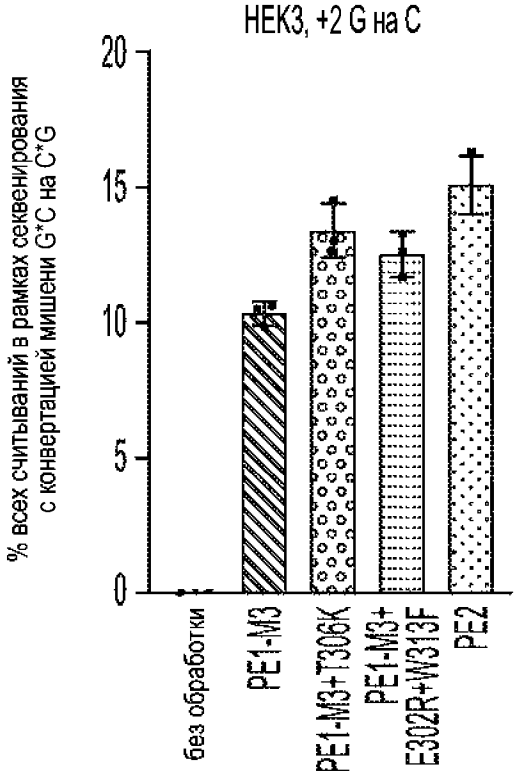
ФИГ.47F



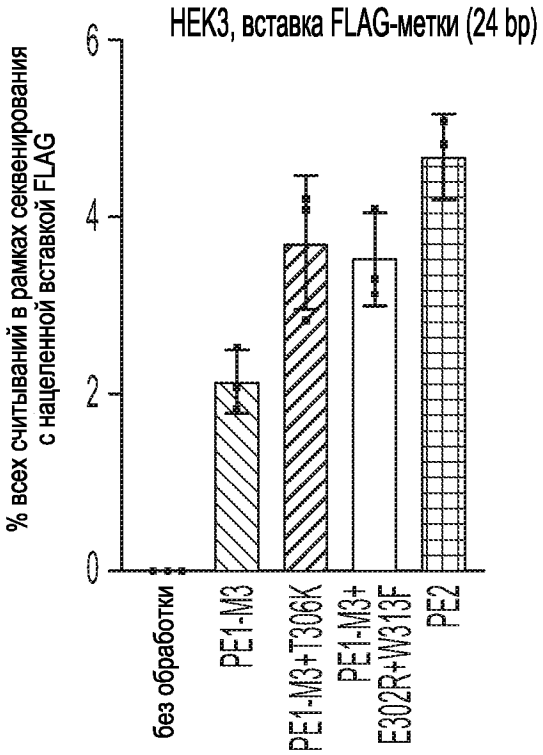
ФИГ.47G



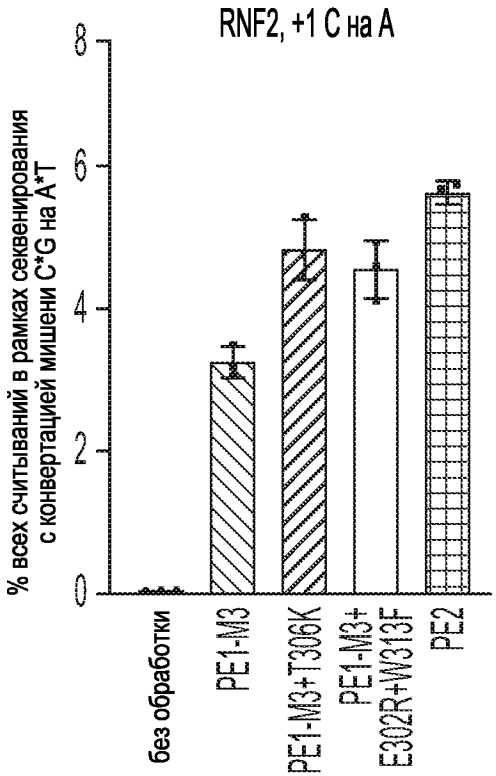
ФИГ.47H



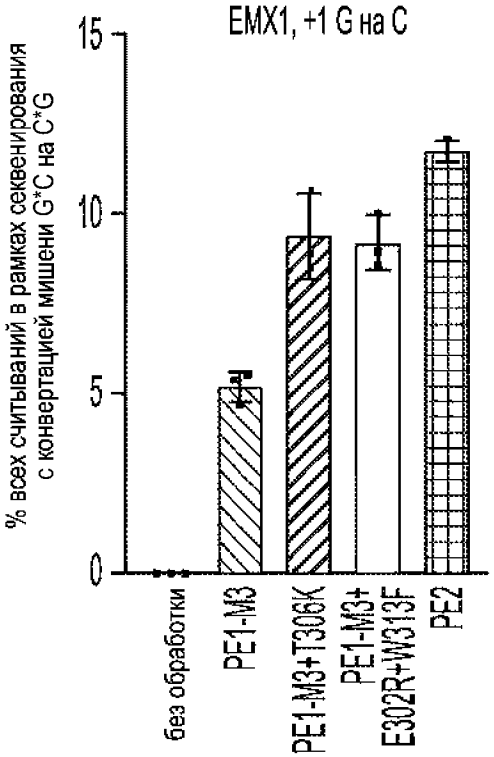
ФИГ.47I



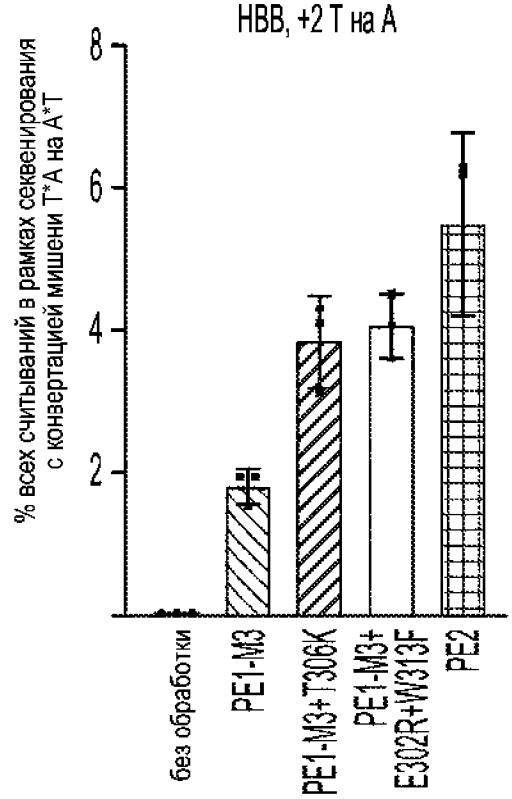
ФИГ.47J



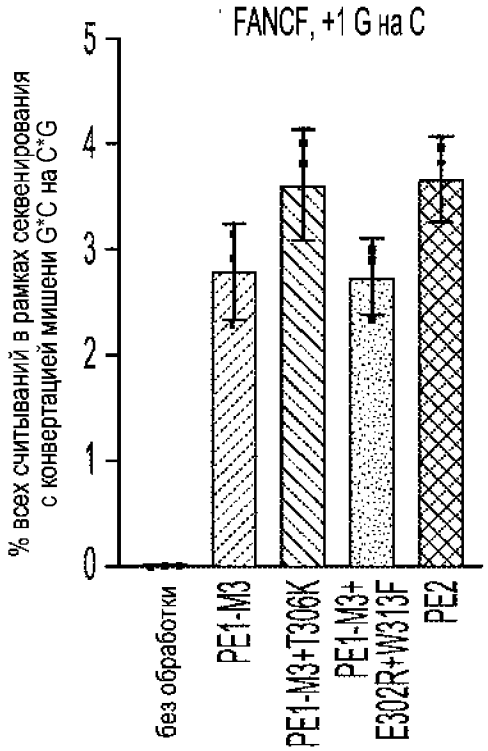
ФИГ.47K



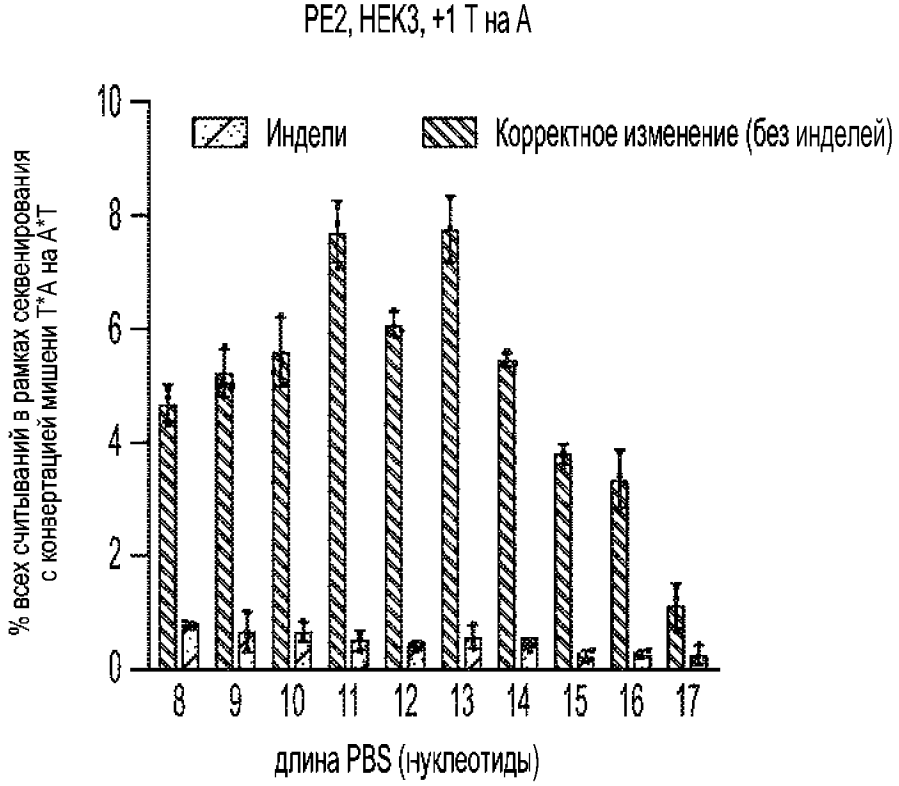
ФИГ.47L



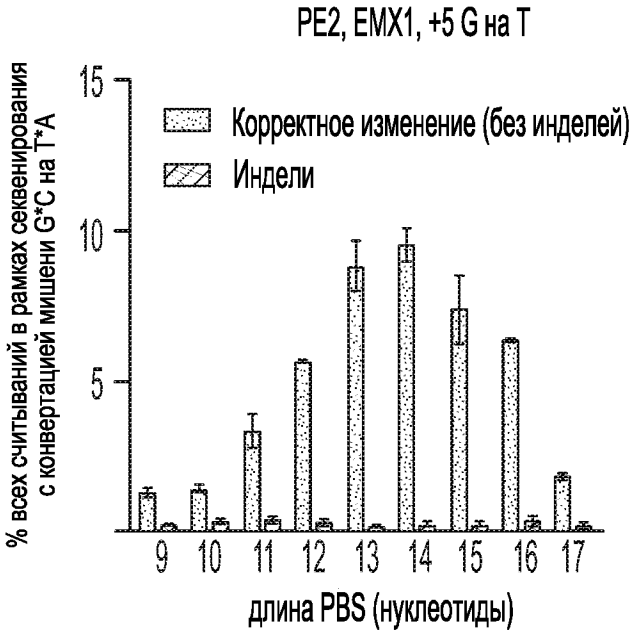
ФИГ.47M



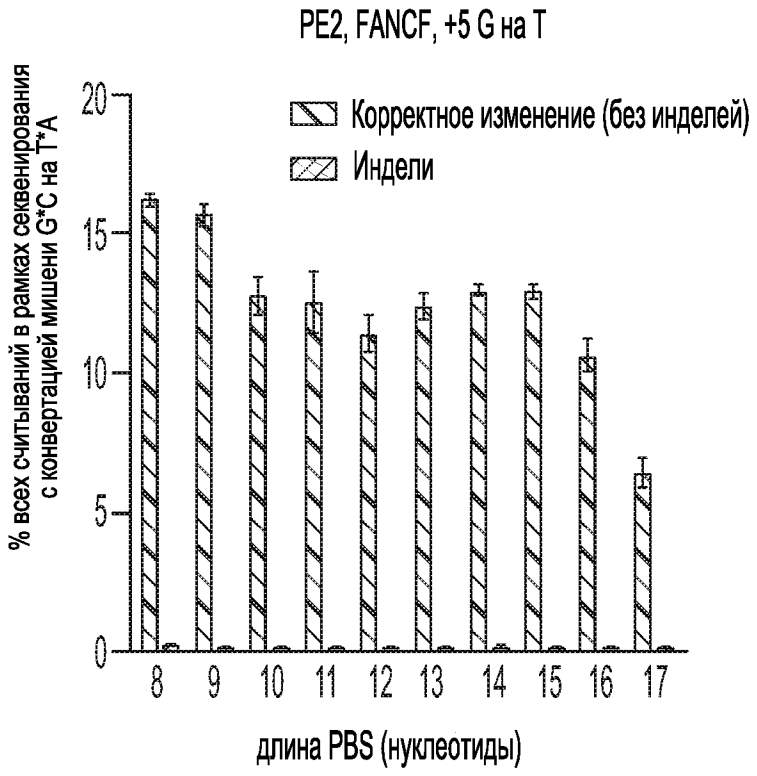
ФИГ.47N



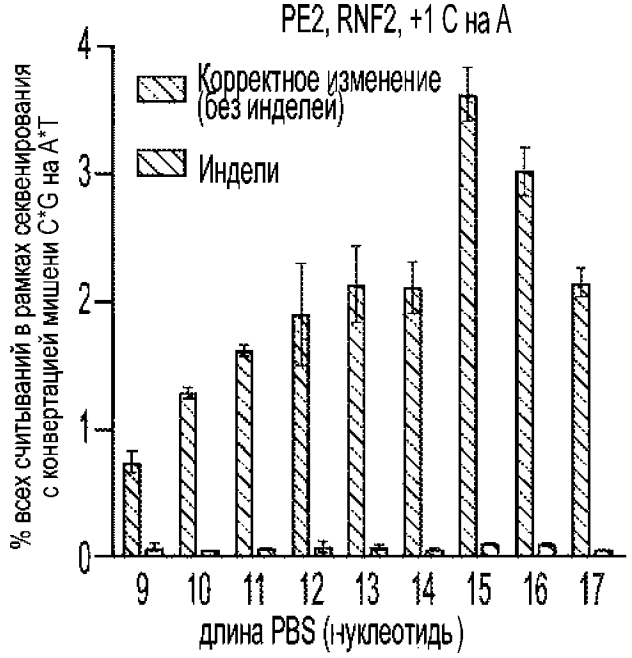
ФИГ.47O



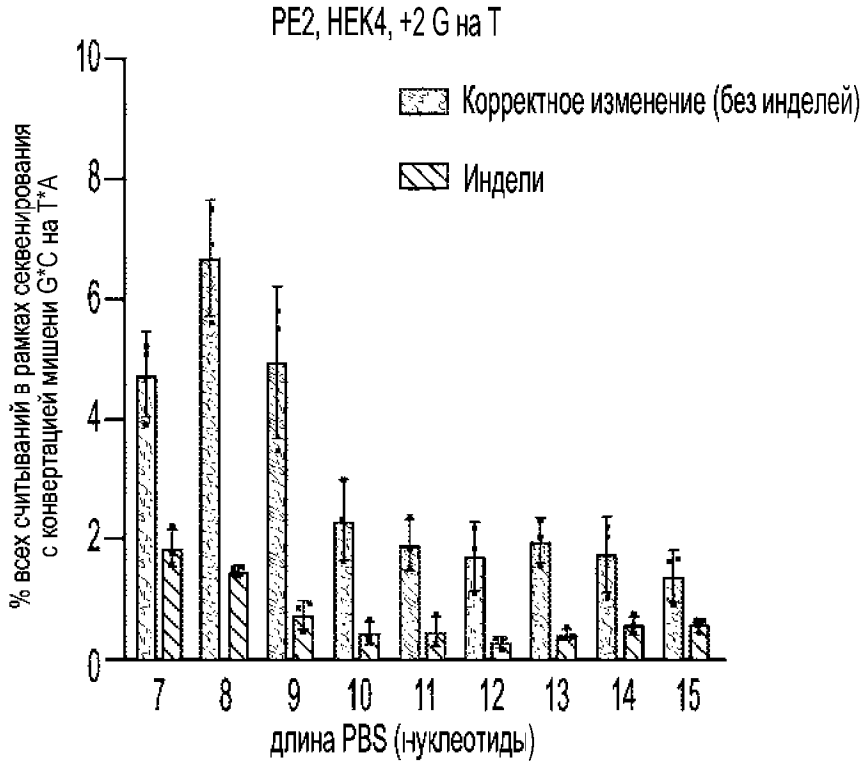
ФИГ.47P



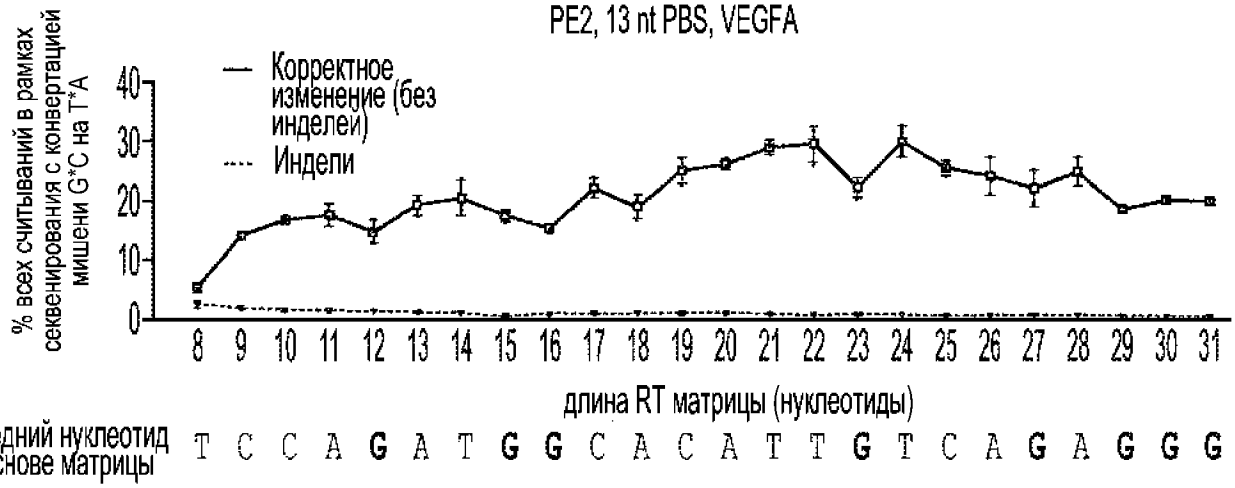
ФИГ.47Q



ФИГ.47R

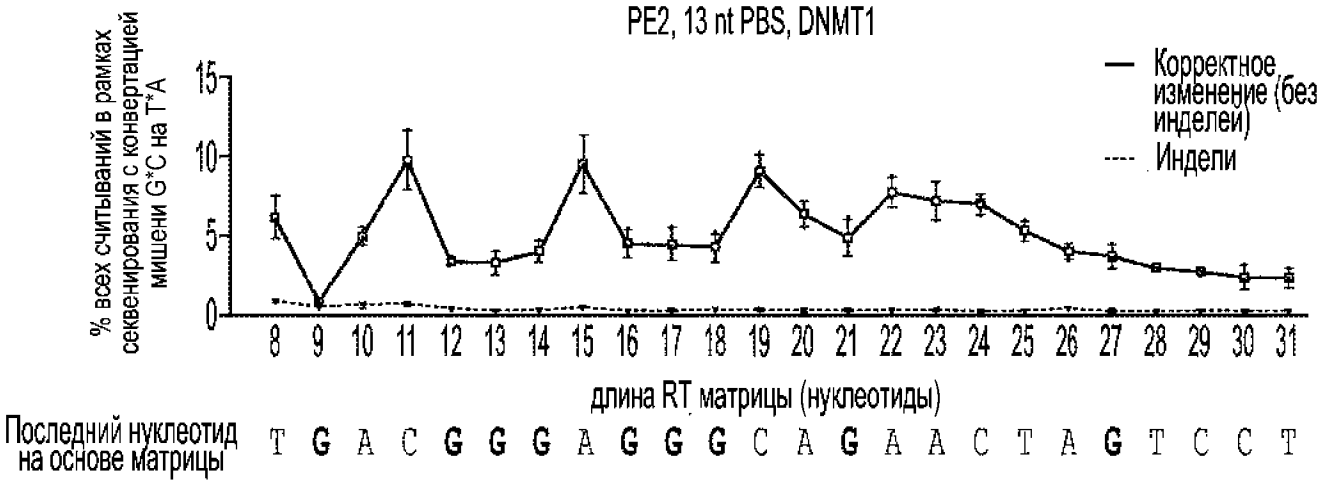


ФИГ.47S



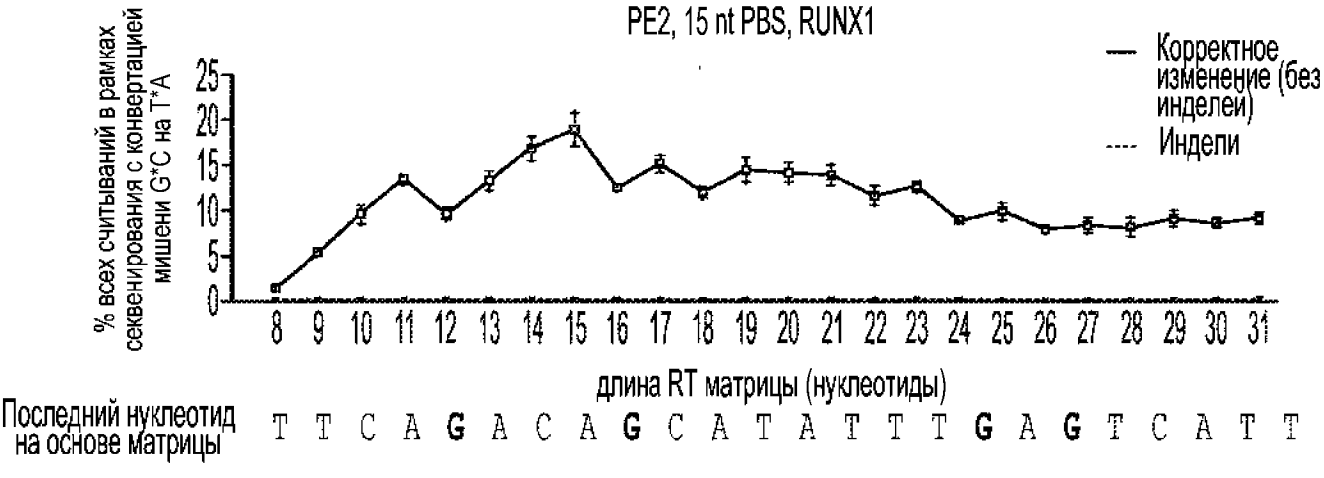
SEQ ID NO: 826

ФИГ.48A



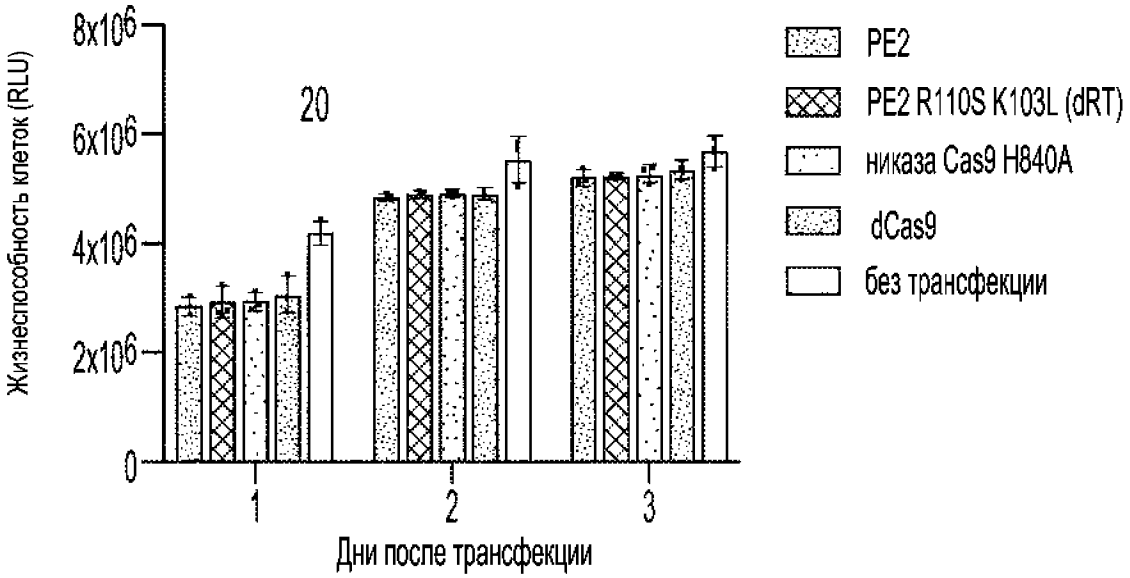
SEQ ID NO: 827

ФИГ.48B

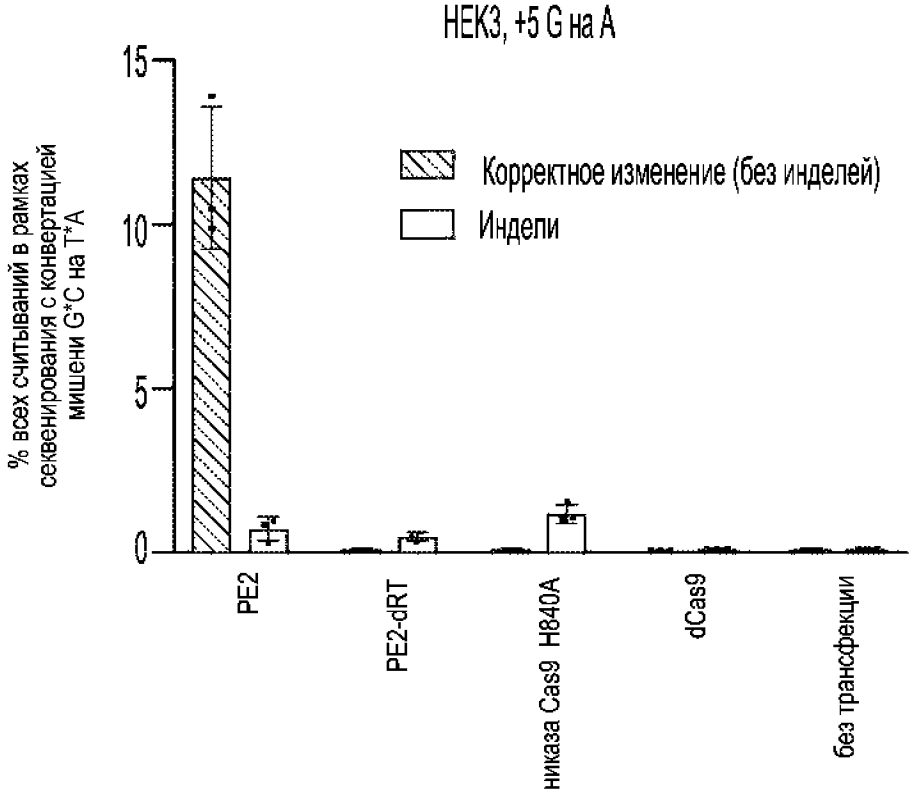


SEQ ID NO: 828

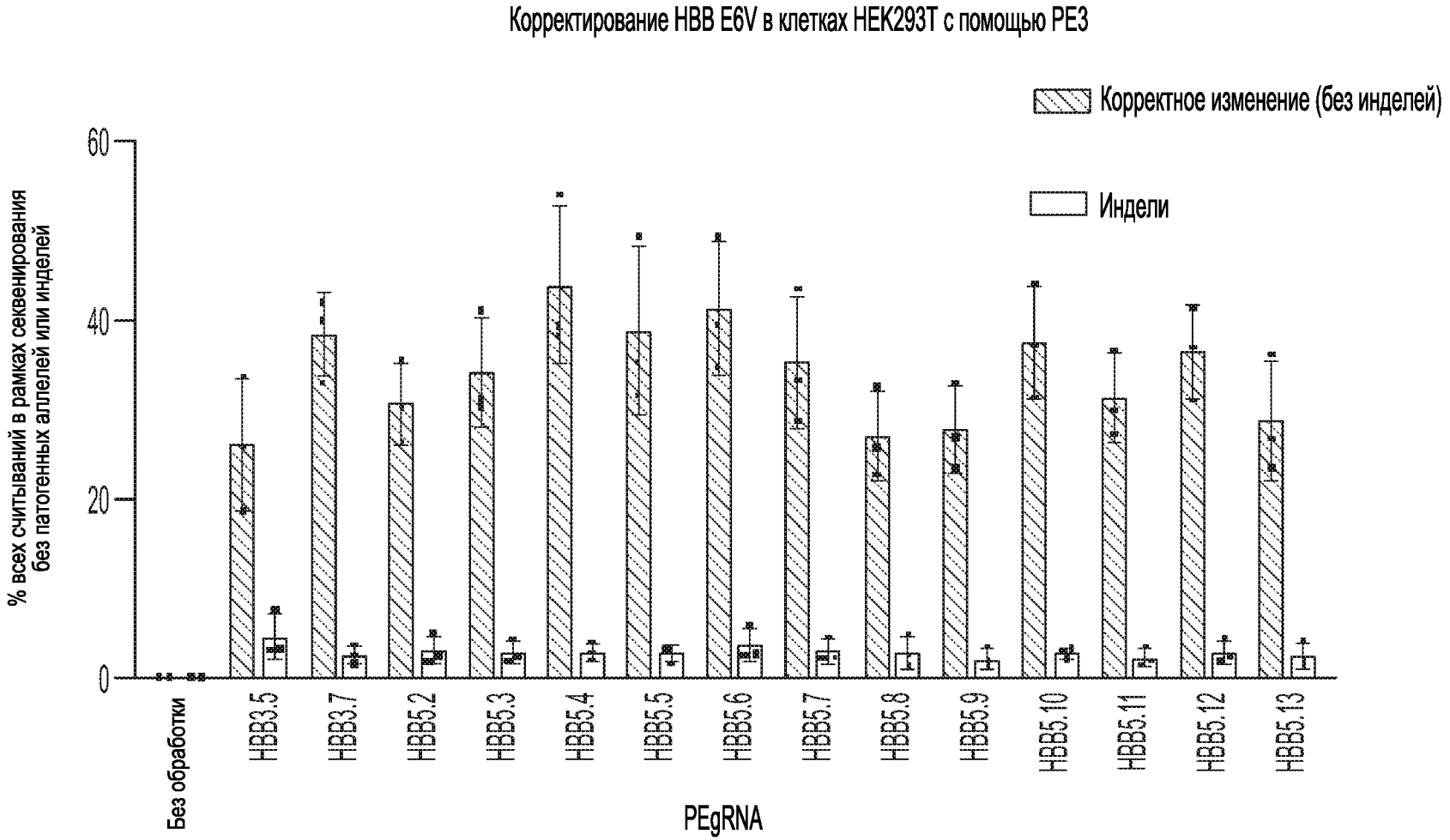
ФИГ.48C



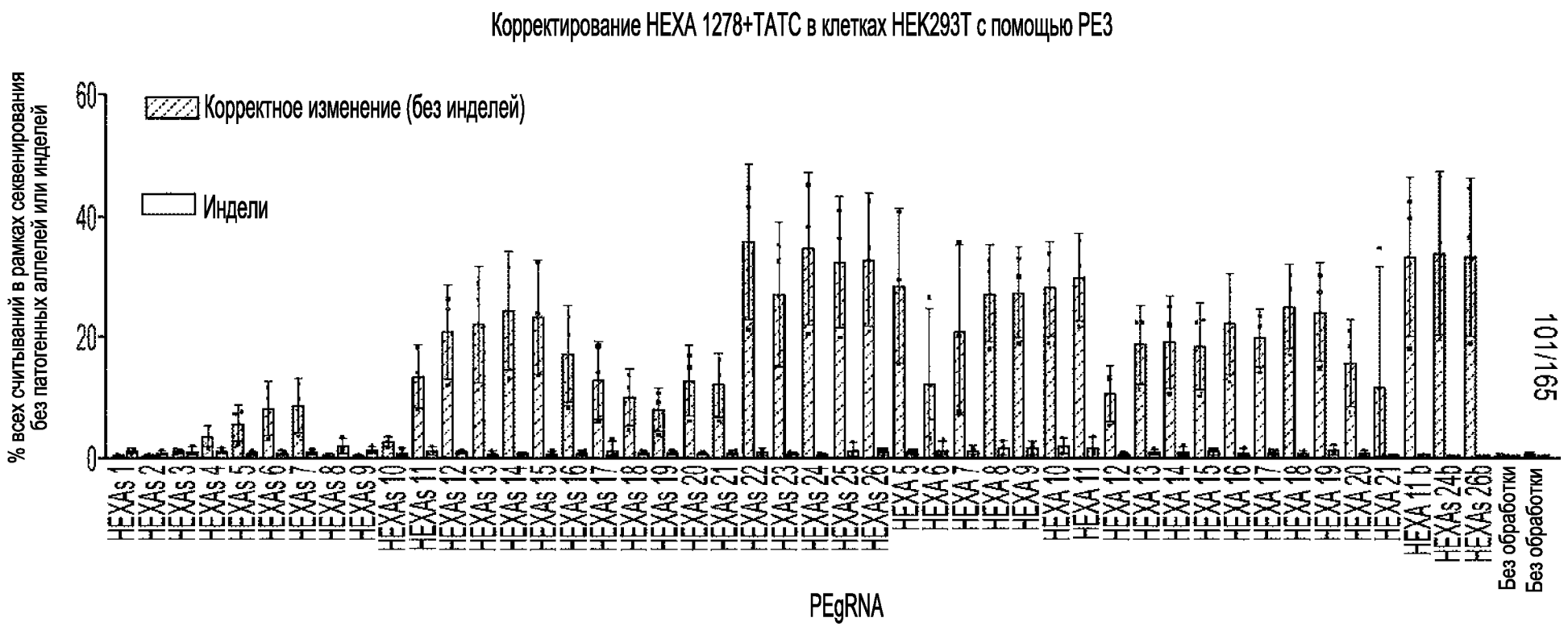
ФИГ.49A



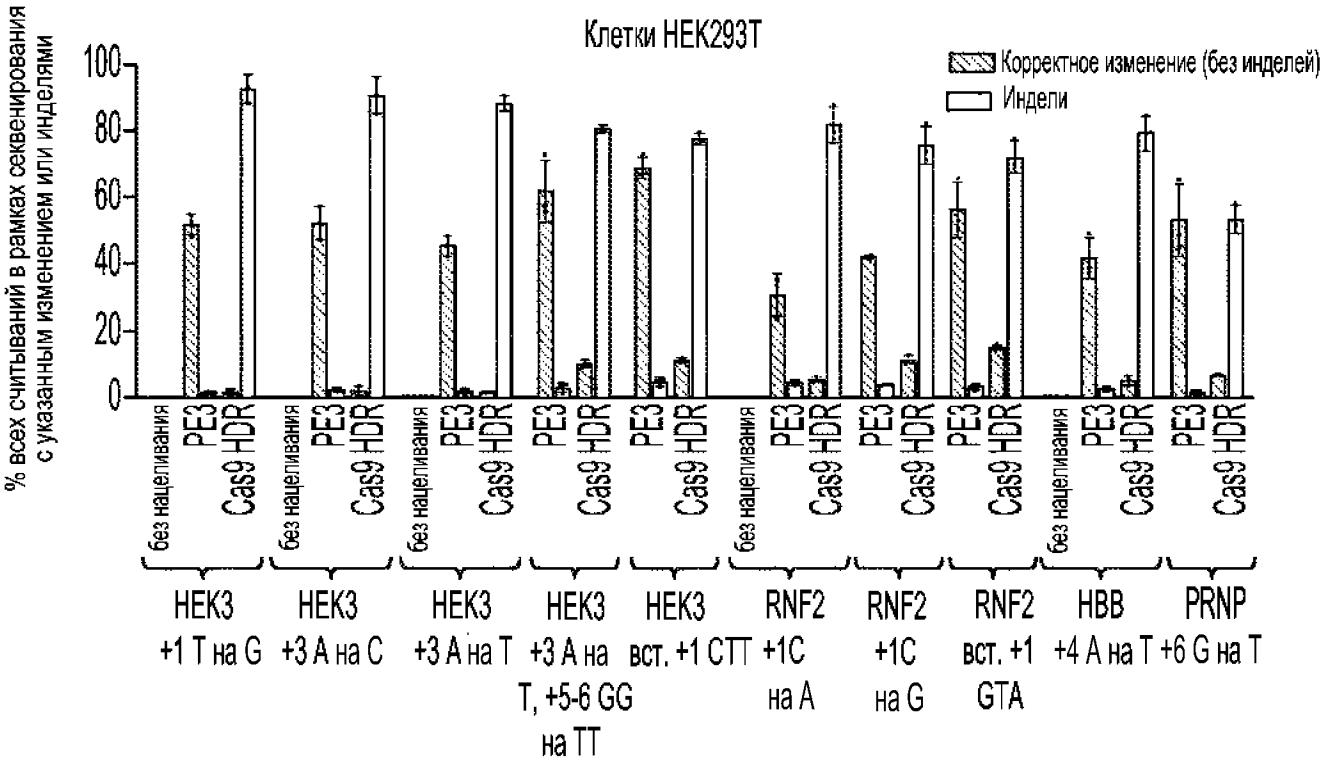
ФИГ.49B


Корректирование HBB E6V в клетках HEK293T с помощью PE3
Корректное изменение (без инделей)
Индели
% всех считываний в рамках секвенирования без патогенных аллелей или инделей
60
40
20
0
Без обработки
HBB3.5
HBB3.7
HBB5.2
HBB5.3
HBB5.4
HBB5.5
HBB5.6
HBB5.7
HBB5.8
HBB5.9
HBB5.10
HBB5.11
HBB5.12
HBB5.13
PEgRNA

