(19) Евразийское патентное ведомство

(21) 202293254 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.03.07

(22) Дата подачи заявки
2021.05.07

(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 ОДНОВРЕМЕННОГО РЕДАКТИРОВАНИЯ ОБЕИХ ЦЕПЕЙ ЦЕЛЕВОЙ ДВУХЦЕПОЧЕЧНОЙ НУКЛЕОТИДНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

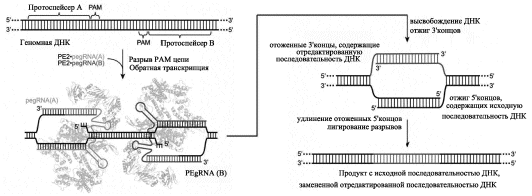
(31) **63/022,397; 63/116,785**
(32) **2020.05.08; 2020.11.20**
(33) **US**
(86) **PCT/US2021/031439**
(87) **WO 2021/226558 2021.11.11**

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

(72) Изобретатель:
Лю Дэвид Р., Анзалоне Эндрю Вито, Леви Джонатан Ма, Гао Синь, Подраки Кристофер Джей (US)

(74) Представитель:
Гизатуллина Е.М., Угрюмов В.М., Строкова О.В., Гизатуллин Ш.Ф., Костюшенкова М.Ю., Джермакян Р.В. (RU)

(57) Изобретение относится к системам, композициям и способам одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию. Согласно некоторым аспектам системы содержат первый и второй комплексы праймированного редактирования, где каждый из первого и второго комплексов праймированного редактирования содержит (1) праймирующий редактор, содержащий (i) программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и (ii) полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и (2) РЕгРНК, содержащую спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера, где матрица для синтеза ДНК кодирует желаемую последовательность ДНК или ее комплемент, где желаемая последовательность ДНК или ее комплемент образуют дуплекс, содержащий отредактированную часть, которая вставляется в целевой участок, подлежащий редактированию. Согласно некоторым аспектам системы содержат первый, второй, третий и четвертый комплексы праймированного редактирования, причем каждый содержит праймирующий редактор и РЕгРНК. Настоящее изобретение также относится к способам одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию. Кроме того, настоящее изобретение относится к фармацевтическим композициям, полинуклеотидам, векторам, клеткам и наборам.



202293254 A1

202293254 A1

СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ ОДНОВРЕМЕННОГО РЕДАКТИРОВАНИЯ ОБЕИХ ЦЕПЕЙ ЦЕЛЕВОЙ ДВУХЦЕПОЧЕЧНОЙ НУКЛЕОТИДНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

ОПИСАНИЕ

Государственная поддержка

[0001] Настоящее изобретение выполнено при государственной поддержке в рамках грантов под номерами U01AI142756, RM1HG009490, R01EB022376 и R35GM118062, присужденных Национальным институтом здравоохранения. Правительство имеет определенные права на изобретение.

Родственные заявки и включение посредством ссылки

[0002] Согласно настоящей заявке испрашивается приоритет в соответствии с предварительной заявкой на выдачу патента США № 63/022397, поданной 8 мая 2020, и предварительной заявкой на выдачу патента США № 63/116785, поданной 20 ноября 2020, содержание каждой из которых включено в настоящий документ посредством ссылки во всей своей полноте.

[0003] Эта предварительная заявка на выдачу патента США также относится к и включает посредством ссылки следующие заявки, а именно, предварительную заявку на выдачу патента США № 62/820813, поданную 19 марта 2019 (Номер дела патентного поверенного B1195.70074US00), предварительную заявку на выдачу патента США № 62/858958 (Номер дела патентного поверенного B1195.70074US01), поданную 7 июня, 2019, предварительную заявку на выдачу патента США № 62/889996 (Номер дела патентного поверенного B1195.70074US02), поданную 21 августа 2019, предварительную заявку на выдачу патента США № 62/922654, поданную 21 августа 2019 (Номер дела патентного поверенного B1195.70083US00), предварительную заявку на выдачу патента США № 62/913553 (Номер дела патентного поверенного B1195.70074US03), поданную 10 октября 2019, предварительную заявку на выдачу патента США № 62/973558 (Номер дела патентного поверенного B1195.70083US01), поданную 10 октября 2019, предварительную заявку на выдачу патента США № 62/931195 (Номер дела патентного поверенного B1195.70074US04), поданную 5 ноября 2019, предварительную заявку на выдачу патента США № 62/944231 (Номер дела патентного поверенного B1195.70074US05), поданную 5 декабря 2019, предварительную заявку на выдачу патента США № 62/974537 (Номер дела патентного поверенного B1195.70083US02), поданную 5 декабря 2019, предварительную

заявку на выдачу патента США № 62/991, 069 (Номер дела патентного поверенного B1195.70074US06), поданную 17 марта 2020, и предварительную заявку на выдачу патента США № (63/100548) (Номер дела патентного поверенного B1195.70083US03), поданную 17 марта 2020. Кроме того, эта предварительная заявка на выдачу патента США также относится к и включает посредством ссылки международные PCT заявки №: PCT/US20/23721, PCT/US20/23730, PCT/US20/23713, PCT/US20/23712, PCT/US20/23727, PCT/US20/23724, PCT/US20/23725, PCT/US20/23728, PCT/US20/23732, PCT/US20/23723, PCT/US20/23553 и PCT/US20/23583, каждая из которых подана 19 марта 2020.

Предшествующий уровень техники настоящего изобретения

[0004] 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 гидовой РНК (гРНК), 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.

[0005] 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]. Репарация, направляемая гомологией, (HDR) долго применялась в клетках человека и других организмов для вставки, коррекции, или замены ДНК-последовательностей в участках двухцепочечных разрывов (DSB) с помощью донорских матриц для репарации ДНК, которые кодируют необходимые редактирования[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].

[0006] 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 стратегии 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.

[0007] 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.

[0008] 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, И, 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 целевом участке значительно расширит диапазон и терапевтический потенциал технологий геномного редактирования на основе CRISPR.

Краткое раскрытие настоящего изобретения

[0009] Настоящее изобретение относится к новой платформе для редактирования генома, называемой «мультилоскутное праймированное редактирование» (включая, например, «двухлоскутное праймированное редактирование» и «четырехлоскутное праймированное редактирование») и представляет собой инновационный подход к «праймированному редактированию» или «классическому праймированному редактированию», как описано авторами настоящего изобретения в Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019), которая включена в настоящий документ посредством ссылки. В то время как классическое праймированное редактирование согласно различным вариантам осуществления полимеризует в участке разрыва один 3′-лоскут, который интегрируется в нуклеиновую кислоту-мишень на той же цепи, описанные согласно настоящему изобретению системы мультилоскутного праймированного редактирования включают различные конструкции, системы и методики, которые согласно различным вариантам осуществления создают пары или множество пар 3′-лоскутов на различных цепях, которые образуют дуплексы, содержащие желаемые редактирования, и которые включаются в молекулы-мишени нуклеиновой кислоты, например, в определенных локусах или сайтах редактирования в геноме. Согласно различным аспектам пары или множество пар 3′-лоскутов образуют дуплексы, поскольку они содержат обратные комплементарные последовательности, которые отжигаются друг с другом после того, как были сгенерированы праймирующими редакторами, описанными в настоящем документе. Дуплексы включаются в целевой участок с помощью клеточных механизмов, которые естественным образом замещают эндогенные последовательности дуплексов, расположенные между соседними участками разрыва. Согласно определенным вариантам осуществления новые дуплексные последовательности могут быть введены в одно или несколько мест (например, в соседние геномные локусы или в два разных участка хромосомы) и могут содержать одну или несколько представляющих интерес последовательностей, например, последовательность, кодирующую белок, последовательность, кодирующую пептид, или РНК-кодирующую последовательность. Согласно одному варианту осуществления новые дуплексные последовательности, установленные системами мультилоскутного праймированного редактирования, могут содержать сайт рекомбиназы, например, сайт *attB* (38 п.о.) и/или *attP* (50 п.о.) рекомбиназы Bxb1, или сайт рекомбиназы, распознаваемый рекомбиназой Hin, рекомбиназой Gin, рекомбиназой Tn3, рекомбиназой β-six, рекомбиназой CinH, рекомбиназой ParA, рекомбиназой γδ, рекомбиназой ϕC31, рекомбиназой TP901, рекомбиназой TG1,

рекомбиназой φBT1, рекомбиназой R4, рекомбиназой φRV1, рекомбиназой φFC1, рекомбиназой MR11, рекомбиназой A118, рекомбиназой U153 и рекомбиназой gp29, рекомбиназой Cre, рекомбиназой FLP, рекомбиназой R, рекомбиназой Lambda, рекомбиназой HK101, рекомбиназой HK022 и рекомбиназой pSAM2.

[0010] Авторы настоящего изобретения недавно разработали праймированное редактирование, которые обеспечивают вставку, делецию или замену в последовательностях геномных ДНК без необходимости двухцепочечных разрывов ДНК сниженной точности. Праймированное редактирование использует сконструированный слитый белок никаза Cas9–обратная транскриптаза (PE1 или PE2) с паре с сконструированной гидовой РНК праймированного редактирования (PEгРНК), которая как направляет Cas9 к целевому участку генома, так и кодирует информацию для установления желаемого редактирования. Праймированное редактирование проводят посредством многостадийного способа редактирования: 1) домен Cas9 связывается с и разрезает целевой участок геномной ДНК, который определяется спейсерной последовательностью PEгРНК, 2) домен обратной транскриптазы использует геномную ДНК с разрывом в качестве праймера для инициации синтеза отредактированной цепи ДНК с использованием сконструированного удлинения PEгРНК в качестве матрицы для обратной транскрипции – это создает одноцепочечный 3'-лоскут, содержащий отредактированную последовательность ДНК, 3) репарация клеточной ДНК устраняет промежуточный 3'-лоскут посредством вытеснения 5'-лоскутов, которое происходит посредством инвазии отредактированным 3'-лоскутом, вырезания 5'-лоскута, содержащего исходную последовательность ДНК, и лигирования нового 3'-лоскута с включением отредактированной цепи ДНК, с образованием гетеродуплекса из одной отредактированной и одной неотредактированной цепи, и 4) репарация клеточной ДНК заменяет неотредактированную цепь в гетеродуплексе с использованием отредактированной цепи в качестве матрицы для репарации, с завершением процесса редактирования.

[0011] Эффективное включение желаемого редактирования требует, чтобы вновь синтезированный 3'-лоскут содержал часть последовательности, которая гомологична участку геномной ДНК. Эта гомология позволяет отредактированному 3'-лоскуту конкурировать с эндогенной цепью ДНК (соответствующим 5'-лоскутом) за включение в дуплекс ДНК. Поскольку отредактированный 3'-лоскут будет содержать меньше гомологии последовательностей, чем эндогенный 5'-лоскут, ожидается, что конкуренция будет благоприятствовать цепи 5'-лоскута. Таким образом, потенциальным фактором, ограничивающим эффективность праймированного редактирования, может быть эффективность инвазии 3'-лоскута эндогенной ДНК и последующее замещение и замена

цепи 5'-лоскута. Более того, успешная инвазия 3'-лоскута и удаление 5'-лоскута включает редактирование только одной цепи генома двухцепочечной ДНК. Постоянная установка редактирования требует восстановления клеточной ДНК для замены неотредактированной комплементарной цепи ДНК с использованием отредактированной цепи в качестве матрицы. В то время как клетка может способствовать замене неотредактированной цепи на отредактированную цепь (стадия 4 выше) путем введения разрыва в неотредактированную цепь, примыкающую к редактированию, с использованием вторичной сгРНК (система PE3), этот способ все еще зависит от второй стадии репарации ДНК. Эти стадии репарации ДНК могут быть особенно неэффективны для редактирований, которые требуют уравновешивания длинных 5'- и 3'-промежуточных лоскутов или содержат длинные негомологичные области, такие как длинные вставки или длинные делеции. Дальнейшие разработки в области праймированного редактирования будут способствовать развитию данной области техники.

[0012] Согласно различным аспектам в настоящем описании раскрыта система мультилоскутного праймированного редактирования (включая, например, системы двухлоскутного праймированного редактирования и системы четырехлоскутного праймированного редактирования), которая решает проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в неотредактированную комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. В системе двухлоскутного праймированного редактирования, например, две РЕгРНК используются для нацеливания на противоположные цепи геномного участка и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 91**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная цепь. Поскольку обе цепи дуплекса синтезируются как отредактированная ДНК, система двухлоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК необходимо только вырезать парные 5'-лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные цепи ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое

содержит минимальную геномную гомологию. Варианты с нуклеазной активностью первичных редакторов, которые разрезают обе цепи ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК. Система четырехлоскутного праймированного редактирования с использованием четырех РЕгРНК обеспечивает подобные преимущества.

[0013] Подобно классическому праймированному редактированию мультилоскутное праймированное редактирование является универсальным и точным способом редактирования генома, который непосредственно записывает новую генетическую информацию в указанный участок ДНК с использованием программируемого нуклеиновой кислотой ДНК-связывающего белка («napDNAbp»), работающего в сочетании с полимеразой (т.е., в форме слитого белка или иным образом представленного *в транс* с napDNAbp), где система праймированного редактирования запрограммирована с помощью гидовой РНК праймированного редактирования (PE), гидовая РНК («РЕгРНК»), которая как определяет целевой участок, так и является матрицей для синтеза желаемого редактирования в виде замещающей цепи ДНК посредством удлинения (либо ДНК, либо РНК), сконструированного на гидовой РНК (например, на 5'- или 3'-конце или во внутренней части гидовой РНК). Цепь замещения, содержащая желаемое редактирование (например, замену одного нуклеотидного основания), имеет ту же последовательность, что и эндогенная цепь целевого участка, подлежащего редактированию (за исключением того, что она включает желаемое редактирование). С помощью механизмов репарации и/или репликации ДНК эндогенная цепь целевого участка заменяется вновь синтезированной замещающей цепью, содержащей желаемое редактирование. В некоторых случаях, праймированное редактирование можно рассматривать как технологию редактирования генома «поиск и замена», поскольку праймирующие редакторы, как описано в настоящем документе, не только ищут и находят желаемый целевой участок для редактирования, но в то же время, кодируют замещающую нить, содержащую желаемое редактирование, которое устанавливают вместо соответствующей эндогенной цепи ДНК целевого участка.

[0014] Мультилоскутные праймирующие редакторы согласно настоящему изобретению относятся в частности к открытию того, что механизм нацеленной-праймированной обратной транскрипции (TPRT) или «праймированного редактирования» может быть использован или адаптирован для проведения точного редактирования генома на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью (например, как показано в различных вариантах осуществления на фиг. 1A-1F). TPRT естественным образом используется мобильными элементами ДНК, такими как не-LTR

ретротранспозоны млекопитающих и бактериальные интроны группы II[28,29]. Авторы настоящего изобретения использовали слияния белок Cas-обратная транскриптаза или родственные системы для нацеливания на конкретную последовательность ДНК с помощью гидовой РНК, создания одноцепочечного разрыва в целевом участке и использования ДНК с разрывом в качестве праймера для обратной транскрипции сконструированной матрицы обратной транскриптазы, интегрированной с гидовой РНК. Однако, хотя концепция начинается с праймирующих редакторов, которые использую обратные транскриптазы в качестве компонента ДНК-полимеразы, мультилоскутные праймирующие редакторы, описанные в настоящем документе, не ограничены обратными транскриптазами, но могут включать использование практически любой ДНК-полимеразы. Действительно, несмотря на то, что настоящее изобретение может относиться к мультилоскутным праймирующим редакторам с «обратными транскриптазами», в настоящем документе указано, что что обратные транскриптазы представляют собой только один тип ДНК-полимеразы, который может работать с мультилоскутным праймированным редактированием. Таким образом, всякий раз, когда в описании упоминаются «обратные транскриптазы», специалисту в данной области должно быть понятно, что вместо обратной транскриптазы можно использовать любую подходящую ДНК-полимеразу. Таким образом, согласно одному аспекту мультилоскутные праймирующие редакторы могут содержать Cas9 (или эквивалентный napDNAbp), который запрограммирован на нацеливание на последовательность ДНК путем связывания ее со специализированной гидовой РНК (т.е., РЕгРНК), содержащей спейсерную последовательность, которая отжигается с компонентом протоспейсерной последовательности в ДНК-мишени. Специализированная гидовая РНК также содержит новую генетическую информацию в форме удлинения, которое кодирует замещающую цепь ДНК, содержащую желаемое генетическое изменение, которая используется для замены соответствующей эндогенной цепи ДНК в целевом участке. Чтобы передать информацию от РЕгРНК к ДНК-мишени, механизм мультилоскутного праймированного редактирования включает разрыв целевого участка в одной цепи ДНК с обнажением 3'-гидроксильной группы. Затем открытая 3'-гидроксильная группа может быть использована для запуска полимеризации ДНК удлинения, кодирующего редактирование, на РЕгРНК непосредственно в целевом участке. Согласно различным вариантам осуществления удлинение, которое обеспечивает матрицу для полимеризации замещающей цепи, содержащей редактирование, может быть образовано из РНК или ДНК. В случае удлинения РНК полимеразой праймирующего редактора может быть РНК-зависимая ДНК-полимераза (например, обратная транскриптаза). В случае удлинения ДНК полимеразой праймирующего редактора может быть ДНК-зависимая ДНК-полимераза.

[0015] При классическом первичном редактировании вновь синтезированная цепь (т.е. замещающая цепь ДНК, содержащая желаемое редактирование), образованная раскрытыми в настоящем документе праймирующими редакторами, будет гомологична геномной целевой последовательности (т.е. иметь ту же последовательность), за исключением включение желаемой замены нуклеотида (например, замены, делеции или вставки одного нуклеотида или их комбинации). Вновь синтезированную (или замещающую) цепь ДНК также можно назвать одноцепочечным лоскутом ДНК, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. Согласно определенным вариантам осуществления система может быть объединена с использованием фермента обратной транскриптазы сниженной точности (например, предоставленного в виде слитого белка с доменом Cas9 или предоставленного *в транс* по отношению к домену Cas9). Фермент обратной транскриптазы сниженной точности может вносить изменения во время синтеза одноцепочечного лоскута ДНК. Таким образом, согласно определенным вариантам осуществления обратная транскриптаза сниженной точности может быть использована для внесения нуклеотидных изменений в ДНК-мишень. В зависимости от обратной транскриптазы сниженной точности, используемой в системе, изменения могут быть случайными или неслучайными.

[0016] При классическом первичном редактировании разрешение гибридизованного промежуточного продукта (содержащего одноцепочечный лоскут ДНК, синтезированный обратной транскриптазой, гибридизованный с цепью эндогенной ДНК) может включать удаление полученного смещенного лоскута эндогенной ДНК (например, с помощью эндонуклеазы 5'-лоскута ДНК, FEN1), лигирование синтезированного одноцепочечного лоскута ДНК с ДНК-мишенью и ассимиляцию желаемого нуклеотидного изменения в результате процессов репарации и/или репликации клеточной ДНК. Поскольку матричный ДНК-синтез предлагает точность до одного нуклеотида для модификации любого нуклеотида, включая вставки и делеции, область применения этого подхода очень широка и, как ожидается, может быть использована для множества применений в фундаментальной науке и терапии.

[0017] Согласно некоторым аспектам настоящее изобретение относится к паре праймирующих редакторов, причем каждый содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и ДНК-полимеразу. Согласно некоторым вариантам осуществления каждый праймирующий редактор способен выполнять редактирование генома путем нацеленной-праймированной обратной транскрипции в присутствии удлиненной гидовой РНК.

[0018] Согласно некоторым аспектам настоящее изобретение относится к паре праймирующих редакторов, причем каждый содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и ДНК-полимеразу, где ДНК-полимераза представлена *в транс* в отношении napDNAbp. Согласно различным вариантам осуществления каждый праймирующий редактор способен выполнять редактирование генома путем нацеленной-праймированной обратной транскрипции в присутствии удлиненной гидовой РНК.

[0019] Согласно некоторым аспектам настоящее изобретение относится к паре праймирующих редакторов, причем каждый содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу. Согласно различным вариантам осуществления каждый праймирующий редактор способен выполнять редактирование генома путем нацеленной-праймированной обратной транскрипции в присутствии удлиненной гидовой РНК.

[0020] Согласно некоторым аспектам настоящее изобретение относится к паре праймирующих редакторов, причем каждый содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, где обратная транскриптаза представлена *в транс* в отношении napDNAbp. Согласно различным вариантам осуществления каждый праймирующий редактор способен выполнять редактирование генома путем нацеленной-праймированной обратной транскрипции в присутствии удлиненной гидовой РНК.

[0021] Согласно определенным вариантам осуществления napDNAbp обладает никазной активностью. NapDNAbp также может представлять собой белок Cas9 или его функциональный эквивалент, такой как Cas9 с нуклеазной активностью, Cas9 без нуклеазной активности (dCas9), или никаза Cas9 (nCas9).

[0022] Согласно определенным вариантам осуществления napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c и Argonaute, и необязательно обладает никазной активностью.

[0023] Согласно другим вариантам осуществления каждый праймирующий редактор двойного праймирующих редакторов в комплексе с удлиненной гидовой РНК способен связываться с целевой последовательностью ДНК.

[0024] Согласно другим вариантам осуществления целевая последовательность ДНК содержит целевую цепь и комплементарную нецелевую цепь.

[0025] Согласно другим вариантам осуществления связывание праймированного редактирования в комплексе с удлиненной гидовой РНК образует R-петлю. R-петля может содержать (i) гибрид РНК-ДНК, содержащий удлиненную гидовую РНК и целевую цепь, и

(ii) комплементарную нецелевую цепь.

[0026] Согласно другим вариантам осуществления комплементарная нецелевая цепь разрывается с образованием праймирующей последовательности обратной транскриптазы, имеющей свободный 3'-конец.

[0027] Согласно различным вариантам осуществления удлиненная гидовая РНК содержит (a) гидовую РНК и (b) удлинение РНК на 5′- или 3′-конце гидовой РНК, или во внутримолекулярном положении гидовой РНК. Удлинение РНК может содержать (i) последовательность матрицы обратной транскрипции, содержащую желаемое изменение нуклеотида, (ii) участок связывания праймера обратной транскрипции и (iii) необязательно линкерную последовательность. Согласно различным вариантам осуществления последовательность матрицы обратной транскрипции может кодировать лоскут одноцепочечной ДНК, комплементарный эндогенной последовательности ДНК, примыкающей к участку разрыва, при этом лоскут одноцепочечной ДНК содержит желаемое изменение нуклеотида.

[0028] Согласно различным вариантам осуществления удлинение РНК составляет 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 в длину.

[0029] Согласно другим вариантам осуществления лоскут одноцепочечной ДНК может гибридизоваться с последовательностью эндогенной ДНК, примыкающей к участку разрыва, тем самым устанавливая желаемое изменение нуклеотида. Согласно другим вариантам осуществления лоскут одноцепочечной ДНК замещает последовательность эндогенной ДНК, примыкающую к участку разрыва и имеющую свободный 5'-конец. Согласно определенным вариантам осуществления замещенная эндогенная ДНК, имеющая 5'-конец, удаляется клеткой.

[0030] Согласно различным вариантам осуществления клеточная репарация лоскута одноцепочечной ДНК приводит к установке желаемого изменения нуклеотида, в результате чего образуется желаемый продукт.

[0031] Согласно различным другим вариантам осуществления желаемое изменение

нуклеотида устанавливается в окне редактирования, которое находится между приблизительно от -4 до +10 последовательности PAM.

[0032] Согласно другим вариантам осуществления, желаемое изменение нуклеотида устанавливается в окне редактирования, которое находится между приблизительно от -5 до +5 от участка разрыва, или приблизительно от -10 до +10 от участка разрыва, или между приблизительно от -20 до +20 от участка разрыва, или между приблизительно от -30 до +30 от участка разрыва, или между приблизительно от -40 до + 40 от участка разрыва, или между приблизительно от -50 до +50 от участка разрыва, или между приблизительно от -60 до +60 от участка разрыва, или между приблизительно от -70 до +70 от участка разрыва, или между приблизительно от -80 до +80 от участка разрыва, или между приблизительно от -90 до +90 от участка разрыва, или между приблизительно от -100 до +100 от участка разрыва, или между приблизительно от -200 до +200 от участка разрыва.

[0033] Согласно различным вариантам осуществления napDNAbp двойных праймирующих редакторов каждый содержит аминокислотную последовательность SEQ ID NO: 18. Согласно различным другим вариантам осуществления napDNAbp содержит аминокислотную последовательность, которая на по меньшей мере 80%, 85%, 90%, 95%, 98% или 99% идентична аминокислотной последовательности согласно любой последовательности из SEQ ID NO: 26-39, 42-61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467 и 482-487 (Cas9), SEQ ID NO: 77-86 (CP-Cas9), SEQ ID NO: 18-25 и 87-88 (SpCas9) и SEQ ID NO: 62-72(Cas12)

[0034] Согласно другим вариантам осуществления обратная транскриптаза раскрытых праймирующих редакторов и/или композиций двойных праймирующих редакторов может содержать любую 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. Согласно другим вариантам осуществления обратная транскриптаза может содержать аминокислотную последовательность, которая на по меньшей мере 80%, 85%, 90%, 95%, 98% или 99% идентична аминокислотной последовательности согласно любой последовательности 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. Эти последовательности могут быть встречающимися в природе последовательностями обратной транскриптазы, например, из ретровируса или ретротранспозона, последовательности могут быть рекомбинантными.

[0035] Согласно различным другим вариантам осуществления праймирующие редакторы двойных праймирующих редакторов, раскрытых в настоящем документе, могут содержать различные структурные конфигурации. Например, согласно вариантам осуществления, согласно которым праймирующие редакторы представляют собой слитый

белок, каждый из слитых белков двойных праймирующих редакторов может содержать структуру NH_2-[napDNAbp]-[обратная транскриптаза]-COOH или NH_2-[обратная транскриптаза]-[napDNAbp]-COOH, где в каждом случае «]-[» указывает на присутствие необязательной линкерной последовательности.

[0036] Согласно различным вариантам осуществления линкерная последовательность содержит аминокислотную последовательность согласно SEQ ID NO: 127, 165-176, 446, 453 и 767-769, или аминокислотную последовательность, которая на по меньшей мере 80%, 85% или 90% или 95% или 99% идентична любой из линкерных аминокислотных последовательностей SEQ ID NO: 127, 165-176, 446, 453 и 767-769.

[0037] Согласно различным вариантам осуществления желаемое изменение нуклеотида, которое включается в ДНК-мишень, может представлять собой изменение одного нуклеотида (например, транзицию или трансверсию), вставку одного или нескольких нуклеотидов или делецию одного или нескольких нуклеотидов.

[0038] В определенных случаях вставка составляет 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 в длину.

[0039] В определенных случаях делеция составляет 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 в длину.

[0040] Согласно другому аспекту настоящее изобретение относится к удлиненной гидовой РНК, содержащей гидовую РНК и по меньшей мере одно удлинение РНК. Удлинение РНК может быть расположено на 3′-конце гидовой РНК. Согласно другим вариантам осуществления удлинение РНК может быть расположено на 5′-конце гидовой

РНК. Согласно другим вариантам осуществления удлинение РНК может быть расположено во внутримолекулярном положении внутри гидовой РНК, однако предпочтительно, чтобы внутримолекулярное расположение удлиненной части не нарушало функционирование протоспейсера.

[0041] Согласно различным вариантам осуществления удлиненная гидовая РНК способна связываться с napDNAbp и направлять napDNAbp на целевую последовательность ДНК. Целевая последовательность ДНК может содержать целевую цепь и комплементарную нецелевую цепь, где гидовая РНК гибридизуется с целевой цепью с образованием гибрида РНК-ДНК и R-петли.

[0042] Согласно различным вариантам осуществления удлиненной гидовой РНК по меньшей мере одно удлинение РНК может содержать последовательность матрицы обратной транскрипции. Согласно другому варианту осуществления удлинение РНК может дополнительно содержит участок связывания праймера обратной транскрипции. Согласно другим вариантам осуществления удлинение РНК может содержать линкер или спейсерную последовательность, которая соединяет удлинение РНК с гидовой РНК.

[0043] Согласно различным вариантам осуществления удлинение РНК может составлять 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 в длину.

[0044] Согласно другим вариантам осуществления последовательность матрицы обратной транскрипции составляет по меньшей мере 3 нуклеотида, по меньшей мере 4 нуклеотидов, 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 в длину.

[0045] Согласно другим вариантам осуществления, где последовательность участка связывания праймера обратной транскрипции составляет по меньшей мере 3 нуклеотидов, 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 в длину.

[0046] Согласно другим вариантам осуществления необязательный линкер или спейсерная последовательность составляет 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 в длину.

[0047] Согласно различным вариантам осуществления удлиненных гидовых РНК, последовательность матрицы обратной транскрипции может кодировать лоскут одноцепочечной ДНК, комплементарный эндогенной последовательности ДНК, примыкающей к участку разрыва, при этом лоскут одноцепочечной ДНК содержит желаемое изменение нуклеотида. Лоскут одноцепочечной ДНК может смещать эндогенную одноцепочечную ДНК в участке разрыва. Смещенная эндогенная одноцепочечная ДНК в участке разрыва может иметь 5'-конец и образовывать эндогенный лоскут, который может быть удален клеткой. Согласно различным вариантам осуществления удаление 5'-концевого эндогенного лоскута может способствовать образованию продукта, поскольку удаление 5'-концевого эндогенного лоскута способствует гибридизации одноцепочечного 3'-лоскута ДНК с соответствующей комплементарной цепью ДНК, а также включению или ассимиляции желаемого изменения нуклеотида, выполняемого одноцепочечным 3'-лоскутом ДНК в ДНК-мишень.

[0048] Согласно различным вариантам осуществления удлиненных гидовых РНК клеточная репарация лоскута одноцепочечной ДНК приводит к установке желаемой замены нуклеотида, в результате чего образуется желаемый продукт.

[0049] Согласно определенным вариантам осуществления РЕгРНК содержит нуклеотидную последовательность согласно SEQ ID NO: 101-104, 181-183, 223-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 757-761, 776-777, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929 и 3972-3989, 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% идентичности последовательности с любой из последовательностей SEQ ID NO: 101-104, 181-183, 223-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 757-761, 776-777, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929 и 3972-3989.

[0050] Согласно другому аспекту настоящее изобретение относится к комплексам, содержащим праймирующий редактор, описанный в настоящем документе, и любую удлиненную гидовую РНК, как описано выше.

[0051] Согласно другим аспектам настоящее изобретение относится к комплексу,

содержащему napDNAbp и удлиненную гидовую РНК. NapDNAbp может представлять собой никазу Cas9, или может представлять собой аминокислотную последовательность согласно SEQ ID NO: 42-57 (никаза Cas9) и 65 (никаза AsCas12a), или аминокислотную последовательность, которая на по меньшей мере 80%, 85%, 90%, 95%, 98% или 99% идентична аминокислотной последовательности согласно любой из SEQ ID NO: 42-57 (никаза Cas9) и 65 (никаза AsCas12a).

[0052] Согласно различным вариантам осуществления, включающих комплекс, удлиненная гидовая РНК способна направлять napDNAbp к целевой последовательности ДНК. Согласно различным вариантам осуществления обратная транскриптаза может быть представлена *в транс*, т.е. получена из другого источника, чем сам комплекс. Например, обратная транскриптаза может быть представлена в той же клетке, содержащей комплекс, путем введения отдельного вектора, отдельно кодирующего обратную транскриптазу.

[0053] Согласно другому аспекту настоящее изобретение относится к системе, содержащей первый и второй комплекс праймированного редактирования, причем каждый комплекс содержит праймирующий редактор и гидовую РНК праймированного редактирования (РЕгРНК). Согласно некоторым вариантам осуществления каждый праймирующий редактор содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и каждый РЕгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК, содержащую отредактированную часть. Две кодируемые последовательности одноцепочечной ДНК могут быть комплементарны друг другу и образовывать дуплекс, который интегрируется в целевой участок, подлежащий редактированию. Согласно некоторым вариантам осуществления две кодируемые последовательности одноцепочечной ДНК могут содержать области, комплементарные друг другу. Согласно определенным вариантам осуществления две кодируемые последовательности одноцепочечной ДНК могут содержать области комплементарности друг, которые составляют по меньшей мере 2 п.о., по меньшей мере 3 п.о., по меньшей мере 4 п.о., по меньшей мере 5 п.о., по меньшей мере 10 п.о., по меньшей мере 20 п.о., по меньшей мере 30 п.о., по меньшей мере 40 п.о., по меньшей мере 50 п.о., по меньшей мере 100 п.о., по меньшей мере 200 п.о., по меньшей мере 300 п.о., по меньшей мере 400 п.о., по меньшей мере 500 п.о., по меньшей мере 600 п.о., по меньшей мере 700 п.о., по меньшей мере 800 п.о., по меньшей мере 900 п.о. или по меньшей мере 1000 п.о. в длину. Согласно некоторым вариантам осуществления настоящее изобретение относится к праймирующему

редактору в виде слитого белка. Согласно определенным вариантам осуществления компоненты праймирующего редактора (*т.е.*, napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы) представлены *в транс*.

[0054] Согласно другому аспекту настоящее изобретение относится к системе, содержащей первый, второй, третий и четвертый комплексы праймированного редактирования, причем каждый комплекс содержит праймирующий редактор и гидовую РНК праймированного редактирования (PEгРНК). Согласно некоторым вариантам осуществления каждый праймирующий редактор содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и каждый PEгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК, содержащую отредактированную часть. Две кодируемые последовательности одноцепочечной ДНК могут быть комплементарны друг другу и образуют дуплекс, который интегрируется в целевой участок, подлежащий редактированию. Согласно некоторым вариантам осуществления две кодируемые последовательности одноцепочечной ДНК могут содержать области, комплементарные друг другу. Согласно определенным вариантам осуществления две кодируемые последовательности одноцепочечной ДНК могут содержать области, комплементарные друг другу, которые составляют по меньшей мере 2 п.о., по меньшей мере 3 п.о., по меньшей мере 4 п.о., по меньшей мере 5 п.о., по меньшей мере 10 п.о., по меньшей мере 20 п.о., по меньшей мере 30 п.о., по меньшей мере 40 п.о., по меньшей мере 50 п.о., по меньшей мере 100 п.о., по меньшей мере 200 п.о., по меньшей мере 300 п.о., по меньшей мере 400 п.о., по меньшей мере 500 п.о., по меньшей мере 600 п.о., по меньшей мере 700 п.о., по меньшей мере 800 п.о., по меньшей мере 900 п.о. или по меньшей мере 1000 п.о. в длину. Согласно некоторым вариантам осуществления праймирующий редактор представлен в виде слитого белка. Согласно определенным вариантам осуществления компоненты праймирующего редактора (*т.е.*, napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы) представлены *в транс*.

[0055] Согласно некоторым вариантам осуществления каждый napDNAbp представляет собой домен Cas9 или его вариант. Согласно некоторым вариантам осуществления каждый napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, или никазный домен Cas9, или их вариант. Согласно определенным вариантам осуществления каждый napDNAbp независимо выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a,

Cas12b1, Cas13a, Cas12c и Argonaute, и необязательно обладает никазной активностью. Согласно различным вариантам осуществления каждый napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 2-65, или аминокислотную последовательность, на по меньшей мере 80%, 85%, 90%, 95% или 99% идентичную любой из последовательностей SEQ ID NO: 2-65.

[0056] Согласно некоторым вариантам осуществления полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу. Согласно определенным вариантам осуществления полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110.

[0057] Согласно некоторым вариантам осуществления каждый праймирующий редактор может содержать линкер, который соединяет napDNAbp и обратную транскриптазу. Согласно определенным вариантам осуществления линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 119-128, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 119-128. Каждый линкер может составлять 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

[0058] Согласно некоторым вариантам осуществления каждый РЕгРНК может независимо содержать нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 192-203, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 192-203.

[0059] Согласно различным вариантам осуществления каждая спейсерная последовательность каждого РЕгРНК может связываться со специфическим участком связывания последовательности двухцепочечной ДНК, примыкающей к целевому участку, подлежащему редактированию. Согласно некоторым вариантам осуществления связывание спейсерной последовательности одного комплекса праймированного редактирования с одной цепью последовательности двухцепочечной ДНК связывание спейсерной последовательности другого комплекса праймированного редактирования с противоположной цепью последовательности двухцепочечной ДНК приводит к разрыву

обеих цепей ДНК в участке разрыва, проксимальном по отношению к последовательностям PAM на каждой цепи.

[0060] Согласно другому аспекту настоящее изобретение относится к полинуклеотидам. Согласно некоторым вариантам осуществления полинуклеотиды могут кодировать любой из комплексов, описанных в настоящем документе. Согласно определенным вариантам осуществления полинуклеотиды могут кодировать любую из PEгРНК, описанных в настоящем документе.

[0061] Согласно другому аспекту настоящее изобретение относится к полинуклеотидам. Согласно определенным вариантам осуществления полинуклеотиды могут кодировать любой из праймирующих редакторов, раскрытых в настоящем документе. Согласно определенным другим вариантам осуществления полинуклеотиды могут кодировать любой из napDNAbp, раскрытых в настоящем документе. Согласно другим вариантам осуществления полинуклеотиды могут кодировать любую из обратных транскриптаз, раскрытых в настоящем документе. Согласно определенным другим вариантам осуществления полинуклеотиды могут кодировать любую из удлиненных гидовых РНК, раскрытых в настоящем документе, любую из последовательностей матрицы обратной транскрипции, или любой из сайтов праймеров обратной транскрипции, или любую из необязательных линкерных последовательностей.

[0062] Согласно другим аспектам настоящее изобретение относится к векторам, содержащим полинуклеотиды, описанные в настоящем документе. Таким образом, согласно определенным вариантам осуществления векторы содержат полинуклеотиды, кодирующие праймирующие редакторы, содержащие napDNAbp и обратную транскриптазу (*т.е.*, в виде слитого белка или экспрессированного *в транс*). Согласно определенным вариантам осуществления векторы содержат полинуклеотиды, кодирующие любой из комплексов, описанных в настоящем документе. Согласно другим вариантам осуществления векторы содержат полинуклеотиды, которые раздельно кодируют napDNAbp и обратную транскриптазу. Согласно другим вариантам осуществления векторы могут содержать полинуклеотиды, которые кодируют удлиненные гидовые РНК. Согласно различным вариантам осуществления векторы могут содержать один или несколько полинуклеотидов, которые кодируют napDNAbp, обратную транскриптазу и удлиненные гидовые РНК в одном и том же или различных векторах. Согласно некоторым вариантам осуществления векторы содержат полинуклеотиды, кодирующие любые PEгРНК описанные в настоящем документе.

[0063] Согласно другим аспектам настоящее изобретение относится к клеткам, содержащим праймирующий редактор, как описано в настоящем документе, и удлиненную

гидовую РНК. Клетки могут быть трансформированы векторами, содержащими праймирующие редакторы, napDNAbp, обратную транскриптазу и удлиненные гидовые РНК. Эти генетические элементы могут быть включены в один и тот же вектор или в разные векторы. Согласно некоторым вариантам осуществления клетки содержат любую из систем или комплексов, описанных в настоящем документе. Клетки могут быть трансформированы полинуклеотидами, кодирующими любую из систем, комплексов и/или РЕгРНК, раскрытых в настоящем документе, или векторами, содержащими полинуклеотиды, кодирующие любую из систем, комплексов или РЕгРНК, раскрытыми в настоящем документе.