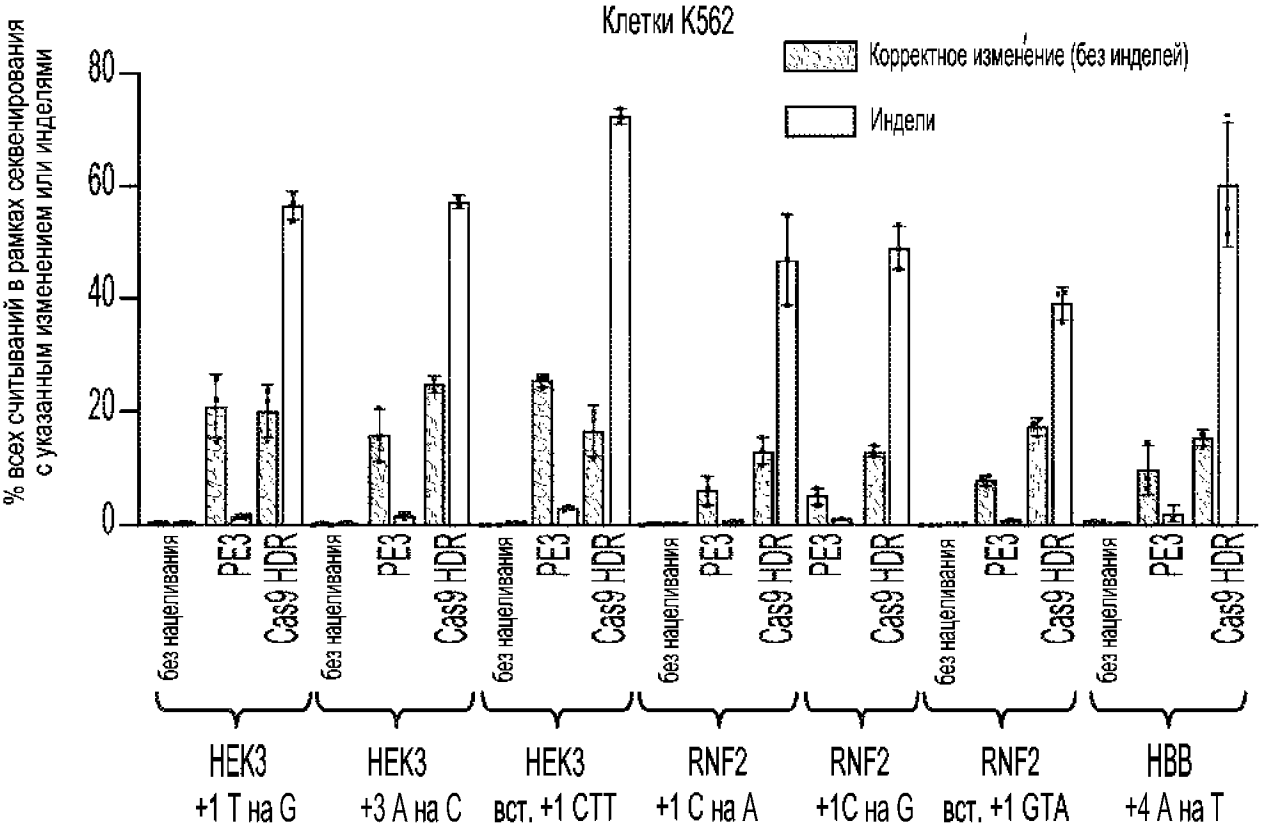
ФИГ.50А


Корректирование HEXA 1278+TATC в клетках HEK293T с помощью PE3
% всех считываний в рамках секвенирования без патогенных аллелей или инделей
Корректное изменение (без инделей)
Индели
60
40
20
0
HEXAs 1
HEXAs 2
HEXAs 3
HEXAs 4
HEXAs 5
HEXAs 6
HEXAs 7
HEXAs 8
HEXAs 9
HEXAs 10
HEXAs 11
HEXAs 12
HEXAs 13
HEXAs 14
HEXAs 15
HEXAs 16
HEXAs 17
HEXAs 18
HEXAs 19
HEXAs 20
HEXAs 21
HEXAs 22
HEXAs 23
HEXAs 24
HEXAs 25
HEXAs 26
HEXA 5
HEXA 6
HEXA 7
HEXA 8
HEXA 9
HEXA 10
HEXA 11
HEXA 12
HEXA 13
HEXA 14
HEXA 15
HEXA 16
HEXA 17
HEXA 18
HEXA 19
HEXA 20
HEXA 21
HEXA 11 b
HEXAs 24b
HEXAs 26b
Без обработки
Без обработки
PEgRNA

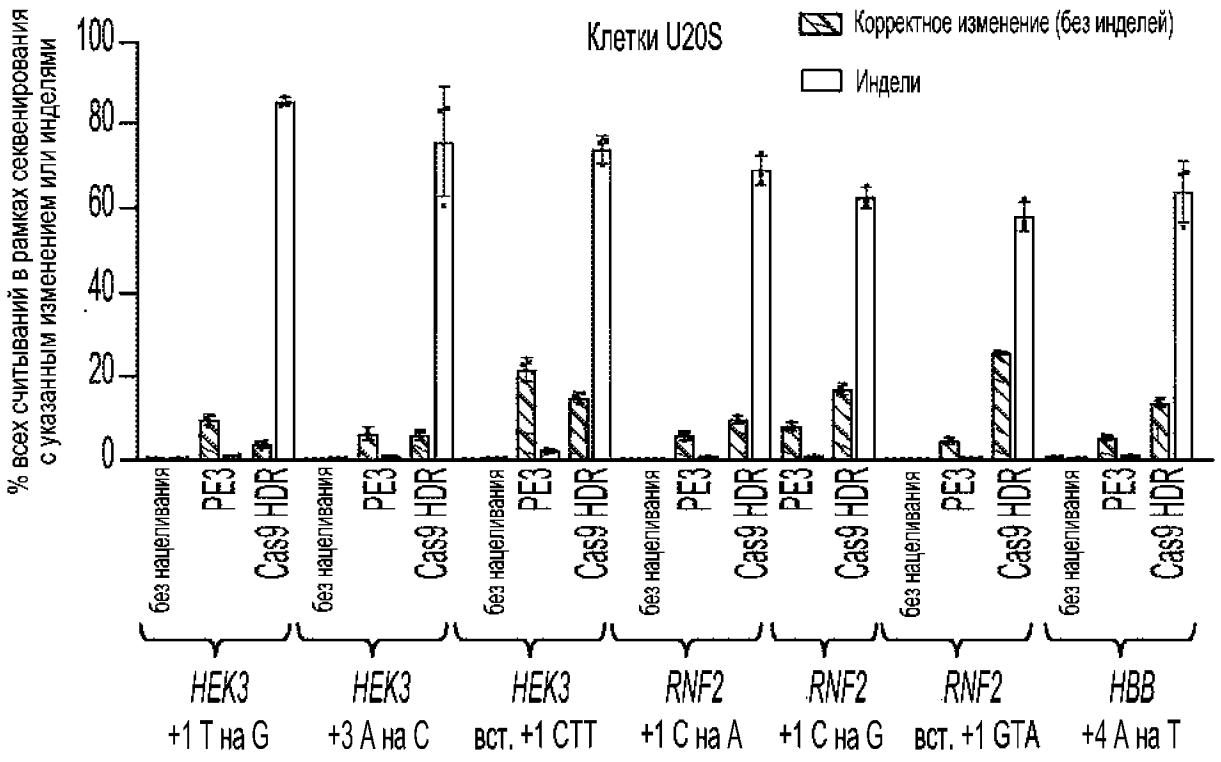
ФИГ.50В



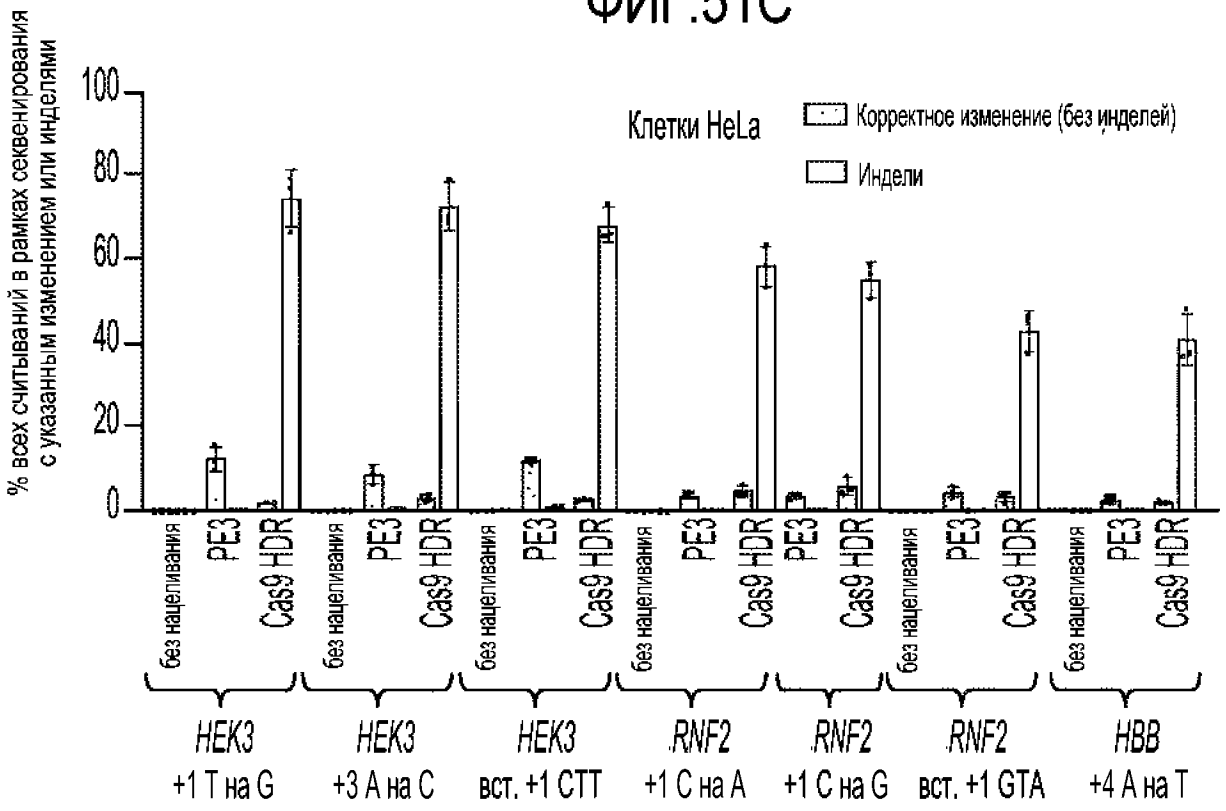
ФИГ.51A



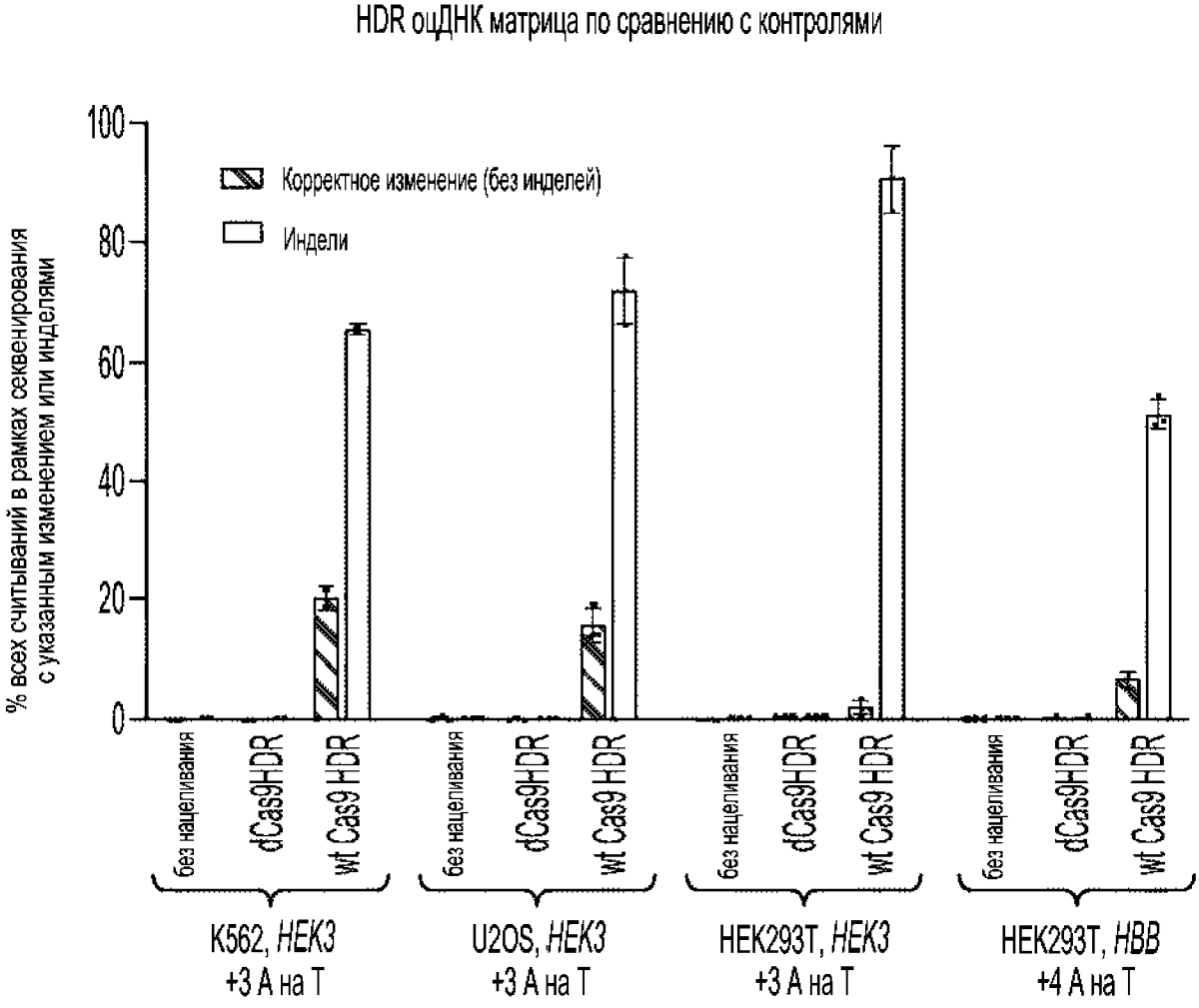
ФИГ.51B



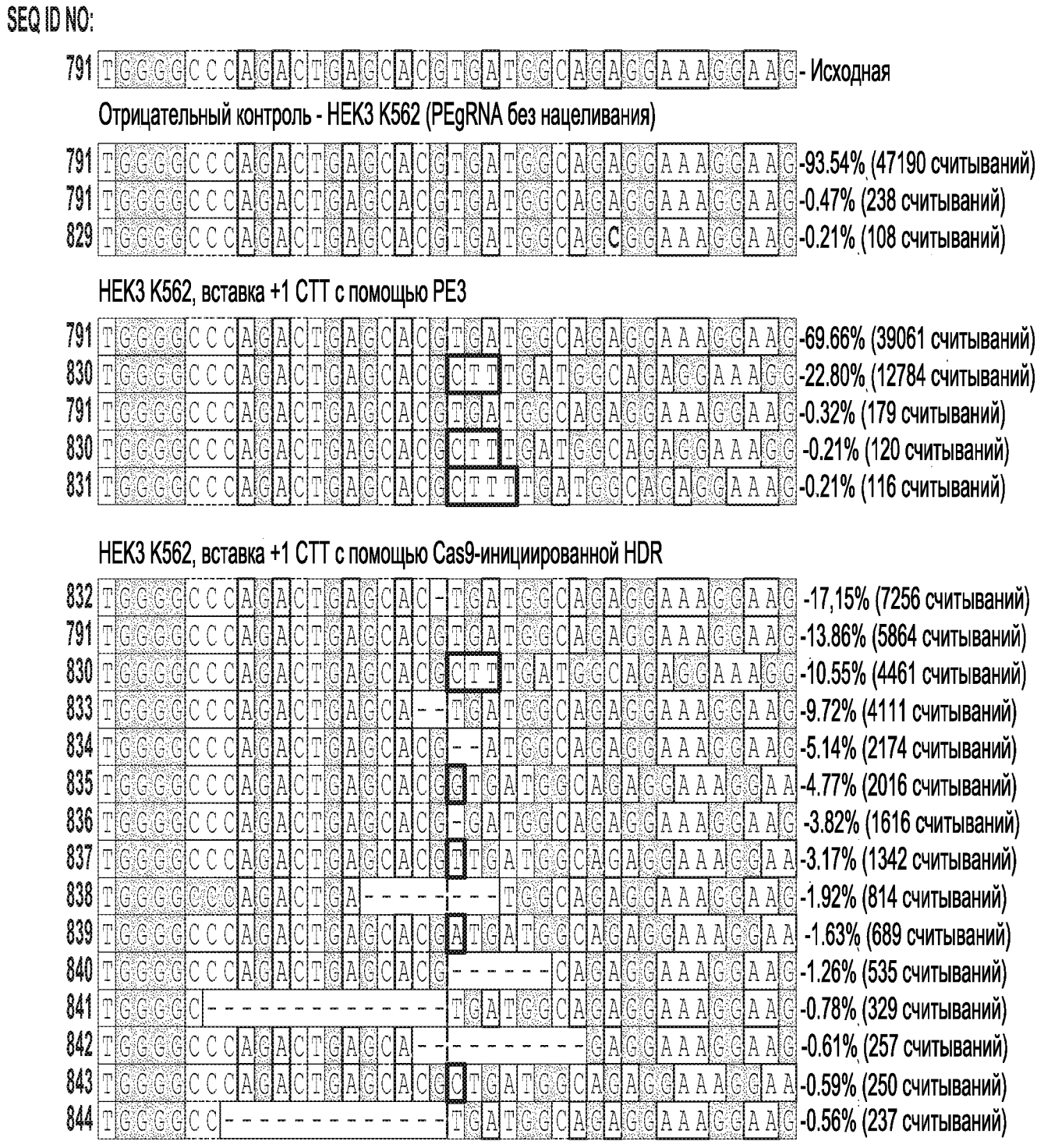
ФИГ.51C



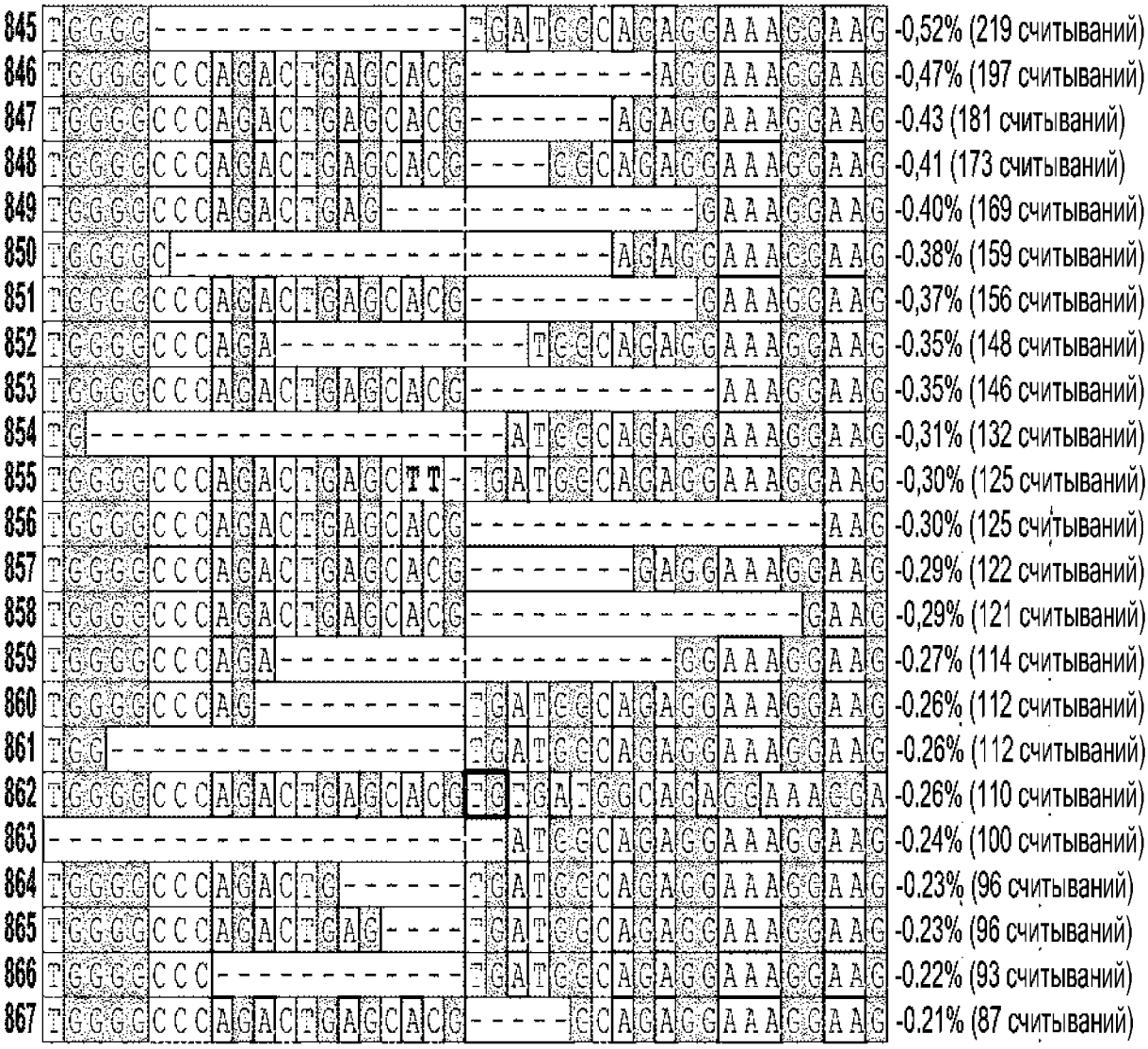
ФИГ.51D


HDR оцДНК матрица по сравнению с контролями
% всех считываний в рамках секвенирования с указанным изменением или инделями
100
80
60
40
20
0
Корректное изменение (без инделей)
Индели
без нацеливания
dCas9HDR
wt Cas9 HDR
K562, HEK3 +3 A на T
U2OS, HEK3 +3 A на T
HEK293T, HEK3 +3 A на T
HEK293T, HBB +4 A на T

ФИГ.51E


SEQ ID NO:
791 - Исходная
Отрицательный контроль - HEK3 K562 (PEgRNA без нацеливания)
791 -93.54% (47190 считываний)
791 -0.47% (238 считываний)
829 -0.21% (108 считываний)
HEK3 K562, вставка +1 CTT с помощью PE3
791 -69.66% (39061 считываний)
830 -22.80% (12784 считываний)
791 -0.32% (179 считываний)
830 -0.21% (120 считываний)
831 -0.21% (116 считываний)
HEK3 K562, вставка +1 CTT с помощью Cas9-инициированной HDR
832 -17,15% (7256 считываний)
791 -13.86% (5864 считываний)
830 -10.55% (4461 считываний)
833 -9.72% (4111 считываний)
834 -5.14% (2174 считываний)
835 -4.77% (2016 считываний)
836 -3.82% (1616 считываний)
837 -3.17% (1342 считываний)
838 -1.92% (814 считываний)
839 -1.63% (689 считываний)
840 -1.26% (535 считываний)
841 -0.78% (329 считываний)
842 -0.61% (257 считываний)
843 -0.59% (250 считываний)
844 -0.56% (237 считываний)

ФИГ.51F



жирный шрифт замещения

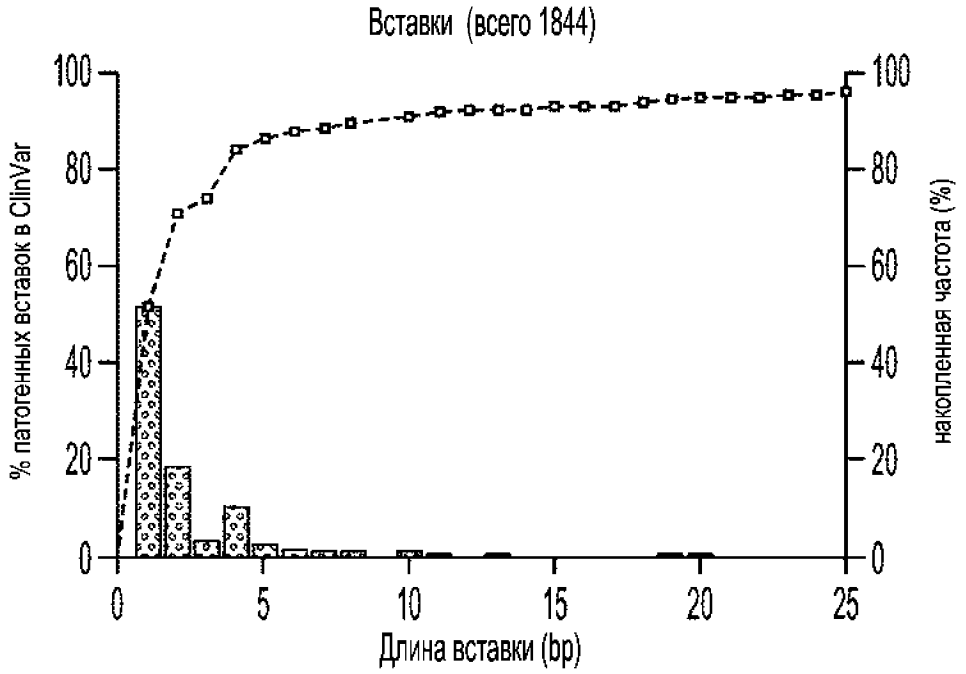
вставки

\- делеции

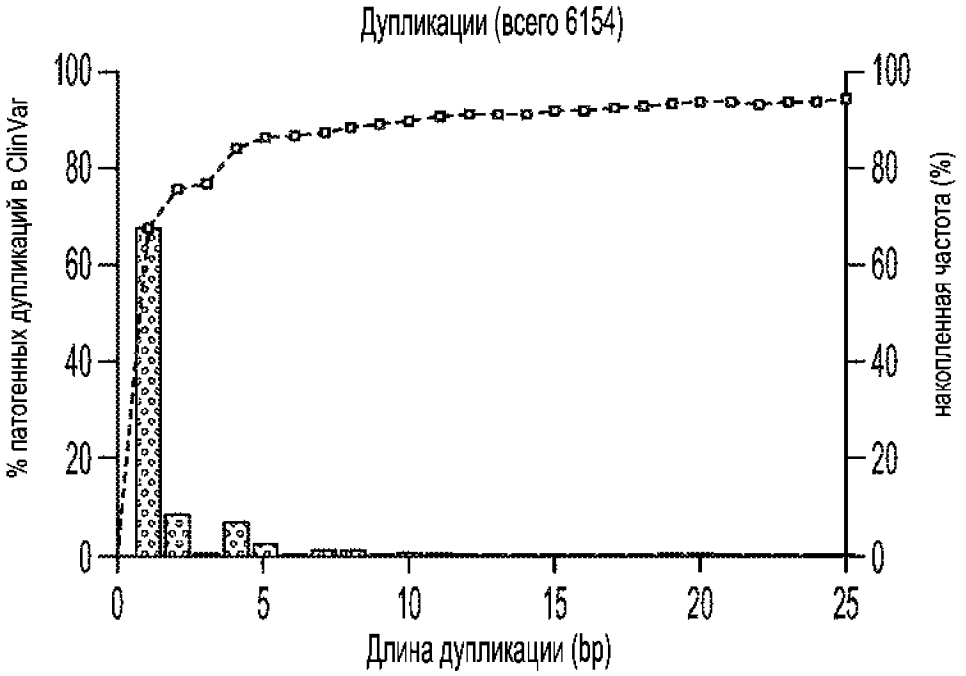
**** предполагаемое положение расщепления

ФИГ.51F

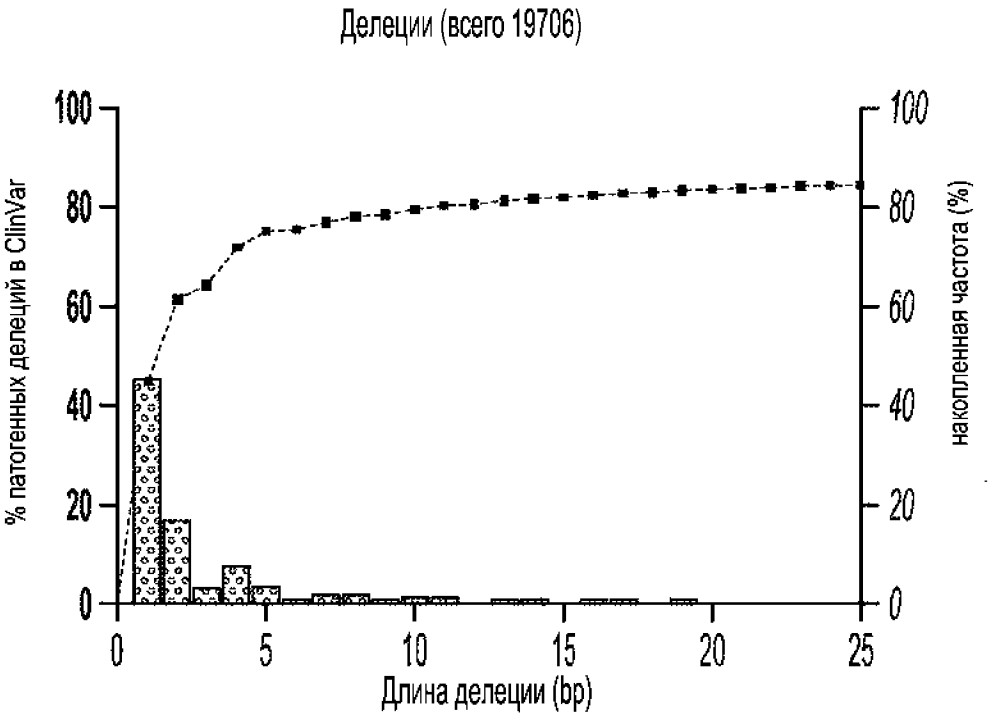
ПРОДОЛЖЕНИЕ



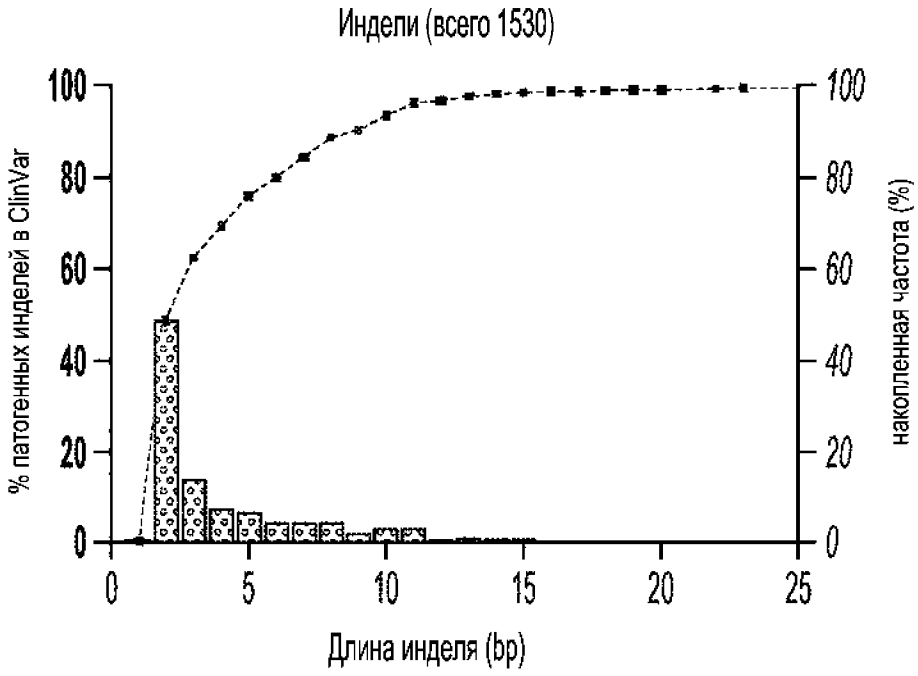
ФИГ.52A



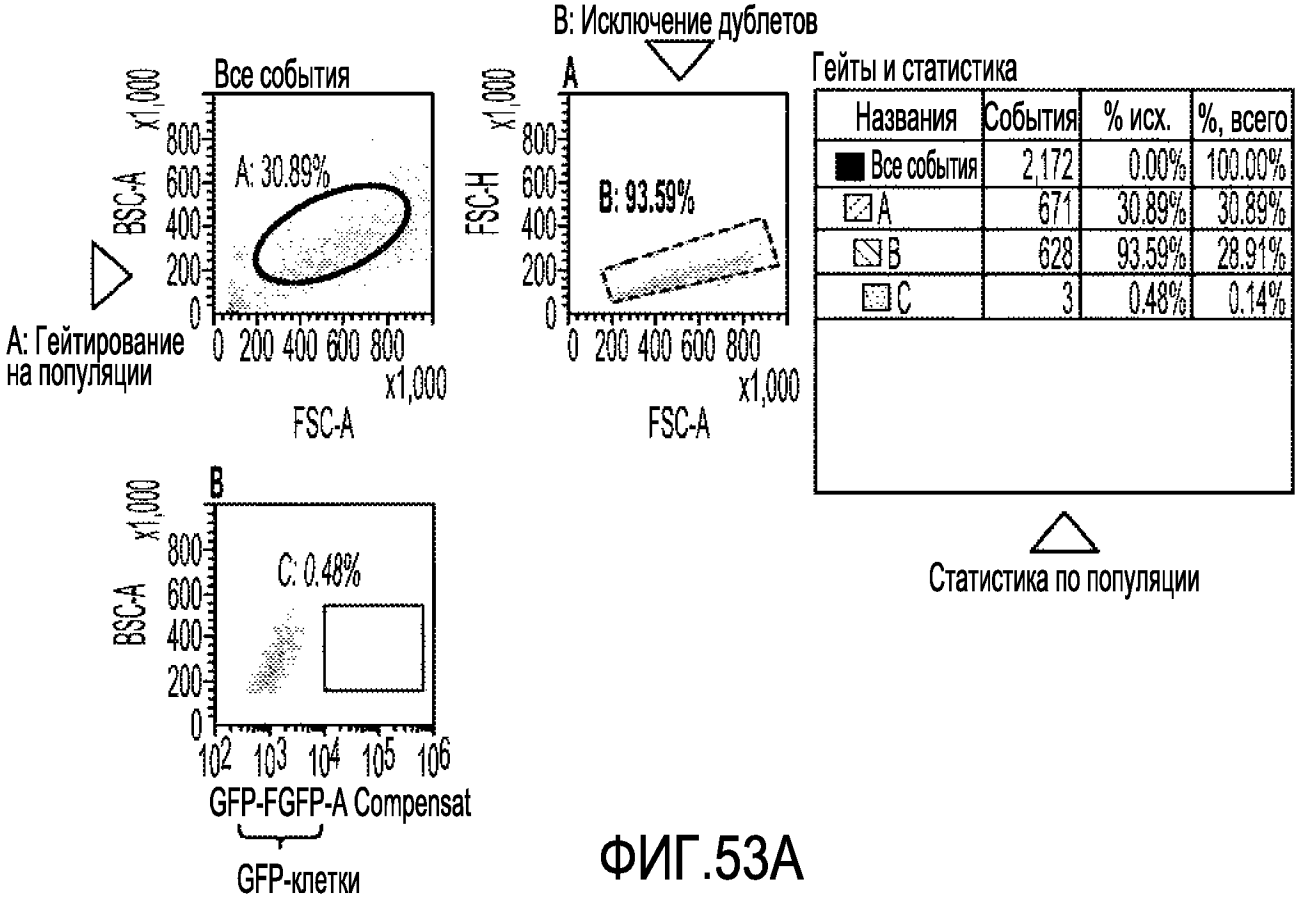
ФИГ.52B



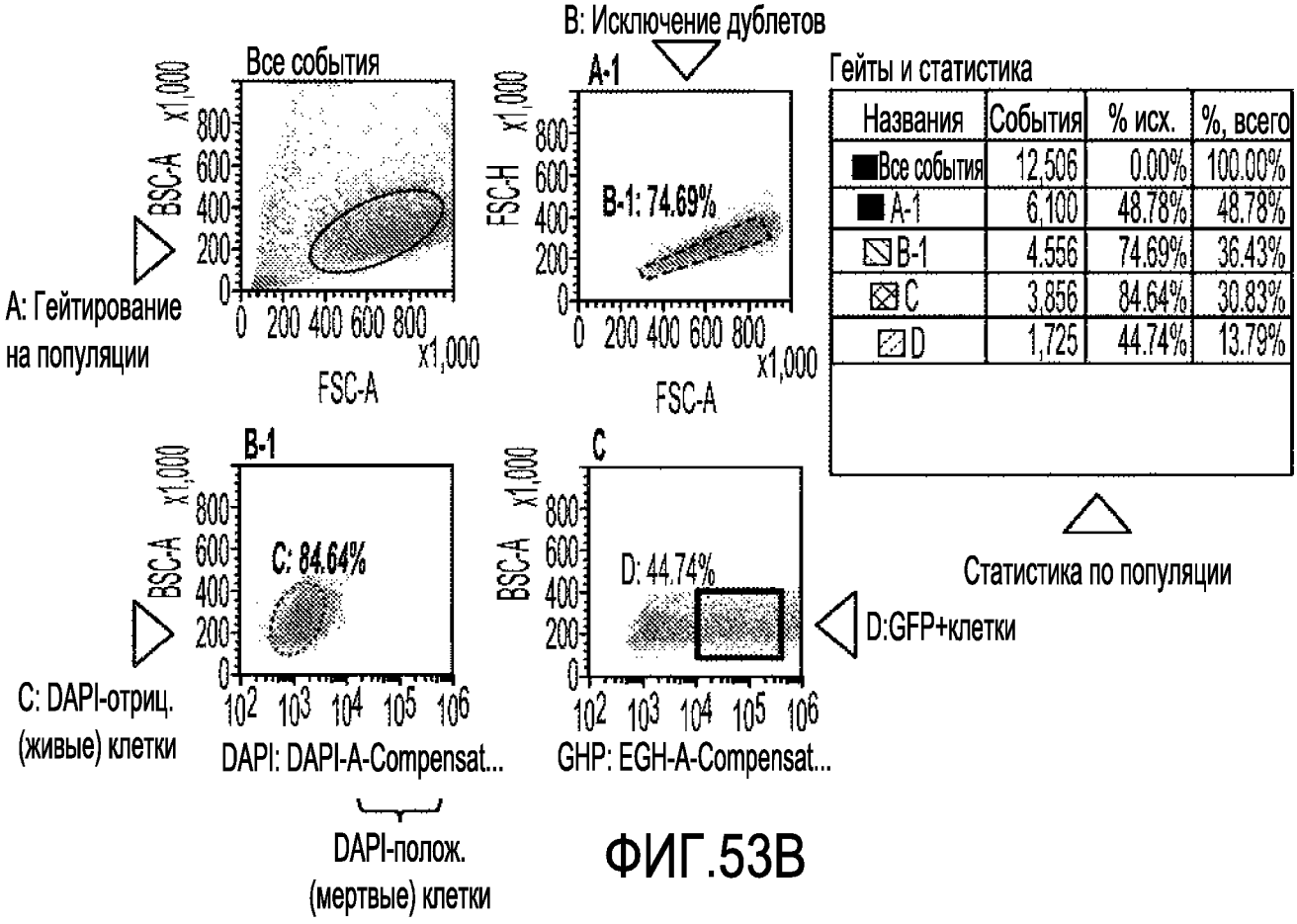
ФИГ.52C



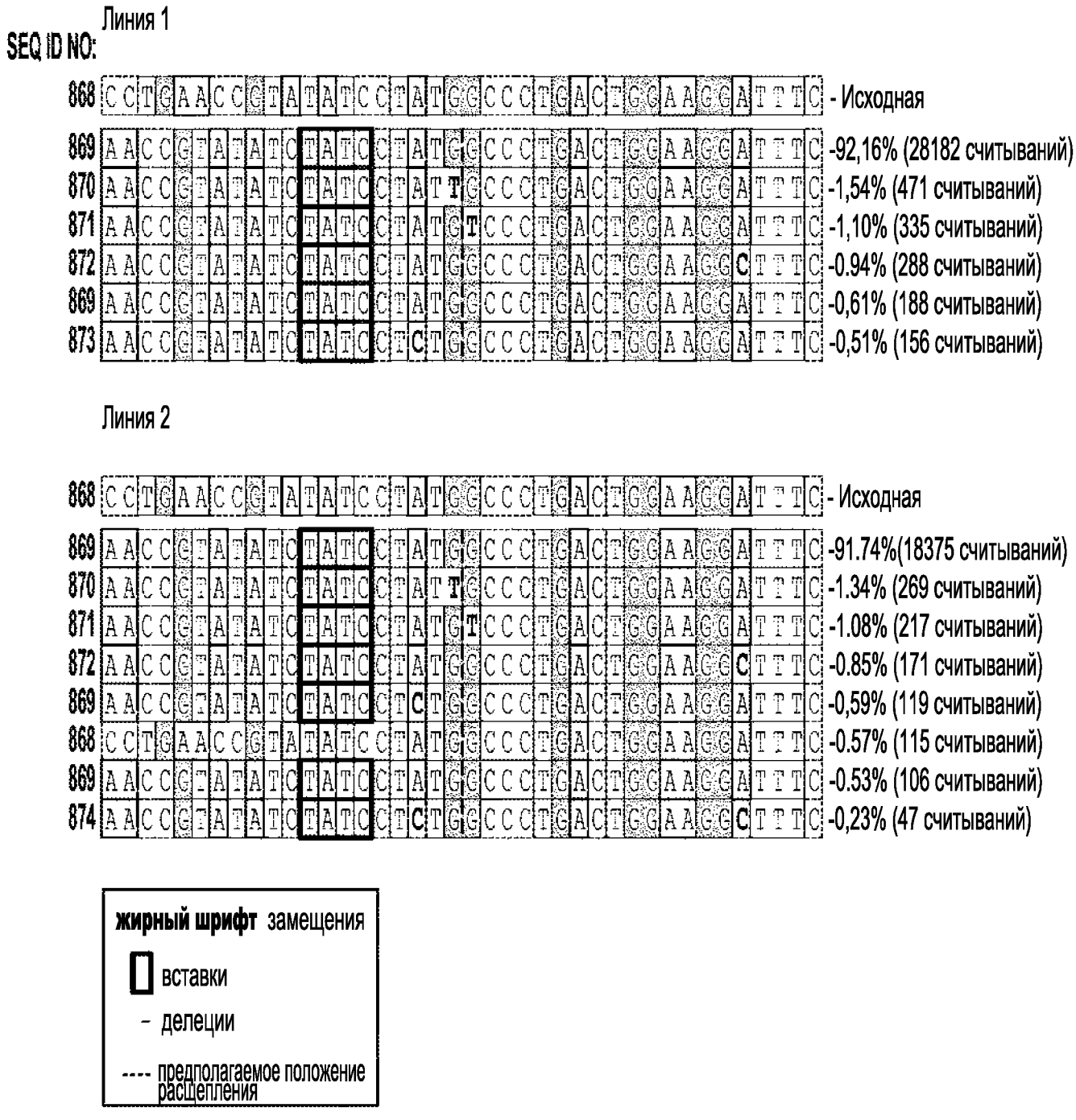
ФИГ.52D



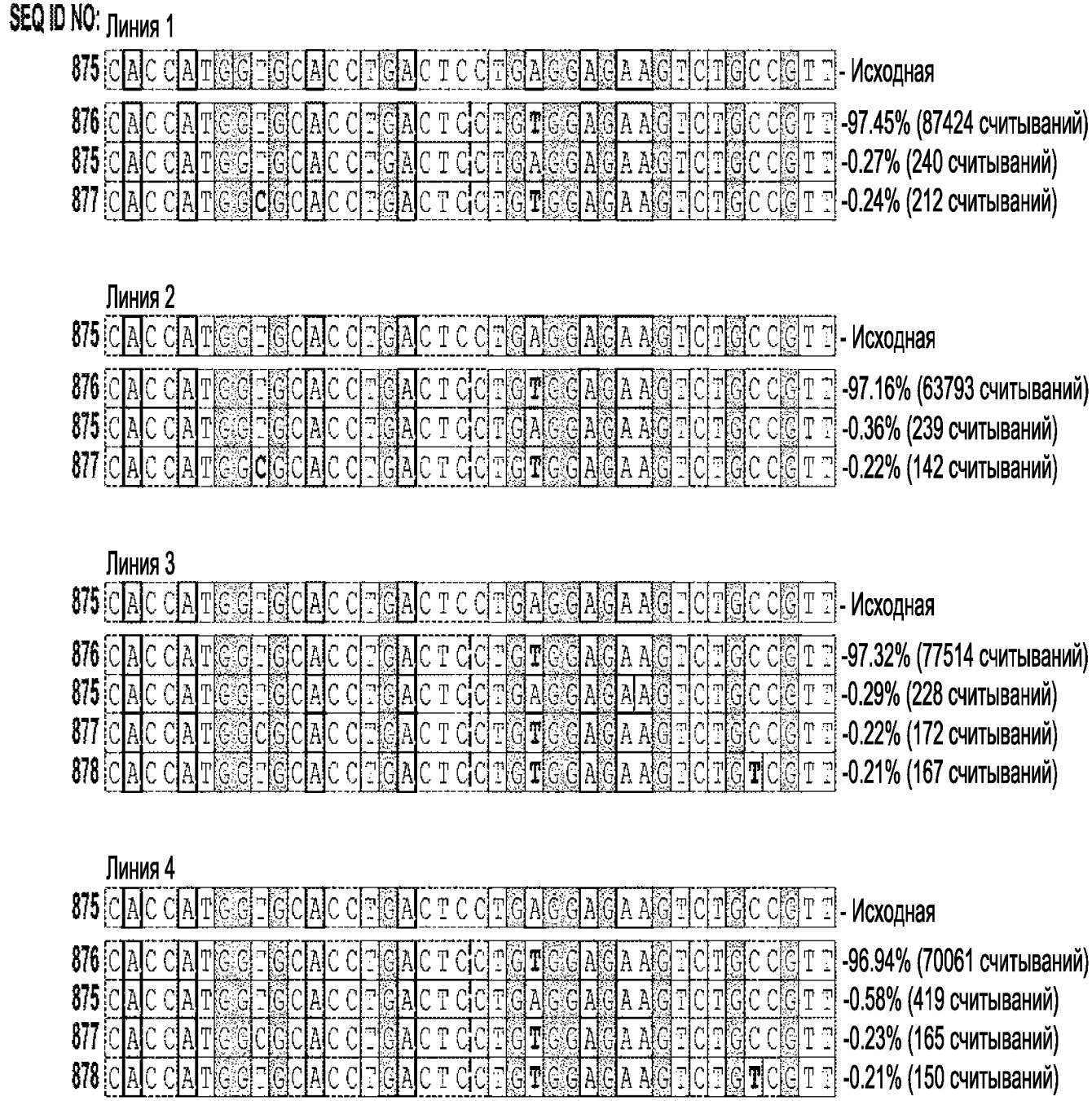
ФИГ.53A



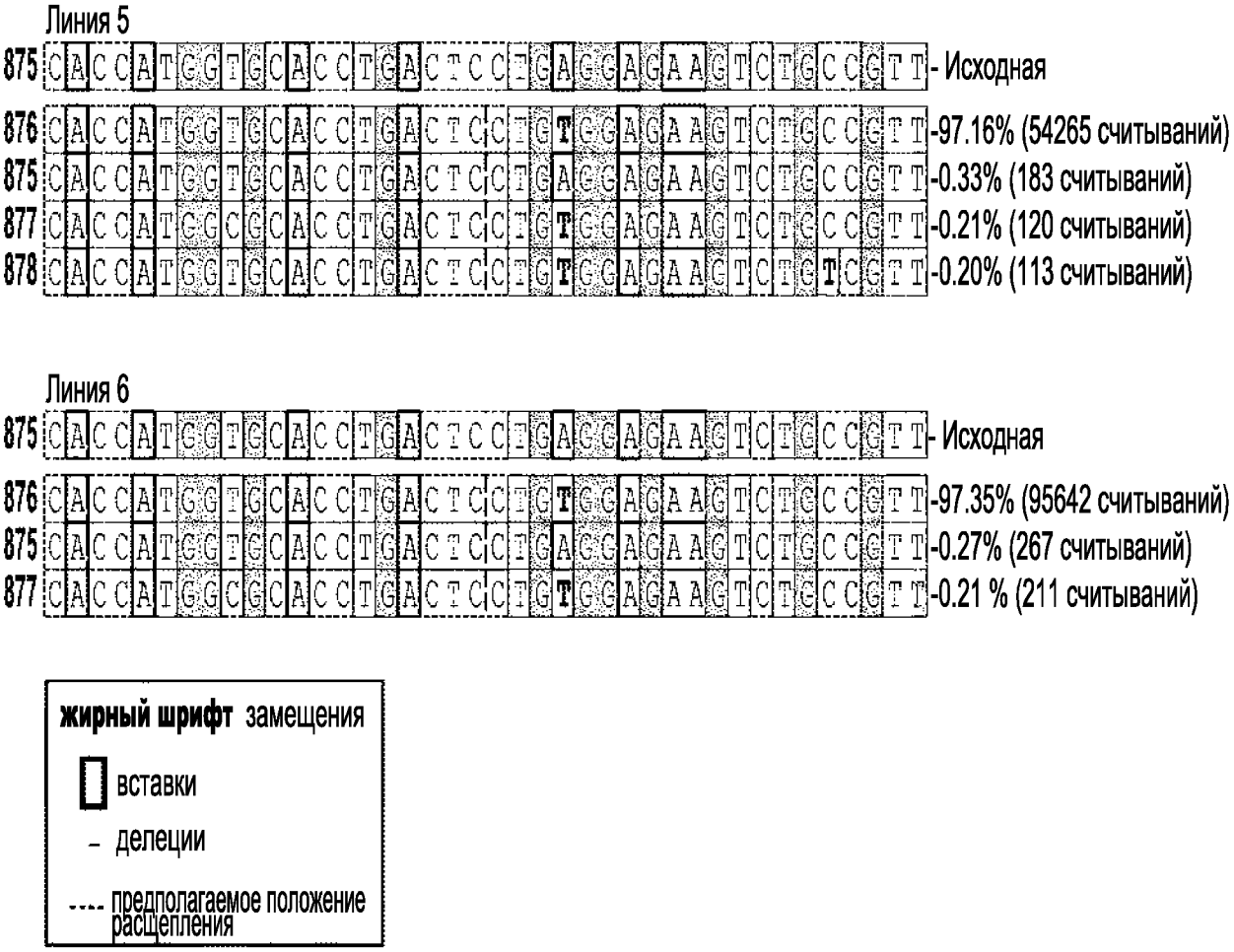
ФИГ.53B


Линия 1
SEQ ID NO:
868 CCTGAACCGTATATCCTATGGCCCTGACTGGAAGGATTTC - Исходная
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGATTTC -92,16% (28182 считываний)
870 AACCGTATATCTATCCTATTGCCCTGACTGGAAGGATTTC -1,54% (471 считываний)
871 AACCGTATATCTATCCTATGTCCCTGACTGGAAGGATTTC -1,10% (335 считываний)
872 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGCTTTC -0.94% (288 считываний)
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGATTTC -0,61% (188 считываний)
873 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGGATTTC -0,51% (156 считываний)
Линия 2
868 CCTGAACCGTATATCCTATGGCCCTGACTGGAAGGATTTC - Исходная
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGATTTC -91.74%(18375 считываний)
870 AACCGTATATCTATCCTATTGCCCTGACTGGAAGGATTTC -1.34% (269 считываний)
871 AACCGTATATCTATCCTATGTCCCTGACTGGAAGGATTTC -1.08% (217 считываний)
872 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGCTTTC -0.85% (171 считываний)
869 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGGATTTC -0,59% (119 считываний)
868 CCTGAACCGTATATCCTATGGCCCTGACTGGAAGGATTTC -0.57% (115 считываний)
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGGATTTC -0.53% (106 считываний)
874 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGGCTTTC -0,23% (47 считываний)
жирный шрифт замещения
вставки
- делеции
---- предполагаемое положение расщепления

ФИГ.53С


SEQ ID NO:
Линия 1
875 - Исходная
876 -97.45% (87424 считываний)
875 -0.27% (240 считываний)
877 -0.24% (212 считываний)
Линия 2
875 - Исходная
876 -97.16% (63793 считываний)
875 -0.36% (239 считываний)
877 -0.22% (142 считываний)
Линия 3
875 - Исходная
876 -97.32% (77514 считываний)
875 -0.29% (228 считываний)
877 -0.22% (172 считываний)
878 -0.21% (167 считываний)
Линия 4
875 - Исходная
876 -96.94% (70061 считываний)
875 -0.58% (419 считываний)
877 -0.23% (165 считываний)
878 -0.21% (150 считываний)

ФИГ.53D

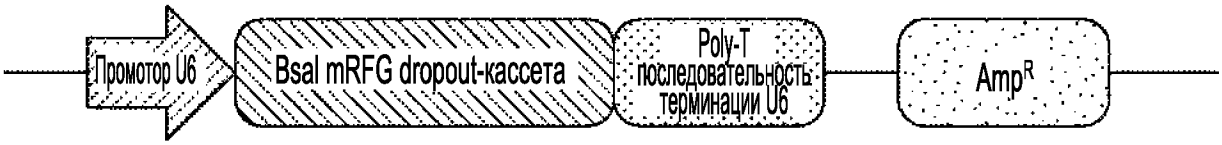


ФИГ.53D

ПРОДОЛЖЕНИЕ

Протокол для одновременного клонирования спейсерной последовательности (участка-мишени) и 3' удлинения PEgRNA, содержащего PBS и RT матрицу, в вектор экспрессии PEgRNA на основе промотора U6 человека:

[Компонент 1] Схема нерасщепленного pU6-PEgRNA-GG-вектора



Плазмида	Описание	Свойства	Разрезается с помощью
pU6-pegRNA-GG-Vector	U6 вектор экспрессии PEgRNA человека	U6_BsaI-mRFP-cassette_U6-Term	BsaI-HFV2

[Компонент 2] Упорядочивание нуклеотидов для получения спейсерной (явл. мишенью) последовательности для векторов PEgRNA человека:

Верхние и нижние олигонуклеотиды должны отжигаться и затем собираться в расщепляемый pU6-PEgRNA-GG-вектор

SEQ ID NO: 879 5′ - CACCNNNNNNNNNNNNNNNNNNNNGTTTT - 3′

SEQ ID NO: 880 3′ - NNNNNNNNNNNNNNNNNNNNCAAAATCTC - 5′

Показанная условная последовательность протоспейсера N20 может иметь различную длину, если это необходимо

[Компонент 3] Упорядочивание нуклеотидов для получения каркаса sgРНК SpCas9 для PEgRNA векторов человека:

Верхние и нижние олигонуклеотиды должны отжигаться и затем собираться в расщепляемый pU6-PEgRNA-GG-вектор

{**Phos**} 5′ - AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG - 3′ SEQ ID NO: 881

SEQ ID NO: 882 3′ - GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCAGC - 5′ {**Phos**}

(Следует обратить внимание, что эти олигонуклеотиды должны быть 5'-фосфорилированными для обеспечения сборки PEgRNA вектора)

[Компонент 4] Упорядочивание нуклеотидов для получения 3' удлинения PEgRNA, содержащего последовательность PBS и RT матрицы для PEgRNA векторов человека:

Верхние и нижние олигонуклеотиды должны отжигаться и затем собираться в расщепляемый pU6-PEgRNA-GG-вектор

SEQ ID NO: 883 5′ - GTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN - 3′

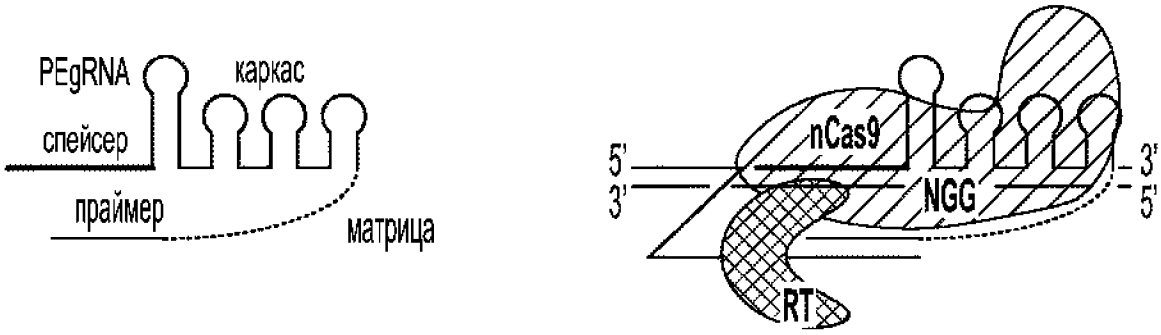
SEQ ID NO: 884 3′ - NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACGAAA - 5′

Показанная условная последовательность матрицы в виде обратной транскриптазы N30 может иметь различную длину, если это необходимо

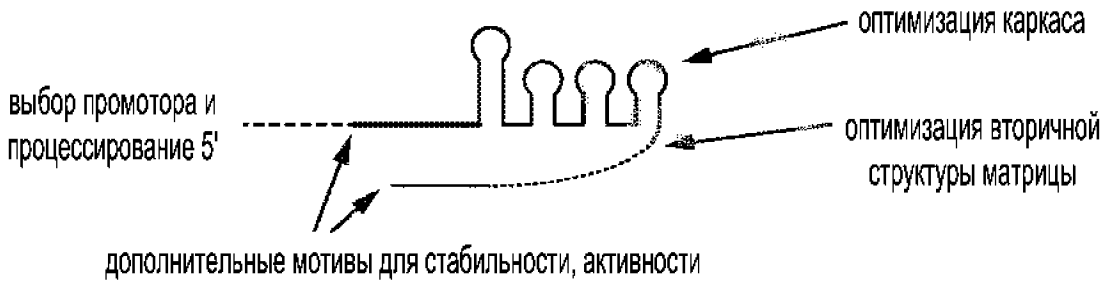
Схема конечного собранного вектора экспрессии PEgRNA на основе U6 из компонентов 1, 2, 3, и 4:



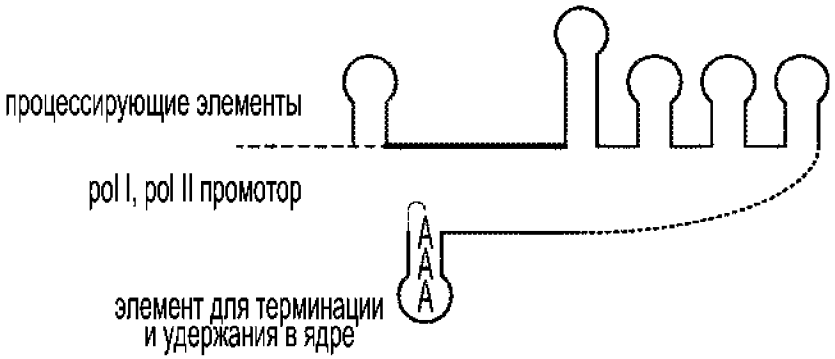
ФИГ.54



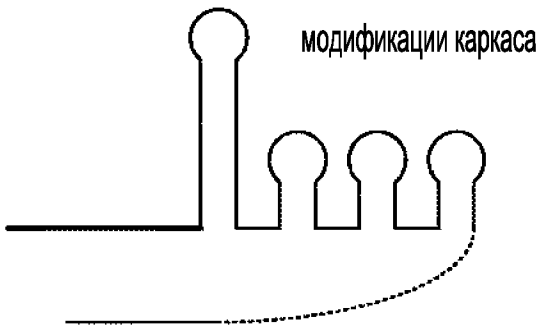
ФИГ.55A



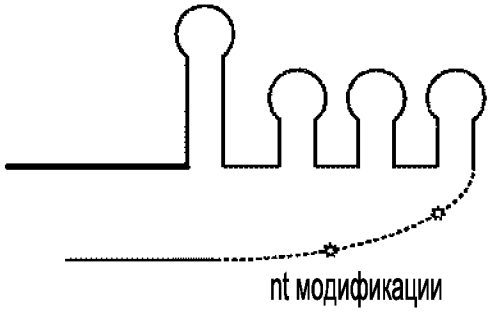
ФИГ.55B



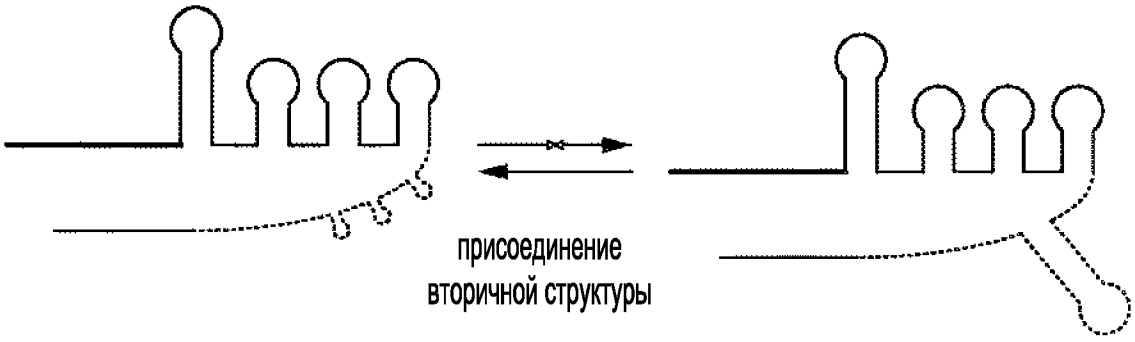
ФИГ.55C



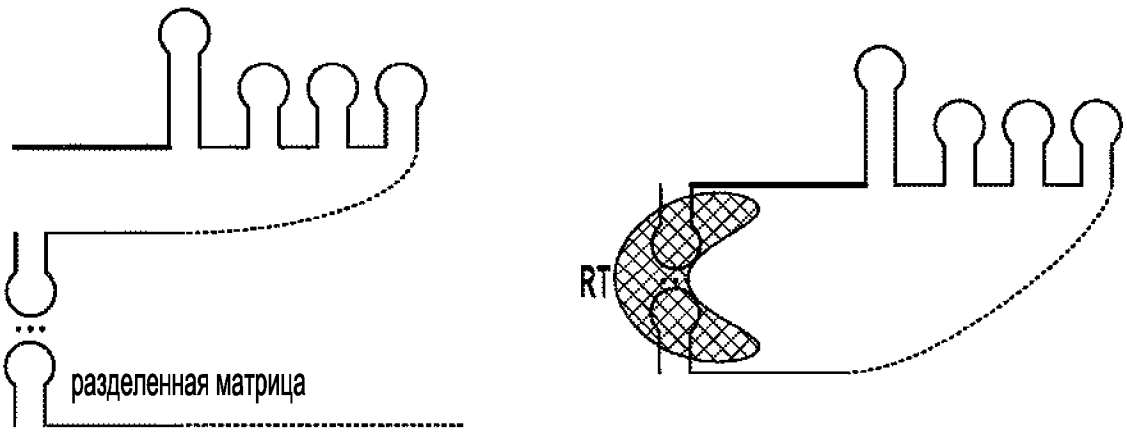
ФИГ.55D



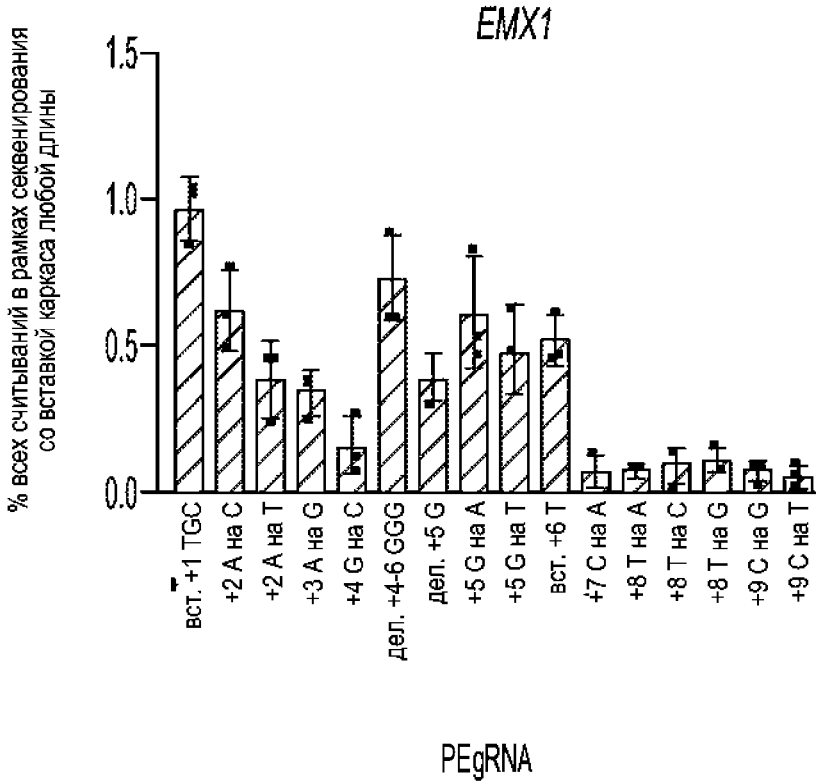
ФИГ.55E

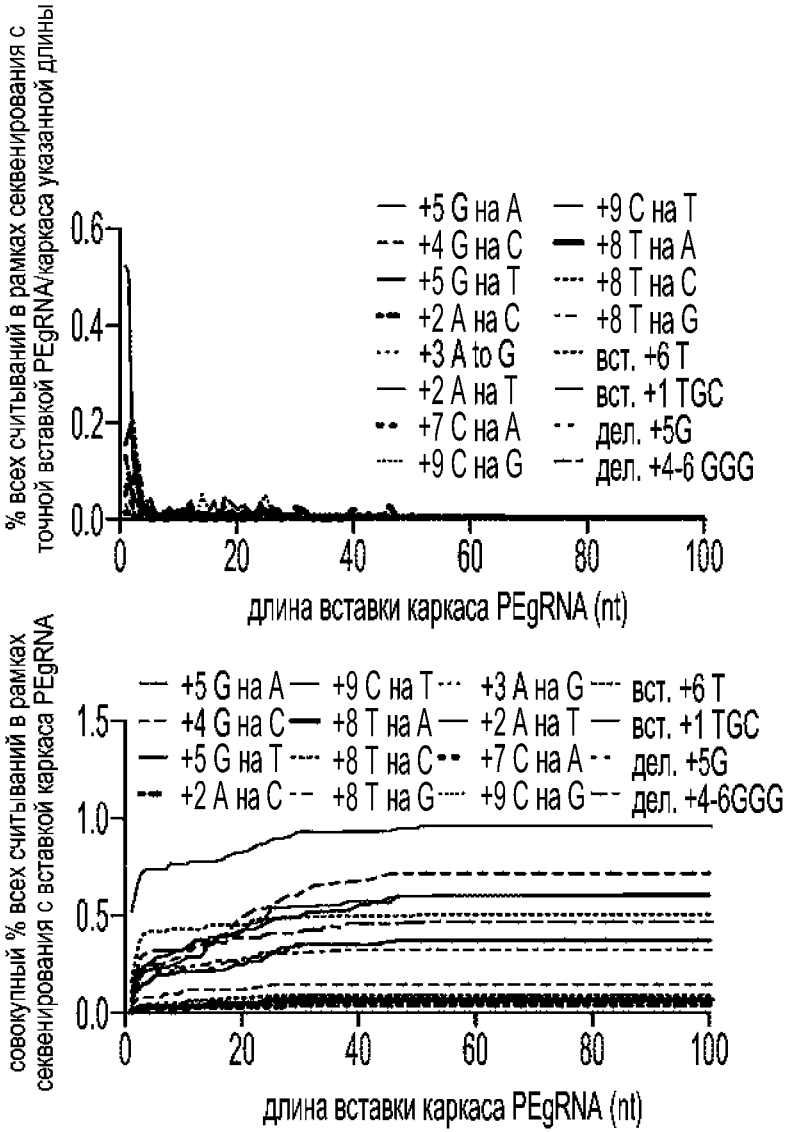


ФИГ.55F

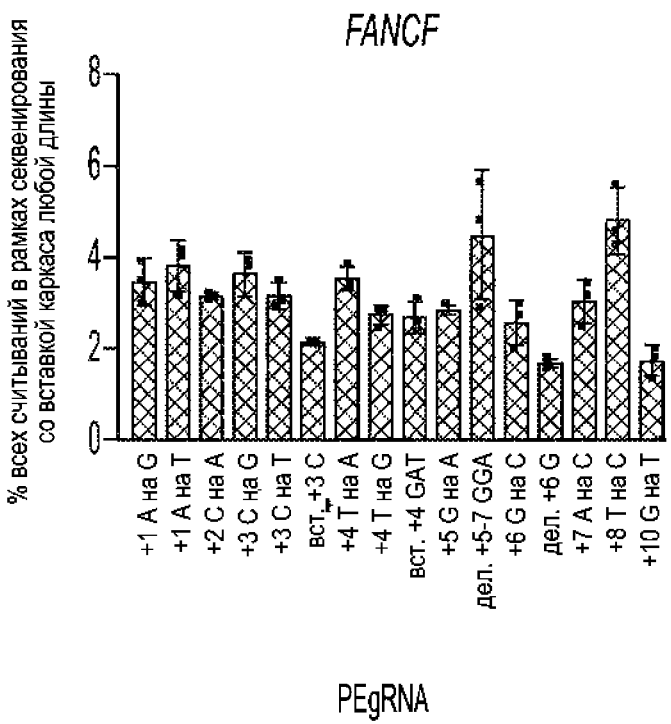


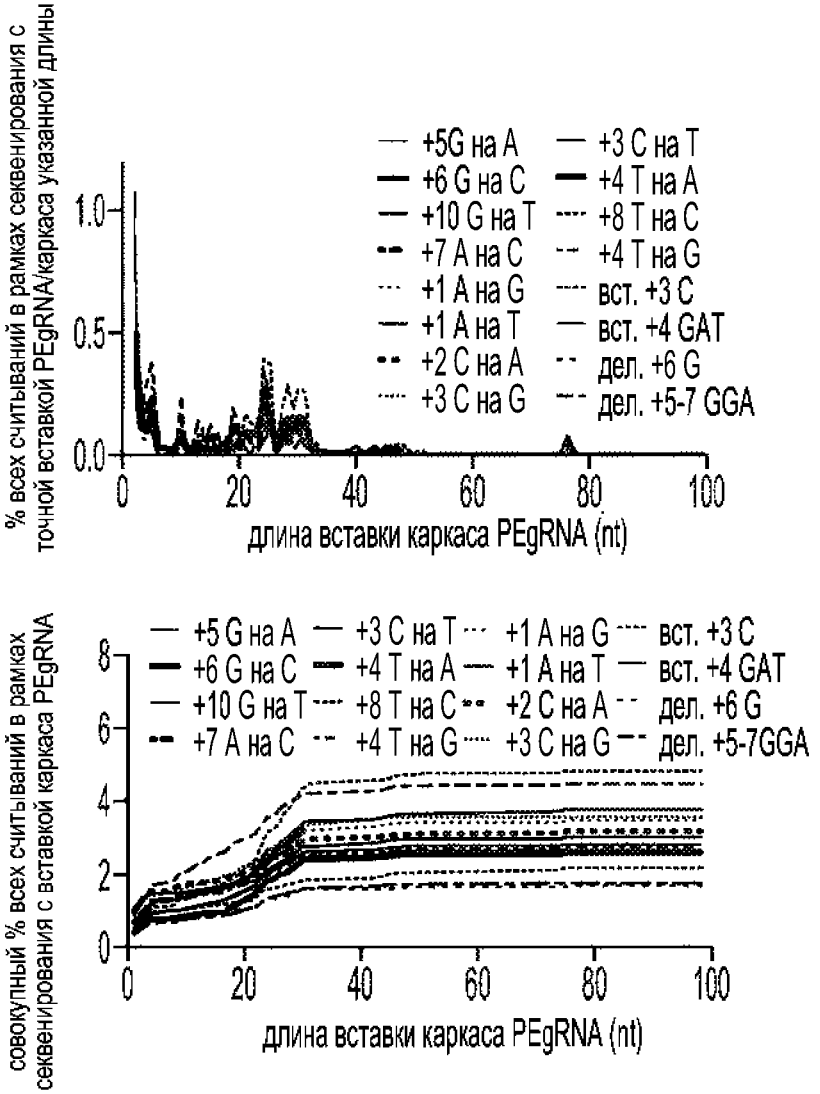
ФИГ.55G


EMX1
% всех считываний в рамках секвенирования со вставкой каркаса любой длины
1.5
1.0
0.5
0.0
вст. +1 TGC
+2 A на C
+2 A на T
+3 A на G
+4 G на C
дел. +4-6 GGG
дел. +5 G
+5 G на A
+5 G на T
вст. +6 T
+7 C на A
+8 T на A
+8 T на C
+8 T на G
+9 C на G
+9 C на T
PEgRNA