[0064] Согласно другим аспектам настоящее изобретение относится к фармацевтическим композициям. Согласно определенным вариантам осуществления фармацевтические композиции содержат один или несколько из napDNAbp, праймирующего редактора, обратной транскриптазы и удлиненной гидовой РНК. Согласно определенным вариантам осуществления фармацевтические композиции содержат любую из систем и/или комплексов, описанных в настоящем документе. Согласно определенным вариантам осуществления фармацевтические композиции содержат любой из праймирующих редакторов, систем или комплексов, описанных в настоящем документе, и фармацевтически приемлемое вспомогательное вещество. Согласно другим вариантам осуществления фармацевтические композиции содержат любую удлиненную гидовую РНК, описанную в настоящем документе, и фармацевтически приемлемое вспомогательное вещество. Согласно другим вариантам осуществления, фармацевтические композиции содержат любую удлиненную гидовую РНК, описанную в настоящем документе, в комбинации с любым праймирующим редактором, описанным в настоящем документе, и фармацевтически приемлемое вспомогательное вещество. Согласно другим вариантам осуществления фармацевтические композиции содержат любую полинуклеотидную последовательность, содержащую один или несколько из napDNAbp, праймирующего редактора, обратной транскриптазы и удлиненной гидовой РНК, или любой из векторов, раскрытых в настоящем документе. Согласно другим вариантам осуществления различные компоненты, раскрытые в настоящем документе, могут быть разделены на одну или несколько фармацевтических композиций. Например, первая фармацевтическая композиция может содержать праймирующий редактор или napDNAbp, вторая фармацевтическая композиция может содержать обратную транскриптазу, и третья фармацевтическая композиция может содержать удлиненную гидовую РНК.

[0065] Согласно другому аспекту настоящее изобретение относится к наборам. Согласно одному варианту осуществления набор содержит один или несколько

полинуклеотидов, кодирующих один или несколько компонентов, включая праймирующий редактор, napDNAbp, обратную транскриптазу и удлиненную гидовую РНК. Наборы могут также содержать векторы, клетки и выделенные препараты полипептидов, включая любой праймирующий редактор, napDNAbp или обратную транскриптазу, раскрытые в настоящем документе.

[0066] Согласно другому аспекту настоящее изобретение относится к способам применения раскрытых композиций вещества, включая способы использования любой из систем, описанных в настоящем документе, для одновременного редактирования обеих комплементарных цепей последовательности двухцепочечной ДНК в целевом участке. Согласно некоторым вариантам осуществления способ предусматривает контакт последовательности двухцепочечной ДНК с любой из систем, раскрытых в настоящем документе.

[0067] Согласно одному аспекту настоящее изобретение относится к способам, предусматривающим контакт последовательности двухцепочечной ДНК в целевом участке с первым и вторым комплексом праймированного редактирования, причем каждый комплекс содержит праймирующий редактор и гидовую РНК праймированного редактирования (PEгРНК). Согласно некоторым вариантам осуществления каждый праймирующий редактор содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и каждый PEгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно некоторым вариантам осуществления каждый праймирующий редактор представлен в виде слитого белка. Согласно некоторым вариантам осуществления компоненты праймирующего редактора представлены *в транс*. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК, содержащую отредактированную часть. Две кодируемые последовательности одноцепочечной ДНК могут быть комплементарны друг другу и образуют дуплекс, который интегрируется в целевой участок, подлежащий редактированию. Различные элементы комплексов праймированного редактирования могут содержать любой из вариантов осуществления систем, раскрытых в настоящем документе.

[0068] Согласно другим аспектам настоящее изобретение относится к способам, предусматривающим контакт последовательности двухцепочечной ДНК в целевом участке с первым, вторым, третьим и четвертым комплексами праймированного редактирования, причем каждый комплекс содержит праймирующий редактор и гидовую РНК праймированного редактирования (PEгРНК). Согласно некоторым вариантам

осуществления каждый праймирующий редактор содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и каждая РЕгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно некоторым вариантам осуществления каждый праймирующий редактор представлен в виде слитого белка. Согласно некоторым вариантам осуществления компоненты праймирующего редактора представлены *в транс*. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК. Две кодируемые последовательности одноцепочечной ДНК могут быть комплементарны друг другу и образуют дуплекс, который интегрируется в целевой участок, подлежащий редактированию. Различные элементы комплексов праймированного редактирования могут содержать любой из вариантов осуществления, раскрытых в настоящем документе.

[0069] Согласно некоторым вариантам осуществления способы, раскрытые в настоящем документе, обеспечивают инверсию целевой последовательности ДНК. Согласно некоторым вариантам осуществления первая последовательность одноцепочечной ДНК, кодируемая первой матрицей для синтеза ДНК, и вторая последовательность одноцепочечной ДНК, кодируемая второй матрицей для синтеза ДНК, находятся на противоположных концах целевой последовательности ДНК, а третья последовательность одноцепочечной ДНК, кодируемая третьей матрицей для синтеза ДНК, и четвертая последовательность одноцепочечной ДНК, кодируемая четвертой матрицей для синтеза ДНК, находятся на противоположных концах той же целевой последовательности ДНК.

[0070] Согласно некоторым вариантам осуществления настоящее изобретение относится к способам, дополнительно предусматривающим предоставление донора кольцевого ДНК. Согласно определенным вариантам осуществления первая последовательность одноцепочечной ДНК, кодируемая первой матрицей для синтеза ДНК, и третья последовательность одноцепочечной ДНК, кодируемая третьей матрицей для синтеза ДНК, находятся на противоположных концах целевой последовательности ДНК, и вторая последовательность одноцепочечной ДНК, кодируемая второй матрицей для синтеза ДНК, и четвертая последовательность одноцепочечной ДНК, кодируемая четвертой матрицей для синтеза ДНК, находятся на доноре кольцевой ДНК. Согласно некоторым вариантам осуществления часть донора кольцевой ДНК между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК замещает целевую последовательность ДНК между первой

последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК.

[0071] Согласно некоторым вариантам осуществления способы, раскрытые в настоящем документе, позволяют транслокацию целевой последовательности ДНК из первой молекулы нуклеиновой кислоты (*например*, первая хромосома) во вторую молекулу нуклеиновой кислоты (*например*, вторая хромосома). Согласно некоторым вариантам осуществления первая последовательность одноцепочечной ДНК, кодируемая первой матрицей для синтеза ДНК, и третья последовательность одноцепочечной ДНК, кодируемая третьей матрицей для синтеза одноцепочечной ДНК, находятся на первой молекуле нуклеиновой кислоты, и вторая последовательность одноцепочечной ДНК, кодируемая второй матрицей для синтеза ДНК, и четвертая последовательность одноцепочечной ДНК, кодируемая четвертой матрицей для синтеза ДНК, находятся на второй молекуле нуклеиновой кислоты. Согласно определенным вариантам осуществления часть первой молекулы нуклеиновой кислоты между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК включена во вторую молекулу нуклеиновой кислоты. Согласно определенным вариантам осуществления часть второй молекулы нуклеиновой кислоты между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК включена в первую молекулу нуклеиновой кислоты.

[0072] Согласно другому аспекту настоящее изобретение относится к паре РЕгРНК для применения в мультилоскутном праймированном редактировании. Согласно некоторым вариантам осуществления пара содержит первую РЕгРНК и вторую РЕгРНК, и каждая РЕгРНК независимо содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК. Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы используются в сочетании с парой РЕгРНК, которые нацелены на отдельные праймирующие редакторы с любой стороны целевого участка, где каждая пара РЕгРНК кодирует 3’-лоскуты нуклеиновой кислоты, которые являются обратными комплементами друг другу. Согласно различным вариантам осуществления 3’-лоскуты, содержащие обратные последовательности комплемента, могут отжигаться друг с другом с образованием дуплекса, содержащего желаемую последовательность редактирования или нуклеиновой кислоты, кодируемую РЕгРНК. Затем дуплекс интегрируется в целевой участок путем замещения соответствующего эндогенного дуплекса, расположенного между соседними участками разрыва.

[0073] Согласно другому аспекту настоящее изобретение относится к множеству РЕгРНК для применения в мультилоскутном праймированном редактировании. Согласно некоторым вариантам осуществления множество содержит первую, вторую, третью и четвертую РЕгРНК. Согласно некоторым вариантам осуществления каждая из четвертой РЕгРНК независимо содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Согласно определенным вариантам осуществления каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК. Две кодируемые последовательности одноцепочечной ДНК могут быть комплементарны друг другу.

[0074] Согласно различным аспектам настоящее изобретение относится к полинуклеотидам, кодирующим любую из пар или множеств РЕгРНК, описанных в настоящем документе. Согласно определенным аспектам настоящее изобретение относится к векторам, кодирующим полинуклеотид, кодирующий любую из пар или множеств РЕгРНК, описанных в настоящем документе. Согласно другому аспекту настоящее изобретение относится к клеткам, содержащим вектор, кодирующий полинуклеотид, кодирующий любую из пар или множеств РЕгРНК, описанных в настоящем документе. Согласно другим аспектам настоящее изобретение относится к фармацевтическим композициям, содержащим любую из пар или множеств РЕгРНК, описанных в настоящем документе, вектор, кодирующий любую из пар или множеств РЕгРНК, описанных в настоящем документе, или клетку, содержащую вектор, кодирующий любую из пар или множеств РЕгРНК, описанных в настоящем документе. Согласно определенным вариантам осуществления фармацевтические композиции содержат фармацевтическое вспомогательное вещество.

[0075] Согласно одному варианту осуществления способы относятся к способу введения желаемого изменения нуклеотида в последовательность двухцепочечной ДНК. Способ первый содержит контакт последовательности двухцепочечной ДНК с комплексом, содержащим праймирующий редактор и удлиненную гидовую РНК, где праймирующий редактор содержит napDNAbp и обратную транскриптазу, и где удлиненная гидовая РНК содержит последовательность матрицы обратной транскрипции, содержащую желаемое изменение нуклеотида. Согласно некоторым вариантам осуществления каждый праймирующий редактор представлен в виде слитого белка. Согласно некоторым вариантам осуществления компоненты праймирующего редактора представлены *в транс*. Далее, способ предусматривает разрыв последовательности двухцепочечной ДНК на нецелевой цепи, с образованием таким образом свободной одноцепочечной ДНК, имеющей 3′-конец. Способ затем предусматривает гибридизацию 3`-конца свободной

одноцепочечной ДНК с последовательностью матрицы обратной транскрипции, тем самым праймируя домен обратной транскриптазы. Способ затем предусматривает полимеризацию цепи ДНК с 3'-конца, тем самым создавая лоскут одноцепочечной ДНК, содержащий желаемую замену нуклеотида. Способ затем предусматривает замещение эндогенной цепи ДНК, примыкающей к участку разреза, одноцепочечным лоскутом ДНК, тем самым устанавливая желаемое изменение нуклеотида в последовательности двухцепочечной ДНК.

[0076] Согласно другим вариантам осуществления настоящее изобретение относится к способу введения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в целевом локусе, предусматривающий контакт молекулы ДНК с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp) и гидовой РНК, которая нацеливает napDNAbp на локус-мишень, где гидовая РНК содержит последовательность матрицы обратной транскриптазы (RT), содержащую по меньшей мере одно желаемое изменение нуклеотида. Затем способ предусматривает образование открытого 3'-конца в цепи ДНК в целевом локусе, а затем гибридизацию открытого 3'-конца с последовательностью матрицы RT, для праймирования обратной транскрипции. Затем одноцепочечный лоскут ДНК, содержащий по меньшей мере одно желаемое изменение нуклеотида, на основе последовательности матрицы RT синтезируют или полимеризуют путем обратной транскриптазы. Наконец, по меньшей мере одно желаемое изменение нуклеотида встраивают в соответствующую эндогенную ДНК, тем самым внося одно или несколько изменений в нуклеотидную последовательность молекулы ДНК в целевом локусе.

[0077] Согласно другим вариантам осуществления настоящее изобретение относится к способу введения одного или нескольких изменений в нуклеотидную последовательность молекулы ДНК в целевом локусе с помощью нацеленной-праймированной обратной транскрипции, причем способ предусматривает: (a) контакт молекулы ДНК в целевом локусе с (i) праймирующим редактором, содержащим программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и обратную транскриптазу, и (ii) гидовой РНК, содержащей матрицу RT, содержащую желаемое изменение нуклеотида, (b) проведение нацеленной-праймированной обратной транскрипции матрицы RT с созданием одноцепочечной ДНК, содержащей желаемое изменение нуклеотида, и (c) включение желаемого изменения нуклеотида в молекулу ДНК в целевом локусе посредством процесса репарации и/или репликации ДНК.

[0078] Согласно определенным вариантам осуществления стадия замещения цепи эндогенной ДНК включает: (i) гибридизацию лоскута одноцепочечной ДНК с цепью эндогенной ДНК, примыкающей к участку разреза, с созданием несовпадения

последовательностей, (ii) удаление эндогенной цепи ДНК, и (iii) исправление несовпадения с образованием желаемого продукта, содержащего желаемое изменение нуклеотида в обеих цепях ДНК.

[0079] Согласно различным вариантам осуществления желаемое изменение нуклеотида может представлять собой замену одного нуклеотида (например, переход или изменение трансверсии), делецию или вставку. Например, желаемой заменой нуклеотида может представлять собой (1) замену G на T, (2) замену G на A, (3) замену G на C, (4) замену T на G, (5) замену T на A, (6) замену T на C, (7) замену C на G, (8) замену C на T, (9) замену C на A, (10) замену A на T, (11) замену A на G или (12) замену A на C.

[0080] Согласно другим вариантам осуществления желаемое изменение нуклеотида может преобразовать (1) пару оснований G:C в пару оснований T:A, (2) пару оснований G:C в пару оснований A:T, (3) пару оснований G:C в пару оснований C:G, (4) пару оснований T:A в пару оснований G:C, (5) пару оснований T:A в пару оснований A:T, (6) пару оснований T:A в пару оснований C:G, (7) пару оснований C:G в пару оснований G:C, (8) пару оснований C:G в пару оснований T:A, (9) пару оснований C:G в пару оснований A:T, (10) A: пару оснований T в пару оснований T:A, (11) пару оснований A:T в пару оснований G:C или (12) пару оснований A:T в пару оснований C:G.

[0081] Согласно другим вариантам осуществления способ вводит желаемую замену нуклеотида, которая представляет собой вставку. В определенных случаях вставка составляет 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 в длину.

[0082] Согласно другим вариантам осуществления способ особ вводит желаемое изменение нуклеотида, которое представляет собой делецию. В некоторых других случаях делеция составляет 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 в длину.

[0083] Согласно различным вариантам осуществления желаемое изменение нуклеотида корректирует ген, ассоциированный с заболеванием. Ген, ассоциированный с заболеванием, может быть связан с моногенетическим нарушением, выбранным из группы, состоящей из: дефицита аденозиндезаминазы (ADA), дефицита альфа-1-антитрипсина, муковисцидоза, мышечная дистрофии Дюшенна, галактоземии, гемохроматоза, болезни Хантингтона, болезни клинового сиропа, синдрома Марфана, нейрофиброматоза 1 типа, врожденной пахионихии, фенилкеотнурии, тяжелого комбинированного иммунодефицита, серповидно-клеточной анемии, синдрома Смита-Лемли-Опица и болезни Тея-Сакса. Согласно другим вариантам осуществления ген, ассоциированный с заболеванием, может быть связан с полигенным заболеванием, выбранным из группы, состоящей из: болезни сердца, высокого кровяного давления, болезни Альцгеймера, артрита, диабета, рака и ожирения.

[0084] Способы, раскрытые в настоящем документе, могут включать слитые белки, имеющие napDNAbp, который представляет собой мертвую нуклеазу Cas9 (dCas9), никазу Cas9 (nCas9) или Cas9 с нуклеазной активностью. Согласно другим вариантам осуществления napDNAbp и обратная транскриптаза не кодируются в виде одного слитого белка, а скорее могут быть предоставлены в виде отдельных конструкций. Таким образом, согласно некоторым вариантам осуществления обратная транскриптаза может быть представлена *в транс* положении относительно napDNAbp (скорее, чем в виде слитого белка).

[0085] Согласно различным вариантам осуществления способов согласно настоящему изобретению napDNAbp может содержать аминокислотную последовательность согласно 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 также может содержать аминокислотную последовательность, которая на по меньшей мере 80%, 85%, 90%, 95%, 98% или 99% идентична аминокислотной последовательности согласно любому 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).

[0086] Согласно различным вариантам осуществления способов согласно настоящему изобретению обратная транскриптаза может содержать любую 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 также может содержать аминокислотную последовательность, которая на по меньшей мере 80%, 85%, 90%, 95%, 98% или 99% идентична аминокислотной последовательности согласно любому 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.

[0087] Способы могут предусматривать применение РЕгРНК, содержащей нуклеотидную последовательность согласно SEQ ID NO: 101-104, 181-183, 223-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 757-761, 776-777, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929 и 3972-3989, 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% идентичности последовательности. Способы могут предусматривать применение удлиненных гидовых РНК, которые содержат удлинение РНК на 3′-конце, где удлинение РНК содержит последовательность матрицы обратной транскрипции.

[0088] Способы могут предусматривать применение удлиненных гидовых РНК, которые содержат удлинение РНК на 5′-конце, где удлинение РНК содержит последовательность матрицы обратной транскрипции.

[0089] Способы могут предусматривать применение удлиненных гидовых РНК, которые содержат удлинение РНК во внутримолекулярном местоположении гидовой РНК, где удлинение РНК содержит последовательность матрицы обратной транскрипции.

[0090] Способы могут предусматривать применение удлиненных гидовых РНК, имеющих одно или несколько удлинений РНК, которые составляют 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 в длину.

[0091] Следует понимать, что вышеупомянутые концепции и дополнительные концепции, обсуждаемые ниже, могут быть организованы в любой подходящей комбинации, поскольку настоящее раскрытие не ограничивается в этом отношении. Кроме

того, другие преимущества и новые признаки настоящего раскрытия станут очевидными из следующего подробного описания различных неограничивающих вариантов осуществления при рассмотрении в сочетании с прилагаемыми чертежами.

Краткое описание чертежей

[0092] Следующие чертежи образуют часть настоящего описания и включены для дополнительной демонстрации определенных аспектов настоящего изобретения, которое может быть лучше понятно со ссылкой на один или несколько из данных чертежей в комбинации с подробным описанием конкретных приведенных в настоящем документе вариантов осуществления.

[0093] На **Фиг.1А** приведена схема иллюстративного способа введения изменения одного нуклеотида, и/или вставки, и/или делеции в молекулу ДНК (например, геном) с помощью слитого белка, содержащего обратную транскриптазу, слитую с белком Cas9, в комплексе с молекулой удлиненной гидовой РНК. Согласно этому варианту осуществления гидовая РНК удлиняется на 3 '-конце с включением последовательности матрицы обратной транскриптазы. Схема показывает, как обратная транскриптаза (RT), слитая с никазой Cas9, в комплексе с гидовой РНК (гРНК) связывается с целевым участком ДНК и осуществляет разрыв PAM-содержащей цепи ДНК, примыкающей к нуклеотиду-мишени. Фермент RT использует разорванную ДНК в качестве праймера для синтеза ДНК из гРНК, которая используется в качестве матрицы для синтеза новой цепи ДНК и кодирует желаемое редактирования. Показанный способ редактирования может быть обозначен как редактирование на основе нацеленной-праймированной обратной транскрипции (TRT редактирование) или эквивалентно как «праймированное редактирование».

[0094] На **Фиг.1В** показано то же, что и на Фиг. 1 А, за исключением того, что комплекс праймированного редактирования обозначен более общим образом как [napDNAbp]-[P]:PEгРНК или [P]-[napDNAbp]:PEгРНК, где «P» обозначает любую полимеразу (например, обратную транскриптазу), «napDNAbp» обозначает программируемый нуклеиновой кислотой ДНК-связывающий белок (например, SpCas9), и «PEгРНК» обозначает гидовую РНК праймированного редактирования, и «]-[» относится к необязательному линкеру. Как описано в других разделах, например, на 3A-3G, PEгРНК содержит 5'-плечо удлинения, содержащий участок связывания праймера и матрицу для синтеза ДНК. Несмотря на то, что это не показано, подразумевается, что плечо удлинения PEгРНК (т.е. которое содержит 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.

[0095] На **Фиг.1С** приведена схема иллюстративного способа введения изменения одного нуклеотида, и/или вставки, и/или делеции в молекулу ДНК (например, геном) с помощью слитого белка, содержащего обратную транскриптазу, слитую с белком Cas9, в комплексе с молекулой удлиненной гидовой РНК. Согласно данному варианту осуществления гидовая РНК удлиняется на 5′-конце для включения последовательности матрицы обратной транскриптазы. Схема показывает, как обратная транскриптаза (RT), слитая с никазой Cas9, в комплексе с гидовой РНК (гРНК) связывается с целевым участком ДНК и осуществляет разрыв PAM-содержащей цепи ДНК, примыкающей к нуклеотиду-мишени. Фермент RT использует разорванную ДНК в качестве праймера для синтеза ДНК из гРНК, которая используется в качестве матрицы для синтеза новой цепи ДНК, которая кодирует желаемое редактирование. Показанный способ редактирования может быть обозначен как редактирование на основе нацеленной-праймированной обратной транскрипции (редактирование TRT) или эквивалентно как «праймированное редактирование».