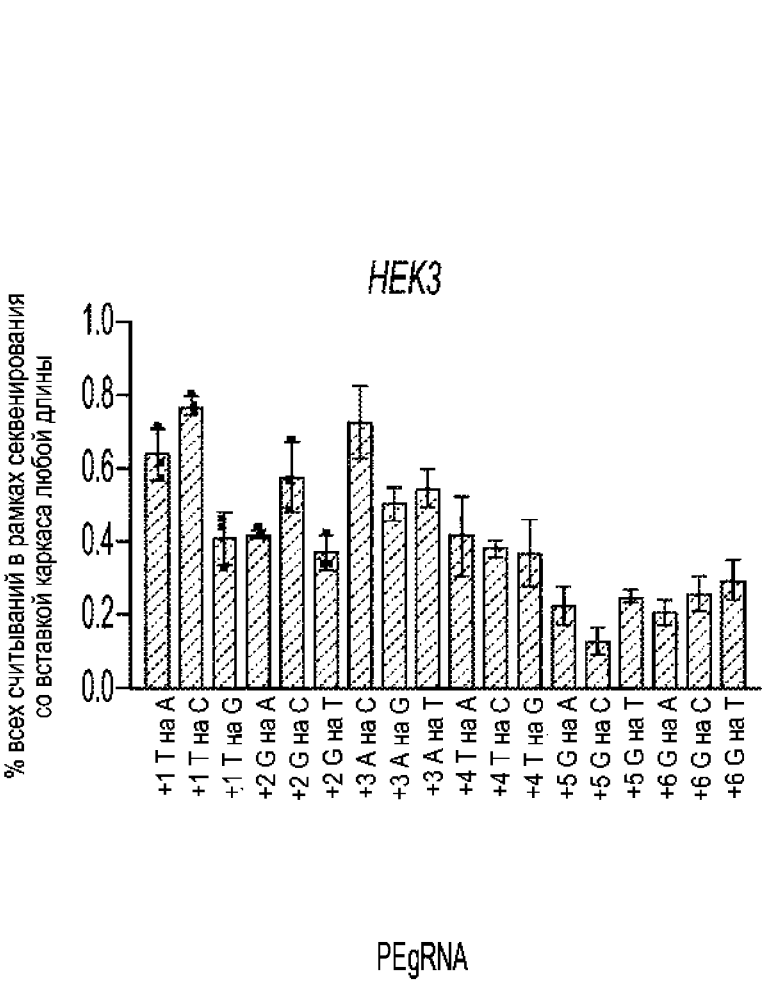

% всех считываний в рамках секвенирования с точной вставкой PEgRNA/каркаса указанной длины
0.6
0.4
0.2
0.0
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+5 G на A
+4 G на C
+5 G на T
+2 A на C
+3 A to G
+2 A на T
+7 C на A
+9 C на G
+9 C на T
+8 T на A
+8 T на C
+8 T на G
вст. +6 T
вст. +1 TGC
дел. +5G
дел. +4-6 GGG
совокупный % всех считываний в рамках секвенирования с вставкой каркаса PEgRNA
1.5
1.0
0.5
0.0
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+5 G на A
+4 G на C
+5 G на T
+2 A на C
+9 C на T
+8 T на A
+8 T на C
+8 T на G
+3 A на G
+2 A на T
+7 C на A
+9 C на G
вст. +6 T
вст. +1 TGC
дел. +5G
дел. +4-6GGG

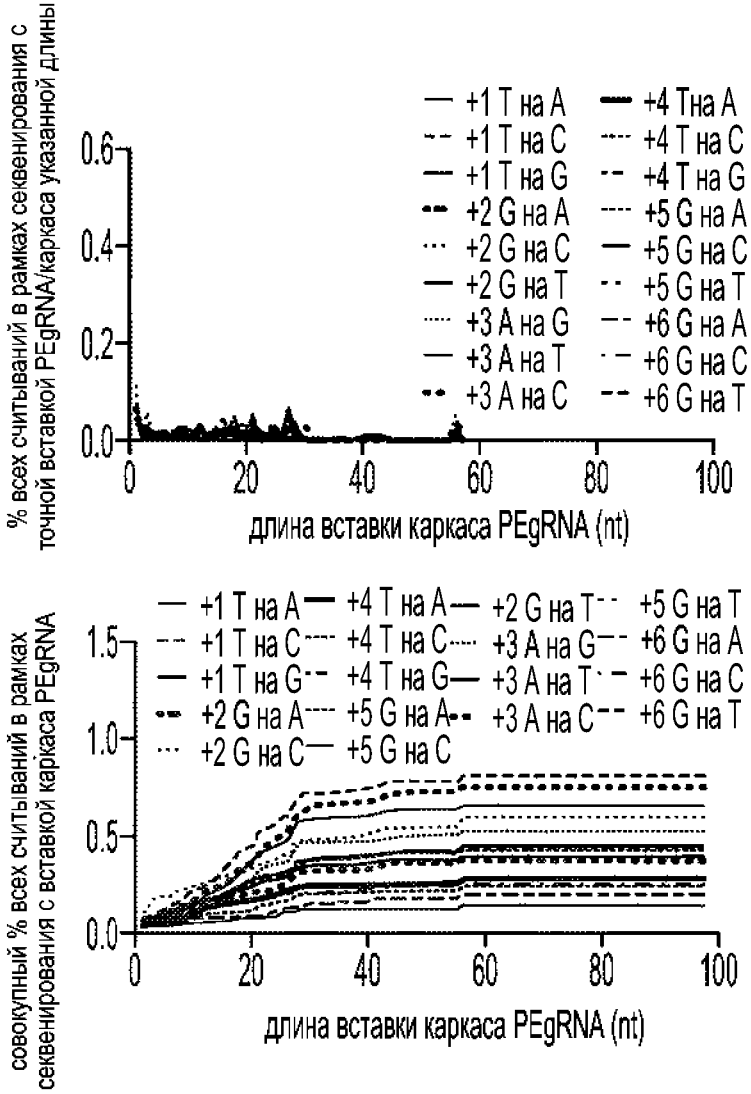
ФИГ.56A


FANCF
% всех считываний в рамках секвенирования со вставкой каркаса любой длины
0
2
4
6
8
+1 A на G
+1 A на T
+2 C на A
+3 C на G
+3 C на T
вст. +3 C
+4 T на A
+4 T на G
вст. +4 GAT
+5 G на A
дел. +5-7 GGA
+6 G на C
дел. +6 G
+7 A на C
+8 T на C
+10 G на T
PEgRNA

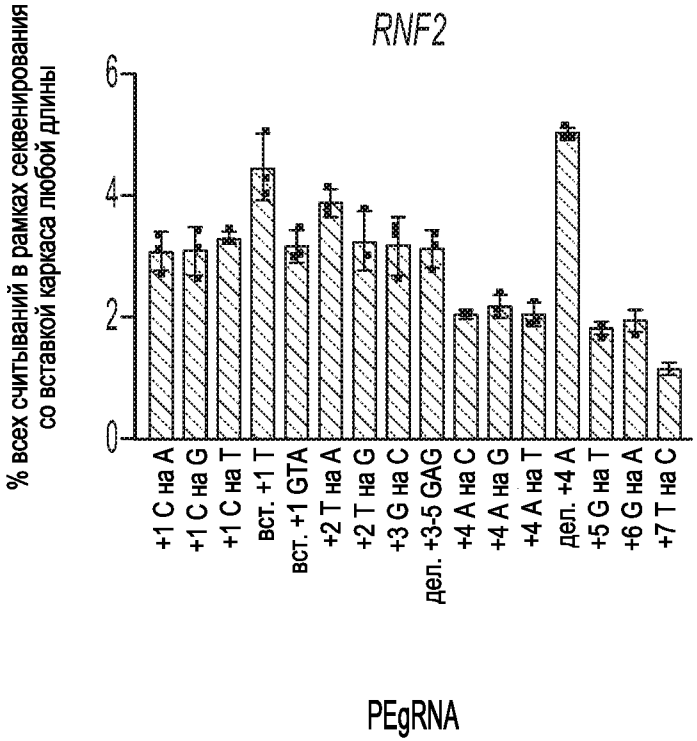

% всех считываний в рамках секвенирования с точной вставкой PEgRNA/каркаса указанной длины
0.0
0.5
1.0
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+5G на A
+6 G на C
+10 G на T
+7 A на C
+1 A на G
+1 A на T
+2 C на A
+3 C на G
+3 C на T
+4 T на A
+8 T на C
+4 T на G
вст. +3 C
вст. +4 GAT
дел. +6 G
дел. +5-7 GGA
совокупный % всех считываний в рамках секвенирования с вставкой каркаса PEgRNA
0
2
4
6
8
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+5 G на A
+6 G на C
+10 G на T
+7 A на C
+3 C на T
+4 T на A
+8 T на C
+4 T на G
+1 A на G
+1 A на T
+2 C на A
+3 C на G
вст. +3 C
вст. +4 GAT
дел. +6 G
дел. +5-7GGA

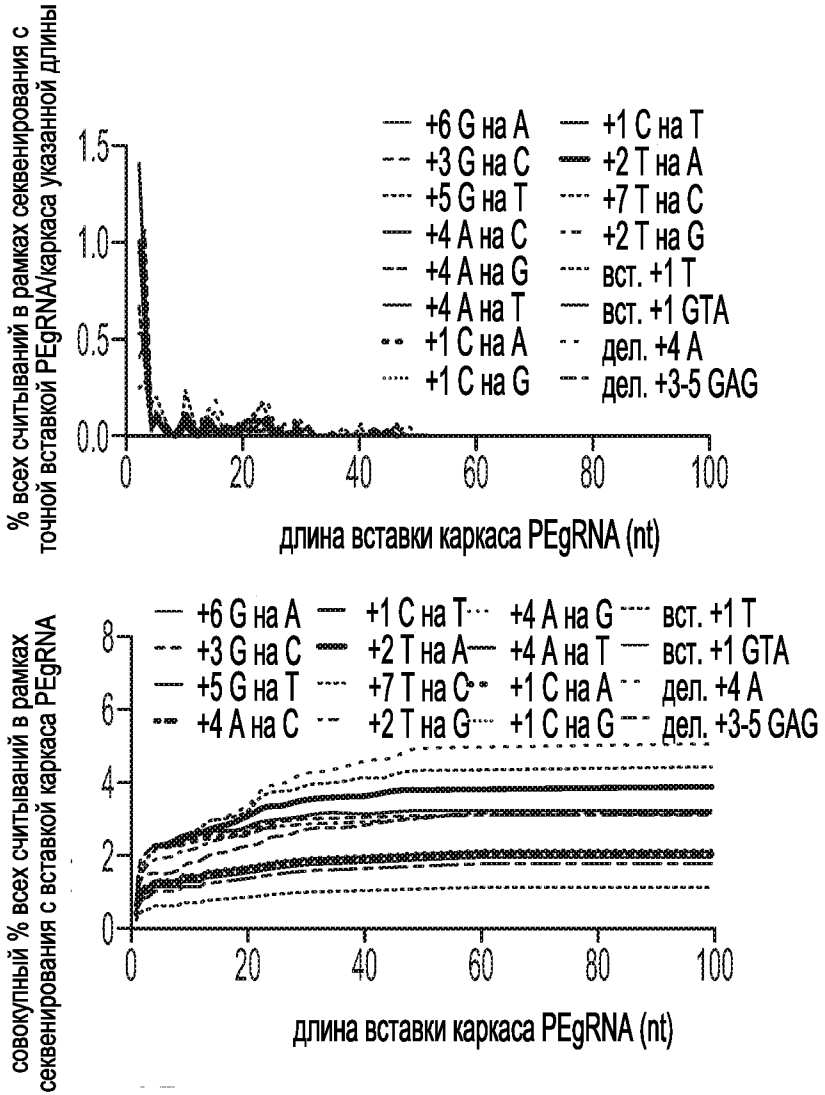
ФИГ.56В


HEK3
% всех считываний в рамках секвенирования со вставкой каркаса любой длины
1.0
0.8
0.6
0.4
0.2
0.0
+1 T на A
+1 T на C
+1 T на G
+2 G на A
+2 G на C
+2 G на T
+3 A на C
+3 A на G
+3 A на T
+4 T на A
+4 T на C
+4 T на G
+5 G на A
+5 G на C
+5 G на T
+6 G на A
+6 G на C
+6 G на T
PEgRNA

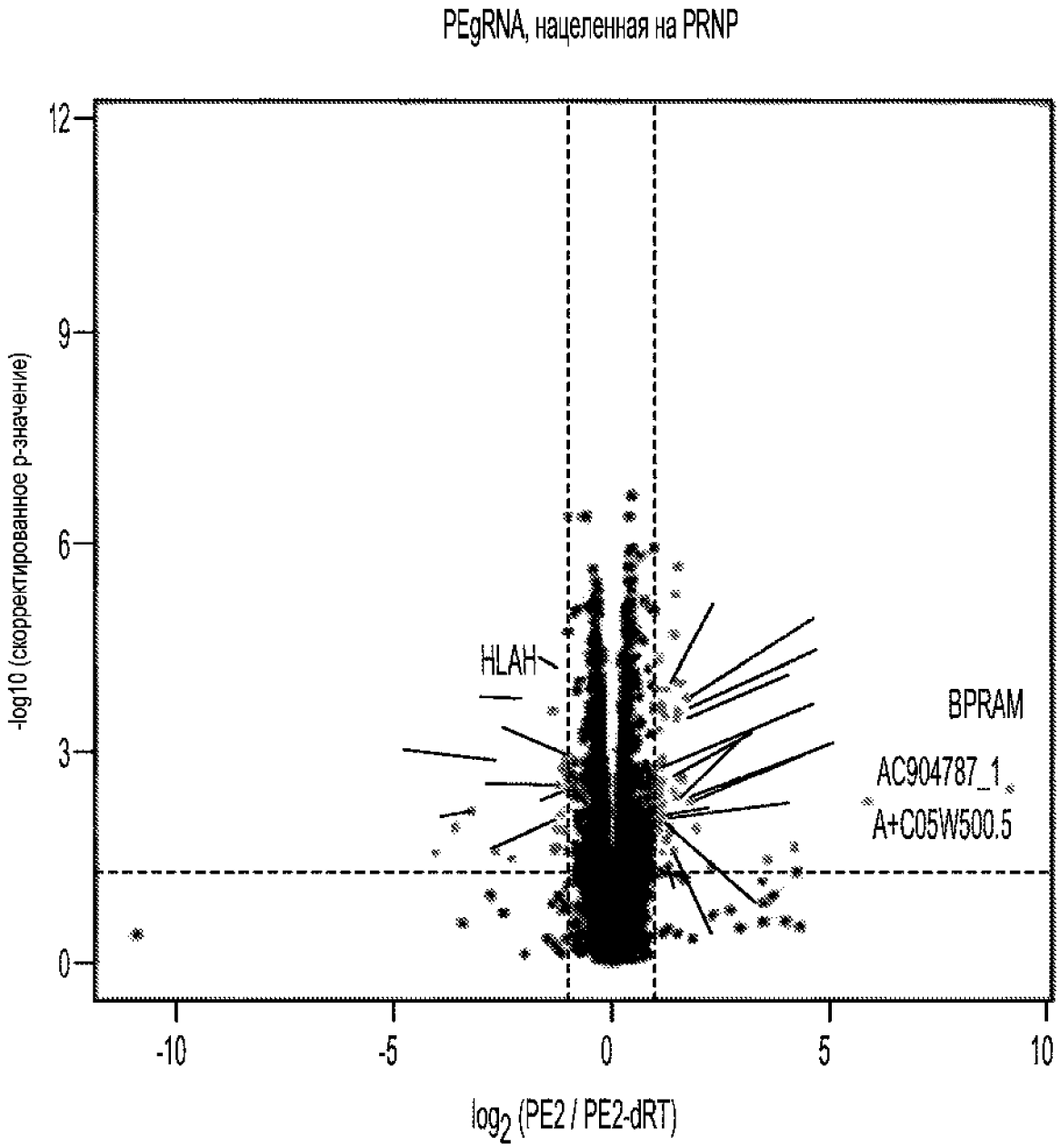

% всех считываний в рамках секвенирования с точной вставкой PEgRNA/каркаса указанной длины
0.6
0.4
0.2
0.0
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+1 T на A
+1 T на C
+1 T на G
+2 G на A
+2 G на C
+2 G на T
+3 A на G
+3 A на T
+3 A на C
+4 Tна A
+4 T на C
+4 T на G
+5 G на A
+5 G на C
+5 G на T
+6 G на A
+6 G на C
+6 G на T
совокупный % всех считываний в рамках секвенирования с вставкой каркаса PEgRNA
1.5
1.0
0.5
0.0
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)

ФИГ.56С

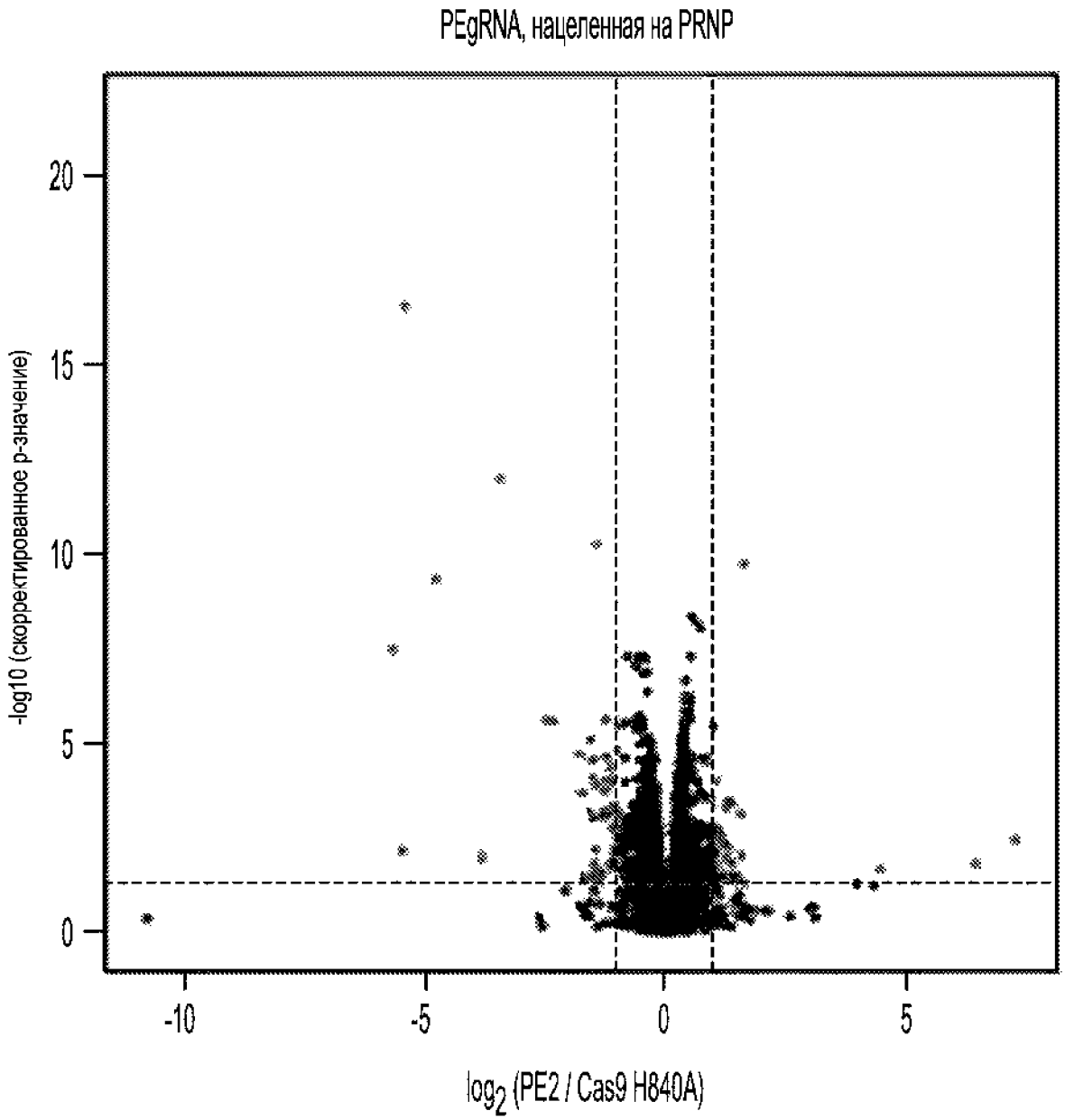

RNF2
% всех считываний в рамках секвенирования со вставкой каркаса любой длины
0
2
4
6
+1 C на A
+1 C на G
+1 C на T
вст. +1 T
вст. +1 GTA
+2 T на A
+2 T на G
+3 G на C
дел. +3-5 GAG
+4 A на C
+4 A на G
+4 A на T
дел. +4 A
+5 G на T
+6 G на A
+7 T на C
PEgRNA


% всех считываний в рамках секвенирования с точной вставкой PEgRNA/каркаса указанной длины
0.0
0.5
1.0
1.5
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)
+6 G на A
+3 G на C
+5 G на T
+4 A на C
+4 A на G
+4 A на T
+1 C на A
+1 C на G
+1 C на T
+2 T на A
+7 T на C
+2 T на G
вст. +1 T
вст. +1 GTA
дел. +4 A
дел. +3-5 GAG
совокупный % всех считываний в рамках секвенирования с вставкой каркаса PEgRNA
0
2
4
6
8
0
20
40
60
80
100
длина вставки каркаса PEgRNA (nt)

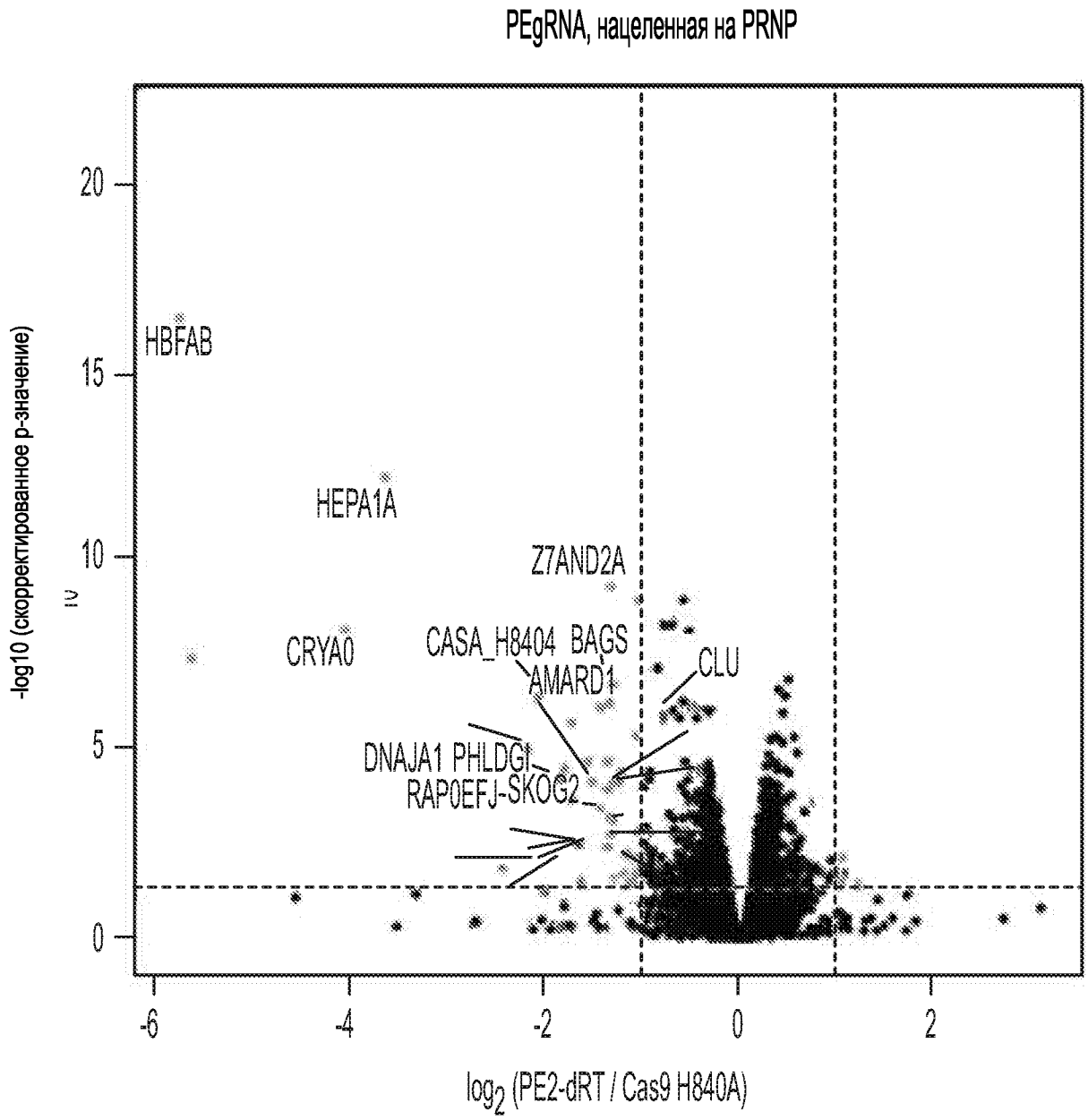
ФИГ.56D


PEgRNA, нацеленная на PRNP
12
9
6
3
0
-log10 (скорректированное p-значение)
HLAH
BPRAM
AC904787_1
A+C05W500.5
-10
-5
0
5
10
log2 (PE2 / PE2-dRT)

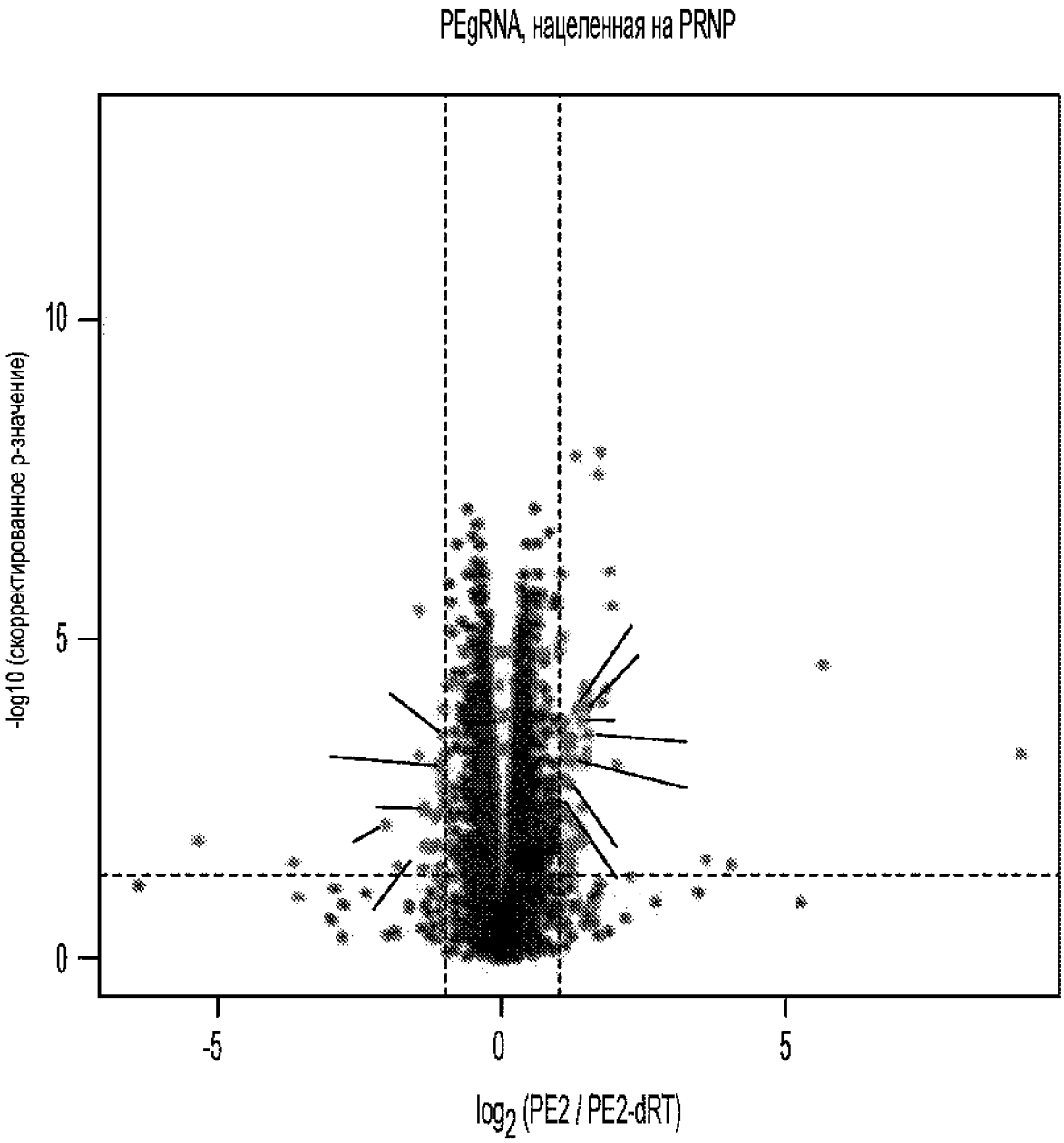
ФИГ.57А


PEgRNA, нацеленная на PRNP
-log10 (скорректированное p-значение)
20
15
10
5
0
-10
-5
0
5
log2 (PE2 / Cas9 H840A)

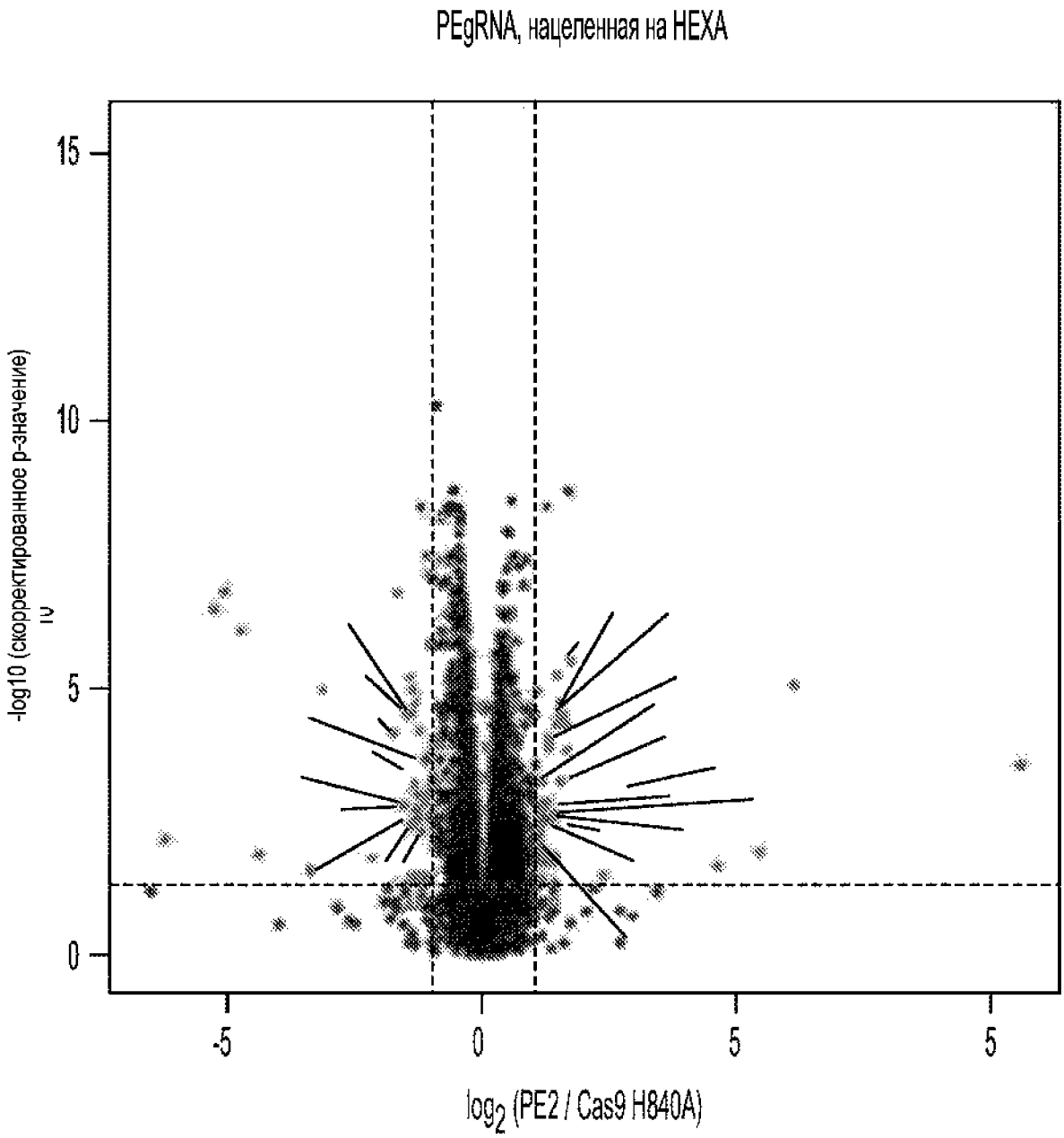
ФИГ.57В


PEgRNA, нацеленная на PRNP
-log10 (скорректированное p-значение)
20
15
10
5
0
HBFAB
HEPA1A
Z7AND2A
CRYAB
CASA_H8404
BAGS
AMARD1
CLU
DNAJA1
PHLDGI
RAP0EFJ-SKOG2
-6
-4
-2
0
2
log2 (PE2-dRT / Cas9 H840A)

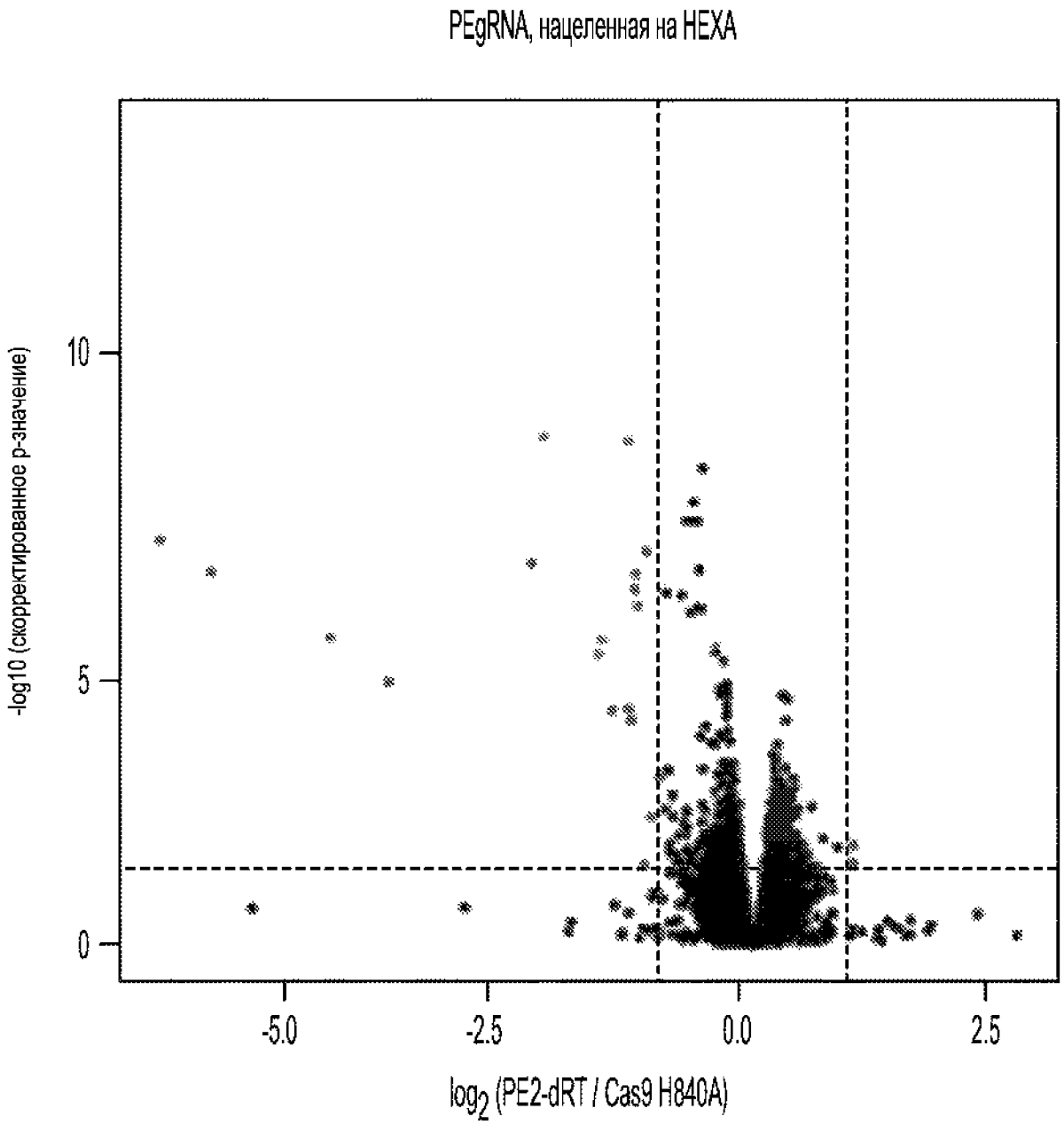
ФИГ.57С


PEgRNA, нацеленная на PRNP
-log10 (скорректированное p-значение)
10
5
0
-5
0
5
log2 (PE2 / PE2-dRT)