[0096] На **Фиг.1В** показано то же, что и на Фиг. 1С, за исключением того, что комплекс праймированного редактирования обозначен более общим образом как [napDNAbp]-[P]:PEгРНК или [P]-[napDNAbp]:PEгРНК, где «P» обозначает любую полимеразу (например, обратную транскриптазу), «napDNAbp» относится к программируемому нуклеиновой кислотой ДНК-связывающему белку (например, SpCas9), и «PEгРНК» относится к гидовой РНК праймированного редактирования, а «]-[» обозначает необязательный линкер. Как описано в других разделах, например, на Фиг.3А-3G, PEгРНК содержит 3'-плечо удлинения, содержащее участок связывания праймера и матрицу для синтеза ДНК. Несмотря на то, что это не показано, подразумевается, что плечо удлинения PEгРНК (т.е. которое содержит 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. Согласно различным вариантам осуществления PEгРНК может быть сконструирована или синтезирована с включением матрицы на основе

для синтеза ДНК.

[0097] На **фиг. 1E** схематически показан иллюстративный способ как синтезированная одноцепочечная ДНК (которая содержит желаемое изменение нуклеотида) разрешается таким образом, что желаемое изменение нуклеотида включается в ДНК. Как показано, после синтеза отредактированной цепи (или «мутагенной цепи») уравновешивание с эндогенной цепью, отщепление лоскута эндогенной цепи и лигирование приводят к включению редактирования ДНК после разрешения несовпадающего ДНК-дуплекса под действием процессов репарации и/или репликации эндогенной ДНК.

[0098] На **фиг. 1F** представлено схематическое изображение, показывающее, что «разрыв противоположной цепи» может быть включен в способ разрешения, показанный на **фиг. 1E**, чтобы помочь управлять образованием желаемого продукта по сравнению с продуктом реверсии. При разрыве противоположной цепи второй комплекс Cas9/гРНК используется для введения второго разрыва на противоположной цепи от исходной цепи с разрывом. Это индуцирует эндогенные процессы репарации и/или репликации клеточной ДНК для предпочтительной замены неотредактированной цепи (т.е. цепи, содержащей второй участок разрыва).

[0099] На **фиг. 1G** представлена еще одна схема иллюстративного процесса введения замены одного нуклеотида, и/или вставки, и/или делеции в молекулу ДНК (например, геном) локуса-мишени с использованием программируемого нуклеиновой кислотой ДНК-связывающего белок (napDNAbp) в комплексе с удлиненной гидовой РНК. Этот процесс можно назвать вариантом осуществления праймированного редактирования. Удлиненная гидовая РНК содержит удлинение на 3'- или 5'-конце гидовой РНК или во внутримолекулярном местоположении гидовой РНК. На стадии (a) комплекс napDNAbp/гРНК связывается с молекулой ДНК, и гРНК направляет napDNAbp для связывания с локусом-мишенью. На стадии (b) вводят разрыв в одну из цепей ДНК (цепь R-петли, или цепь, содержащую PAM, или нецелевую цепь ДНК, или цепь протоспейсера) локуса-мишени (например, с помощью нуклеазы или химического агента), тем самым создавая доступный 3'-конец в одной из цепей локуса-мишени. Согласно определенным вариантам осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, т.е. в цепи, которая не гибридизуется с последовательностью гидовой РНК. На стадии (c) 3'-концевая цепь ДНК взаимодействует с удлиненной частью гидовой РНК с праймированием обратной транскрипции. Согласно определенным вариантам осуществления 3’-концевая цепь ДНК гибридизуется со специфической праймирующей последовательностью RT на удлиненной части гидовой РНК. На стадии (d) вводят

обратную транскриптазу, которая синтезирует одну цепь ДНК от 3'-конца праймированного сайта к 3'-концу гидовой РНК. Это образует лоскут одноцепочечной ДНК, содержащий желаемую замену нуклеотида (например, замену, вставку или делецию одного основания, или их комбинацию). На стадии (e) высвобождаются napDNAbp и гидовая РНК. Стадии (f) и (g) относятся к разрешению лоскута одноцепочечной ДНК таким образом, чтобы желаемое изменение нуклеотида было включено в целевой локус. Этот способ можно привести к образованию желаемого продукта путем удаления соответствующего 5'-лоскута эндогенной ДНК, который образуется, когда 3'-лоскут одноцепочечной ДНК вставляется и гибридизуется с комплементарной последовательностью на другой цепи. Способ также может быть направлен на образование продукта с разрывом второй нити, как показано на **фиг. 1F**. Этот способ может привнести по меньшей мере одно или несколько из следующих генетических изменений: трансверсии, транзиции, делеции и инверсии.

[0100] На **фиг. 1H** приведено схематическое изображение типов генетических изменений, которые возможны при описанных в настоящем документе способах праймированного редактирования. Типы изменений нуклеотидов, достижимых путем праймированного редактирования, включают делеции (включая короткие и длинные делеции), однонуклеотидные изменения (включая транзиции и трансверсии), инверсии и вставки (включая короткие и длинные делеции).

[0101] На **фиг. 1I** представлена схема, изображающая временные разрывы второй цепи, на примере PE3b (PE3b = слитый белок праймирующего редактора PE2 + PEгРНК + гидовая РНК для разрыва второй цепи). Временный разрыв второй цепи представляет собой вариант разрыва второй цепи для облегчения образования желаемого отредактированного продукта. Термин «временной» относится к тому факту, что разрыв второй цепи с неотредактированной цепью происходит только после того, как желаемое редактирование установлено в отредактированной цепи. Это позволяет избежать одновременных разрывов на обеих цепях, которые приводят к разрывам двухцепочечной ДНК.

[0102] На **фиг. 1J-1K** изображен вариант праймированного редактирования, рассматриваемого в настоящем документе, который заменяет napDNAbp (например, никазу SpCas9) любым программируемым нуклеазным доменом, таким как нуклеазы цинковых пальцев (ZFN) или эффекторные нуклеазы, подобные активаторам транскрипции (TALEN). Таким образом, предполагается, что подходящие нуклеазы не обязательно должны быть «запрограммированы» молекулой нацеливания нуклеиновой кислоты (такой как гидовая РНК), а скорее могут быть запрограммированы путем определения специфичности ДНК-связывающего домена, такого как и в частности нуклеаза. Так же, как и при праймированном редактировании с фрагментами napDNAbp, предпочтительно, чтобы

такие альтернативные программируемые нуклеазы были модифицированы таким образом, чтобы разрезалась только одна цепь ДНК-мишени. Другими словами, программируемые нуклеазы предпочтительно должны функционировать как никазы. После выбора программируемой нуклеазы (например, ZFN или TALEN) в систему могут быть встроены дополнительные функциональные возможности, позволяющие ей работать в соответствии с механизмом, подобным праймированному редактированию. Например, программируемые нуклеазы можно модифицировать путем присоединения (например, через химический линкер) к ним удлиняющего плеча РНК или ДНК, где удлиняющее плечо содержит участок связывания праймера (PBS) и матрицу для синтеза ДНК. Программируемая нуклеаза также может быть связана (например, через химический или аминокислотный линкер) с полимеразой, природа которой будет зависеть от того, представляет ли собой удлиняющее плечо ДНК или РНК. В случае удлиняющего плеча РНК полимераза может представлять собой РНК-зависимую ДНК-полимеразу (например, обратную транскриптазу). В случае удлиняющего плеча ДНК полимераза может быть ДНК-зависимой ДНК-полимеразой (например, прокариотической полимеразой, включая Pol I, Pol II или Pol III, или эукариотической полимеразой, включая Pol a, Pol b, Pol g, Pol d, Pol e или Pol z). Система может также включать другие функции, добавленные в виде слияний с программируемыми нуклеазами или добавленные *в транс* для облегчения реакции в целом (например, (a) геликазу для раскручивания ДНК в месте разреза, чтобы сделать разрезанную цепь с 3'-концом, доступным в качестве праймера, (b) лоскутную эндонуклеазу (например, FEN1), помогающую удалить эндогенную цепь на разрезанной цепи, чтобы запустить реакцию в направлении замены эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:гРНК для создания второго участка разрыва на противоположной цепи, что может способствовать интеграции репарации синтеза посредством предпочтительной клеточной репарации неотредактированной цепи). Аналогично праймированному редактированию с помощью napDNAbp, такой комплекс с программируемой нуклеазой можно использовать для синтеза, а затем встраивания вновь синтезированной замещающей цепи ДНК, несущей представляющее интерес редактирование, перманентно, в целевой участок ДНК.

[0103] На **фиг. 1L** показаны согласно одному варианту осуществления анатомические признаки ДНК-мишени, которые могут быть отредактированы путем праймированного редактирования. ДНК-мишень включает «нецелевую цепь» и «целевую цепь». Целевая цепь представляет собой цепь, которая отжигается со спейсером РЕгРНК комплекса праймированного редактирования, который распознает сайт PAM (в данном случае NGG, который распознается каноническими праймирующими редакторами на

основе SpCas9). Целевая цепь также можно назваться «цепью без PAM» или «цепью без редактирования». Напротив, нецелевая цепь (то есть цепь, содержащая протоспейсер и PAM-последовательность NGG) может называться «PAM-цепью» или «цепью редактирования». Согласно различным вариантам осуществления сайт разрыва комплекса PE будет находиться в протоспейсере на цепи PAM (например, с PE на основе SpCas9). Расположение разрыва будет характерным для конкретного Cas9, который формирует PE. Например, в случае PE на основе SpCas9 сайт разрыва в фосфодисложноэфирной связи между основаниями три («-3» положение относительно положения 1 последовательности PAM) и четыре («-4» положение относительно положения 1 последовательности PAM). Сайт разрыва в протоспейсере образует свободную 3'-гидроксильную группу, которая, как видно на следующих чертежах, образует комплексы с участком связывания праймера плеча удлинения PEгРНК и обеспечивает субстрат для начала полимеризации одиночной цепи ДНК, кодируемой матрицей синтеза ДНК, плеча удлинения PEгРНК. Эта реакция полимеризации катализируется полимеразой (например, обратной транскриптазой) слитого белка PE в направлении от 5' к 3'. Полимеризация завершается до достижения ядра гРНК (например, посредством включения сигнала терминации полимеризации или вторичной структуры, которая функционирует для прекращения полимеризационной активности PE), образуя лоскут одноцепочечной ДНК, который удлинен от исходной 3'-гидроксильной группы разорванной цепи PAM. Матрица синтеза ДНК кодирует одноцепочечную ДНК, которая гомологична эндогенной одноцепочечной ДНК с 5'-концом, которая следует непосредственно за участком разрыва на цепи PAM и включает желаемое изменение нуклеотида (например, замену, вставку, делецию, инверсию одного основания). Положение желаемого редактирования может быть в любом положении, следующем за сайтом разрыва на цепи PAM, 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 по ходу транскрипции от участка разрыва). Как только 3'-концевая одноцепочечная ДНК (содержащая представляющее интерес редактирование)

заменяет эндогенную 5'-концевую одноцепочечную ДНК, процессы репарации и репликации ДНК приведут к постоянной установке сайта редактирования на цепи PAM, а затем к коррекции несоответствия в цепи, отличной от PAM, которое будет существовать в участке редактирования. Таким образом, редактирование будет распространяться на обе цепи ДНК на целевом участке ДНК. Следует понимать, что ссылка на «отредактированную цепь» и «неотредактированную» цепь предназначена только для обозначения цепей ДНК, участвующих в механизме PE. «Отредактированная цепь» представляет собой цепь, которая сначала подвергается редактированию путем замены одноцепочечной ДНК с 5'-концом непосредственно по ходу транскрипции от участка разрыва на синтезированную одноцепочечную ДНК с 3'-концом, содержащую желаемое редактирование. «Неотредактированная» цепь представляет собой пару цепей с отредактированной цепью, но которая сама также редактируется посредством репарации и/или репликации, чтобы стать комплементарной отредактированной цепью и, в частности, включить представляющее интерес редактирование.

[0104] На **фиг. 1M** изображен механизм праймированного редактирования, показывающий анатомические признаки ДНК-мишени, комплекс праймированного редактора и взаимодействие между PEгРНК и ДНК-мишенью. Во-первых, праймирующий редактор, содержащий слитый белок, содержащий полимеразу (например, обратную транскриптазу) и napDNAbp (например, никазу SpCas9, например, SpCas9, имеющую деактивирующую мутацию в нуклеазном домене HNH (например, H840A) или деактивирующую мутацию в домен нуклеазы RuvC (D10A)) образует комплекс с PEгРНК и ДНК, имеющей ДНК-мишень, подлежащую редактированию. PEгРНК содержит спейсер, ядро гРНК (также известное как каркас гРНК или основная цепь гРНК) (которое связывается с napDNAbp) и плечо удлинения. Плечо удлинения может быть на 3'-конце, 5'-конце или в любом месте внутри молекулы PEгРНК. Как показано, плечо удлинения находится на 3'-конце PEгРНК. Плечо удлинения содержит в направлении от 3' к 5' участок связывания праймера и матрицу для синтеза ДНК (содержащую как представляющее интерес редактирование, так и области гомологии (т.е. плечи гомологии), которые гомологичны 5'-концу одноцепочечной ДНК непосредственно после сайта разрыва на цепи PAM. Как показано, как только разрыв вводится, в результате чего образуется свободная 3'-гидроксильная группа непосредственно против хода транскрипции от участка разрыва, область непосредственно против хода транскрипции от участка разрыва на цепи PAM отжигается с комплементарной последовательностью при 3'-концк плеча удлинения, называемого «участком связывания праймера», создавая короткую двухцепочечную область с доступным 3'-гидроксильным концом, который образует субстрат для

полимеразы комплекса праймированного редактирования. Полимераза (например, обратная транскриптаза), затем полимеризует в виде цепи ДНК от 3'-гидроксильного конца до конца плеча удлинения. Последовательность одноцепочечной ДНК кодируется матрицей для синтеза ДНК, которая является частью плеча удлинения (т.е. за исключением участка связывания праймера), которое «считывается» полимеразой для синтеза новой ДНК. Эта полимеризация эффективно удлиняет последовательность исходного 3'-гидроксильного конца начального участка разрыва. Матрица для синтеза ДНК кодирует одну цепь ДНК, которая включает не только желаемое редактирование, но и области, гомологичные эндогенной одиночной цепи ДНК, расположенные непосредственно по ходу транскрипции от участка разрыва на цепи PAM. Затем кодируемая 3'-концевая одноцепочечная ДНК (т.е. 3'-концевой одноцепочечный лоскут ДНК) вытесняет соответствующую гомологичную эндогенную 5'-концевую одноцепочечную ДНК непосредственно по ходу транскрипции от участка разрыва на цепи PAM, образуя промежуточную ДНК, имеющую 5'-концевой одноцепочечный лоскут ДНК, который удаляется клеткой (например, с помощью лоскутной эндонуклеазы). 3'-концевой одноцепочечный лоскут ДНК, который гибридизуется с комплементом эндогенного 5'-концевого одноцепочечного лоскута ДНК, лигируют с эндогенной цепью после удаления 5'-лоскута ДНК. Желаемое редактирование на 3'-концевом одноцепочечном лоскуте ДНК, теперь подвергнутом отжигу и лигированию, образует несовпадение с цепью комплемента, которая подвергается репарации ДНК и/или циклу репликации, таким образом постоянно устанавливая желаемую редактирование на обеих цепях.

[0105] На **фиг. 2** показаны три комплекса Cas (SpCas9, SaCas9 и LbCas12a), которые можно использовать в описанных в настоящем документе праймирующих редакторах, и их особенности расщепления PAM, гРНК и ДНК. На чертеже показаны конструкции комплексов, включающих SpCas9, SaCas9 и LbCas12a.

[0106] На **фиг. 3A-3F** показаны конструкции сконструированной гРНК праймирующего редактора на 5'-конце (**фиг. 3A**), гРНК праймирующего редактора на 3'-конце (**фиг. 3B**) и внутримолекулярного удлинения (**фиг. 3C**). Удлиненная гидовая РНК (или удлиненная гРНК) также может упоминаться в настоящем документе как РЕгРНК или «гидовая РНК праймированного редактирования». На **фиг. 3D** и **фиг. 3E** представлены дополнительные варианты осуществления гРНК праймирующего редактора на 3'- и 5'-конце (РЕгРНК), соответственно. На **фиг. 3F** проиллюстрировано взаимодействие между гидовой РНК праймирующего редактор на 3'-конце и целевой последовательностью ДНК. Варианты осуществления на **фиг. 3A-3C** показывают иллюстративные расположения последовательности матрицы обратной транскрипции (т.е. или в более широком смысле

матрицы для синтеза ДНК, как указано, поскольку RT представляет собой только один тип полимеразы, который можно использовать в контексте праймирующих редакторов), участок связывания праймера и необязательную линкерную последовательность в удлиненных частях 3'-, 5'- и внутримолекулярных вариантов, а также общее расположение спейсерных и ядерных областей. Описанный способ праймированного редактирования не ограничивается этими конфигурациями удлиненных гидовых РНК. Вариант осуществления согласно **фиг. 3D** показывает структуру иллюстративной РЕгРНК, рассматриваемой в настоящем документе. РЕгРНК содержит три основных составных элемента, упорядоченных в направлении от 5' к 3', а именно: спейсер, ядро гРНК и плечо удлинения на 3'-конце. Удлинение может быть дополнительно разделено на следующие структурные элементы в направлении от 5' к 3', а именно: участок связывания праймера (A), матрица редактирования (B) и плечо гомологии (C). Кроме того, РЕгРНК может содержать необязательную область модификатора 3'-конца (e1) и необязательную область модификатора 5'-конца (e2). Кроме того, РЕгРНК может содержать сигнал терминации транскрипции на 3'-конце РЕгРНК (не показано). Эти структурные элементы дополнительно определены в настоящем документе. Изображение структуры РЕгРНК не является ограничивающим и охватывает варианты расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми другими показанными областями и без ограничения расположены на 3'- и 5'-концах. Согласно определенным вариантам осуществления РЕгРНК может содержать вторичную структуру РНК, такую как без ограничения шпильки, петля на стебле, петли, домены рекрутирования РНК-связывающего белка (например, аптамер MS2, который рекрутирует и связывается с белком MS2cp). Например, такие вторичные структуры могут располагаться внутри спейсера, ядра гРНК или плеча удлинения и, в частности, внутри областей модификатора e1 и/или e2. В дополнение к вторичным структурам РНК РЕгРНК могут содержать (например, в областях модификатора e1 и/или e2) химический линкер или поли(N)-линкер или хвост, где «N» может быть любым нуклеотидным основанием. Согласно некоторым вариантам осуществления (например, как показано на фиг. 72(c)), химический линкер может предотвращать обратную транскрипцию каркаса или ядра сгРНК. Кроме того, согласно определенным вариантам осуществления (например, см. фиг. 72(c)), плечо удлинения (3) может состоять из РНК или ДНК и/или может включать один или несколько аналогов нуклеотидных оснований (например, которые могут добавлять функциональность, например, термостойкость). Более того, ориентация плеча удлинения (3) может быть в естественном направлении от 5' к 3' или синтезирована в противоположной ориентации в направлении от 3' к 5' (относительно

ориентации молекулы РЕгРНК в целом). Также следует отметить, что специалист в данной области способен выбрать подходящую ДНК-полимеразу в зависимости от природы материалов нуклеиновой кислоты в плече удлинения (т.е. ДНК или РНК) для использования в праймированном редактировании, которое может быть реализовано либо в виде слияния с napDNAbp, либо представлено *в транс* в виде отдельного фрагмента для синтеза желаемого кодируемого матрицей 3'-лоскута одноцепочечной ДНК, который включает желаемое редактирование. Например, если плечо удлинения представляет собой РНК, то ДНК-полимераза может быть обратной транскриптазой или любой другой подходящей РНК-зависимой ДНК-полимеразой. Однако, если плечо удлинения представляет собой ДНК, то ДНК-полимераза может быть ДНК-зависимой ДНК-полимеразой. Согласно различным вариантам осуществления ДНК-полимераза может находиться *в транс*, например, посредством использования домена рекрутирования РНК-белок (например, шпильки MS2, установленной на РЕгРНК (например, в области e1 или e2 или в другом месте, и белок MS2cp, слитый с ДНК-полимеразой, тем самым совместно локализуя ДНК-полимеразу с РЕгРНК. Также необходимо отметить, что участок связывания праймера обычно не образует часть матрицы, используемой ДНК-полимеразой (например, обратной транскриптазой) для кодирования полученного 3'-лоскута одноцепочечной ДНК, который включает желаемое редактирование. Таким образом, обозначение «матрица для синтеза ДНК» относится к области или части плеча удлинения (3), которое используется ДНК-полимеразой в качестве матрицы для кодирования желаемого 3'-лоскута одноцепочечной ДНК, содержащего редактирование и участки гомологии с 5'-эндогенным одноцепочечным лоскутом ДНК, который замещен 3'-цепью одноцепочечной ДНК, продуктом синтеза ДНК с праймированным редактированием. Согласно некоторым вариантам осуществления матрица для синтеза ДНК включает «матрицу редактирования» и «плечо гомологии», или одно или несколько плечей гомологии, например, до и после матрицы редактирования. Матрица редактирования может представлять собой замену в пределах одного нуклеотида или может представлять собой вставку или инверсию ДНК. Кроме того, матрица редактирования может также включать делецию, которая может быть сконструирована путем кодирования плеча гомологии, содержащей желаемую делецию. Согласно другим вариантам осуществления матрица для синтеза ДНК может также включать область e2 или ее часть. Например, если область e2 содержит вторичную структуру, которая вызывает прекращение активности ДНК-полимеразы, то возможно, что функция ДНК-полимеразы будет прекращена до того, как какая-либо часть области e2 будет фактически закодирована в ДНК. Также возможно, что часть или даже вся область e2 будет закодирована в ДНК. Как много e2 фактически используется в качестве матрицы, будет зависеть от ее состава и от

того, прерывает ли этот состав функцию ДНК-полимеразы.