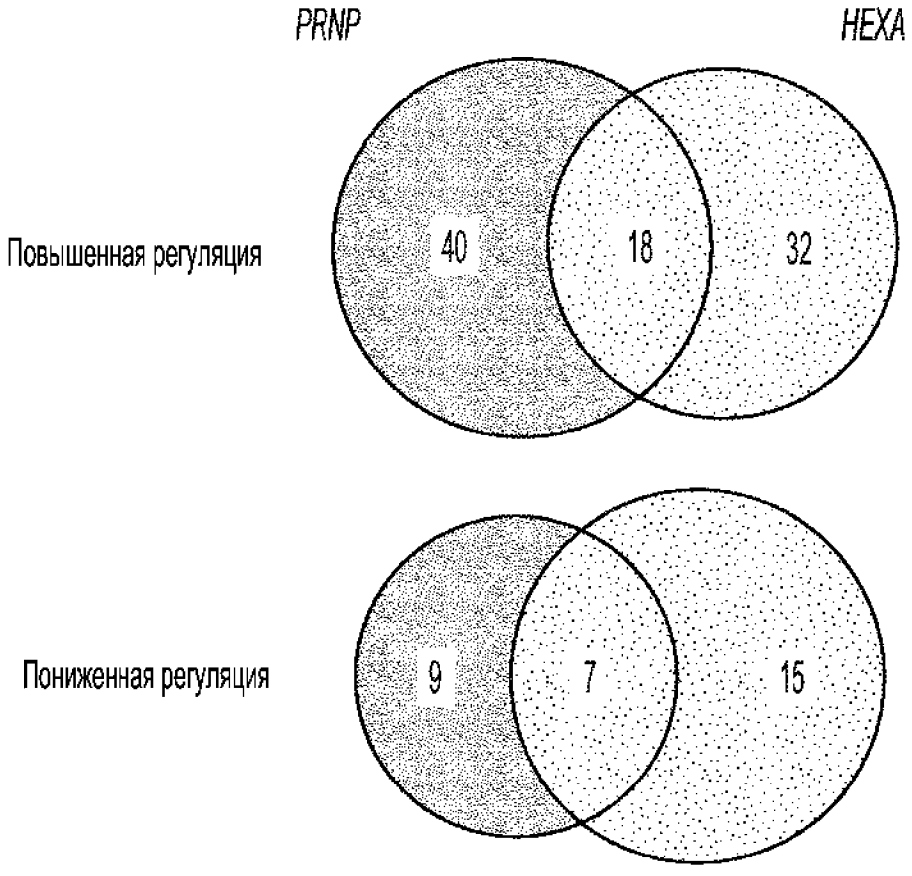
ФИГ.57D


PEgRNA, нацеленная на HEXA
15
10
5
0
-log10 (скорректированное p-значение)
-5
0
5
5
log2 (PE2 / Cas9 H840A)

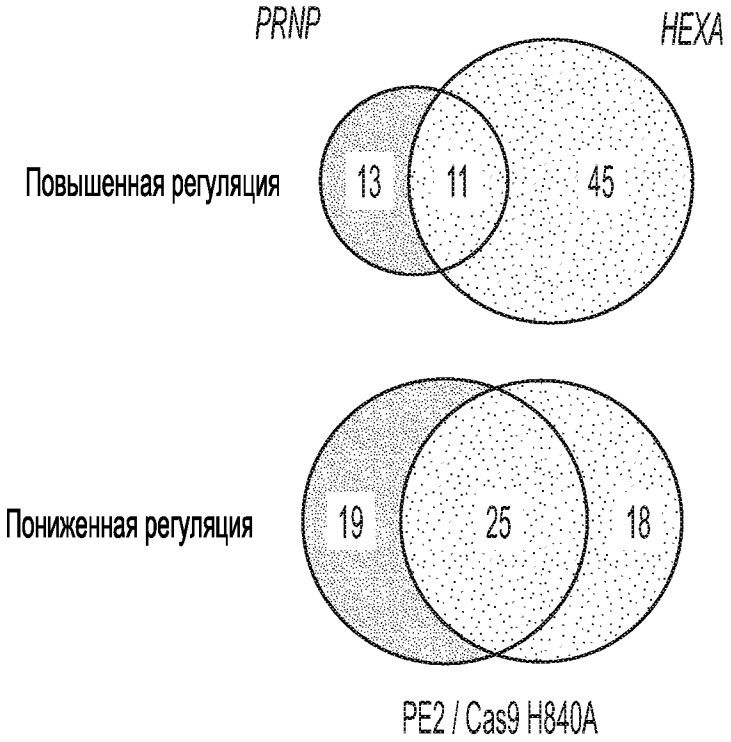
ФИГ.57E


PEgRNA, нацеленная на HEXA
-log10 (скорректированное p-значение)
10
5
0
-5.0
-2.5
0.0
2.5
log2 (PE2-dRT / Cas9 H840A)

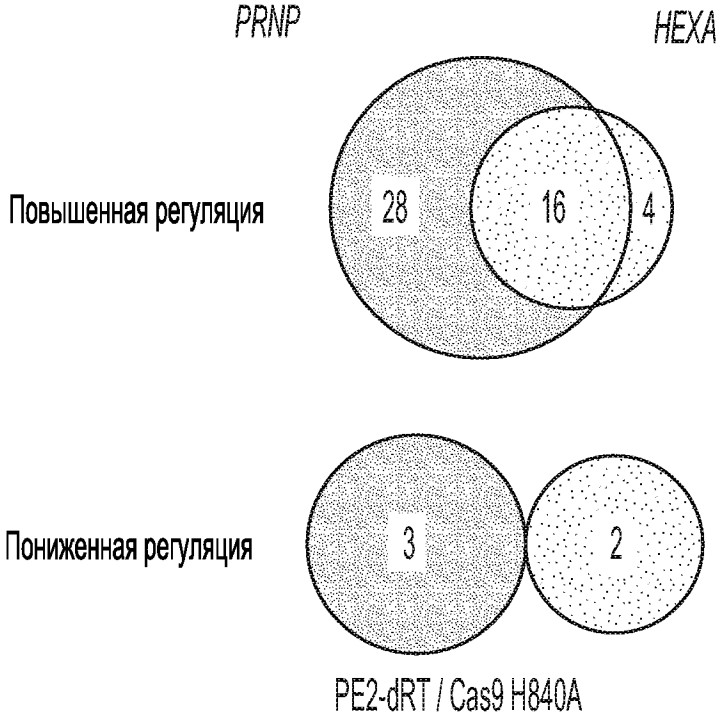
ФИГ.57F


PRNP
HEXA
Повышенная регуляция
40
18
32
Пониженная регуляция
9
7
15
PE2 / PE2-dRT

ФИГ.57G

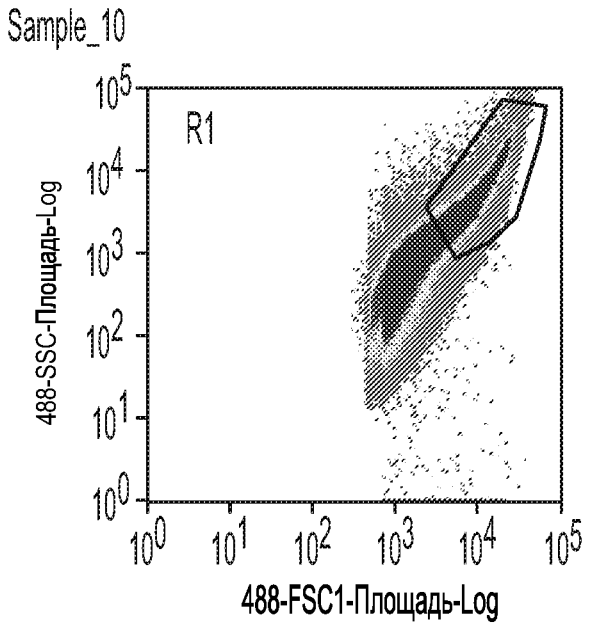


ФИГ.57H

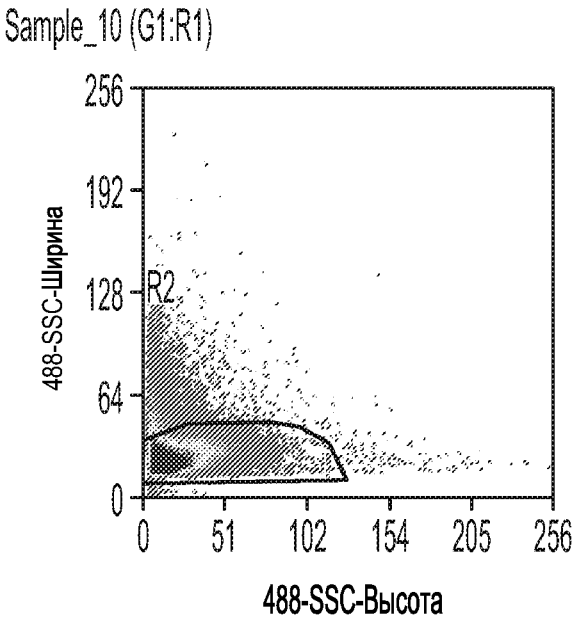


ФИГ.57I

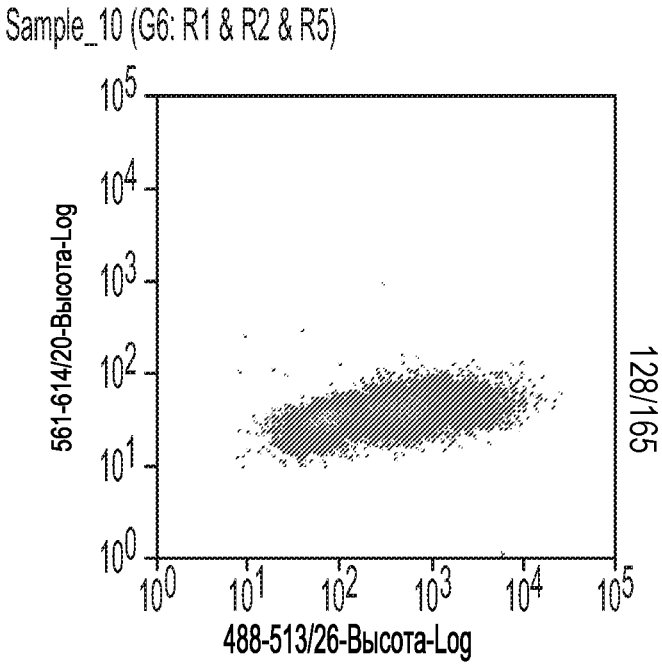
TdT DAPI планшеты Sample_10 D:---\091119\1.1 pool.fcs



Область	Кол-во	Гист. %	Всего, %
Всего	271204	100.00	100.00
RI	86850	32.02	32.02

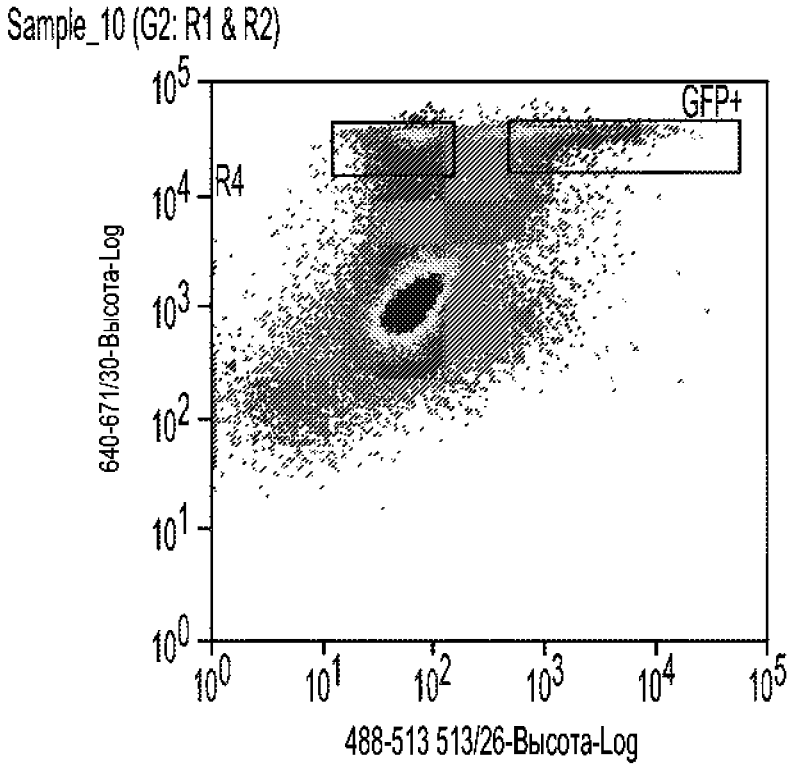


Область	Кол-во	Гист., %	Всего, %
Всего	86850	100.00	32.02
R2	81289	93.60	29.97



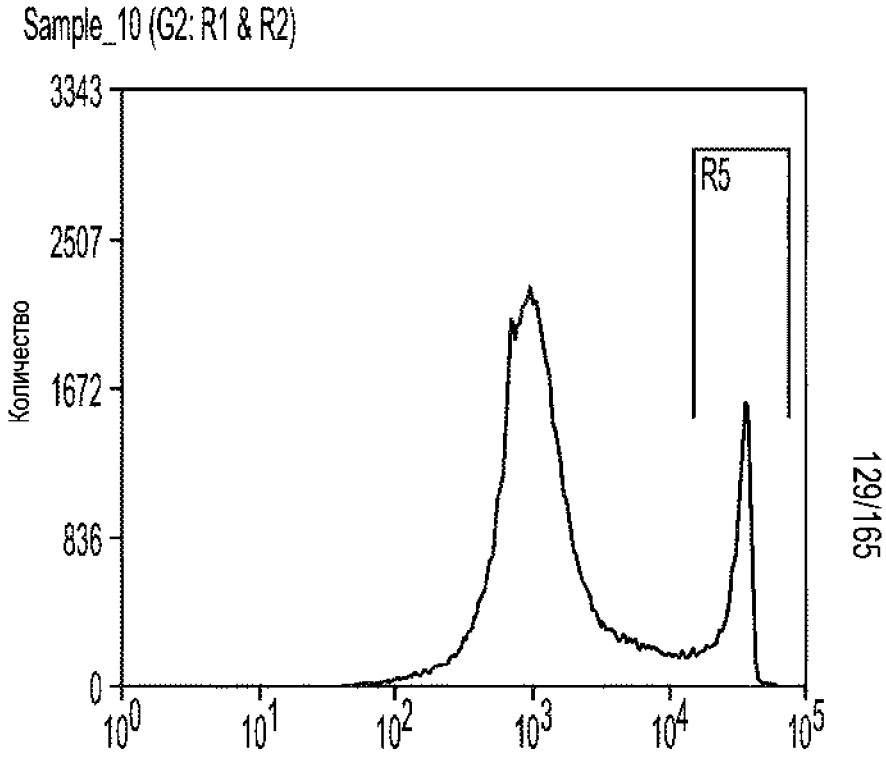
Область	Кол-во	Гист.,%	Всего, %
Всего	12569	100.00	4.63

ФИГ.58



Область	Кол-во	Гист., %	Всего, %	Связано	Количество мод
Total	81289	100.00	29.97 (1.00,1000000.04)	(100000.04,1...	128
GFP +	4418	5.43	1.63 (485.60, 48559.55)	(58170.94,1...	40
R4	4709	5.79	1.74 (11.98,46415.91)	(150.13,1501...	34

zniziolek Wednesday, September 11,2019 01:36:41PM Page 1



Область	Кол-во	Гист., %	Средн.	CV
Всего	81289	100.00	6521.68	179.64
R5	12569	15.46	32516.62	24.27

ФИГ.58
ПРОДОЛЖЕНИЕ

Обзор клонирования

1. Расщепить плазмиду в виде pU6-PEgRNA-GG-вектора (компонент 1) с помощью BsaI и выделить фрагмент плазмиды (~2,2 kb), содержащий ориджин репликации, промотор U6, poly-T последовательность терминации U6, и ген AmpR.
2. Упорядочить олигонуклеотиды для:
   a. Требуемой последовательности спейсера (мишени), фланкированной указанными выступами (компонент 2)
   i. Использовать 5'-3' последовательность требуемой мишени для олигонклеотида верхней цепи (включая 5' CACC и 3' GTTTT выступы) и использовать обратный комплемент последовательности-мишени для олигонуклеотида из нижней цепи (включая 5' CTCTAAAAC выступ). Спейсерные последовательности должны начинаться на нуклеотид G для эффективного инициирования транскрипции
   b. Требуемой матрицы в виде 3' удлинения PEgRNA, фланкированной указанными выступами (компонент 3)
   i. Использовать смысловую последовательность РНК в качестве олигонуклеотида верхней цепи (включая 5' GTGC выступ) и использовать обратный комплемент этой последовательности для олигонуклеотида нижней цепи (включая 5' AAAA выступ)
   c. Матрицы каркаса sgPHK SpCas9, включающей совместимые выступы Golden Gate (компонент 4)
   i. Эти олигонуклеотиды не являются полной последовательностью каркаса, поскольку выступы из остальных компонентов обеспечивают внесение нескольких недостающих нуклеотидов
   ii. Следует обратить внимание: эти олигонуклеотиды должны быть 5'-фосфорилированными. Олигонуклеотиды могут быть 5'-фосфорилированными производителем или 5-фосфорилированы ферментативно с помощью T4 PNK (см. протокол ниже)
3. Отжиг верхних и нижних нуклеотидов для компонентов 2, 3, и 4 в отдельных реакциях отжига в соответствии с протоколом ниже. Если последовательность каркаса sgPHK SpCas9 (компонент 4) не фосфорилирована — фосфорилировать с помощью T4 PNK.
4. Сборка Golden Gate выделенного 2,2 kb фрагмента из компонента 1 с компонентами 2, 3 и 4
5. Трансформация лигированного продукта в E.coli. Резистентность к антибиотикам, обеспеченная компонентом 1 из плазмиды на основе pU6-PEgRNA-GG-вектора, представляет собой резистентность к ампициллину и карбенициллину.
6. Выделить и секвенировать плазмиды из полученных клональных трансформантов.

ФИГ.59A

Протокол клонирования PEgRNA

Этап 1: Расщепить плазмиду в виде pU6-PEgRNA-GG-вектора (компонент 1)

Соединить следующие компоненты в пробирке для ПЦР:

2000 нг pU6-PEgRNA-GG-вектора (компонент 1)	X мкл
Bsa1-HFv2 (NEB)	1.0 мкл
10x буфера Cutsmart	3.0 мкл
H_2O	до 30.0 мкл
Общий объем реакции	30.0 мкл

Инкубировать при 37 °C в течение 4-16 ч

Выделить ~2.2-kb фрагмент из разрезанной плазмиды

ФИГ.59B

Этапы 2 и 3: Упорядочить и осуществить отжиг частей олигонуклеотидов (компоненты 2, 3, и 4)

Материалы

Буфер для отжига: H_2O с добавлением 10 мМ Трис-Cl pH 8.5 и 50 мМ NaCl
Комплементарные олигонуклеотидные части

Протокол

Соединить следующие компоненты в пробирке для ПЦР

Верхний олигонуклеотид, 100 мкМ	1.0 мкл
Нижний олигонуклеотид, 100 мкМ	1.0 мкл
Буфер для отжига (компоненты 2,3, и 4)	23.0 мкл
Общий объем реакции	25.0 мкл

В термоциклере; нагреть до 95 °C в течение 3 минут, затем постепенно остужать (0,1 °C/сек) до 22 °C
Разбавить подвергнутые отжигу олигонуклеотиды 1:4 путем добавления 75 мкл H_2O. Конечная концентрация каждого олигонуклеотида будет составлять 1 мкМ после данного разбавления. Не разбавлять каркас sgРНК (компонент 4), если осуществлялось фосфорилирование с помощью PNK на этапе 2.5.

ФИГ.59C

Этап 2.b.ii.: фосфорилирование каркаса sgRNA (не требуется, если олигонуклеотиды были закуплены в фосфорилированном виде)

Протокол

Соединить следующие компоненты в пробирке для ПЦР

4 мкМ олигонуклеотидный дуплекс с этапа 1	6.25 мкл
10xT4 ДНК-лигазный буфер (NEB)	2.50 мкл
T4 PNK (NEB)	0.50 мкл
H_2O	15.75 мкл
Общий объем реакции	20.0 мкл

Инкубировать в термоциклере при 37 °C в течение 60 минут
После данного фосфорилирования подвергнутые отжигу олигонуклеотиды каркаса достигают концентрации 1 мкМ. Перейти к этапу 3.

ФИГ.59D

Этап 4: сборка PEgRNA

PEgRNA. реакция сборки Golden Gate

Расщепленная плазмида в виде pU6-PEgRNA-GG-вектора - Предварительно вырезанный, выделенный 2.2 kb фрагмент	1.00 мкл при 30 нг/мкл
Подвергнутые отжигу протоспейсерные олигонуклеотиды (компонент 2)	1.00 мкл при 1мкМ
Подвергнутые отжигу олигонуклеотиды 3' удлинения PEgRNA (компонент 4)	1.00 мкл при 1мкМ
Подвергнутые отжигу олигонуклеотиды каркаса sgРНК (компонент 3) - Олигонуклеотиды должны быть фосфорилированы	1.00 мкл при 1мкМ
BsaI-HFv2 (NEB)	0.25 мкл
T4 ДНК-лигаза (NEB)	0.50 мкл
10xT4 ДНК-лигазный буфер(NEB)	1.00 мкл
H_2O	4.25 мкл
Общий объем реакции:	10.0 мкл

Инкубировать при комнатной температуре в течение 10 минут
После инкубирования осуществить следующую программу в термициклере:
5 минут при 37 °C, затем 5 минут при 80 °C, затем удерживать при 12 °C

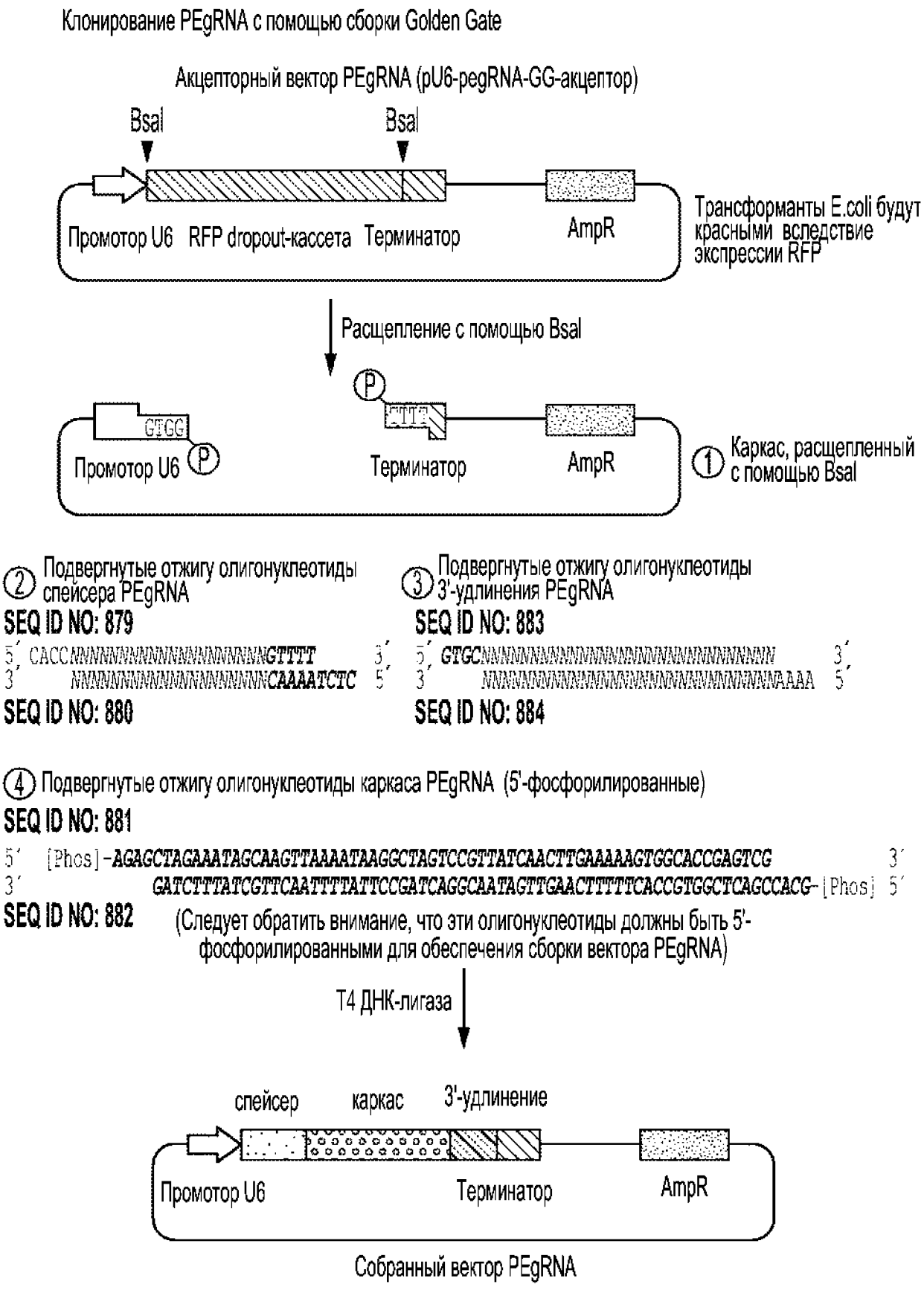
ФИГ.59E

Этапы 5 и 6: Трансформация собранных плазмид

Трансформировать 1 мкл из 10 мкл реакции сборки в подходящие клетки. Требуемые трансформанты будут устойчивыми к ампициллину и карбенициллину.

На следующей диаграмме показан обзор протокола клонирования PEgRNA.

ФИГ.59F


Клонирование PEgRNA с помощью сборки Golden Gate
Акцепторный вектор PEgRNA (pU6-pegRNA-GG-акцептор)
BsaI
BsaI
Промотор U6
RFP dropout-кассета
Терминатор
AmpR
Трансформанты E.coli будут красными вследствие экспрессии RFP
Расщепление с помощью BsaI
GTGG
Промотор U6
P
CTTA
Терминатор
AmpR
① Каркас, расщепленный с помощью BsaI
② Подвергнутые отжигу олигонуклеотиды спейсера PEgRNA
SEQ ID NO: 879
5′ CACCNNNNNNNNNNNNNNNNNNGTTTT 3′
3′ NNNNNNNNNNNNNNNNNNCAAAATCTC 5′
SEQ ID NO: 880
③ Подвергнутые отжигу олигонуклеотиды 3'-удлинения PEgRNA
SEQ ID NO: 883
5′ GTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3′
3′ NNNNNNNNNNNNNNNNNNNNNNNNNNNNAAA 5′
SEQ ID NO: 884
④ Подвергнутые отжигу олигонуклеотиды каркаса PEgRNA (5'-фосфорилированные)
SEQ ID NO: 881
5′ [Phos]-AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG 3′
3′ GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCACG-[Phos] 5′
SEQ ID NO: 882
(Следует обратить внимание, что эти олигонуклеотиды должны быть 5'-фосфорилированными для обеспечения сборки вектора PEgRNA)
T4 ДНК-лигаза
спейсер
каркас
3'-удлинение
Промотор U6
Терминатор
AmpR
Собранный вектор PEgRNA

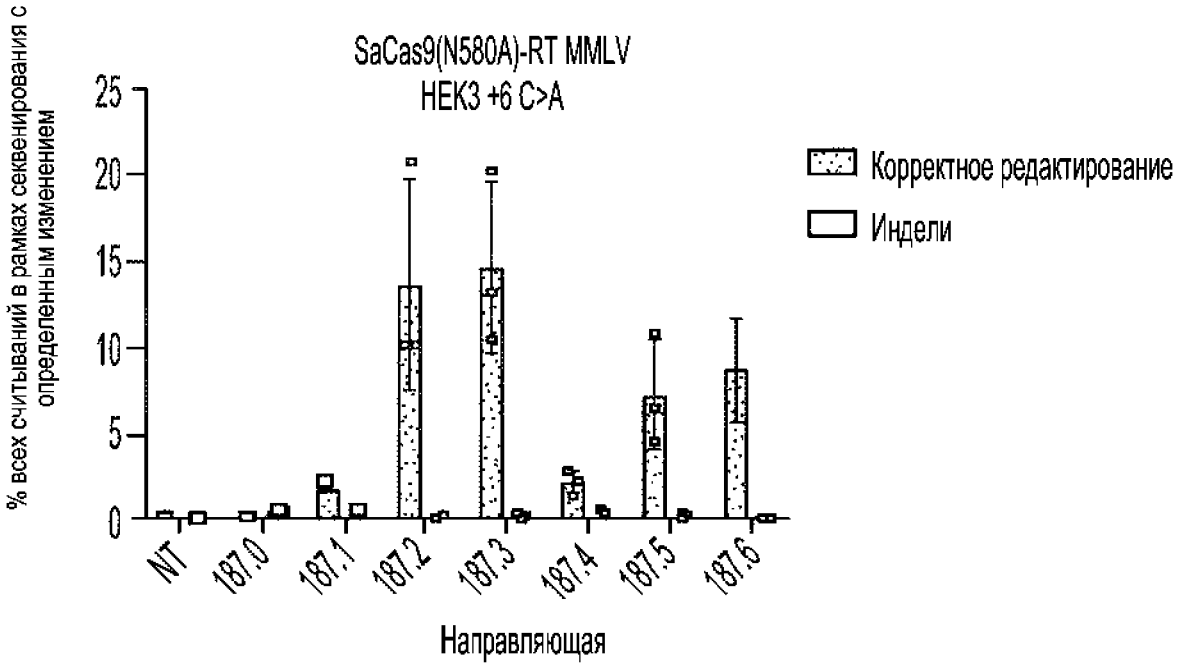
ФИГ.59G

```
## поиск последовательности каркаса sgРНК

import pandas as pd
import Bio as bio
from Bio import SeqIO
import glob

generates list of fastq files to analyze
sources = glob.glob('*.fastq')

# чтение файлов fastq в словарь с именами файлов в качестве ключей
fastqdict = {}
for i in range(len(sources)):
        temp = list(SeqIO.parse(source[i],"fastq"))
        fastqdict[sources[i]]= [str(temp[k].seq) for k in range(len(temp))]


# искомая последовательность для поиска вводится в следующий словарь с подходящим ключом
scaffdict =
{'HEK3':'CAGAGGACCGACTCGGTCCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACTCACGTGCTCAGTCTGGCCGGTG', (SEQ ID NO: 885)
'EMX1':'ATCACGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TGCTATTTCTAGCTCTAAAACTTCTTCTTCTGCTCGGACTCGGTG', (SEQ ID NO: 886)
'FANCF':'TTTCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAAC
TTGCTATTTCTAGCTCTAAAACGGTGCTGCAGAAGGATTCCGGTG', (SEQ ID NO: 887)
'RNF2':'TCGTTGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACCAGGTAATGACTAAGATGACGGTG'} (SEQ ID NO: 888)

# сравнение и подсчет итеративных последовательностей искомой строки в подходящих файлах fastq
# искомый ключ должен содержаться в имени файла fastq
# полученные значения обозначают совокупное количество минимальной степени встраивания sgРНК
# то есть значение x обозначает, что x считываний содержат y или более оснований каркаса
resultdict = dict.fromkeys(sources)
for key in fastqdict:
   for scaffold in scaffdict:
      if scaffold in str(key):
         resultlist = []
         for j in range(len(scaffdict[scaffold])):
```

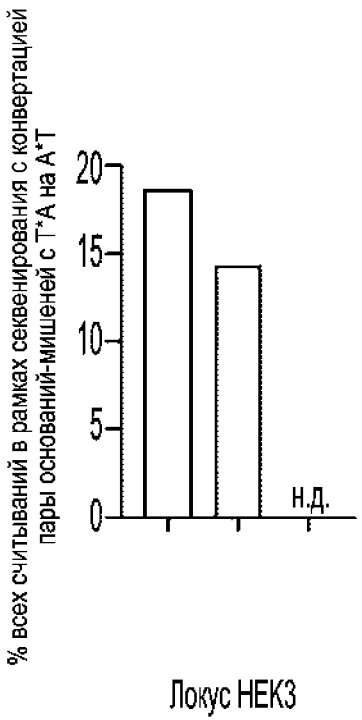
ФИГ.60А

```
                extent = scaffdict[scaffold][0:(j+1)]
                counter = 0
                for i in range(len(fastqdict[key])):
                    if extent in fastqdict[key][i]:
                        counter = counter + 1
                resultlist.append(counter)
            resultdict[key]=resultlist

# запись результатов во фрейм данных с нумерацией с 1
resultdf = pd.DataFrame.from_dict(resultdict)
resultdf = resultdf.reindex(sorted(resultdf.columns), axis=1)
resultdf.index = range(1,len(resultdf)+1)

# конвертация совокупных значений количества в конкретные количества
# то есть указанное значение x обозначает, что x считываний содержат точно у оснований каркаса
resultdf2=resultdf.copy()
for entry in resultdf:
    for i in range(1,len(resultdf[entry])+1):
        try:
            resultdf2[entry][i] = resultdf[entry][i]-resultdf[entry][i+1]
        except:
            resultdf2[entry][i] = resultdf[entry][i]

# конвертация конкретных значений количества в частоты
resultdf3=resultdf2.copy()
for entry in resultdf3:
    resultdf3[entry]=resultdf2[entry].div(resultdf[entry][1])*100

# запись результатов в файлы Excel
resultdf.to_excel('cumulativecounts.xlsx')
resultdf2.to_excel('specificcounts.xlsx')
resultdf3.to_excel('specificfrequencies.xlsx')
```