[0107] Вариант осуществления согласно **фиг. 3E** показывает структуру другой охватываемой настоящим изобретением РЕгРНК. РЕгРНК содержит три основных составных элемента, упорядоченных в направлении от 5' к 3', а именно спейсер, ядро гРНК 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), матрица редактирования (B) и плечо гомологии (C). Кроме того, РЕгРНК может содержать необязательную область модификатора 3'-конца (e1) и необязательную область модификатора 5'-конца (e2). Кроме того, РЕгРНК может содержать сигнал терминации транскрипции на 3'-конце РЕгРНК (не показан). Эти структурные элементы дополнительно определены в настоящем документе. Изображение структуры РЕгРНК не является ограничивающим и охватывает варианты расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми другими показанными областями и без ограничения они расположены на 3'- и 5'-концах. Согласно определенным вариантам осуществления РЕгРНК может содержать вторичные РНК-структуры, такие как без ограничения шпильки, петля на стебле, петли, домены рекрутирования РНК-связывающего белка (например, аптамер MS2, который рекрутирует и связывается с белком MS2cp). Эти вторичные структуры могут располагаться в любом месте в молекуле РЕгРНК. Например, такие вторичные структуры могут располагаться внутри спейсера, ядра гРНК или плеча удлинения и, в частности, внутри областей модификатора e1 и/или e2. В дополнение к вторичным структурам РНК РЕгРНК могут содержать (например, в областях модификатора e1 и/или e2) химический линкер или поли(N)-линкер или хвост, где «N» может быть любым нуклеотидным основанием. Согласно некоторым вариантам осуществления (например, как показано на фиг. 72(c)), химический линкер может предотвращать обратную транскрипцию каркаса или ядра сгРНК. Кроме того, согласно определенным вариантам осуществления (например, см. фиг. 72(c)), плечо удлинения (3) может состоять из РНК или ДНК и/или может включать один или несколько аналогов нуклеотидных оснований (например, которые могут добавлять функциональность, например, термостойкость). Более того, ориентация плеча удлинения (3) может быть в естественном направлении от 5' к 3' или синтезирована в противоположной ориентации в направлении от 3' к 5' (относительно ориентации молекулы РЕгРНК в целом). Также следует отметить, что специалист в данной области может выбрать подходящую ДНК-полимеразу в зависимости от природы материалов нуклеиновой кислоты в плече удлинения (т.е. ДНК или РНК) для использования в праймированном редактировании, которое может быть реализовано либо в виде слияния с napDNAbp, либо

представлено в транс в виде отдельного фрагмента для синтеза желаемого кодируемого матрицей 3'-лоскута одноцепочечной ДНК, который включает желаемое редактирование. Например, если плечо удлинения представляет собой РНК, то ДНК-полимераза может быть обратной транскриптазой или любой другой подходящей РНК-зависимой ДНК-полимеразой. Однако, если плечо удлинения представляет собой ДНК, то ДНК-полимераза может быть ДНК-зависимой ДНК-полимеразой. Согласно различным вариантам осуществления ДНК-полимераза может находиться в транс, например, посредством использования домена рекрутирования РНК-белок (например, шпильки MS2, установленной на РЕгРНК (например, в области e1 или e2 или в другом месте, и белок MS2cp, слитый с ДНК-полимеразой, тем самым совместно локализующий ДНК-полимеразу с РЕгРНК). Также отмечено, что участок связывания праймера обычно не образует часть матрицы, используемой ДНК-полимеразой (например, обратной транскриптазой) для кодирования полученного 3'-лоскута одноцепочечной ДНК, который включает желаемое редактирование. Таким образом, обозначение «матрица для синтеза ДНК» относится к области или части плеча удлинения (3), которые используются ДНК-полимеразой в качестве матрицы для кодирования желаемого 3'-лоскута одноцепочечной ДНК, содержащего редактирование и участки гомологии с 5'-эндогенным лоскутом одноцепочечной ДНК, который заменен 3'-цепью одноцепочечной ДНК, продуктом синтеза ДНК с праймированным редактированием. Согласно некоторым вариантам осуществления матрица для синтеза ДНК включает «матрицу редактирования» и «плечо гомологии», или одно или несколько плечей гомологии, например, до и после матрицы редактирования. Матрица редактирования может представлять собой замену в пределах одного нуклеотида или может представлять собой вставку или инверсию ДНК. Кроме того, матрица редактирования может также включать делецию, которая может быть сконструирована путем кодирования плеча гомологии, содержащей желаемую делецию. Согласно другим вариантам осуществления матрица для синтеза ДНК может также включать область e2 или ее часть. Например, если область e2 содержит вторичную структуру, которая вызывает прекращение активности ДНК-полимеразы, то возможно, что функция ДНК-полимеразы будет прекращена до того, как какая-либо часть области e2 будет фактически закодирована в ДНК. Также возможно, что часть или даже вся область e2 будет закодирована в ДНК. Как много e2 фактически используется в качестве матрицы, будет зависеть от ее состава и от того, прерывает ли этот состав функцию ДНК-полимеразы.

[0108] Схематическое изображение на **фиг. 3F** показывает взаимодействие типичной РЕгРНК с целевым участком двухцепочечной ДНК и сопутствующее образование 3'- одноцепочечного лоскута ДНК, содержащего представляющее интерес

генетическое изменение. Двухцепочечная ДНК показана с верхней цепью (т.е. целевой цепью) в направлении от 3' к 5', а нижней цепью (т.е. цепью PAM или нецелевой цепью) в направлении от 5' к 3'. Верхняя цепь содержит комплемент «протоспейсера» и комплемент последовательности PAM и называется «целевой цепью», поскольку это цепь, которая является мишенью для спейсера РЕгРНК и отжигается со спейсером. Комплементарная нижняя цепь называется «нецелевой цепью» или «цепью PAM», или «цепью протоспейсера», поскольку она содержит последовательность PAM (например, NGG) и протоспейсер. Хотя это не показано, изображенная РЕгРНК будет образовывать комплекс с Cas9 или эквивалентным доменом слитого белка праймирующего редактора. Как показано на схеме, спейсер РЕгРНК отжигается с комплементарной областью протоспейсера на целевой цепи. Это взаимодействие образуется как гибрид ДНК/РНК между спейсерной РНК и комплементом ДНК протоспейсера и индуцирует образование R-петли в протоспейсере. Как указано в другом месте настоящего документа, белок Cas9 (не показан) затем индуцирует разрыв в нецелевой цепи, как показано. Затем это приводит к образованию 3'-лоскута одноцепочечной ДНК непосредственно против хода транскрипции от участка разрыва, который, в соответствии с *z*, взаимодействует с 3'-концом РЕгРНК в участке связывания праймера. 3'-конец лоскута одноцепочечной ДНК (т.е. последовательность праймера обратной транскриптазы) отжигается с участком связывания праймера (A) на РЕгРНК, тем самым праймируя обратную транскриптазу. Затем обратная транскриптаза (например, представленная *в транс-* или *цис–* как слитый белок, присоединенный к конструкции Cas9) затем полимеризует одну цепь ДНК, которая кодируется матрицей для синтеза ДНК (включая матрицу редактирования (B) и плечо гомологии (C)). Полимеризация продолжается в направлении к 5'-концу плеча удлинения. Полимеризованная цепь оцДНК (одноцепочечной ДНК) образует 3'-концевой лоскут оцДНК, который, как описано в другом месте (например, как показано на **фиг. 1G**), включается в эндогенную ДНК, замещая соответствующую эндогенную цепь (которая удаляется как 5'-концевой лоскут ДНК эндогенной ДНК) и вводя желаемое редактирования нуклеотидов (изменение одной пары нуклеотидов, делеции, вставки (включая целые гены) с помощью естественных циклов репарации/репликации ДНК.

[0109] На **фиг. 3G** представлен другой вариант осуществления праймированного редактирования, охватываемого настоящим изобретением. В частности, на верхней схеме изображен один из вариантов осуществления праймирующего редактора (PE), который включает слитый белок napDNAbp (например, SpCas9) и полимеразу (например, обратную транскриптазу), которые соединены линкером. PE образует комплекс с РЕгРНК, связываясь с ядром гРНК РЕгРНК. В показанном варианте осуществления РЕгРНК снабжена 3'-плечом

удлинения, которое содержит, начиная с 3'-конца, участок связывания праймера (PBS), за которым следует матрица для синтеза ДНК. На нижней схеме показан вариант праймирующего редактора, называемый «*транс*-праймирующим редактором (tPE)». Согласно этому варианту осуществления матрица для синтеза ДНК и PBS отделены от РЕгРНК и представлены на отдельной молекуле, называемой матрицей РНК *транс*-праймирующего редактора («tPERT»), которая содержит домен рекрутирования РНК-белок (например, шпилька MS2). Само PE дополнительно модифицируют для включения слияния с рекрутирующим белком rPERT («RP»), который представляет собой белок, который специфически распознает и связывается с рекрутинговым доменом РНК-белок. В примере, где рекрутирующий домен РНК-белок представляет собой шпильку MS2, соответствующий рекрутирующий белок rPERT может представлять собой MS2cp системы мечения MS2. Система мечения MS2 основана на естественном взаимодействии белка оболочки бактериофага MS2 («MCP» или «MS2cp») со структурой петля на стебле или шпилька, присутствующей в геноме фага, т.е. «шпилькой MS2» или «аптамером MS2». В случае *транс*-праймированного редактирования комплекс RP-PE:гРНК «рекрутирует» tPERT, имеющий соответствующий домен рекрутирования РНК-белок, для совместной локализации с комплексом PE:гРНК, тем самым обеспечивая PBS и матрицу для синтеза ДНК *в транс* для использования при праймированном редактировании, как показано в примере, изображенном на фиг. 3H.

[0110] На **фиг. 3H** показан способ *транс*-праймированного редактирования. В этом варианте осуществления *транс*-праймирующий редактор содержит праймирующий редактор «PE2» (т.е. слияние Cas9(H840A) и вариант MMLV RT), слитый с белком MS2cp (т.е. тип рекрутирующего белка, который распознает и связывается с аптамером MS2) и который образует комплекс с сгРНК (т.е. со стандартной гидовой РНК в отличие от РЕгРНК). *Транс*-праймирующий редактор связывается с ДНК-мишенью и разрывает нецелевую цепь. Белок MS2cp рекрутирует tPERT *в транс* посредством специфического взаимодействия с рекрутинговым доменом РНК-белок на молекуле tPERT. tPERT становится совместно локализованным с *транс*-праймирующим редактором, тем самым обеспечивая функции PBS и матрицы для синтеза ДНК в транс для использования полимеразой обратной транскриптазы для синтеза лоскута одноцепочечной ДНК, имеющего 3'-конец и содержащего желаемую генетическую информацию, кодируемую матрицей для синтеза ДНК.

[0111] На **фиг. 4A-4E** продемонстрированы *in vitro* TPRT анализы (т.е., анализы праймированного редактирования). На **фиг. 4A** представляет собой схему матричного удлинения гРНК флуоресцентно меченых ДНК-субстратов с помощью фермента RT,

PAGE. На **фиг. 4B** показано TPRT (т.е. праймированное редактирование) с субстратами с предварительным разрывом, dCas9 и 5'-удлиненными гРНК с различной длиной матрицы для синтеза. На **фиг. 4C** показана реакция RT с ДНК-субстратами с предварительным разрывом в отсутствие Cas9. На **фиг. 4D** показана TPRT (т.е. праймированное редактирование) на субстратах полной двухцепочечной ДНК с Cas9(H840A) и 5'-удлиненных гРНК. На **фиг. 4E** показана матрица 3'-удлиненной гРНК с предварительно разрезанными и полными субстратами двухцепочечной ДНК. Все реакции проводят с RT M-MLV.

[0112] На **фиг. 5** показаны валидации *in vitro* с использованием 5'-удлиненных гРНК с матрицами для синтеза различной длины. В качестве субстратов использовали флуоресцентно меченные (Cy5) ДНК-мишени, которые предварительно разрезали в этой серии экспериментов. Cas9, используемый в этих экспериментах, представляет собой каталитически мертвый Cas9 (dCas9), а используемая RT представляет собой Superscript III, коммерческую RT, полученную из вируса мышиного лейкоза Молони (M-MLV). Комплексы dCas9:гРНК формировали из очищенных компонентов. Затем добавляли флуоресцентно меченный ДНК-субстрат вместе с dNTP и ферментом RT. После 1 часа инкубации при 37 °C продукты реакции анализировали с помощью денатурирующего электрофореза в полиакриламидном геле с мочевиной (PAGE). Изображение геля показывает удлинение исходной цепи ДНК до длины, которая соответствует длине матрицы обратной транскрипции.

[0113] На **фиг. 6** показаны валидации *in vitro* с использованием 5'-удлиненных гРНК с матрицами для синтеза различной длины, что близко соответствует матрицам, показанным на **фиг. 5**. Однако в этой серии экспериментов ДНК-субстраты не подвергали предварительному разрыву. Cas9, используемый в этих экспериментах, представляет собой никазу Cas9 (мутант SpyCas9 H840A), а используемая RT представляет собой Superscript III, коммерческую RT, полученную из вируса мышиной лейкемии Молони (M-MLV). Продукты реакции анализировали с помощью денатурирующего электрофореза в полиакриламидном геле с мочевиной (PAGE). Как показано на геле, никаза эффективно расщепляет цепь ДНК при использовании стандартной гРНК (gRNA_0, дорожка 3).

[0114] На **фиг.** 7 показано, что 3'-удлинения поддерживают синтез ДНК и не оказывают существенного влияния на активность никазы Cas9. Предварительно надрезанные субстраты (черная стрелка) практически количественно превращаются в продукты RT при использовании либо dCas9, либо никазы Cas9 (дорожки 4 и 5). Конверсия более 50% в продукт RT (красная стрелка) наблюдается при использовании полных субстратов (дорожка 3). Используют никазу Cas9 (мутант SpyCas9 H840A), каталитически

мертвый Cas9 (dCas9) и Superscript III, коммерческую RT, полученную из вируса мышиной лейкемии Молони (M-MLV).

[0115] На **фиг. 8** продемонстрированы эксперименты с двумя цветами, которые использовались для определения того, происходит ли реакция RT преимущественно с гРНК в цис-положении (связанной в одном комплексе). Были проведены два отдельных эксперимента для 5'-удлиненных и 3'-удлиненных гРНК. Продукты анализировали с помощью PAGE. Соотношение продукта рассчитывали как (Cy3цис/Cy3транс)/(Cy5транс/Cy5цис).

[0116] На **фиг. 9A-9D** продемонстрирована лоскутная модель субстрата. На **фиг. 9A** показан двухлоскутный репортер для лоскут-направленного мутагенеза. На **фиг. 9B** показана репарация стоп-кодона в клетках HEK. На **фиг. 9C** показаны секвенированные клоны дрожжей после репарации лоскута. На **фиг. 9D** показано тестирование различных характеристик лоскута в клетках человека.

[0117] На **фиг. 10** продемонстрировано праймированное редактирование на плазмидных субстратах. Двух-флуоресцентная репортерная плазмида была сконструирована для экспрессии дрожжей (S. cerevisiae). Экспрессия этой конструкции в дрожжах продуцирует только GFP. Реакция праймированного редактирования *in vitro* вводит точечную мутацию и трансформирует исходную плазмиду или плазмиду с разрывом *in vitro* Cas9(H840A) в дрожжи. Колонии визуализируют с помощью флуоресцентной визуализации. Показаны двухлоскутные дрожжевые плазмидные трансформанты. Трансформация исходной плазмиды или плазмиды с разрывом *in vitro* Cas9(H840A) приводит к образованию колоний, экспрессирующих только зеленый GFP. Реакция праймированного редактирования с 5'-удлиненными или 3'-удлиненными гРНК дает смесь зеленых и желтых колоний. Последние экспрессируют как GFP, так и mCherry. Больше желтых колоний наблюдается с 3'-удлиненной гРНК. Также показан положительный контроль, не содержащий стоп-кодона.

[0118] На **фиг. 11** показано праймированное редактирование на плазмидных субстратах, аналогичное эксперименту на **фиг. 10**, но вместо установки точечной мутации в стоп-кодоне при праймированном редактировании устанавливается вставка одного нуклеотида (слева) или делеция (справа), которая приводит к репарации мутации сдвига рамки считывания и позволяет синтезировать mCherry по ходу транскрипции. В обоих экспериментах использовали 3'-удлиненные гРНК.

[0119] На **фиг. 12** показаны продукты праймированного редактирования на плазмидных субстратах, охарактеризованные секвенированием по Сэнгеру. Отдельные колонии из трансформаций TRT отбирали и анализировали с помощью секвенирования по

Сэнгеру. Точные редактирования наблюдали путем секвенирования выбранных колоний. Зеленые колонии содержали плазмиды с исходной последовательностью ДНК, в то время как желтые колонии содержали точную мутацию, созданную с помощью праймированного редактирования гРНК. Других точечных мутаций или вставок-делеций не наблюдали.

[0120] На **фиг. 13** показан потенциальный объем применения новой технологии праймированного редактирования в сравнении с технологиями редактирования оснований, опосредованного дезаминазой.

[0121] На **фиг. 14** показана схема редактирования в клетках человека.

[0122] На **фиг. 15** показано удлинение участка связывания праймера в гРНК.

[0123] На **фиг. 16** показаны усеченные гРНК для нацеливания на соседние участки.

[0124] На **фиг. 17A-17C** представлены графики, отображающие % конверсии T в A на нуклеотиде-мишени после трансфекции компонентов в клетках эмбриональной почки человека (HEK). На **фиг. 17A** показаны данные, которые представляют результаты с использованием N-концевого слияния обратной транскриптазы MLV дикого типа с никазой Cas9(H840A) (линкер из 32 аминокислот). **Фиг. 17B** аналогичен **фиг. 17A**, но для C-концевого слияния фермента RT. **фиг. 17C** аналогичен **фиг. 17A**, но линкер между MLV RT и Cas9 имеет длину 60 аминокислот вместо 32 аминокислот.

[0125] На **фиг. 18** показано редактирование T в A высокой чистоты в сайте HEK3 с помощью высокопроизводительного секвенирования ампликона. Результат анализа секвенирования отображает наиболее распространенные генотипы отредактированных клеток.