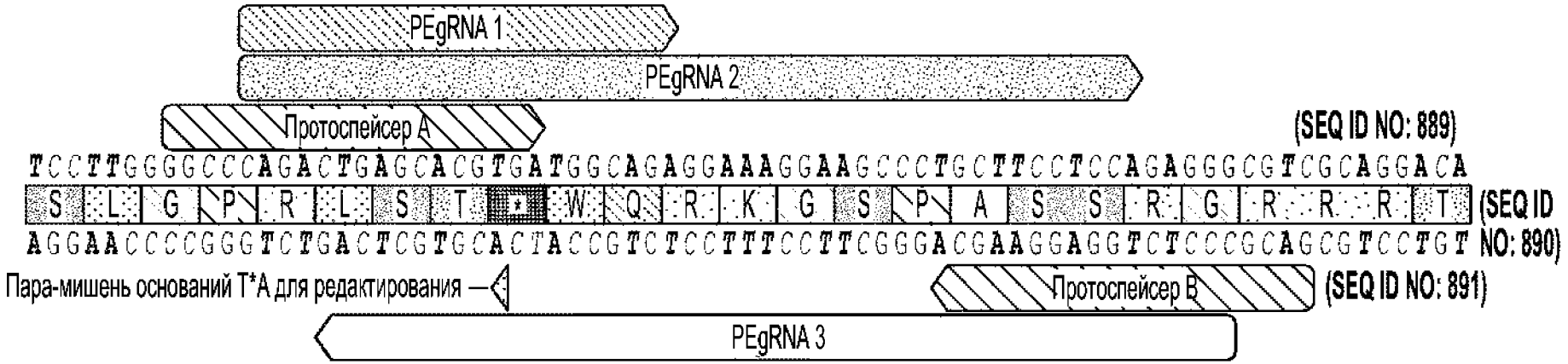
ФИГ.60B



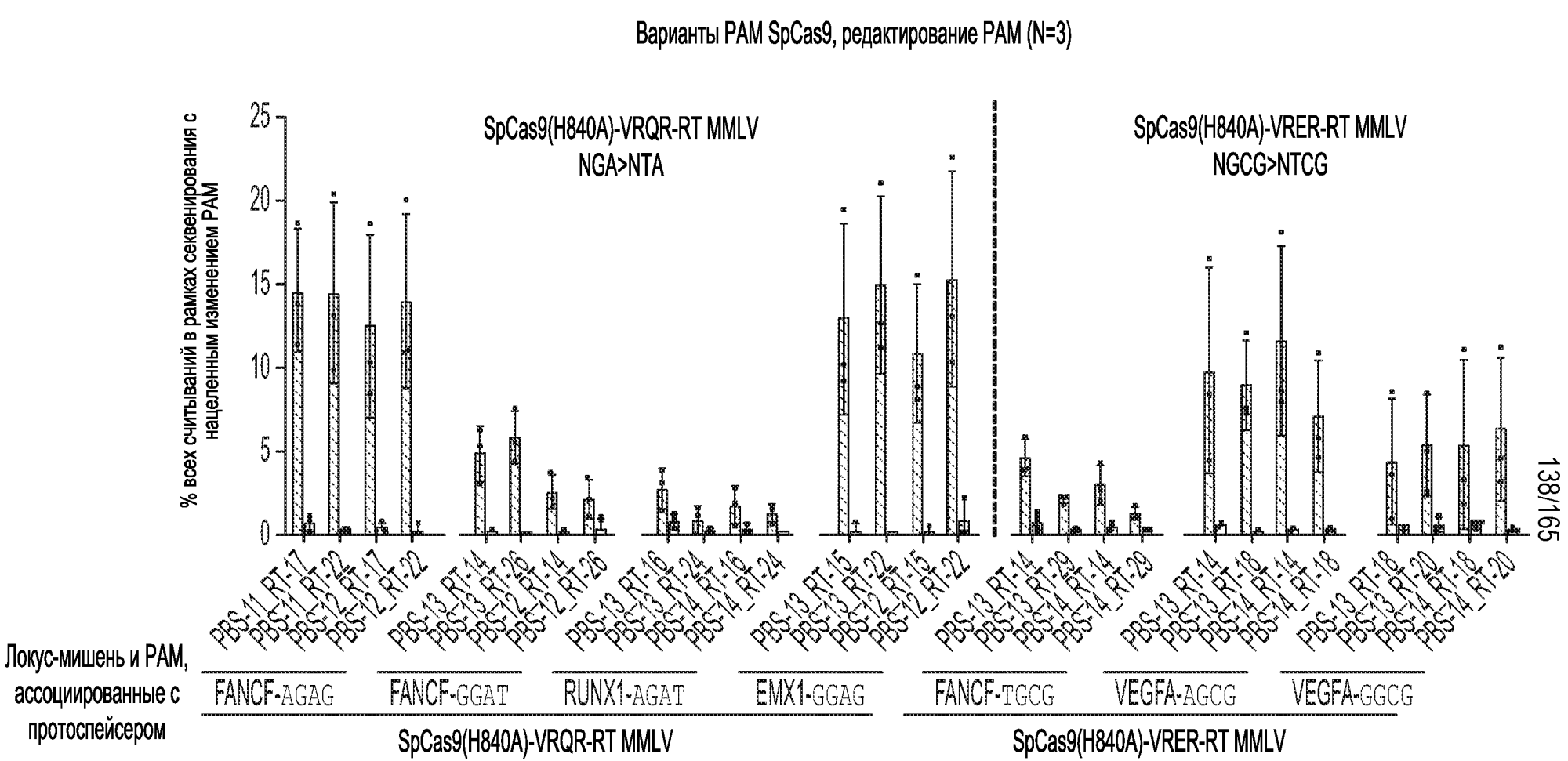
ФИГ.61



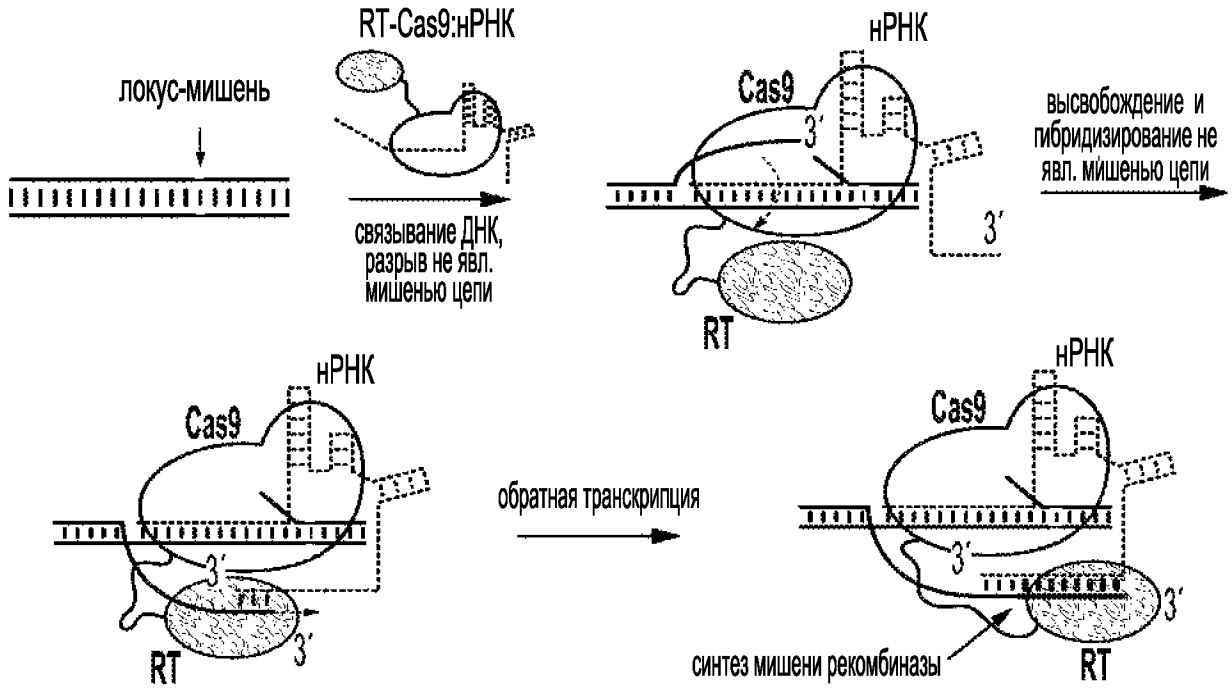
ФИГ.62А


PEgRNA 1
PEgRNA 2
Протоспейсер A
(SEQ ID NO: 889)
TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCAGAGGGCGTCGCAGGACA
S L G P R L S T * W Q R K G S P A S S R G R R R T
(SEQ ID NO: 890)
AGGAACCCCGGGTCTGACTCGTGCACTACCGTCTCCTTTCCTTCGGGACGAAGGAGGTCTCCCGCAGCGTCCTGT
Пара-мишень оснований T*A для редактирования
Протоспейсер B
(SEQ ID NO: 891)
PEgRNA 3

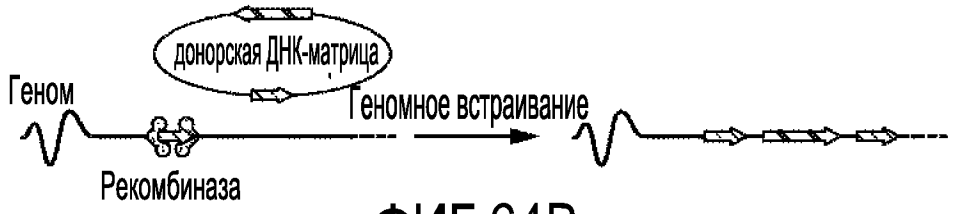
ФИГ.62B


Варианты PAM SpCas9, редактирование PAM (N=3)
% всех считываний в рамках секвенирования с нацеленным изменением PAM
SpCas9(H840A)-VRQR-RT MMLV
NGA>NTA
SpCas9(H840A)-VRER-RT MMLV
NGCG>NTCG
25
20
15
10
5
0
PBS-11_RT-17
PBS-11_RT-22
PBS-12_RT-17
PBS-12_RT-22
PBS-13_RT-14
PBS-13_RT-26
PBS-12_RT-14
PBS-12_RT-26
PBS-13_RT-16
PBS-13_RT-24
PBS-14_RT-16
PBS-14_RT-24
PBS-13_RT-15
PBS-13_RT-22
PBS-12_RT-15
PBS-12_RT-22
PBS-13_RT-14
PBS-13_RT-29
PBS-14_RT-14
PBS-14_RT-29
PBS-13_RT-14
PBS-13_RT-18
PBS-14_RT-14
PBS-14_RT-18
PBS-13_RT-18
PBS-13_RT-20
PBS-14_RT-18
PBS-14_RT-20
Локус-мишень и PAM, ассоциированные с протоспейсером
FANCF-AGAG
FANCF-GGAT
RUNX1-AGAT
EMX1-GGAG
FANCF-TGCG
VEGFA-AGCG
VEGFA-GGCG
SpCas9(H840A)-VRQR-RT MMLV
SpCas9(H840A)-VRER-RT MMLV
Длина участка связывания праймера PEgRNA (PBS), длина RT матрицы,
и использованная система праймерного редактирования
Процент G>T изменений относительно всех считываний
Индели из всех считываний

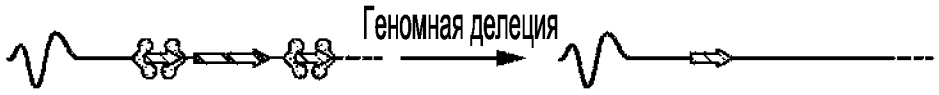
ФИГ.63



ФИГ.64A



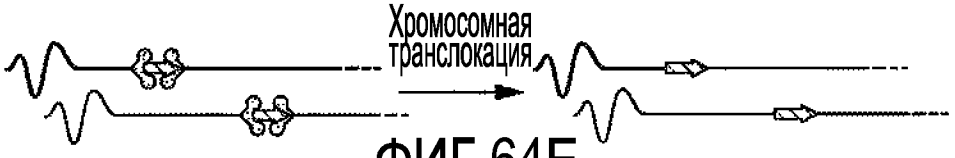
ФИГ.64B



ФИГ.64C



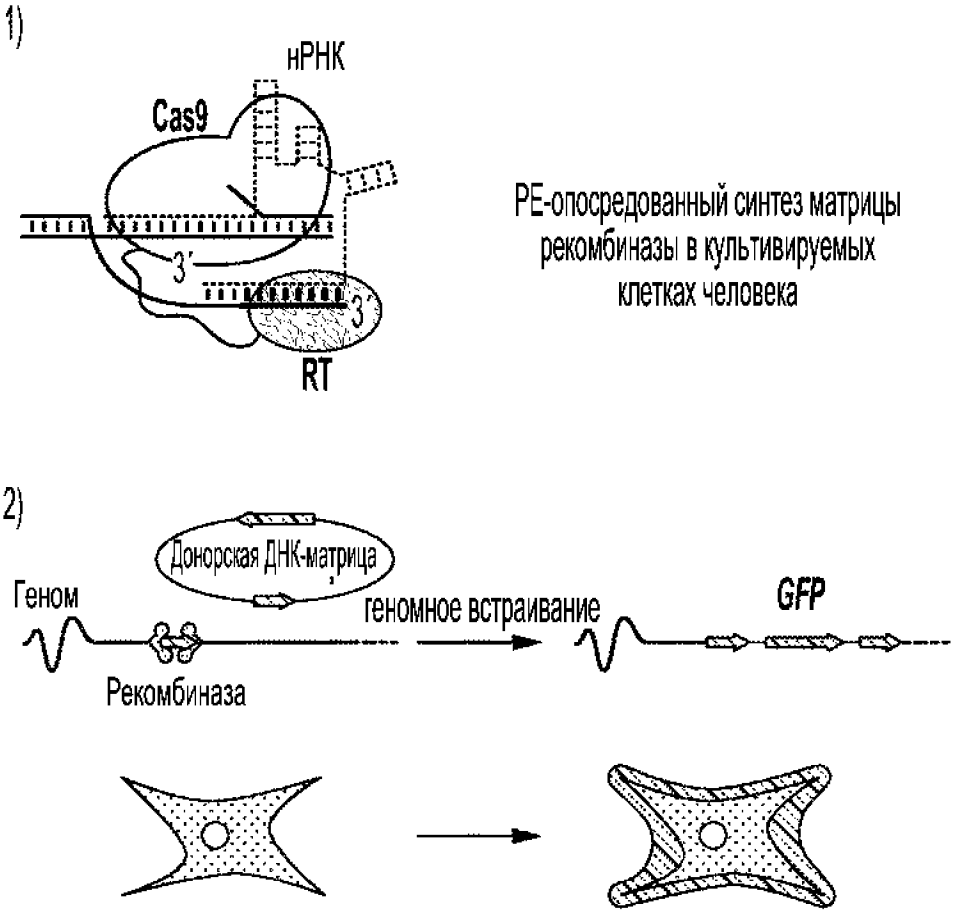
ФИГ.64D



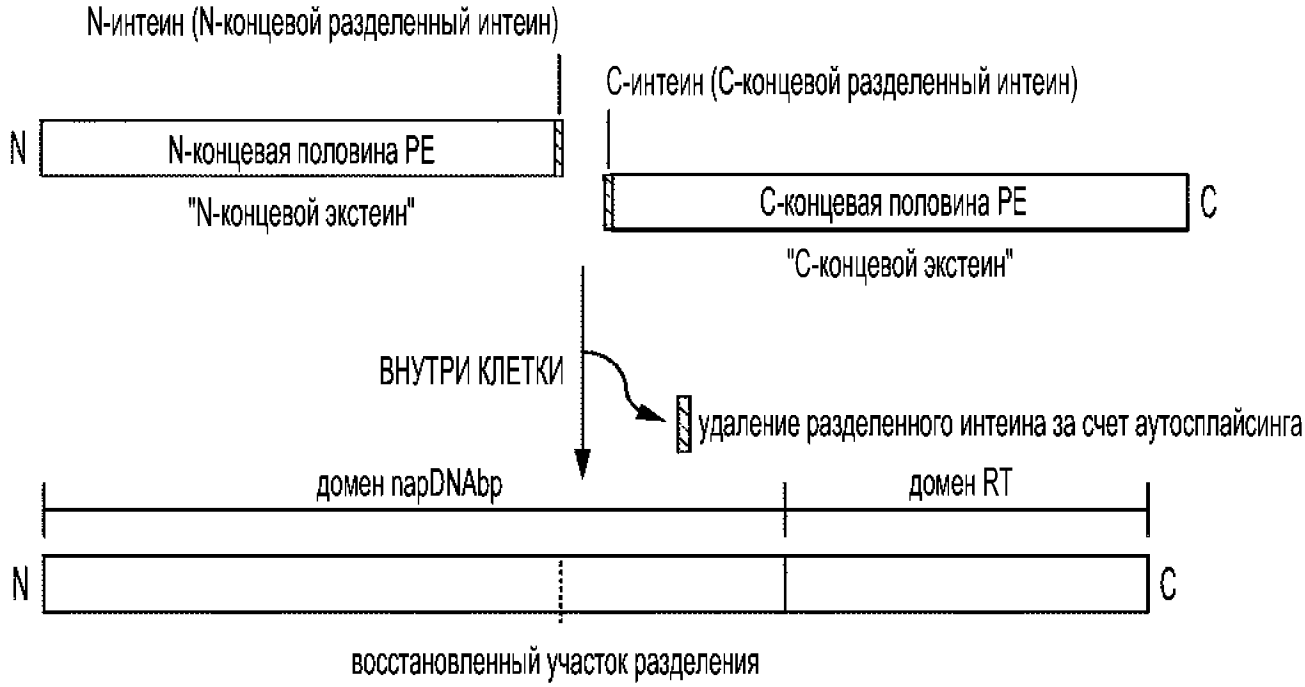
ФИГ.64E



ФИГ.64F

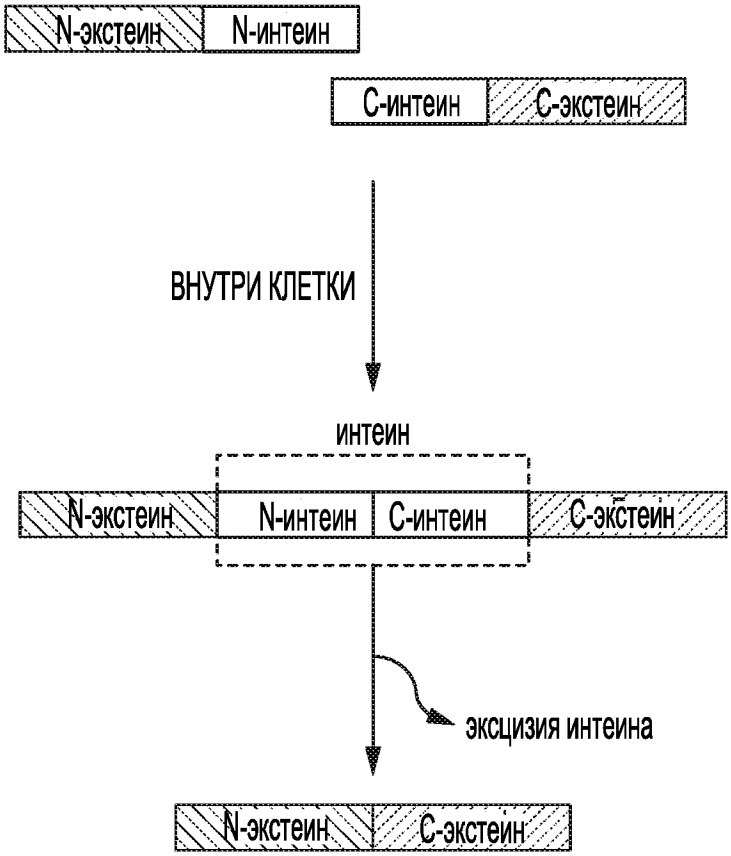

1)
нРНК
Cas9
3'
3'
RT
PE-опосредованный синтез матрицы рекомбиназы в культивируемых клетках человека
2)
Донорская ДНК-матрица
Геном
геномное встраивание
GFP
Рекомбиназа

ФИГ.65



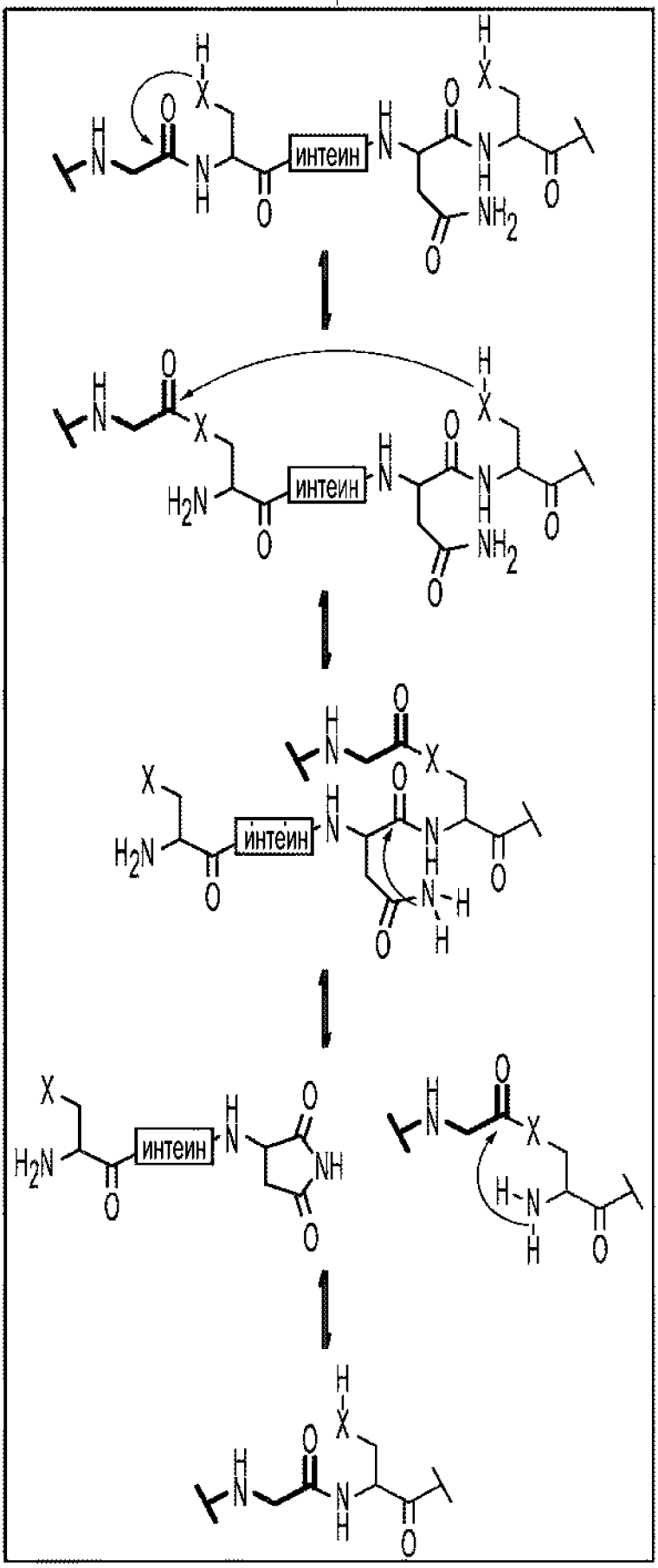
Праймерный редактор

ФИГ.66

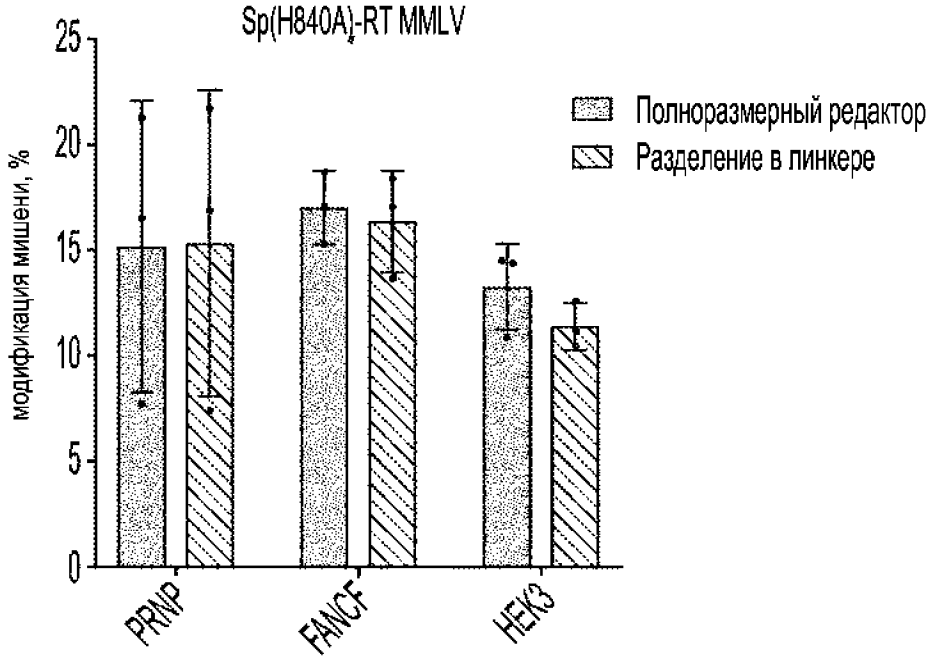

N-экстеин
N-интеин
C-интеин
C-экстеин
ВНУТРИ КЛЕТКИ
интеин
N-экстеин
N-интеин
C-интеин
C-экстеин
эксцизия интеина
N-экстеин
C-экстеин

ФИГ.67А

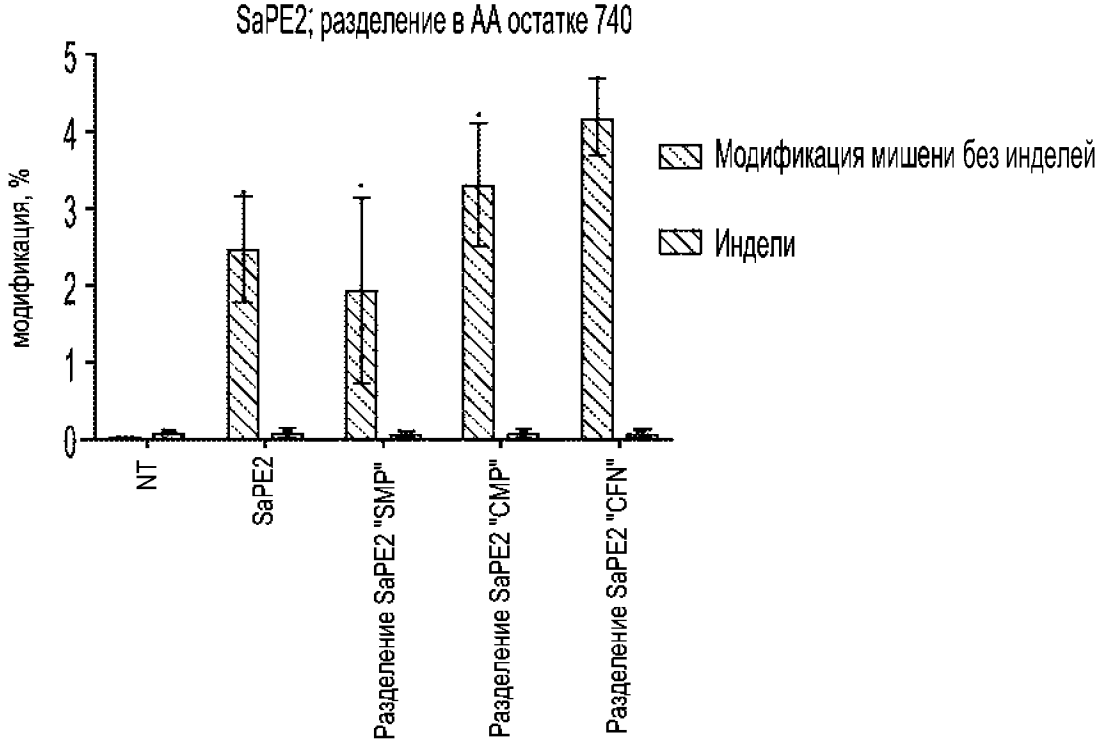
Механизм эксцизии интеина



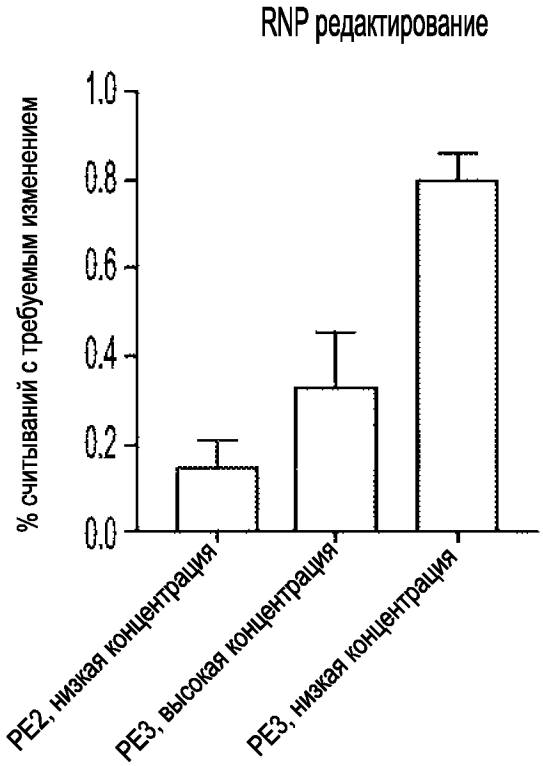
ФИГ.67В



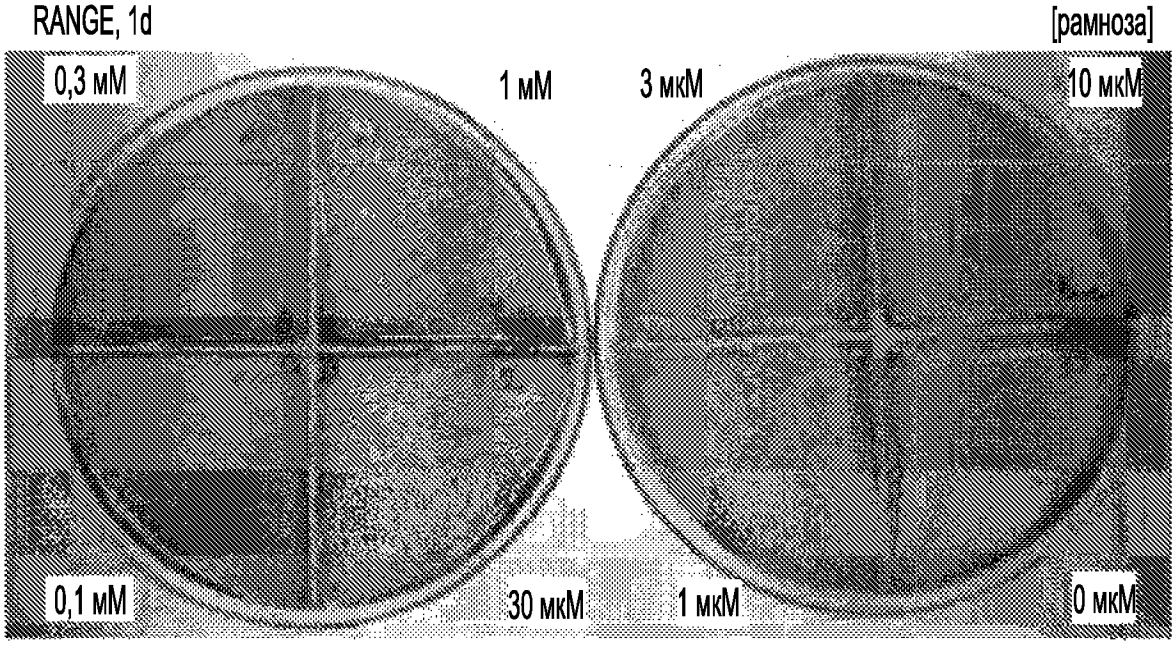
ФИГ.68A



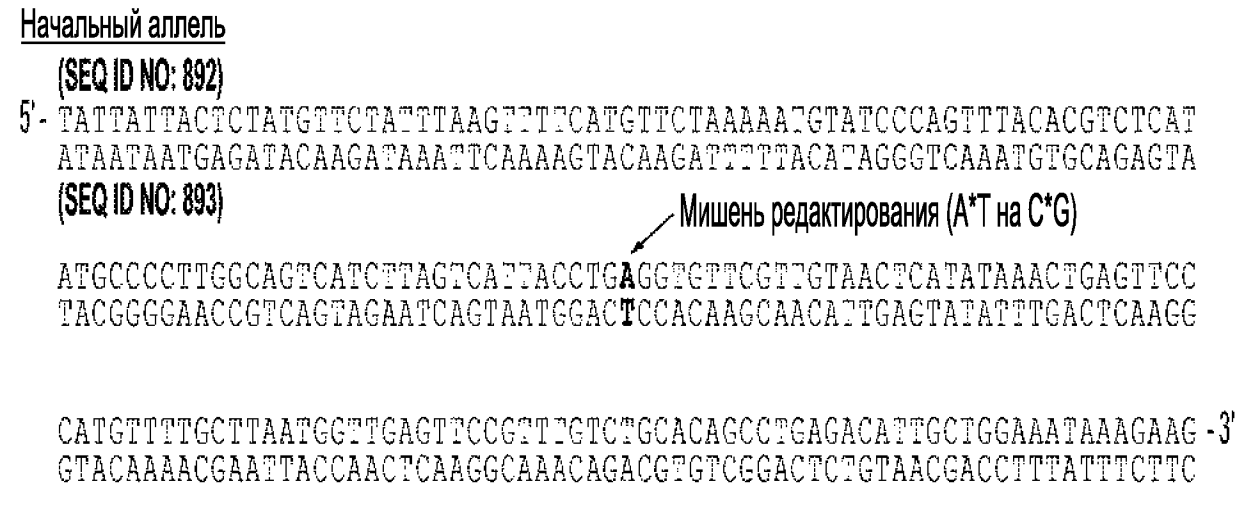
ФИГ.68B


RNP редактирование
% считываний с требуемым изменением
1.0
0.8
0.6
0.4
0.2
0.0
PE2, низкая концентрация
PE3, высокая концентрация
PE3, низкая концентрация

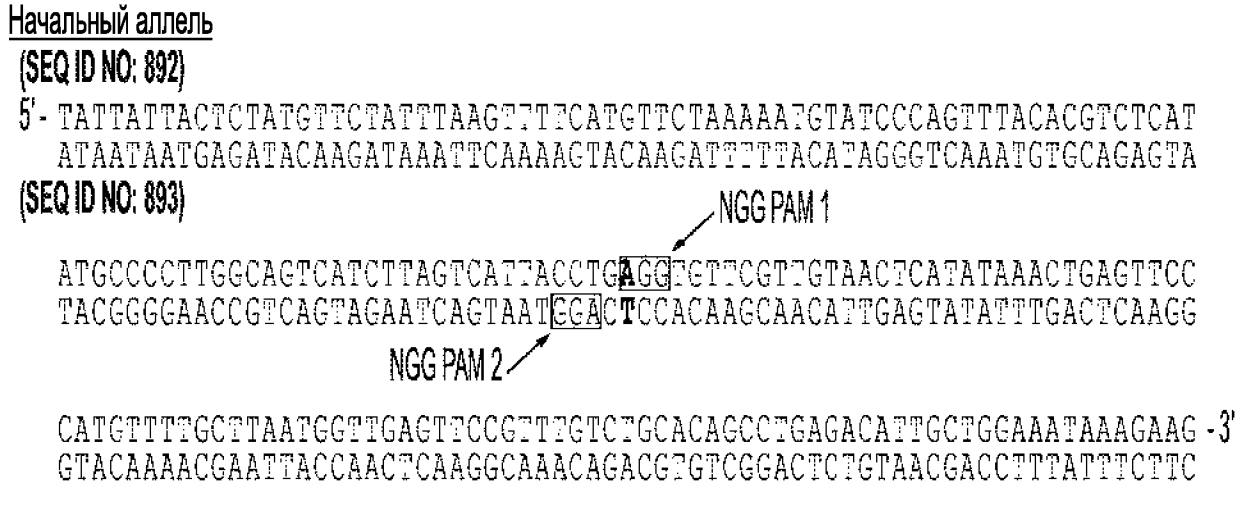
ФИГ.68C



ФИГ.69



ФИГ.70A



ФИГ.70B



ФИГ.70C

Начальный аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

Спейсер

ATGCCCCTTGGCAGTCATCTTAGTCATTACCTGAGCGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACTCCACAAGCAACATTGAGTATATTTGACTCAAGG

CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG -3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

Спейсер= GTCATCTTAGTCATTACCTG

ФИГ.70D

Начальный аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

комплемент PBS

ATGCCCCTTGGCAGTCATCTTAGTCATTACCTGAGCGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACTCCACAAGCAACATTGAGTATATTTGACTCAAGG

CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG -3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

PBS= GTAATGACTAAGATG

ФИГ.70E

Требуемый аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

комплемент RT матрицы

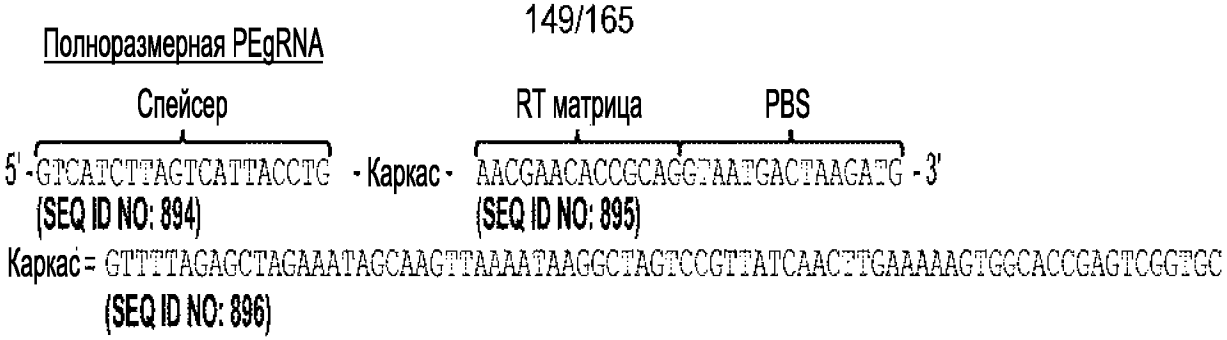
ATGCCCCTTGGCAGTCATCTTAGTCATTACCTGCGCGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACGCCACAAGCAACATTGAGTATATTTGACTCAAGG

Изменение

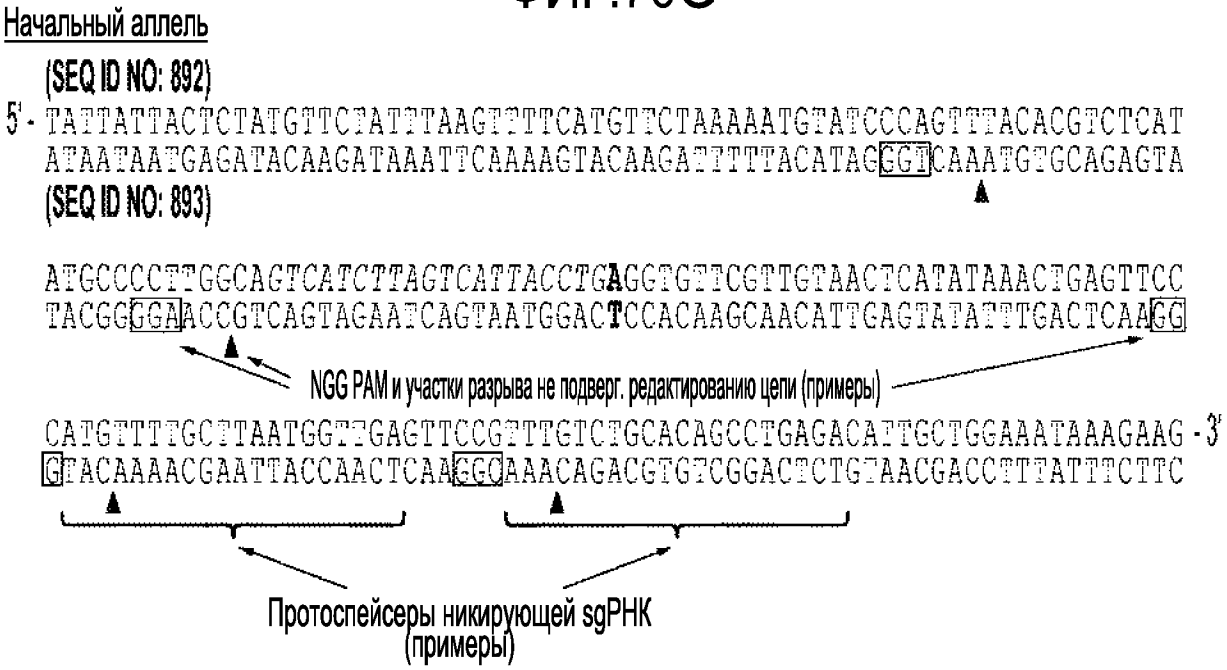
CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG -3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

RT матрица = AACGAACACCGCAG

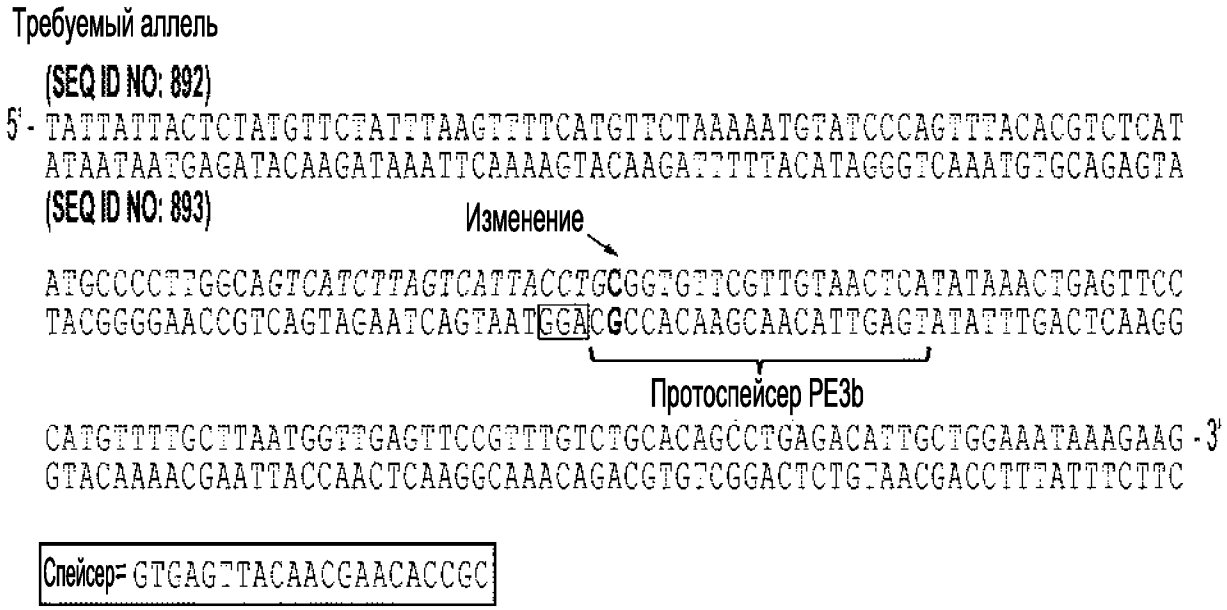
ФИГ.70F



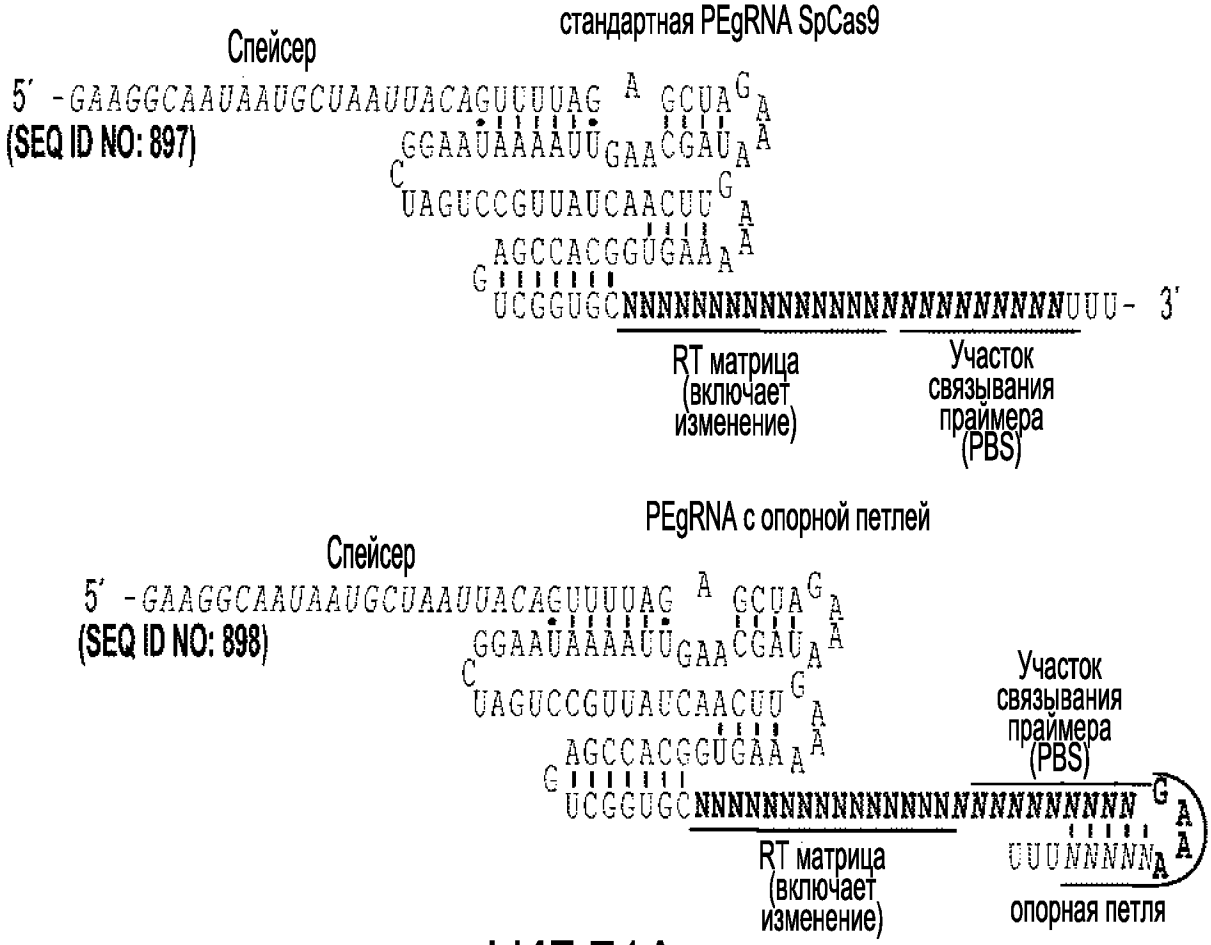
ФИГ.70G



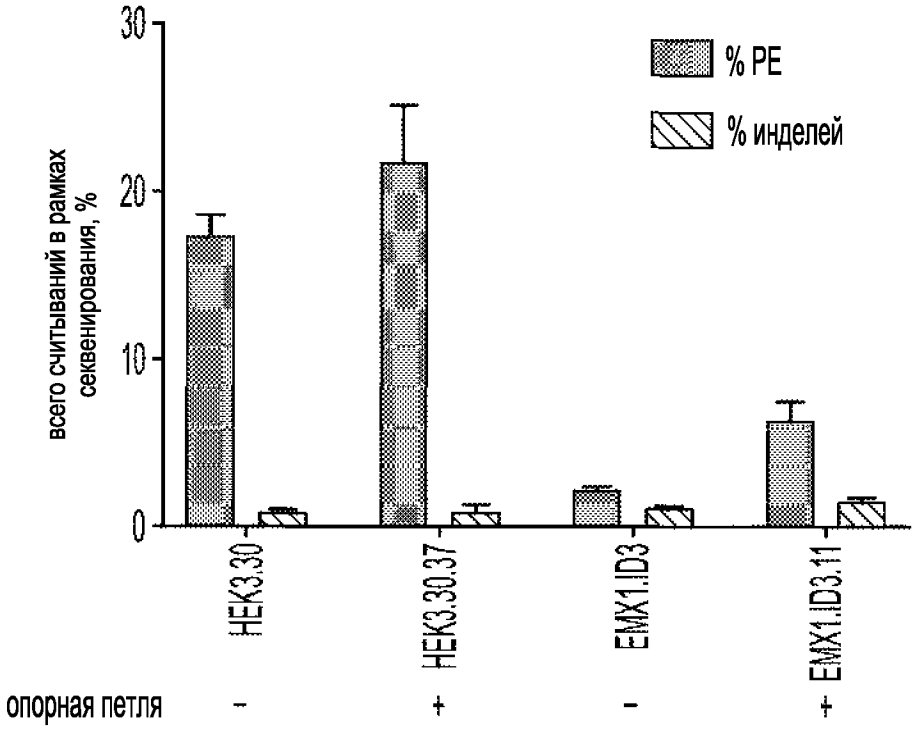
ФИГ.70H



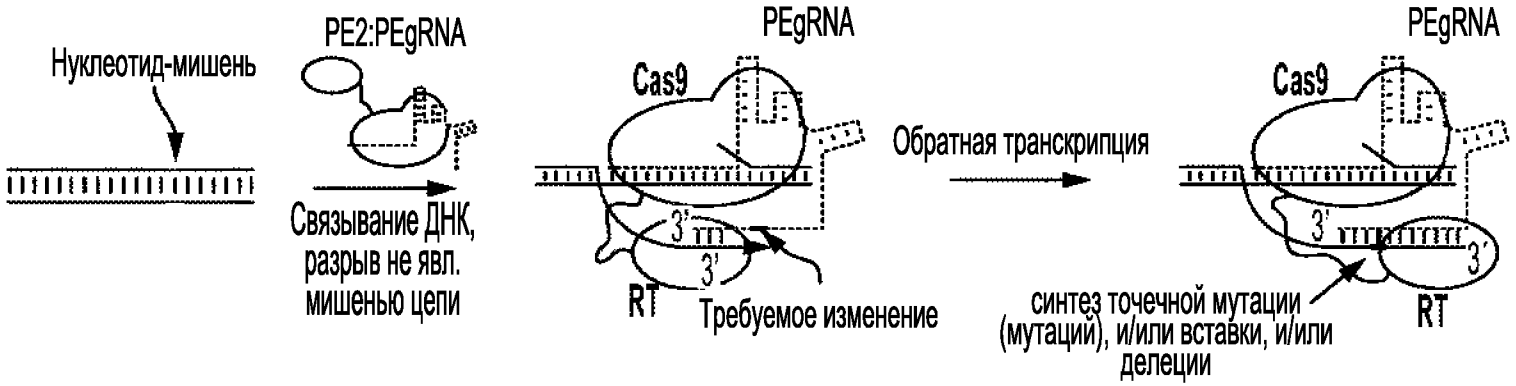
ФИГ.70I



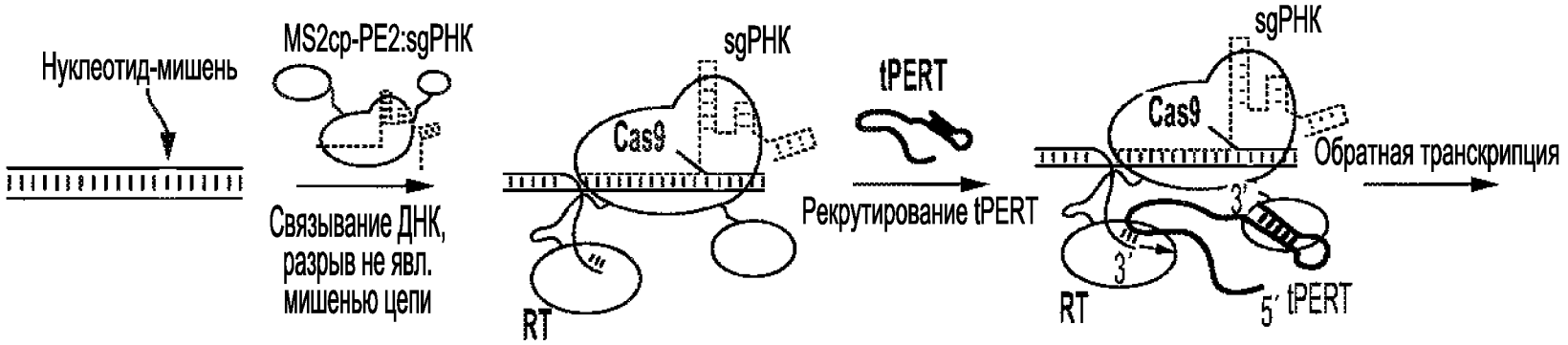
ФИГ.71A



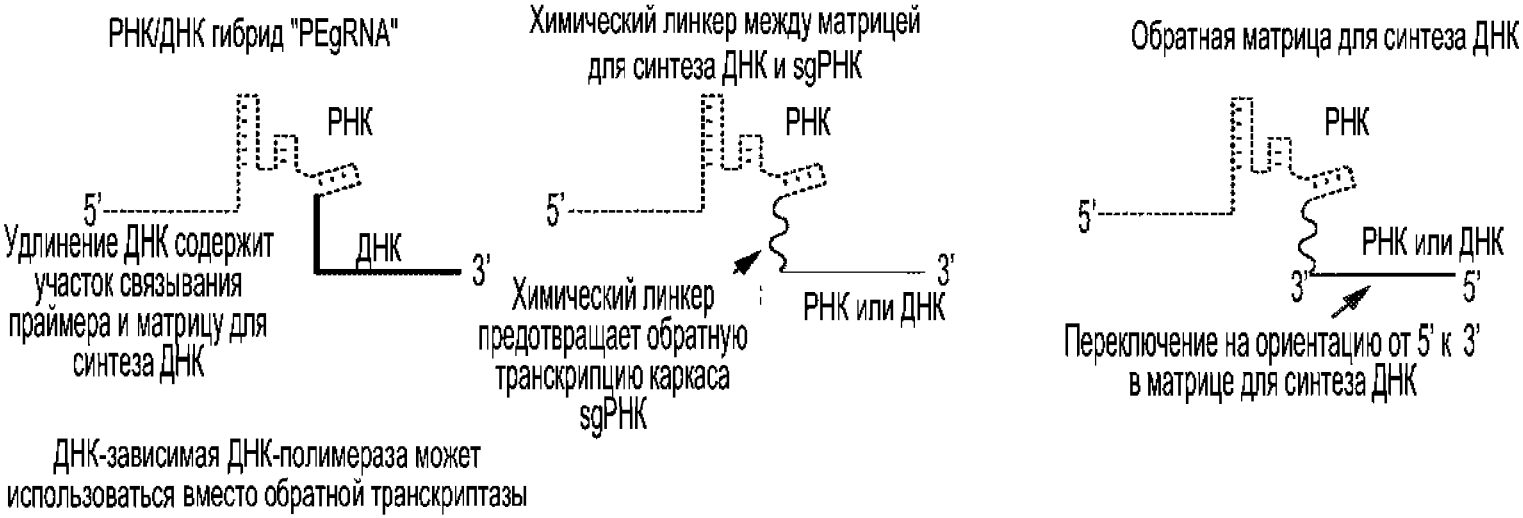
ФИГ.71B



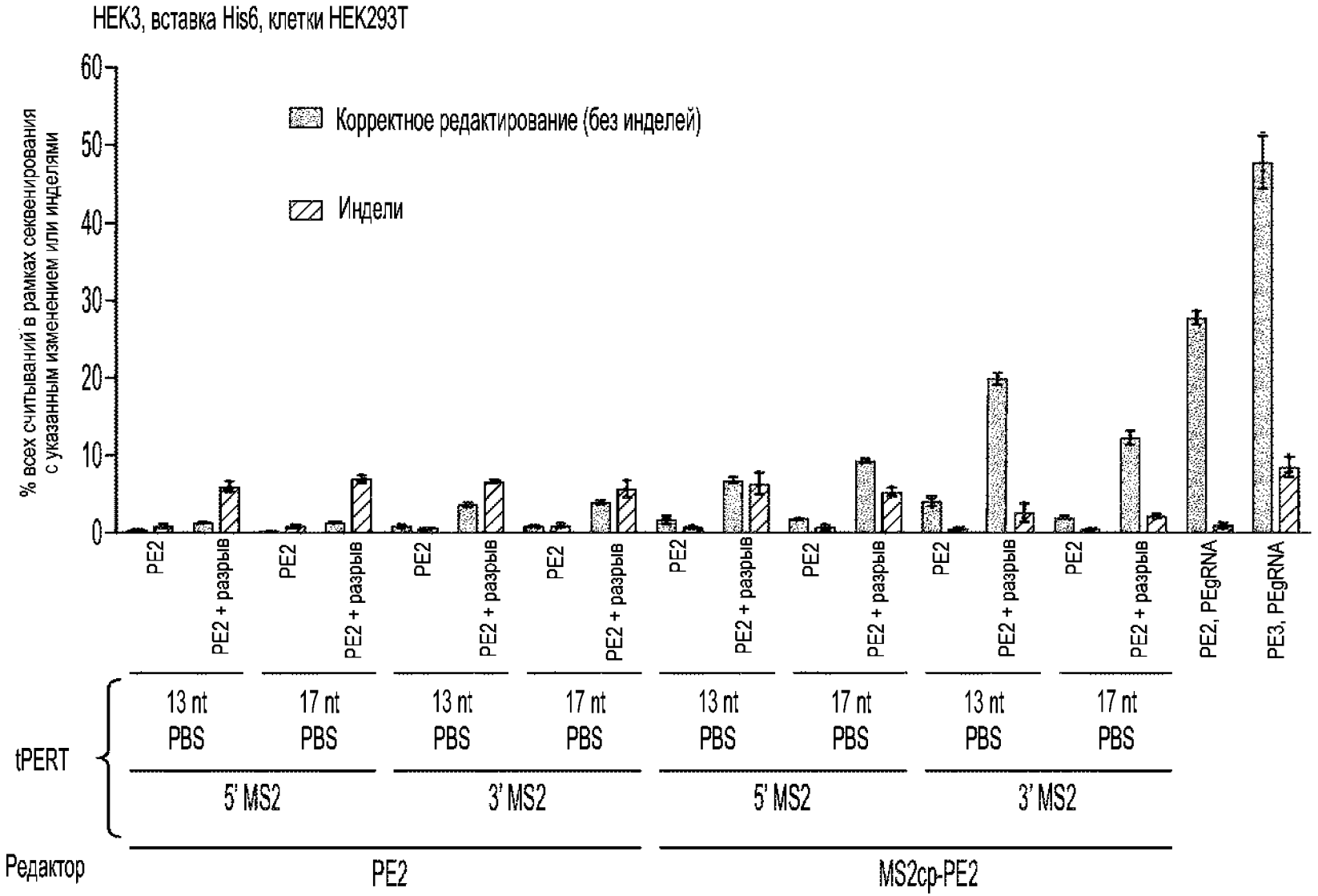
ФИГ.72A



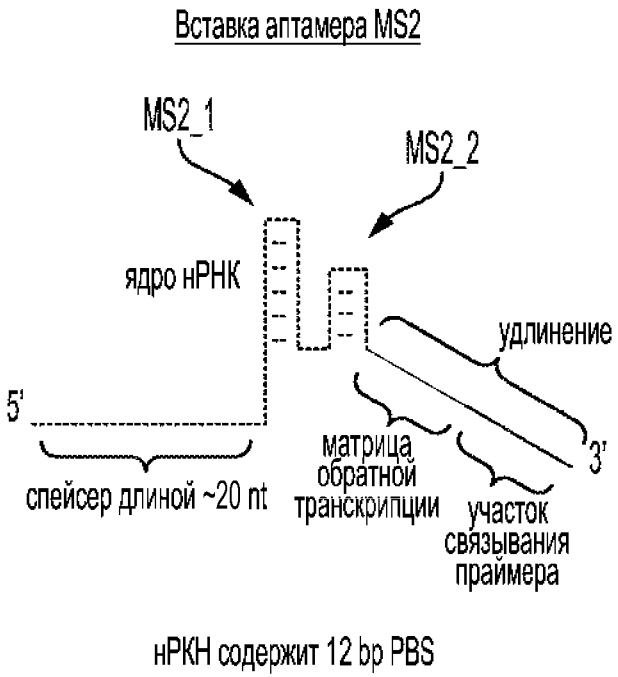
ФИГ.72B

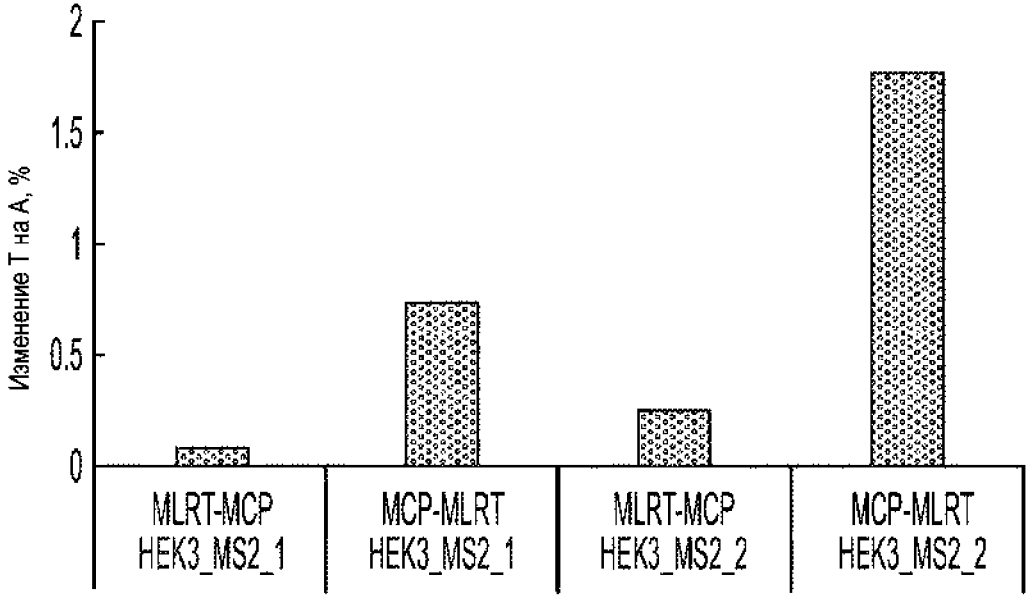

РНК/ДНК гибрид "PEgRNA"
РНК
5'
Удлинение ДНК содержит участок связывания праймера и матрицу для синтеза ДНК
ДНК
3'
ДНК-зависимая ДНК-полимераза может использоваться вместо обратной транскриптазы
Химический линкер между матрицей для синтеза ДНК и sgРНК
РНК
5'
Химический линкер предотвращает обратную транскрипцию каркаса sgРНК
3'
РНК или ДНК
Обратная матрица для синтеза ДНК
РНК
5'
РНК или ДНК
3'
5'
Переключение на ориентацию от 5' к 3' в матрице для синтеза ДНК

ФИГ.72C

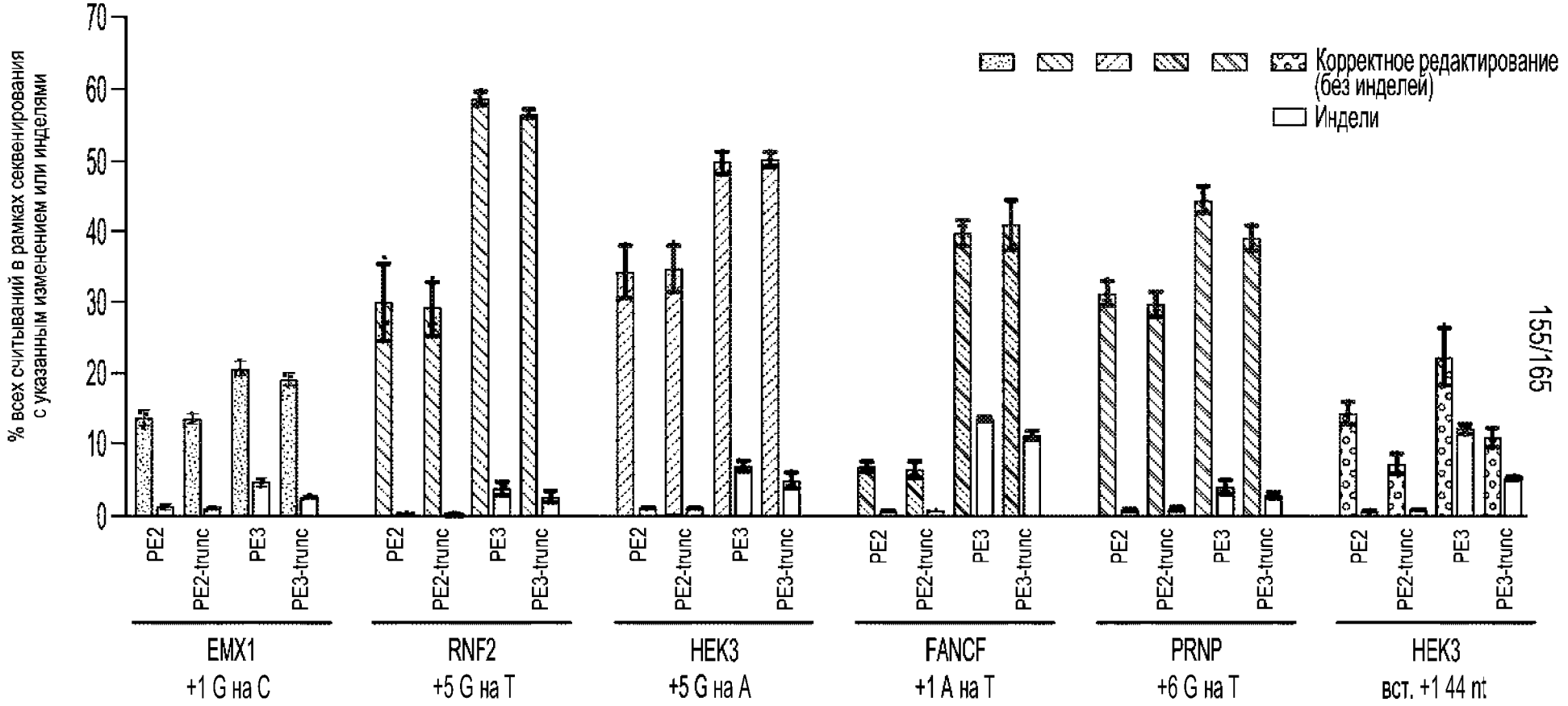

HEK3, вставка His6, клетки HEK293T
% всех считываний в рамках секвенирования с указанным изменением или инделями
Корректное редактирование (без инделей)
Индели
0
10
20
30
40
50
60
PE2
PE2 + разрыв
PE2, PEgRNA
PE3, PEgRNA
tPERT
13 nt PBS
17 nt PBS
5' MS2
3' MS2
Редактор
PE2
MS2cp-PE2

ФИГ.73

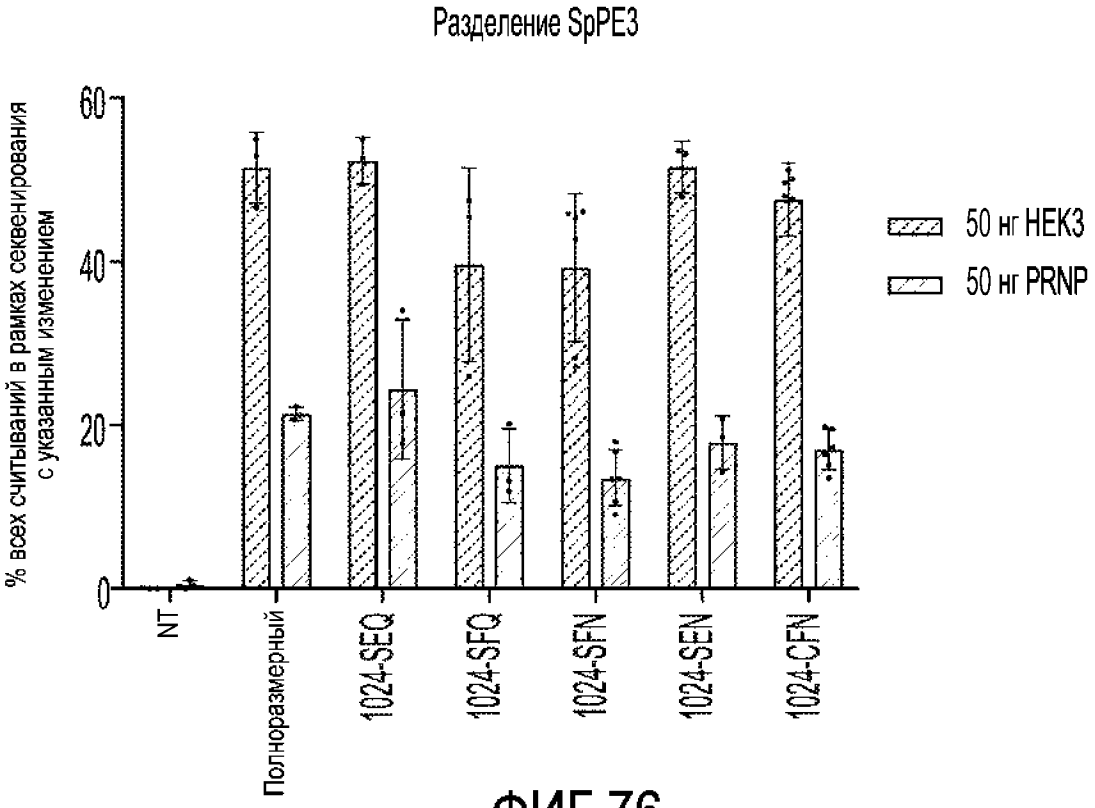

Вставка аптамера MS2
MS2_1
MS2_2
ядро нРНК
удлинение
5'
3'
матрица обратной транскрипции
участок связывания праймера
спейсер длиной ~20 nt
нРКН содержит 12 bp PBS


Изменение T на A, %
2
1.5
1
0.5
0
MLRT-MCP HEK3_MS2_1
MCP-MLRT HEK3_MS2_1
MLRT-MCP HEK3_MS2_2
MCP-MLRT HEK3_MS2_2

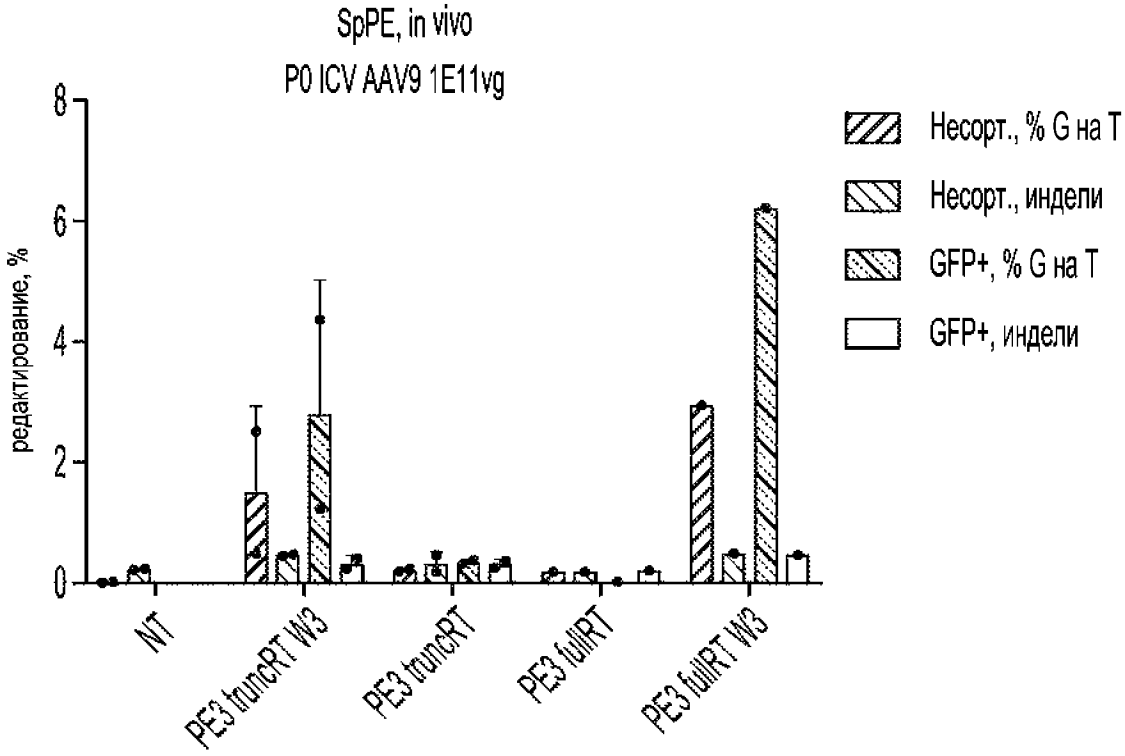
ФИГ.74


% всех считываний в рамках секвенирования с указанным изменением или инделями
70
60
50
40
30
20
10
0
Корректное редактирование (без инделей)
Индели
PE2
PE2-trunc
PE3
PE3-trunc
EMX1
+1 G на C
RNF2
+5 G на T
HEK3
+5 G на A
FANCF
+1 A на T
PRNP
+6 G на T
HEK3
вст. +1 44 nt

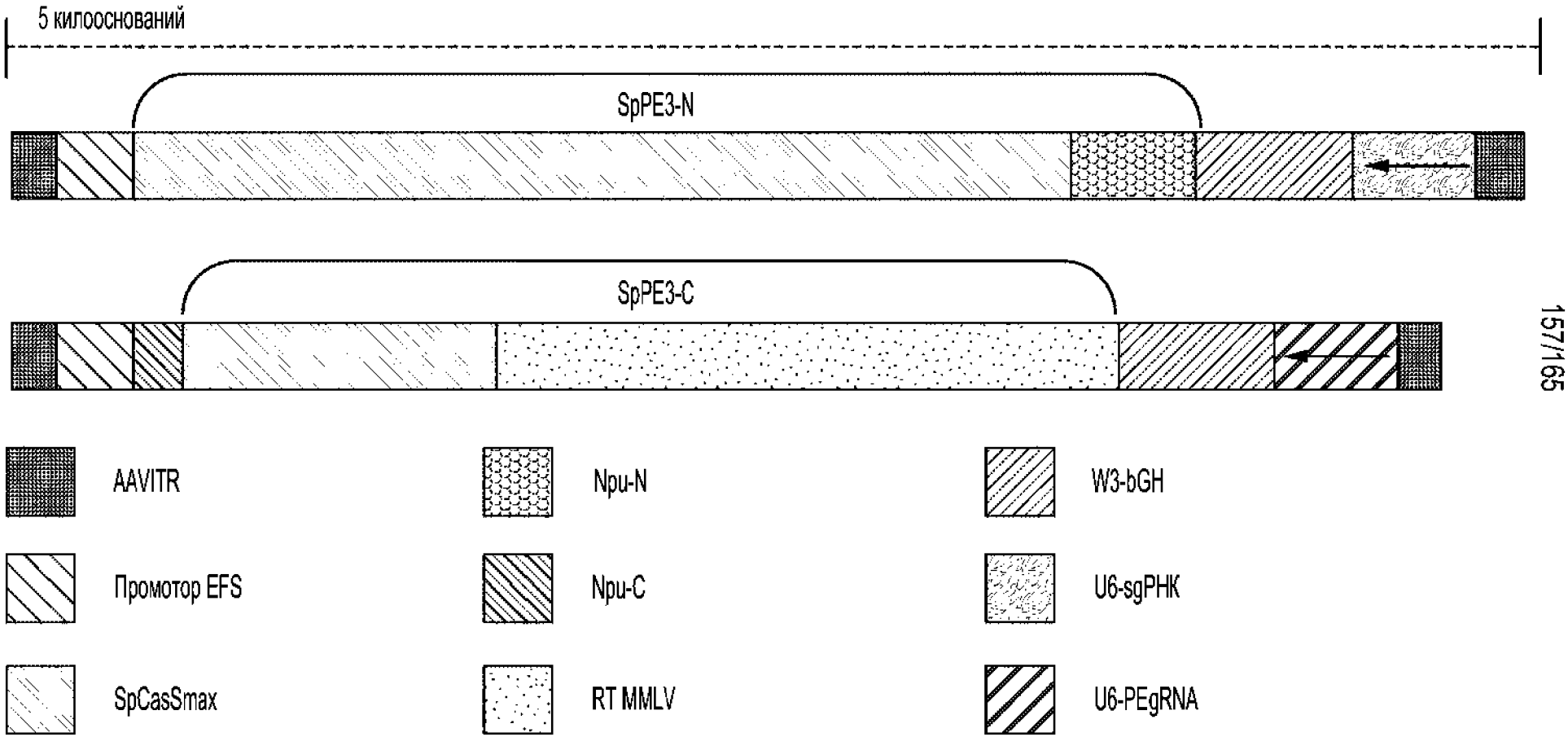
ФИГ.75


Разделение SpPE3
% всех считываний в рамках секвенирования с указанным изменением
60
40
20
0
50 нг HEK3
50 нг PRNP
NT
Полноразмерный
1024-SEQ
1024-SFQ
1024-SFN
1024-SEN
1024-CFN