[0126] На **фиг. 19** показана эффективность редактирования нуклеотида-мишени (синие столбцы) наряду со скоростью вставки-делеции (оранжевые столбцы). WT относится к ферменту MLV RT дикого типа. Мутантные ферменты (M1 - M4) содержат мутации, перечисленные справа. Скорость редактирования была количественно определена с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[0127] На **фиг. 20** показана эффективность редактирования нуклеотида-мишени, когда одноцепочечный разрыв вводится в комплементарную цепь ДНК вблизи нуклеотида-мишени. Тестировали разрывы на различных расстояниях от нуклеотида-мишени (треугольники). Эффективность редактирования в целевой паре оснований (синие столбцы) показана наряду со скоростью образования вставок-делеций (оранжевые столбцы). Пример «нет» не содержит гидовой РНК для разрыва комплементарной цепи. Скорость редактирования была количественно определена с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[0128] На **фиг. 21** показаны обработанные данные высокопроизводительного

секвенирования, показывающие желаемую мутацию трансверсии T в A и общее отсутствие других основных побочных продуктов редактирования генома.

[0129] На **фиг. 22** представлена схема иллюстративного способа проведения направленного мутагенеза с обратной транскриптазой сниженной точности в локусе-мишени с использованием программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) в комплексе с удлиненной гидовой РНК, т.е. праймированного редактирования с RT сниженной точности. Этот способ можно назвать вариантом праймированного редактирования для направленного мутагенеза. Удлиненная гидовая РНК содержит удлинение на 3'- или 5'-конце гидовой РНК или во внутримолекулярном месте гидовой РНК. На стадии (a) комплекс napDNAbp/гРНК связывается с молекулой ДНК, и гРНК направляет napDNAbp для связывания с локусом-мишенью, подлежащим мутагенезу. На стадии (b) в одну из цепей ДНК целевого локуса вводят разрыв (например, с помощью нуклеазы или химического агента), тем самым создавая доступный 3'-конец в одной из цепей целевого локуса. Согласно определенным вариантам осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, т.е. в цепи, которая не гибридизуется с последовательностью гидовой РНК. На стадии (c) 3'-концевая цепь ДНК взаимодействует с удлиненной частью гидовой РНК, чтобы инициировать обратную транскрипцию. Согласно определенным вариантам осуществления 3'-концевая цепь ДНК гибридизуется со специфической праймирующей последовательностью RT на удлиненной части гидовой РНК. На стадии (d) вводят обратную транскриптазу сниженной точности, которая синтезирует мутагенизированную одноцепочечную ДНК от 3'-конца праймированного сайта к 3'-концу гидовой РНК. Примеры мутаций отмечены звездочкой «*». Это образует лоскут одноцепочечной ДНК, содержащий желаемую мутагенизированную область. На стадии (e) высвобождаются napDNAbp и гидовая РНК. Стадии (f) и (g) относятся к разрешению лоскута одноцепочечной ДНК (содержащего мутагенизированную область) таким образом, чтобы желаемая мутагенизированная область была включена в целевой локус. Этот способ можно направлять к образованию желаемого продукта путем удаления соответствующего 5'-лоскута эндогенной ДНК, который образуется, когда 3'-лоскут одноцепочечной ДНК вторгается и гибридизуется с комплементарной последовательностью на другой цепи. Способ также может быть направлен на образование продукта с разрывом второй цепи, как показано на **фиг. 1F**. После способов эндогенной репарации и/или репликации ДНК мутагенизированная область включается в обе цепи ДНК локуса ДНК.

[0130] На **фиг. 23** приведена схема конструирования гРНК для сокращения последовательностей тринуклеотидных повторов и сокращения тринуклеотидных повторов

с помощью редактирования генома TPRT (т.е. праймированного редактирования). Экспансия тринуклеотидных повторов связана с рядом заболеваний человека, включая болезнь Гентингтона, синдром фрагильной Х-хромосомы и атаксию Фридрейха. Наиболее распространенный тринуклеотидный повтор содержит триплеты CAG, хотя также встречаются триплеты GAA (атаксия Фридрейха) и триплеты CGG (синдром фрагильной Х-хромосомы). Наследование предрасположенности к экспансии или приобретение уже экспансивного родительского аллеля увеличивает вероятность приобретения заболевания. Патогенные расширения тринуклеотидных повторов гипотетически могут быть исправлены с помощью праймированного редактирования. Область против хода транскрипции от области повтора может быть разорвана РНК-управляемой нуклеазой, а затем использована для запуска синтеза новой цепи ДНК, которая содержит здоровое количество повторов (которое зависит от конкретного гена и заболевания). После повторяющейся последовательности добавляется короткий участок гомологии, который соответствует идентичности последовательности, примыкающей к другому концу повтора (красная цепь). Инвазия вновь синтезированной цепи и последующая замена эндогенной ДНК вновь синтезированным лоскутом приводит к аллелю с сокращенными повторами.

[0131] На **фиг. 24** представлена схема, показывающая точную делецию 10 нуклеотидов с праймированным редактированием. Гидовая РНК, нацеленная на локус HEK3, была сконструирована с помощью матрицы обратной транскрипции, которая кодирует 10-нуклеотидную делецию после участка разрыва. Эффективность редактирования в трансфицированных клетках HEK оценивали с помощью секвенирования ампликонов.

[0132] На **фиг. 25** представлена схема, показывающая конструкцию гРНК для пептидного мечения генов в эндогенных геномных локусах и пептидного мечения с помощью редактирования генома TPRT (т.е. праймированного редактирования). Системы мечения FlAsH и ReAsH состоят из двух частей: (1) флуорофорно-мышьяковый зонд и (2) генетически кодируемый пептид, содержащий тетрацистеиновый мотив, примером которого является последовательность FLNCCPGCCMEP (SEQ ID NO: 1). При экспрессии внутри клеток белки, содержащие мотив тетрацистеина, могут быть флуоресцентно помечены флуорофорно-мышьяковыми зондами (см. ссылку J. Am. Chem. Soc., 2002, 124 (21), pp 6063–6076. DOI: 10.1021/ja017687n). В системе «сортировки» используются бактериальные ферменты сортазы, которые ковалентно конъюгируют меченые пептидные зонды с белками, содержащими подходящие пептидные субстраты (см. ссылку: 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-tag (EQKLISEEDL (SEQ ID NO: 6)) обычно используются в качестве эпитопных меток для иммунологических анализов. Pi-кольцо кодирует пептидную последовательность (FCPF (SEQ ID NO: 622)), которая может быть помечена пентафторароматическими субстратами (ссылка: Nat. Chem. 2016 Feb,8(2):120-8. doi: 10.1038/nchem.2413).

[0133] На **фиг. 26A** показано точное включение His_6-метки и FLAG-метки в геномную ДНК. Гидовая РНК, нацеленная на локус HEK3, была сконструирована с помощью матрицы обратной транскрипции, которая кодирует либо вставку His-метки из 18 нуклеотидов, либо вставку FLAG-метки из 24 нуклеотидов. Эффективность редактирования в трансфицированных клетках HEK оценивали с помощью секвенирования ампликонов. Необходимо отметить, что полная 24-нуклеотидная последовательность FLAG-метки находится за пределами рамки просмотра (секвенирование подтвердило полную и точную вставку). На **фиг. 26B** схематично показаны различные применения, включающие мечение белков/пептидов, включая (a) превращение белков в растворимые или нерастворимые, (b) изменение или отслеживание клеточной локализации белка, (c) увеличение периода полужизни белка, (d) облегчение очистки белков и (e) облегчение обнаружения белков.

[0134] На **фиг. 27** показан обзор праймированного редактирования путем установки защитной мутации в *PRNP*, которая предотвращает или останавливает прогрессирование прионной болезни. Последовательности PEгРНК соответствуют остаткам с номерами 1-20 SEQ ID NO: 810 слева (т.е. 5' каркаса сгРНК) и остаткам с номерами 21-43 SEQ ID NO: 810 справа (т.е. 3' конца каркас сгРНК).

[0135] На **фиг. 28A** представляет собой схему основанной на PE вставки последовательностей, кодирующих мотивы РНК. На **фиг. 28B** представлен перечень (не исчерпывающий) некоторых примеров мотивов, которые потенциально могут быть вставлены, и их функций.

[0136] На **фиг. 29A** изображен праймирующий редактор. На **фиг. 29B** показаны возможные модификации геномной, плазмидной или вирусной ДНК под действием PE. На **фиг. 29C** показан пример схемы встраивания библиотеки пептидных петель в указанный белок (в данном случае GFP) посредством библиотеки PEгРНК. На **фиг. 29D** показан пример возможных программируемых делеций кодонов или N- или C-концевых усечений белка с использованием различных PEгРНК. Предсказано, что делеции произойдут с минимальным образованием мутаций сдвига рамки считывания.

[0137] На **фиг. 30** показана возможная схема итеративной вставки кодонов в системе

непрерывной эволюции, такой как PACE.

[0138] На **фиг. 31** представлена сконструированная гРНК, показывающая ядро гРНК, спейсер ~20 нуклеотидов, соответствующий последовательности гена-мишени, матрицу обратной транскрипции с нуклеотидной последовательностью иммуногенного эпитопа и участок связывания праймера, соответствующий последовательности гена-мишени.

[0139] На **фиг. 32** представлена схема, показывающая использование праймированного редактирования в качестве средства для вставки эпитопов известной иммуногенности в эндогенную или чужеродную геномную ДНК, что приводит к модификации соответствующих белков.

[0140] На **фиг. 33** представлена схема, показывающая конструкцию РЕгРНК для вставок последовательности связывания праймеров и вставки связывания праймеров в геномную ДНК с использованием праймированного редактирования для определения нецелевого редактирования. Согласно этому варианту осуществления праймированное редактирование проводится внутри живой клетки, ткани или модели животного. В качестве первой стадии разрабатывается соответствующая РЕгРНК. На верхней схеме показана типичная РЕгРНК, которую можно использовать в этом аспекте. Спейсер в РЕгРНК (обозначенный как «протоспейсер») комплементарен одной из цепей геномной мишени. Комплекс РЕ:РЕгРНК (т.е. комплекс РЕ) устанавливает одноцепочечный 3'-концевой лоскут в участке разрыва, который содержит кодируемую последовательность связывания праймера и область гомологии (кодируемую плечом гомологии РЕгРНК), которая комплементарна область сразу по ходу транскрипции от участка разреза (красный). В результате инвазии лоскута и процессов репарации/репликации ДНК синтезированная цепь включается в ДНК, тем самым устанавливая участок связывания праймера. Этот процесс может происходить в желаемой геномной мишени, а также в других участках генома, которые могут взаимодействовать с РЕгРНК нецелевым образом (т.е. РЕгРНК направляет комплекс РЕ к другим нецелевым участкам из-за комплементарности области спейсера к другим участкам генома, которые не являются предполагаемым участком генома). Таким образом, связывающая последовательность праймера может быть установлена не только в желаемой геномной мишени, но и в нецелевых участках генома в других частях генома. Чтобы обнаружить вставку этих участков связывания праймеров как в предполагаемые целевые участки генома, так и в нецелевые участки генома, геномную ДНК (пост-РЕ) можно выделить, фрагментировать и лигировать с адаптерными нуклеотидами (показаны красным). Затем может быть проведена ПЦР с олигонуклеотидами ПЦР, которые отжигаются с адаптерами и со встроенной последовательностью связывания праймера для

амплификации целевых и нецелевых участков геномной ДНК, в которые с помощью PE был вставлен участок связывания праймера. Затем можно проводить высокопроизводительное секвенирование и выравнивать последовательности для идентификации точек вставки последовательностей связывания PE-вставленных праймеров либо в участке-мишени, либо в нецелевых участках.

[0141] На **фиг. 34** представлена схема, показывающая точную вставку гена с PE.

[0142] На **фиг. 35A** представлена схема, показывающая естественный путь передачи сигнала инсулина. На **фиг. 35B** представлена схема, показывающая активацию меченного FKBP12 рецептора инсулина, контролируемую FK1012.

[0143] На **фиг. 36** показаны низкомолекулярные мономеры. Ссылки: миметик FK506 (**2**)[107]

[0144] На **фиг. 37A-37B** показаны низкомолекулярные димеры. Ссылки: FK1012 4[95,96], FK1012 5[108], FK1012 6[107], AP1903 7[107], димер циклоспорина A 8[98], димер FK506–циклоспорина A (FkCsA) 9[100].

[0145] На **фиг. 38A-38F** представлен обзор праймированного редактирования и исследований *in vitro* и на дрожжевых клетках. На **фиг. 38A** показаны 75122 известных патогенных генетических варианта человека в ClinVar (по состоянию на июль 2019 г.), классифицированных по типам. На **фиг. 38B** показано, что комплекс праймированного редактирования состоит из белка праймирующего редактора (PE), содержащего домен разрыва ДНК, управляемый РНК, такой как никаза Cas9, слитый со сконструированным доменом обратной транскриптазы и образующий комплекс с гидовой РНК праймированного редактирования (PEгРНК). Комплекс PE:PEгРНК связывается с целевым участком ДНК и позволяет выполнять большое количество точных изменений ДНК в широком диапазоне положений ДНК до или после мотива, примыкающего к протоспейсеру (PAM) целевого участка. На **фиг. 38C** показано, что при связывании с ДНК-мишенью комплекс PE:PEгРНК разрывает цепь ДНК, содержащую PAM. Образовавшийся свободный 3'-конец гибридизуется с участком связывания праймеров PEгРНК. Домен обратной транскриптазы катализирует удлинение праймера с использованием матрицы RT PEгРНК, в результате чего образуется вновь синтезированная цепь ДНК, содержащая желаемое редактирование (3'-лоскут). Уравновешивание между отредактированным 3'-лоскутом и неотредактированным 5'-лоскутом, содержащим исходную ДНК, с последующим расщепление и лигирование клеточного 5'-лоскута, а также репарация или репликация ДНК для разрешения гетеродуплексной ДНК приводит к стабильно отредактированной ДНК. На **фиг. 38D** показаны анализы удлинения праймера 5'-удлиненной PEгРНК *in vitro* с субстратами двухцепочечной ДНК с предварительным разрывом, содержащими 5'-Cy5-

меченые цепи PAM, dCas9 и коммерческий вариант M-MLV RT (RT, индекс III). dCas9 была в комплексе с РЕгРНК, содержащими RT-матрицу различной длины, затем добавляли к ДНК-субстратам вместе с указанными компонентами. Реакции инкубировали при 37 °C в течение 1 часа, затем анализировали с помощью денатурирующего PAGE с мочевиной и визуализировали на флуоресценцию Cy5. На **фиг. 38E** показаны анализы удлинения праймера, выполненные, как на **фиг. 38D**, с использованием 3'-удлиненных РЕгРНК, предварительно образовавших комплекс с никазой dCas9 или Cas9 H840A, и субстратов двухцепочечной ДНК, меченных 5'-Cy5 с предварительным разрывом или без него. На **фиг. 38F** показаны дрожжевые колонии, трансформированные репортерными плазмидами слияния GFP-mCherry, отредактированными *in vitro* с помощью РЕгРНК, никазы Cas9 и RT. Плазмиды, содержащие несмысловые мутации или мутации со сдвигом рамки между GFP и mCherry, редактировали с помощью 5'-удлиненных или 3'-удлиненных РЕгРНК, которые восстанавливали трансляцию mCherry посредством мутации трансверсии, вставки 1 п.о. или делеции 1 п.о. Двойные положительные клетки GFP и mCherry (желтые) отражают успешное редактирование.

[0146] На **фиг. 39A-39D** показано праймированное редактирование геномной ДНК в клетках человека с помощью PE1 и PE2. На **фиг. 39A** показано, что РЕгРНК содержат спейсерную последовательность, каркас сгРНК и 3'-удлинение, содержащее участок связывания праймера (зеленый) и матрицу обратной транскрипции (RT) (фиолетовый), которая содержит отредактированное (отредактированные) основание (основания) (красный). Участок связывания праймера гибридизуется с цепью ДНК, содержащей PAM, непосредственно против хода транскрипции от участка разрыва. Матрица RT гомологична последовательности ДНК по ходу транскрипции от разрыва, за исключением кодируемого редактирования. На **фиг. 39B** показана установка редактирования трансверсии T•A в A•T в участке HEK3 в клетках HEK293T с использованием никазы Cas9 H840A, слитой с обратной транскриптазой M-MLV дикого типа (PE1) и РЕгРНК с участками связывания праймеров различной длины. На **фиг. 39** C показано использование сконструированной пентамутантной обратной транскриптазы M-MLV (D200N, L603W, T306K, W313F, T330P) в PE2, что существенно улучшает эффективность трансверсии при праймированном редактировании в пяти участках генома в клетках HEK293T, а также редактирования с небольшими вставками и небольшими делециями в HEK3. На **фиг. 39D** приведено сравнение эффективности редактирования PE2 с различной длиной матрицы RT в пяти участках генома в клетках HEK293T. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов.

[0147] На **фиг. 40A-40C** показано, что системы PE3 и PE3b разрывают

неотредактированную цепь с увеличением эффективности праймированного редактирования. На **фиг. 40A** приведен обзор праймированного редактирование посредством PE3. После первоначальный синтез отредактированной цепи, репарация ДНК удаляет либо вновь синтезированную цепь, содержащую редактирование (вырезание 3′-лоскута) или исходную цепь геномной ДНК (вырезание 5′-лоскута). Вырезание 5′-лоскута оставляет ДНК гетеродуплекс, содержащий одну отредактированную цепь, а также одну неотредактированную цепь. Несовпадающая репарация или репликация ДНК могут привести к разрешению гетеродуплекса к получением либо отредактированных, либо не отредактированных продуктов. Разрыв неотредактированной цепи способствует репарации этой цепи, в результате чего предпочтительно образуется стабильная дуплексная ДНК, содержащая желаемое редактирование. На **фиг. 40B** показано влияние разрыва комплементарной цепи на эффективность PE3-опосредованного праймированного редактирования и образование вставок-делеций. «Нет» относится к PE2 контролям, которые не приводят к разрыву комплементарной цепи. На **фиг. 40C** приведено сравнение эффективности редактирования с PE2 (нет разрыва комплементарной цепи), PE3 (разрыв общей комплементарной цепи) и PE3b ((разрыв специфической для редактирования комплементарной цепи). Все результаты редактирования отражают процент от общего числа чтений секвенирования, которые содержат предполагаемое редактирование и не содержат вставок среди всех обработанных клеток, без сортировки. Значения и ошибка отражают среднее и стандартное отклонение для трех независимых биологических репликатов.

[0148] На **фиг. 41A-41K** показаны целевые вставки, делеции и все 12 типов точечных мутаций с PE3 в семи эндогенных геномных локусах человека в клетках HEK293T. На **фиг. 41A** приведен график, показывающий все 12 типов однонуклеотидных редактирований переходов и трансверсии от положения +1 до +8 (считая положение PEгРНК-индуцированного разрыва между положениями +1 и -1) сайта HEK3 с использованием 10-нуклеотидной матрицы RT. На **фиг. 41B** приведен график, показывающий редактирования трансверсии PE3 длинного диапазона в сайте HEK3 с использованием матрицы RT из 34 нуклеотидов. На **фиг. 41C-41H** приведены графики, показывающие все 12 типов редактирования переходов и трансверсий в различных положениях основного окна редактирования для (**фиг. 41C**) RNF2, (**фиг. 41D**) FANCF, (**фиг. 41E**) EMX1, (**фиг. 41F**) RUNX1, (**фиг. 41G**) VEGFA и (**фиг. 41H**) DNMT1. На **фиг. 41I** приведен график, показывающий целевые вставки 1 и 3 п.о. и делеции 1 и 3 п.о. с PE3 в семи эндогенных геномных локусах. На **фиг. 41J** приведен график, показывающий целевые точные делеции от 5 до 80 п.о. в целевом участке HEK3. На **фиг. 41K** приведен

график, показывающий комбинированное редактирование вставок и делеций, вставок и точечных мутаций, делеций и точечных мутаций и двойных точечных мутаций в трех эндогенных геномных локусах. Все результаты редактирования отражают процент от общего числа чтений секвенирования, которые содержат предполагаемое редактирование и не содержат вставок среди всех обработанных клеток, без сортировки. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов.

[0149] На **фиг. 42A-42H** показано сравнение праймированного редактирования и редактирования оснований, а также нецелевого редактирования с помощью Cas9 и PE3 на известных нецелевых участках Cas9. На **фиг. 42A** показана общая эффективность редактирования C•G в T•A при одних и 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 делеций от обработок на **ФИГ. 42A.** На **фиг. 42C** показана эффективность редактирования точным редактированием C•G в T•A (без побочных редактирований или вставок-делеций) для PE2, PE3, BE2max и BE4max в HEK3, FANCF и EMX1. Для EMX1 также показаны точные комбинированные редактирования PE всех возможных комбинаций преобразования C•G в T•A в трех целевых нуклеотидах. На **фиг. 42D** показана общая эффективность редактирования A•T в G•C для PE2, PE3, ABEdmax и ABEmax в HEK3 и FANCF. На **фиг. 42E** показана точная эффективность редактирования A•T в G•C без побочных редактирований или вставок-делеций для HEK3 и FANCF. На **фиг. 42F** показана частота вставок-делеций от обработок на **фиг. 42D**. На **фиг. 42G** показана средняя эффективность редактирования в трех повторах (процент считываний секвенирования с вставками-делециями) в клетках HEK293T для нуклеазы Cas9 в четырех целевых и 16 известных нецелевых сайтах. Проанализированные 16 нецелевых сайтов были первыми четырьмя нецелевыми сайтами, о которых ранее сообщалось[118,159] для каждого из четырех целевых участков. Для каждого целевого участка Cas9 был соединен с сгРНК или с каждой из четырех РЕгРНК, которые распознают один и тот же протоспейсер. На **фиг. 42H** показаны средние показатели эффективности редактирования «целевого» и «нецелевого» и эффективности исправлений с тремя повторениями (ниже в скобках) в клетках HEK293T для PE2 или PE3, в паре с каждой РЕгРНК (**фиг. 42G**). Выходы при целевом редактировании отражают процент от общего числа чтений секвенирования, которые содержат предполагаемое редактирование и не содержат вставок-делеций среди всех обработанных клеток без сортировки. Выходы нецелевого редактирования отражают модификацию нецелевого локуса в соответствии с праймированным редактированием. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых

биологических репликатов.

[0150] На **фиг. 43A-43I** показано праймированное редактирование в различных линиях клеток человека и первичных кортикальных нейронах мыши, установка и коррекция патогенных мутаций трансверсий, вставок или делеций, а также сравнение праймированного редактирования и HDR. На **фиг. 43A** приведен график, показывающий установку (посредством трансверсии T•A-в-A•T) и коррекцию (посредством трансверсии A•T-в-T•A) патогенной мутации E6V в HBB в клетках HEK293T. Показана коррекция либо в HBB дикого типа, либо в HBB, содержащий молчащую мутацию, которая разрушает PAM PEгРНК. На **фиг. 43B** приведен график, показывающий установку (посредством вставки 4 п.о.) и коррекцию (посредством делеции 4 п.о.) патогенного аллеля HEXA 1278+TATC в клетках HEK293T. Показана коррекция либо к HEXA дикого типа, либо к HEXA, содержащей молчащую мутацию, которая разрушает PAM PEгРНК. На **фиг. 43C** представлен график, показывающий установку защитного варианта G127V в PRNP в клетках HEK293T посредством трансверсии G•C-в-T•A. На **фиг. 43D** представлен график, показывающий праймированное редактирование в других линиях клеток человека, включая K562 (лейкемические клетки костного мозга), U2OS (клетки остеосаркомы) и HeLa (клетки рака шейки матки). На **фиг. 43E** представлен график, показывающий установку трансверсионной мутации G•C-в-T•A в DNMT1 праймированных кортикальных нейронов мыши с использованием лентивирусной системы с PE3двойным расщепленным интеином, в которой N-концевая половина представляет собой Cas9 (1-573), слитый с N-интеином и через саморасщепляющийся пептид P2A с GFP-KASH, а C-концевая половина представляет собой C-интеин, слитый с остатком PE2. Половины PE2 экспрессируются промотором синапсинов человека, который является высокоспецифичным для зрелых нейронов. Отсортированные значения отражают редактирование или вставки-делеции из GFP-положительных ядер, тогда как несортированные значения относятся ко всем ядрам. На **фиг. 43F** приведено сравнение эффективности редактирования HDR, опосредованного PE3 и Cas9, в эндогенных геномных локусах в клетках HEK293T. На **фиг. 43G** приведено сравнение эффективности редактирования HDR, опосредованного PE3 и Cas9, в эндогенных геномных локусах в клетках K562, U2OS и HeLa. На **фиг. 43H** приведено сравнение образования побочных продуктов инделирования HDR, опосредованного PE3 и Cas9, в клетках HEK293T, K562, U2OS и HeLa. На **фиг. 43I** показана направленная вставка метки His6 (18 п.о.), эпитопной метки FLAG (24 п.о.) или удлиненного сайта LoxP (44 п.о.) в клетки HEK293T с помощью PE3. Все результаты редактирования отражают процент от общего числа чтений секвенирования, которые содержат предполагаемое редактирование и не содержат вставок-делеций среди всех обработанных клеток. Значения и планки

погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов.

[0151] На **фиг. 44A-44G** показаны валидационные исследования праймированного редактирования *in vitro* с флуоресцентно меченными ДНК-субстратами. На **фиг. 44A** показаны анализы сдвига электрофоретической подвижности с dCas9, 5'-удлиненными РЕгРНК и 5'-Cy5-мечеными ДНК-субстратами. РЕгРНК 1 - 5 содержат линкерную последовательность из 15 нуклеотидов (линкер A для РЕгРНК 1, линкер B для РЕгРНК 2 - 5) между спейсером и PBS, последовательность PBS из 5 нуклеотидов и матрицы RT из 7 нуклеотидов (РЕгРНК 1 и 2), 8 нуклеотидов (РЕгРНК 3), 15 нуклеотидов (РЕгРНК 4) и 22 нуклеотидов (РЕгРНК 5). РЕгРНК используют на **фиг. 44E** и **44F**, полные последовательности перечислены в таблицах 2A-2C. На **фиг. 44B** показаны анализы *in* vitro разрыва для Cas9 H840A с использованием 5'-удлиненных и 3'-удлиненных РЕгРНК. На **фиг. 44C** показано Cas9-опосредованное образование вставок-делеций в клетках HEK293T в HEK3 с использованием 5'-удлиненных и 3'-удлиненных РЕгРНК. На **фиг. 44D** показан обзор биохимических анализов *in vitro* с праймированным редактированием. Были протестированы 5'-Cy5-меченые субстраты двухцепочечной ДНК с предварительно внесенным разрывом и без разрыва. сгРНК, 5'-удлиненные РЕгРНК или 3'-удлиненные РЕгРНК были предварительно объединены в комплекс с никазой dCas9 или Cas9 H840A, а затем объединены с субстратом дцДНК, M-MLV RT и dNTP. Реакциям давали возможность протекать при 37 ° C в течение 1 часа перед разделением с помощью денатурирующего PAGE с мочевиной и визуализации с помощью флуоресценции Cy5. На **фиг. 44E** показаны реакции удлинения праймера с использованием 5'-удлиненных РЕгРНК, ДНК-субстратов с предварительным разрывом и dCas9, приводящие к значительному превращению в продукты RT. На **фиг. 44F** показаны реакции удлинения праймера с использованием 5'-удлиненных РЕгРНК, как на **фиг. 44B**, с ДНК-субстратом без разрывов и никазой Cas9 H840A. Выход продукта значительно снижаются по сравнению с субстратом с предварительным разрывом. На **фиг. 44G** показана реакция удлинения праймера *in vitro* с использованием 3'-РЕгРНК, которая приводит к получению одного продукта, выявленного путем денатурирующего PAGE с мочевиной. Полосу продукта RT вырезали, элюировали из геля, затем подвергали присоединение гомополимера с помощью терминальной трансферазы (TdT) с использованием либо dGTP, либо dATP. Продукты с присоединением гомополимера удлиняли поли-T или поли-C праймерами, и полученную ДНК секвенировали. Следы Сэнгера показывают, что три нуклеотида, полученные из каркаса гРНК, были подвергнуты обратной транскрипции (добавлены в качестве последних 3'-нуклеотидов к продукту ДНК). Следует отметить, что в экспериментах по

праймированному редактированию клеток млекопитающих вставка каркаса РЕгРНК происходит гораздо реже, чем in vitro (**фиг. 56A-56D**), возможно, из-за неспособности связанной обратной транскриптазы получить доступ к Cas9-связанному каркасу гидовой РНК и/или клеточного вырезания несовпадающих 3'-концов 3'-лоскутов, содержащих каркасные последовательности РЕгРНК.

[0152] На **фиг.45A-45C** показана клеточная репарация 3'-лоскутов ДНК в дрожжах из *in vitro* реакций праймированного 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 праймированного редактирования осуществляли *in vitro с* использованием никазы Cas9 H840A, РЕгРНК, 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 праймированного редактирования. На **фиг.45C** показана коррекция стоп-кодона в виде трансверсии T•A в A•T с использованием 3'-удлиненной РЕгРНК или 5'-удлиненной РЕгРНК, как показано на **фиг. 45D**. **FIG. 45E** показана коррекция +1 сдвига рамки считывания в виде делеции 1 п.о. с использованием 3'-удлиненной РЕгРНК. На **фиг. 45F** показана коррекция -1 сдвига рамки считывания в виде вставки 1 п.о. с использованием 3'-удлиненной РЕгРНК. 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**.

[0153] 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 в A•T и образование инделей при использовании PE1 в положении +1 НЕК3 при использовании РЕгРНК, содержащих 10-нуклеотидной матрицы RT и последовательности PBS в диапазоне 8-17 нуклеотидов. На **фиг. 46B** показана эффективность редактирования в виде трансверсии G•C в T•A и образование

инделей при использовании PE1 в положении +5 EMX1 при использовании РЕгРНК, содержащих 13-нуклеотидной матрицы RT и последовательности PBS в диапазоне 9-17 нуклеотидов. На **фиг. 46C** показана эффективность редактирования в виде трансверсии G•C в T•A и образование инделей при использовании PE1 в положении +5 FANCF при использовании РЕгРНК, содержащих 17-нуклеотидной матрицы RT и последовательности PBS в диапазоне 8-17 нуклеотидов. На **фиг. 46D** показана эффективность редактирования в виде трансверсии G•C в T•A и образование инделей при использовании PE1 в положении +1 RNF2 при использовании РЕгРНК, содержащих 11-нуклеотидной матрицы RT и последовательности PBS в диапазоне 9-17 нуклеотидов. На **фиг. 46E** показана эффективность редактирования в виде трансверсии G•C в T•A и образование инделей при использовании PE1 в положении +2 HEK4 при использовании РЕгРНК, содержащих 13-нуклеотидной матрицы RT и последовательности PBS в диапазоне 7-15 нуклеотидов. На **фиг. 46F** показана опосредованная PE1 +1 делеция T, +1 вставка A, и +1 вставка CTT в участке HEK3 с использованием 13-нуклеотидноой PBS и 10-нуклеотидной матрицы RT. Последовательности РЕгРНК соответствуют тем, которые использованы на **фиг.39C** (см. Таблицы 3A-3R). Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических репликатов.

[0154] На **фиг. 47A-47S** показан анализ вариантов RT M-MLV для праймированного редактирования. На **фиг.47A** показаны аббревиатуры для вариантов праймирующего редактора на данном чертеже. На **фиг.47B** показаны изменения в виде направленной вставки и делеции с использованием PE1 в локусе HEK3. На **фиг.47C- 47H** показано сравнение 18 конструкций праймирующего редактора, содержащих варианты RT M-MLV для их способности вносить изменение в виде +2 трансверсии G•C в C•G в HEK3, как показано на **фиг. 47C**, вставки 24 п.о. FLAG в HEK3 как показано на **фиг. 47D**, редактирования путем трансверсии +1 C•G в A•T в RNF2 как показано на **фиг. 47E**, редактирования путем трансверсии +1 G•C в C•G в EMX1 как показано на **фиг FIG. 47F**, редактирования путем трансверсии +2 T•A в A•T в *HBB* как показано на **фиг. 47G**, и редактирование путем трансверсии +1 G•C в C•G в FANCF как показано на **фиг. 47H**. На **фиг. 47I-47N** показано сравнение четырех конструкций праймирующего 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 вариация +1 T•A в A•T в HEK3. На **фиг. 47P** показана вариация +5 G•C в T•A в EMX1. На **фиг. 47Q** показана вариация +5 G•C в T•A в FANCF. На **фиг. 47R** показана вариация +1 C•G A•T в RNF2. На **фиг. 47S** показана вариация +2 G•C в T•A в

HEK4. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических репликатов.

[0155] На **фиг. 48A-48C** показаны свойства конструкции PBS PEгРНК и последовательности матрицы RT. На **фиг. 48A** показана эффективность опосредованного PE2 редактирования в виде +5 трансверсии G•C в T•A (синяя линия) в VEGFA в клетках HEK293T в виде функции от длины матрицы RT. Для сравнения указаны индели (серая линия). Последовательность ниже графика показывает последний нуклеотид на основе матрицы для синтеза PEгРНК. Выделены G нуклеотиды (которым соответствуют матричные C в PEгРНК), для максимизации эффективности праймированного редактирования следует избегать матрицы RT, оканчивающиеся на C. На **фиг. 48B** показано редактирование в виде трансверсии +5 G•C в T•A и инделей для DNMT1 как на **фиг. 48A**. На **фиг. 48C** показано редактирование в виде трансверсии +5 G•C в T•A и инделей для RUNX1 как на **фиг. 48A**. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических репликатов.

[0156] На **фиг. 49A-49B** показаны эффекты of PE2, PE2 R110S K103L, Cas9 H840A никазы и dCas9 на жизнеспособность клеток. HEK293T клетки трансфицировали плазмидами, кодирующими PE2, R110S K103L, Cas9 H840A никазу или dCas9, совместно с нацеленной на HEK3 PEгРНК 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 репликатов, каждый из которых был реализован в трех технических повторах. На **фиг. 49B** показано процентное значение редактирования и инделей для PE2, PE2 R110S K103L, Cas9 H840A никазы или dCas9, совместно с нацеленной на HEK3 PEгРНК плазмидой, которая кодирует редактирование +5 G в 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 репликатов.

[0157] На **фиг. 50A-50B** показана опосредованная PE3 коррекция *HBB* E6V и коррекция *HEXA* 1278+TATC с использованием различных PEгРНК. На **фиг. 50A** показан скрининг 14 PEгРНК на предмет коррекции аллеля *HBB* E6V в клетках HEK293T с использованием PE3. Все проанализированные PEгРНК обеспечивают конвертацию аллеля *HBB* E6V обратно в *HBB* дикого типа без внесения какой-либо молчащей мутации PAM. На **фиг. 50B** показан скрининг 41 PEгРНК на предмет коррекции аллеля *HEXA* 1278+TATC в

клетках HEK293T с использованием PE3 или PE3b. Меченные PEгРНК *HEXA* обеспечивают коррекцию патогенного аллеля посредством делеции 4 п.о. со сдвигом, которая разрушает PAM и оставляет молчащую мутацию. Данные меченные PEгРНК *HEXA* обеспечивают коррекцию патогенного аллеля обратно в дикий тип. В записях, оканчивающихся на «b», используется специфичная для редактирования вносящая разрыв сгРНК в комбинации с PEгРНК (система PE3b). Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических репликатов.

[0158] На **фиг. 51A-51G** показана активность PE3 в линиях клеток человека и сравнение PE3 и Cas9-инициированной HDR. Эффективность осуществления корректного редактирования (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 PEгРНК, которые нацелены на не являющийся мишенью локус. На **Фиг. 51E** показаны контрольные эксперименты с использованием PEгРНК+PE3 без нацеливания и с использованием dCas9+сгРНК 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-51G** 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 PEгРНК 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 репликатов.

[0159] 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.

[0160] На **Фиг. 53A-53E** 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-53E** показаны таблицы аллелей для линий гомозиготных клеток HBB E6V HEK293T.

[0161] На **фиг. 54** представляет собой схему, показывающую процедуру клонирования PEгРНК.

[0162] На **фиг. 55A-55G** представляют собой схемы конструкций PEгРНК. На **фиг.55A** показана простая диаграмма PEгРНК с доменами, отмеченными (слева) и связанными с nCas9 в участке генома (справа). На **Фиг.55B** показаны различные типы модификаций PEгРНК, которые, как ожидается, будут повышать активность. На **Фиг.55C** показаны модификации PEгРНК для повышения транскрипции более длинных РНК посредством выбора промотора и 5', 3' процессинга и терминации. На **фиг.55B** 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 РЕгРНК, 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 РЕгРНК с использованием второй последовательности матрицы, соединенной с РНК элементом на 3'-конце РЕгРНК (слева). Встраивание элементов на 5' или 3'-концах РЕгРНК может усиливать связывание RT.

[0163] На **фиг. 56A-56D** показано встраивание последовательности каркаса РЕнНРК в локусы-мишени. HTS данные анализировали на предмет вставки последовательности каркаса РЕгРНК, 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 РЕгРНК в рамках вставки, примыкающей к RT матрице (слева); процентная величина общих считываний в рамках секвенирования, содержащих вставку последовательности каркаса РЕгРНК указанной длины (в середине); и кумулятивная общая процентная величина вставки РЕгРНК 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.56B** показано то же, что и на **фиг.56A,** но для RNF2. Значения и планки ошибок отражают среднее значение и стандартное отклонение трех независимых биологических репликатов.

[0164] На **Фиг.57A-571** 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 РЕгРНК. РНК, соответствующие 14410 генам и 14368 генам детектировали в образцах PRNP и HEXA, соответственно. На **фиг.57A-57E** показана диаграмма Volcano, демонстрирующая -loglO 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 РЕгРНК, **(фиг.57B)** PE2 против Cas9 H840A с нацеленной на PRNP РЕгРНК, **(фиг.57C)** PE2-dRT против Cas9 H840A с нацеленной на PRNP РЕгРНК, **(фиг.57D)** PE2 против PE2-dRT с нацеленной на HEXA РЕгРНК, **(фиг.57E)** PE2 против Cas9 H840A с нацеленной на HEXA РЕгРНК, **(фиг.57E)** PE2-dRT против Cas9 H840A с нацеленной на HEXA РЕгРНК. Красные точки обозначают гены, которые демонстрируют >2-кратное изменение относительного содержания, которое является статистически значимым (FDR-скорректированное p-значение < 0,05). На **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.

[0165] На **фиг. 58A-58B** показано репрезентативное FACS гейтирование для сортинга ядер нейронов. Ядра последовательно гейтировали на основе сигнала DyeCycle Ruby, соотношения FSC/SSC. соотношения ширина SSC/высота SSC, и соотношения GFP/DyeCycle.

[0166] На **фиг. 59A-59G** показан протокол клонирования 3'-удлиненных PEгРНК в векторы экспрессии U6 млекопитающих с помощью сборки Golden Gate. На **фиг.59A** показан обзор клонирования. На **фиг.59B** показана «стадия 1: Осуществить расщепление pU6-PEгРНК-GG-плазмидного вектора (компонент 1)». На **Фиг.59C** показаны «стадии 2 и 3: осуществить упорядочивание и отжиг нуклеотидных частей (компоненты 2, 3, и 4)». На **Фиг. 59D** показана «стадия 2.b.ii.: фосфорилирование каркаса сгРНК (не обязательно, если нуклеотиды были приобретены в фосфорилированном виде)». На **фиг.59E** показана «стадия 4: сборка PEгРНК». На **фиг. 59F** показаны «стадии 5 и 6: Трансформация собранных плазмид». На **Фиг. 59G** показана диаграмма с обзором протокола клонирования PEгРНК.

[0167] На **фиг. 60A-60B** показан скрипт на языке Python для количественного анализа интеграции каркаса PEгРНК. Был написан специальный скрипт на языке Python для описания и количественного анализа вставок PEгРНК в геномных локусах-мишенях. Скрипт в интерактивном режиме осуществляет сравнение текстовых срок увеличивающейся длины, взятых из исходной последовательности (последовательности каркаса гидовой 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 гидовой 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 сгРНК.