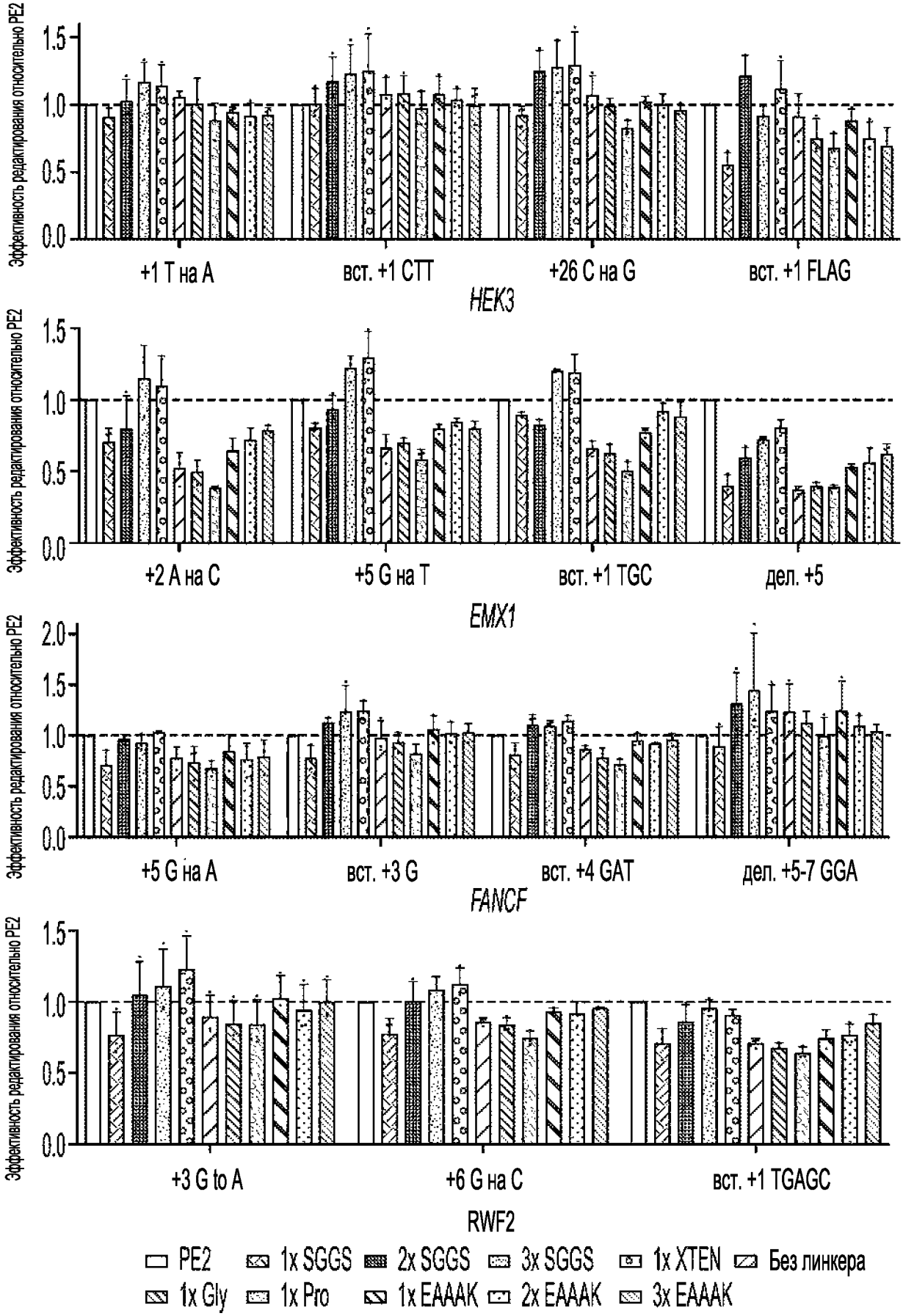
ФИГ.76


SpPE, in vivo
P0 ICV AAV9 1E11vg
редактирование, %
8
6
4
2
0
Несорт., % G на T
Несорт., индели
GFP+, % G на T
GFP+, индели
NT
PE3 truncRT W3
PE3 truncRT
PE3 fullRT
PE3 fullRT W3

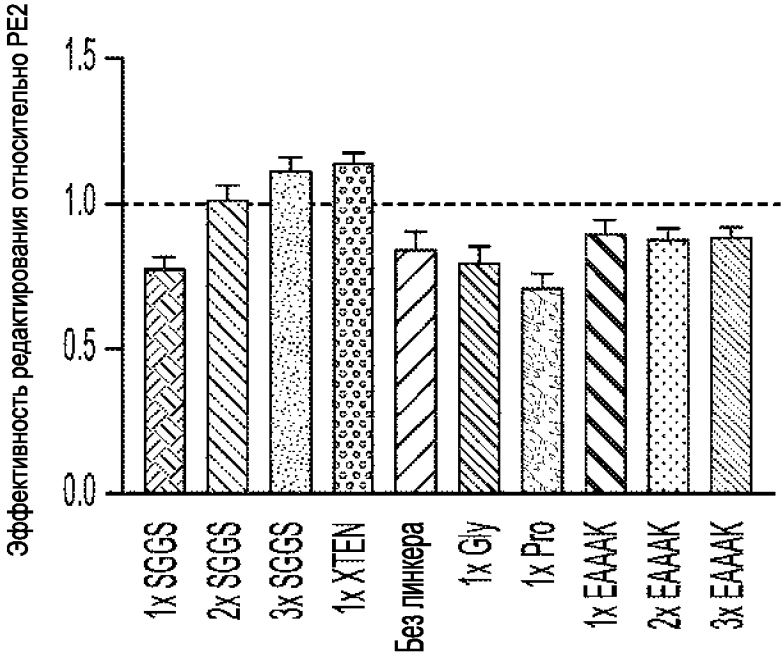
ФИГ.77


5 килооснований
SpPE3-N
SpPE3-C
AAVITR
Промотор EFS
SpCasSmax
Npu-N
Npu-C
RT MMLV
W3-bGH
U6-sgPHK
U6-PEgRNA

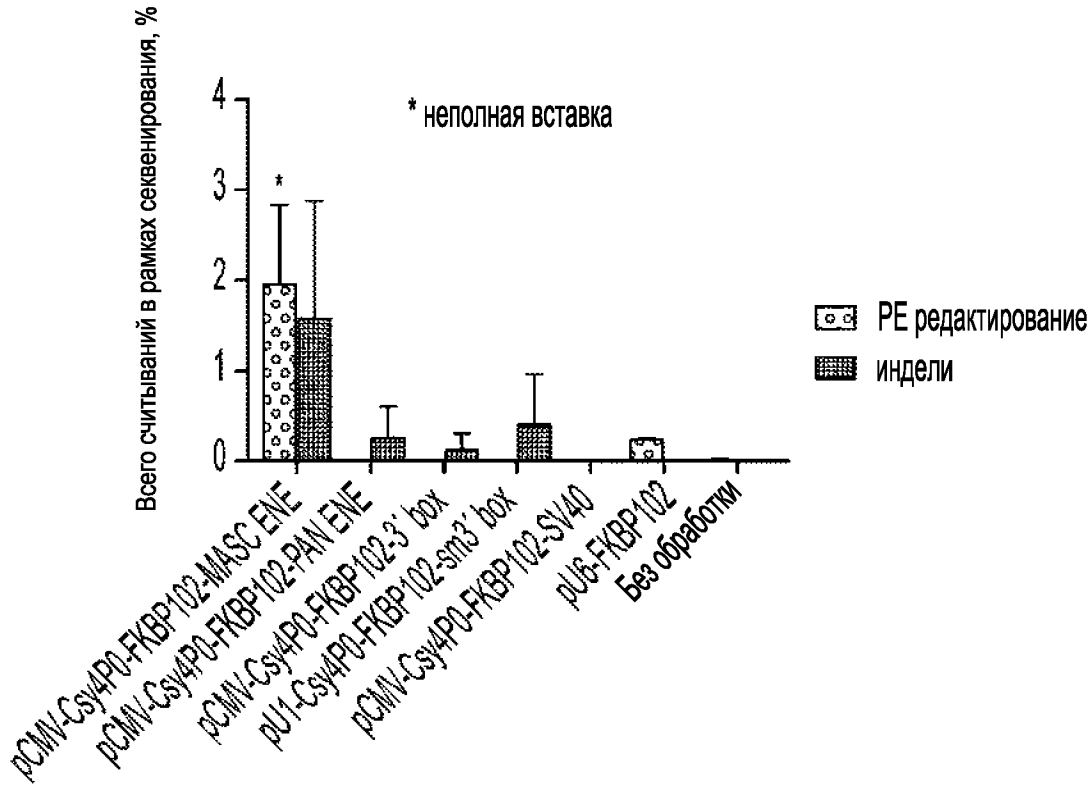
ФИГ.78


Эффективность редактирования относительно PE2
1.5
1.0
0.5
0.0
+1 T на A
вст. +1 CTT
+26 C на G
вст. +1 FLAG
HEK3
+2 A на C
+5 G на T
вст. +1 TGC
дел. +5
EMX1
2.0
+5 G на A
вст. +3 G
вст. +4 GAT
дел. +5-7 GGA
FANCF
+3 G to A
+6 G на C
вст. +1 TGAGC
RWF2
PE2
1x SGGS
2x SGGS
3x SGGS
1x XTEN
Без линкера
1x Gly
1x Pro
1x EAAAK
2x EAAAK
3x EAAAK

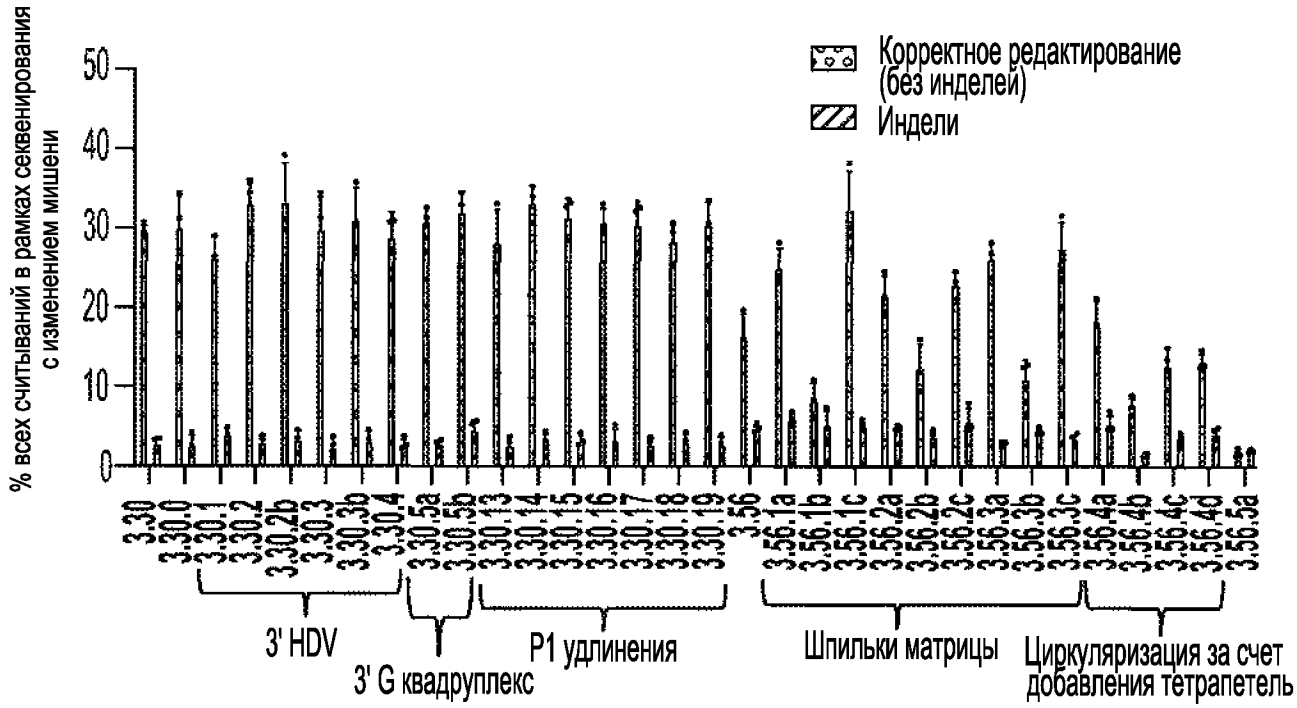
ФИГ.79


Эффективность редактирования относительно PE2
1.5
1.0
0.5
0.0
1x SGGS
2x SGGS
3x SGGS
1x XTEN
Без линкера
1x Gly
1x Pro
1x EAAAK
2x EAAAK
3x EAAAK

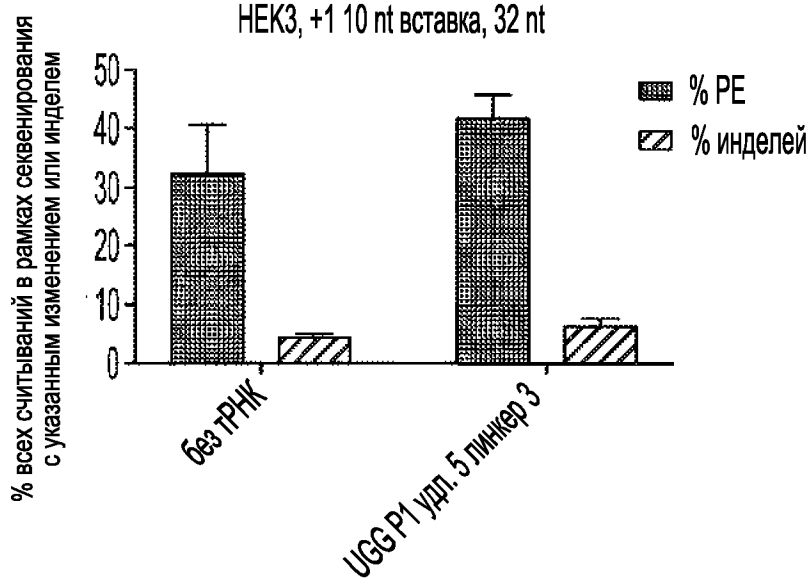
ФИГ.80


Всего считываний в рамках секвенирования, %
* неполная вставка
4
3
2
1
0
PE редактирование
индели
pCMV-Csy4P0-FKBP102-MASC ENE
pCMV-Csy4P0-FKBP102-PAN ENE
pCMV-Csy4P0-FKBP102-3' box
pU1-Csy4P0-FKBP102-sm3' box
pCMV-Csy4P0-FKBP102-SV40
pU6-FKBP102
Без обработки

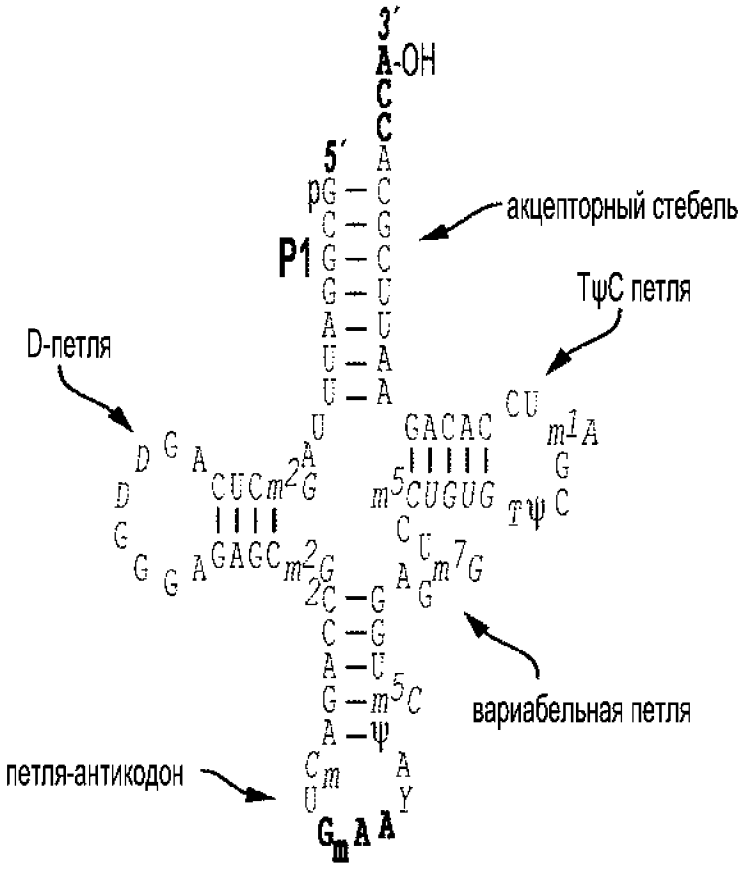
ФИГ.81


% всех считываний в рамках секвенирования с изменением мишени
Корректное редактирование (без инделей)
Индели
3' HDV
3' G квадруплекс
P1 удлинения
Шпильки матрицы
Циркуляризация за счет добавления тетрапетель

ФИГ.82


HEK3, +1 10 nt вставка, 32 nt
% всех считываний в рамках секвенирования с указанным изменением или инделем
% PE
% инделей
без тРНК
UGG P1 удл. 5 линкер 3'

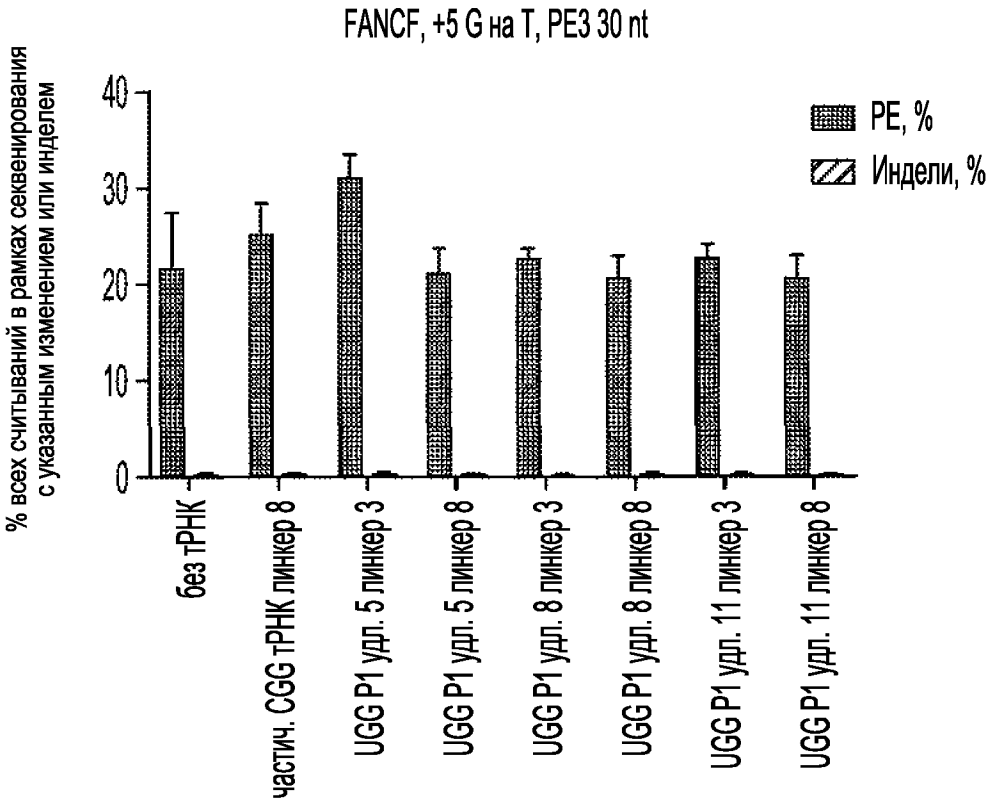
ФИГ.83



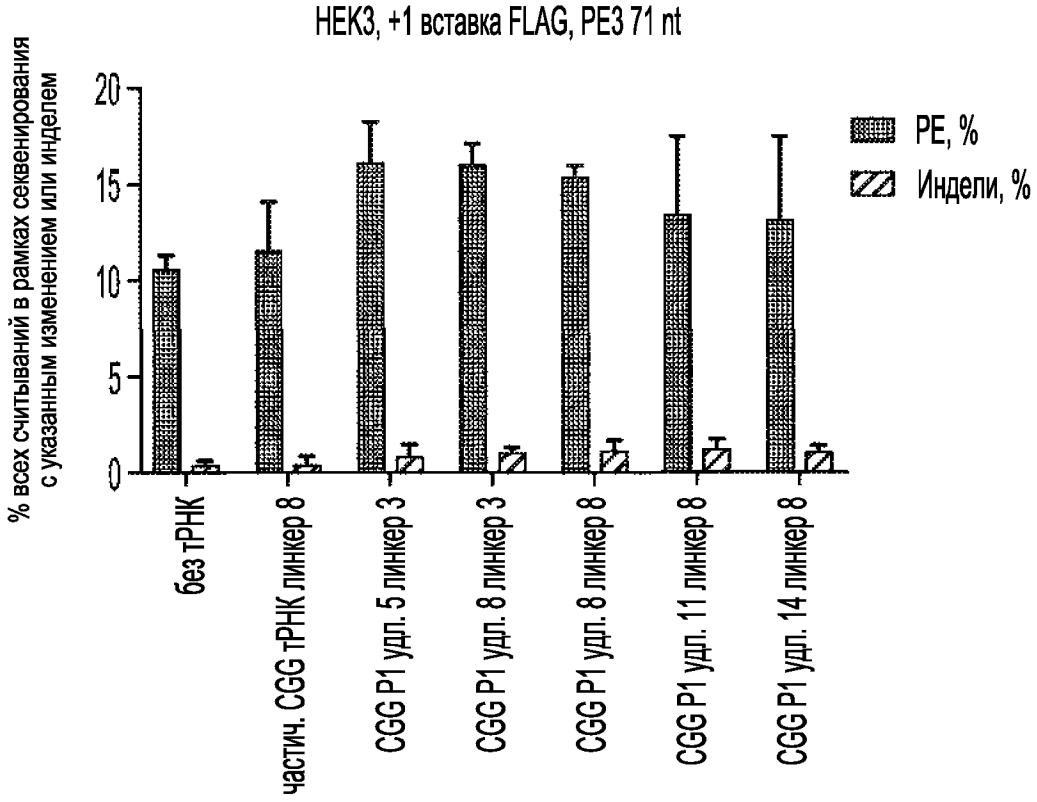
ФИГ.84A



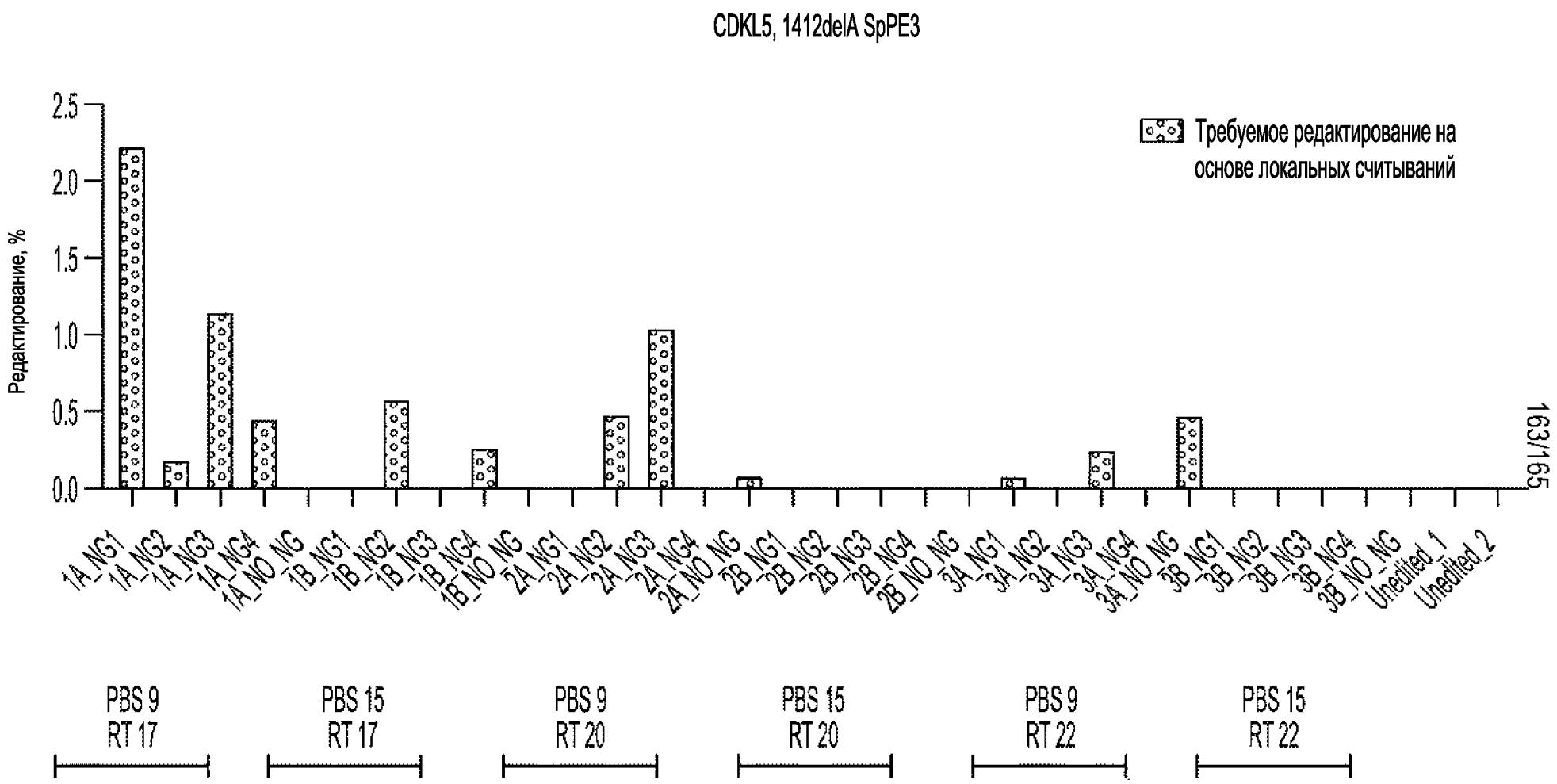
ФИГ.84B


FANCF, +5 G на T, PE3 30 nt
% всех считываний в рамках секвенирования с указанным изменением или инделем
40
30
20
10
0
PE, %
Индели, %
без тРНК
частич. CGG тРНК линкер 8
UGG P1 удл. 5 линкер 3
UGG P1 удл. 5 линкер 8
UGG P1 удл. 8 линкер 3
UGG P1 удл. 8 линкер 8
UGG P1 удл. 11 линкер 3
UGG P1 удл. 11 линкер 8

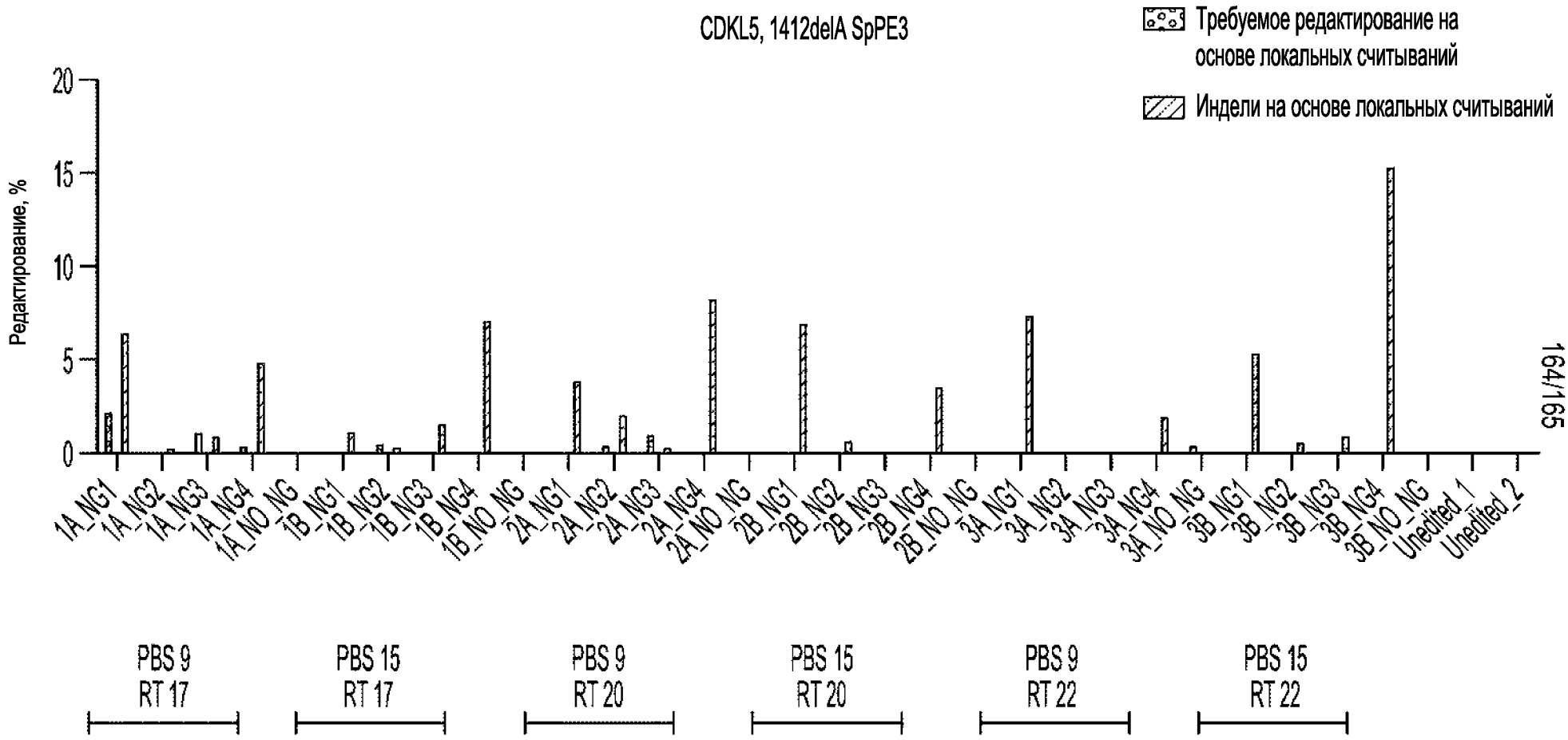
ФИГ.85


HEK3, +1 вставка FLAG, PE3 71 nt
% всех считываний в рамках секвенирования с указанным изменением или инделем
20
15
10
5
0
PE, %
Индели, %
без тРНК
частич. CGG тРНК линкер 8
CGG P1 удл. 5 линкер 3
CGG P1 удл. 8 линкер 3
CGG P1 удл. 8 линкер 8
CGG P1 удл. 11 линкер 8
CGG P1 удл. 14 линкер 8

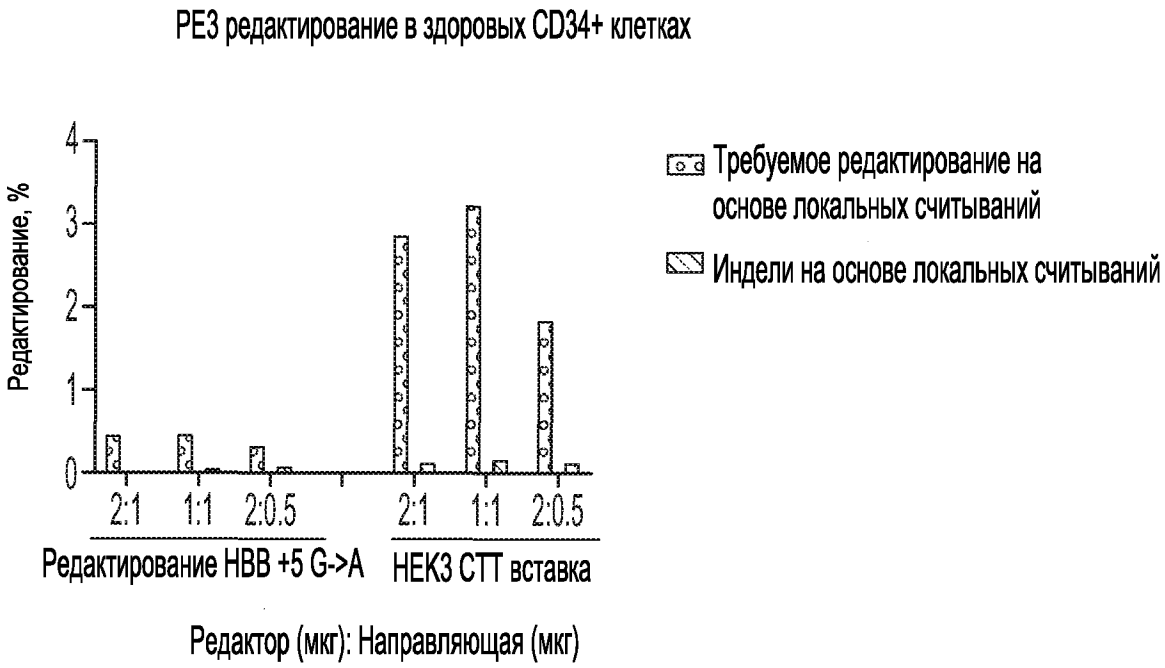
ФИГ.86


CDKL5, 1412delA SpPE3
Требуемое редактирование на основе локальных считываний
Редактирование, %
2.5
2.0
1.5
1.0
0.5
0.0
163/165
1A_NG1
1A_NG2
1A_NG3
1A_NG4
1A_NO_NG
1B_NG1
1B_NG2
1B_NG3
1B_NG4
1B_NO_NG
2A_NG1
2A_NG2
2A_NG3
2A_NG4
2A_NO_NG
2B_NG1
2B_NG2
2B_NG3
2B_NG4
2B_NO_NG
3A_NG1
3A_NG2
3A_NG3
3A_NG4
3A_NO_NG
3B_NG1
3B_NG2
3B_NG3
3B_NG4
3B_NO_NG
Unedited_1
Unedited_2
PBS 9
RT 17
PBS 15
RT 17
PBS 9
RT 20
PBS 15
RT 20
PBS 9
RT 22
PBS 15
RT 22

ФИГ.87


CDKL5, 1412delA SpPE3
Требуемое редактирование на основе локальных считываний
Индели на основе локальных считываний
Редактирование, %
20
15
10
5
0
1A_NG1
1A_NG2
1A_NG3
1A_NG4
1A_NO_NG
1B_NG1
1B_NG2
1B_NG3
1B_NG4
1B_NO_NG
2A_NG1
2A_NG2
2A_NG3
2A_NG4
2A_NO_NG
2B_NG1
2B_NG2
2B_NG3
2B_NG4
2B_NO_NG
3A_NG1
3A_NG2
3A_NG3
3A_NG4
3A_NO_NG
3B_NG1
3B_NG2
3B_NG3
3B_NG4
3B_NO_NG
Unedited_1
Unedited_2
PBS 9
RT 17
PBS 15
RT 17
PBS 9
RT 20
PBS 15
RT 20
PBS 9
RT 22
PBS 15
RT 22

ФИГ.88


PE3 редактирование в здоровых CD34+ клетках
Редактирование, %
4
3
2
1
0
2:1
1:1
2:0.5
2:1
1:1
2:0.5
Редактирование HBB +5 G->A
HEK3 CTT вставка
Редактор (мкг): Направляющая (мкг)
Требуемое редактирование на основе локальных считываний
Индели на основе локальных считываний

ФИГ.89