[0168] На **фиг. 61** представляет собой график, показывающий процентное значение всех считываний секвенирования с указанным редактированием для SaCas9(N580A)-MMLV RT HEK3 +6 C>A. Показаны значения для корректных изменений, а также.

[0169] На **фиг. 62A-62B** продемонстрирована важность протоспейсера для эффективного внесения требуемого редактирования в точном положении с использованием праймированного редактирования. На **фиг. 62A** представляет собой график, показывающий процентное значение всех считываний секвенирования с целевыми парами оснований T•A

для различных локусов HEK3. На **фиг. 62B** представляет собой анализ последовательности, показывающий то же самое.

[0170] На **фиг. 63** представлен график, показывающий варианты PAM SpCas9 при редактировании PAM (N=3). Процентное значение всех считываний секвенирования с нацеленным редактированием 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) PEгРНК, длина матрицы RT (RT), и использованная система PE.

[0171] На **фиг. 64A-64F** представлена схема, показывающая внесение различных сайт-специфических мишеней рекомбиназы (SSR) в геном с использованием PE. На **фиг. 64A** показана общая схема вставки целевой последовательности рекомбиназы с помощью праймирующего редактора. На **фиг. 64B** показано, как отдельная SSR-мишень, вставленная с помощью PE, может применяться в качестве участка для геномного встраивания донорской матрицы ДНК. На **фиг. 64C** показано, как тандемная вставка в целевой участок SSR может использоваться для делеции части генома. На **фиг. 64D** показано, как тандемная вставка целевых участков SSR может использоваться для инверсии части генома. На **фиг. 64E** показано, как вставка двух целевых участков SSR в двух удаленных хромосомных участках может приводить к хромосомной транслокации. На **фиг. 64F** показано, как вставка двух различных целевых участков SSR в геноме может использоваться для обмена кассетами из донорской матрицы ДНК.

[0172] На **фиг. 65** показан 1) PE-опосредованный синтез целевого участка SSR в геноме клетки человека и 2) использование данного целевого участка SSR для интеграции донорской матрицы ДНК, включающей маркер экспрессии GFP. После успешной интеграции GFP заставляет клетку флуоресцировать.

[0173] На **фиг. 66** изображен один вариант осуществления праймирующего редактора, представленного в виде двух половин белка PE, которые соединяются в виде цельного праймирующего редактора посредством аутосплайсингового действия половин расщепленного интеина, расположенных в конце или начале каждой из половин белка праймирующего редактора.

[0174] На **фиг. 67A-67B** изображен механизм удаления интеина из полипептидной последовательности и восстановление пептидной связи между N-концевыми и C-концевыми последовательностями экстеина. На фиг. 67A 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 С-концевой половиной белка (или «С экстеином») с образованием цельного единого полипептида, включающего части в виде N экстеина и С экстеина. Согласно различным вариантам осуществления N-экстеин может соответствовать N-концевой половине расщепленного слитого белка праймирующего редактора, а С экстеин может соответствовать С-концевой половине расщепленного праймирующего 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 С экстеина (обозначенная синим половина). Эксцизия разделенных интеинов (т.е. N интеина и С интеина в рамках конфигурации с расщепленным интеином) также может быть обозначена как «транс-сплайсинг», поскольку она включает действие в виде сплайсинга двух отдельных компонентов, представленных *в транс*.

[0175] На **фиг. 68A** продемонстрировано, что доставка обеих половин расщепленного интеина SpPE (SEQ ID NO: 762) в линкере обеспечивает сохранение активности в трех тестовых локусах при совместной трансфекции в клетки HEK293T.

[0176] На **фиг. 68B** продемонстрировано, что доставка обеих половин расщепленного 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 SaCas9 (первые остатки С-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 процента праймирующего редактирования.

[0177] 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.

[0178] 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 редактирования.

[0179] На **фиг. 70A-70I** представлен пример подвергнутой редактированию целевой

последовательности в качестве иллюстрации пошаговой инструкции для конструирования РЕгРНК и вносящих разрыв сгРНК для праймированного редактирования. На **фиг. 70A: Стадия 1. Определение целевой последовательности и редактирования**. Получить последовательность целевого участка ДНК (~200 п.о.) вокруг месторасположения желаемого редактирования (точечной мутации, вставки, делеции, или их комбинации). На **фиг. 70B: Стадия 2. Положение 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 $\geq$ 30 нуклеотидов от положения редактирования. На **фиг. 70C: Стадия 3. Положение участков разрыва**. Для каждого рассматриваемого PAM идентифицировать соответствующий участок разрыва. Для никазы SpCas9 H840A расщепление осуществляется в PAM-содержащей цепи между 3-ми 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 целевого редактирования на PAM-содержащей цепи. В показанном ниже примере присутствуют два возможных PAM. Для простоты остальные этапы продемонстрируют конструирование РЕгРНК с использованием только PAM 1. На **фиг. 70D: Стадия 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 РЕгРНК будет просто представлять собой протоспейсерную последовательность. Если первый нуклеотид протоспейсера не представляет собой G, то спейсерная последовательность РЕгРНК представляет собой G с последующей протоспейсерной последовательностью. На **фиг. 70E: Стадия 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 SpCas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера (PBS) РЕгРНК, содержащий от 12 до 13 нуклеотидов комплементарности относительно праймера ДНК, может использоваться для последовательностей, которые содержат ~40-60% содержимого GC. Для последовательностей с низким содержанием GC должны быть

протестированы более длинные PBS (от 14 до 15 нуклеотидов). Для последовательностей с высоким содержанием GC должны быть протестированы более короткие PBS (от 8 до 11 нуклеотидов). 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. На **фиг. 70F: Стадия 6. Конструирование матрицы RT**. RT матрица кодирует сконструированное редактирование 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 нуклеотидов), среднюю (от 13 до 16 нуклеотидов) и длинную (от 17 до 20 нуклеотидов) матрицу RT. Для длинных редактирований (положения +7 и более) рекомендуется использовать матрицы RT, которые имеют длину по меньшей мере 5 нуклеотидов (предпочтительно 10 или более нуклеотидов) после положения редактирования для обеспечения достаточной гомологии 3'-лоскута ДНК. Для длинных изменений несколько матриц RT должны быть подвергнуты скринингу для идентификации функциональных конструкций. Для более крупных вставок и делеций (>5 нуклеотидов) рекомендуется встраивание большей 3 ' гомологии (~20 нуклеотидов 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 PEгРНК). Поскольку многие RT матрицы поддерживают эффективное праймированное 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. Необходимо отметить, что по сравнению с редактированием SNP, редактирования в виде вставки или делеции с использованием RT матриц той же длины не будут содержать идентичную гомологию. На **фиг. 70G: Стадия 7. Сборка полной последовательности PEгРНК**. Следует соединить компоненты PEгРНК в следующем порядке (от 5' к 3'): спейсер, каркас, RT матрица и PBS. На **фиг. 70H: Стадия 8. Конструирование вносящих разрыв гРНК для PE3**. Идентифицировать PAM на не подвергаемой редактированию цепи против хода транскрипции и по ходу транскрипции от редактирования. Оптимальные положения разрыва имеют высокую зависимость от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии

от 40 до 90 нуклеотидов 5' относительно положения РЕгРНК- индуцированного разрыва, ведут к более высоким выходам редактирования и меньшему количеству инделей. Вносящая разрыв сгРНК имеет спейсерную последовательность, которая совпадает с протоспейсером длиной 20 нуклеотидов в *начальном* аллеле, с добавлением 5' G, если протоспейсер не начинается на G. На **фиг. 70I: Стадия 9. Конструирование вносящей разрыв сгРНК для 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 вносящей разрыв сгРНК совпадает с последовательностью подвергаемого желаемому редактированию аллеля, но не начального аллеля. Система PE3b работает эффективно, когда подвергаемый редактированию нуклеотид (нуклеотиды) попадают в участок затравки (~10 нуклеотидов от PAM) протоспейсера вносящей разрыв сгРНК. Это предотвращает разрыв комплементарной цепи до готовности подвергнутой редактированию цепи, предотвращая конкурирование между РЕгРНК и сгРНК 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. PE3b сгРНК должны иметь спейсерную последовательность, которая совпадает с протоспейсером длиной 20 нуклеотидов в *требуемом* аллеле, с добавлением 5' G, если это необходимо.

[0180] На **фиг. 71A** показана нуклеотидная последовательность молекулы РЕгРНК SpCas9 (сверху), которая терминируется на 3 '-конце в «UUU» и не содержит элемент в виде петли. Нижняя часть фигуры изображает ту же молекулу РЕгРНК SpCas9, но которая дополнительно модифицирована для включения элемента в виде петли, имеющего последовательность 5'-«GAAANNNNN»-3', вставленную непосредственно перед «UUU» 3'-концом. «N» может представлять собой любое нуклеотидное основание.

[0181] На **фиг. 71B** продемонстрировано, что эффективность праймированного редактирования в клетках HEK или клетках EMX повышается с применением РЕгРНК, содержащей элементы в виде петли, в то время как процентная величина образования инделей в значительной степени не меняется.

[0182] На **фиг. 72A-72C** изображены альтернативные конфигурации РЕгРНК, которые могут применяться в праймированном редактировании. На **фиг. 72A** изображен вариант осуществления PE2:РЕгРНК для праймированного редактирования. Данный вариант осуществления включает PE2 (слитый белок, содержащий Cas9 и обратную транскриптазу) в виде комплекса с РЕгРНК (как также описано на фиг. 1A-1I и/или фиг.

3A-3E). Согласно данному варианту осуществления матрица для обратной транскрипции встраивается в 3'-плечо удлинения на сгРНК с получением РЕгРНК, а фермент в виде ДНК-полимеразы представляет собой обратную транскриптазу (RT), слитую напрямую с Cas9. На **фиг. 72B** изображен вариант осуществления MS2cp-PE2:сгРНК + tPERT. Данный вариант осуществления содержит продукт слияния PE2 (Cas9 + обратная транскриптаза), который дополнительно слит с белком оболочки бактериофага MS2 (MS2cp) с образованием слитого белка MS2cp-PE2. Для реализации праймированного редактирования слитый белок MS2cp-PE2 образует комплекс с сгРНК, которая нацеливает комплекс на определенный целевой участок в ДНК. Затем данный вариант осуществления предусматривает внесение РНК матрицы для *транс*-праймированного редактирования («tPERT»), которая действует вместо РЕгРНК посредством предоставления участка связывания праймера (PBS) и матрицы для синтеза ДНК на отдельной молекуле, т.е. tPERT, которая также имеет аптамер MS2 (структуру типа «петля на стебле»). Белок MS2cp рекрутирует tPERT посредством связывания с аптамером MS2 молекулы. На **фиг. 72C** Изображены альтернативные варианты конструирования РЕгРНК, 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 РЕгРНК для применения в рамках праймированного редактирования, при этом плечо удлинения гибридной РЕгРНК представляет собой ДНК вместо РНК. В подобном варианте осуществления ДНК- зависимая ДНК-полимераза может использоваться вместо обратной транскриптазы для синтеза 3'-лоскута ДНК, включающего требуемое генетическое изменение, который образован с помощью праймированного редактирования. Согласно другому варианту осуществления плечо удлинения может быть синтезировано для включения химического линкера, который предотвращает использование ДНК-полимеразой (например, обратной транскриптазой) каркаса или основы сгРНК в качестве матрицы. Согласно другому варианту осуществления плечо удлинения может включать матрицу для синтеза ДНК, которая имеет обратную ориентацию относительно общей ориентации молекулы РЕгРНК. В качестве примера и как показано для РЕгРНК в ориентации от 5' к 3' и с удлинением, присоединенным к 3'-концу каркаса сгРНК, матрица для синтеза ДНК ориентируется в противоположном направлении, то есть в направлении от 3' к 5'. Этот вариант осуществления может представлять собой преимущество для вариантов осуществления РЕгРНК с плечами удлинения, расположенными на 3'-конце гРНК. Посредством 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 гРНК в качестве матрицы.

[0183] На **фиг. 73** показано праймированное редактирование с использованием tPERT и системы рекрутирования MS2 (также известной как способ мечения MS2). сгРНК, нацеливающая белок праймирующего редактора (PE2) на локус-мишень, экспрессируется в комбинации с tPERT, содержащей участок связывания праймера (PBS длиной 13 нуклеотидов или 17 нуклеотидов), RT матрицей, кодирующей вставку в виде His6 метки и плечо гомологии, и аптамером MS2 (расположенным на 5' или 3'-конце молекулы tPERT). Используется белок праймирующего редактора (PE2) или MS2cp, слитый с N-концом PE2. Редактирование осуществляли с использованием сгРНК, вносящей разрыв в комплементарную цепь, или без нее, как в ранее разработанной системе PE3 (обозначенной на оси X метками «PE2+разрыв» или «PE2», соответственно). Это также обозначается и определяется в настоящем документе как «разрыв второй цепи.»

[0184] На **фиг. 74** показана экспрессия аптамера MS2 обратной транскриптазы в транс и его рекрутирование системой аптамера MS2. РЕгРНК содержит РНК аптамер MS2, вставленный в любую одну из двух шпилек каркаса сгРНК. 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.

[0185] На **фиг**. **75** приведена гистограмма для сравнения эффективности (то есть «% общих считываний секвенирования с указанным редактированием или инделями») PE2, PE2- усеченный, PE3, и PE3-усеченный в различных целевых участках в различных линиях клеток. Данные показывают, что праймирующие редакторы, содержащие усеченные варианты RT, были приблизительно настолько же эффективны, что и праймирующие редакторы, содержащие неусеченные RT белки.

[0186] На **фиг**.**76** продемонстрирована эффективность редактирования праймирующих редакторов на основе расщепленного интеина. Клетки HEK239T трансфицировали плазмидами, кодирующими полноразмерный PE2 или PE2 на основе расщепленного интеина, РЕгРНК и вносящей разрыв гидовой РНК. Указана консенсусная последовательность (большая часть амино-концевых остатков C- концевого экстеина). Показана процентная величина редактирования в двух участках: вставка CTT HEK3 +1 и G на T PRNP +6. Повторы независимых трансфекций n=3.

[0187] На **фиг.** 77 демонстрируется эффективность редактирования праймирующих редакторов на основе расщепленного интеина. 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.

[0188] На **фиг. 78** показана эффективность редактирования праймирующих редакторов на основе расщепленного интеина. В частности, на чертежах приведены конструкции AAV-SpPE3 с расщеплением. 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-сгРНК кассету, экспрессирующую вносящую разрыв сгРНК, а С-концевой геном содержит U6-РЕгРНК кассету, экспрессирующую РЕгРНК.

[0189] На **фиг. 79** показана эффективность редактирования определенных оптимизированных линкеров. В частности, данные показывают эффективность редактирования конструкции 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 РЕгРНК для редактирований транзиции, трансверсии, вставки и делеции. Замещающие линкеры обозначены как «1x SGGS»(SEQ ID NO: 174), «2x SGGS»(SEQ ID NO: 446), «3x SGGS» (SEQ ID NO: 3889), «1x XTEN» (SEQ ID NO: 171), «нет линкера», «1x Gly», «1x Pro», «1x EAAAK»(SEQ ID NO: 3968), «2x EAAAK» «(SEQ ID NO: 3969) и «3x EAAAK» (SEQ ID NO: 3970). Эффективность редактирования измеряется в формате гистограммы относительно «контрольной» эффективности редактирования PE2. Линкер PE2 представляет собой SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127). Все редактирование осуществляли в контексте системы PE3, то есть которая соответствует конструкции для редактирования PE2 плюс добавлению оптимальной вторичной гидовой сгРНК, вносящей разрыв.

[0190] 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).

[0191] На **фиг. 81** изображен уровень транскрипции РЕгРНК из различных промоторов.

[0192] На **фиг. 82** показано влияние различных типов модификаций структуры РЕгРНК на эффективность редактирования по сравнению с немодифицированной РЕгРНК.

[0193] На **фиг. 83** Показан эксперимент в виде PE, целью которого являлось

редактирование гена HEK3, в частности вставка длиной 10 нуклеотидов в положении +1 относительно участка разрыва, и использование PE3.

[0194] На **фиг. 84A** изображен пример PEгРНК, имеющей спейсер, ядро гРНК 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.

[0195] На **фиг. 84B** изображена структура тРНК, которая может использоваться для модификации структур PEгРНК. P1 может иметь различную длину. P1 может быть удлинен для обеспечения предотвращения процессинга РНКазой P продукта слияния в виде PEгРНК-тРНК.

[0196] На **фиг. 85** показан эксперимент с PE, который нацелен на редактирование гена FANCF, в частности конверсию G на T в положении +5 относительно участка разрыва, и использование конструкции PE3.

[0197] На **фиг. 86** показан эксперимент с PE, который нацелен на редактирование гена HEK3, в частности вставка FLAG метки длиной 71 нуклеотидов в положении +1 относительно участка разрыва, и использование конструкции PE3.

[0198] На **фиг. 87** приведены результаты скрининга в клетках N2A, где PEгРНК обеспечивает внесение 1412Adel, с подробной информацией о длине участка связывания праймера (PBS) и длине матрицы в виде обратной транскриптазы (RT). (Показаны с инделями и без низ).

[0199] На **фиг. 88** приведены результаты скрининга в клетках N2A, где PEгРНК обеспечивает внесение 1412Adel, с подробной информацией о длине участка связывания праймера (PBS) и длине матрицы в виде обратной транскриптазы (RT). (Показаны с инделями и без низ).

[0200] На **фиг. 89** приведены результаты редактирования в промежуточном локусе в гене ß-глобина и в HEK3 в здоровых HSC при различной концентрации редактора относительно PEгРНК и вносящей разрыв гРНК.

[0201] На **фиг. 90** представлена схема варианта осуществления двухлоскутного праймированного редактирования. На целевую последовательность ДНК воздействуют два комплекса праймированного редактирования (управляемые PEгРНК-A и PEгРНК-B). Две PEгРНК нацелены на противоположные цепи двойной спирали. Каждый праймирующий редактор (PE2•PEгРНК) разрезает одну цепь ДНК, затем синтезирует 3'-лоскут ДНК, используя PEгРНК в качестве матрицы. Действие двух комплексов праймированного редактирования приводит к образованию промежуточного соединения, содержащего два 3'-

лоскута на противоположных цепях ДНК. Два 3'-лоскута комплементарны друг другу на своих 3'-концах. Отжиг 3'-концов 3'-лоскутов приводит к образованию двойной дуплексной структуры, при этом один дуплекс состоит из парных 3'-лоскутов, содержащих новую последовательность ДНК (красный), а другой дуплекс состоит из парных 5'-лоскутов. содержащий исходную последовательность ДНК (черный). Вырезание промежуточного исходного дуплекса ДНК (черные парные 5'-лоскуты) дает вид ДНК с двойным разрывом, содержащий желаемую новую последовательность ДНК (красный), заменяющую исходную последовательность ДНК. Связывание обоих разрывов завершает процесс редактирования.

[0202] На **фиг. 91A-91B** приведены результаты Примера 7. На **фиг. 91A** показана таблица выходных аллелей Crispresso2 (выровненная по желаемому продукту) для замены последовательности из 90 п.о. новой последовательностью из 22 п.о. в сайте HEK3 в клетках HEK293T с использованием двух праймирующих редакторов. На желаемый продукт приходится более 80% прочтений секвенирования. Эталонный исходный аллель показан над секвенированными аллелями для сравнения. На **фиг. 91B** показаны последовательности РЕгРНК, использованные для достижения замены последовательности, показанной на фиг. 91A. РЕгРНК1 и РЕгРНК 2 нацелены на разные цепи двойной спирали ДНК и генерируют 5'-смещенные разрывы, как показано на фиг. 90.

[0203] На **фиг. 92** показаны варианты конструкции конструкций РЕгРНК для двухлоскутного праймированного редактирования. Для двухлоскутного праймированного редактирования используются две РЕгРНК, показанные на чертеже как РЕгРНК А и РЕгРНК В. Каждая РЕгРНК содержит спейсерную последовательность (темно-синяя), которая направляет комплекс праймированного редактирования к целевому участку ДНК. Две РЕгРНК нацелены на противоположные цепи двойной спирали ДНК. Как и другие РЕгРНК, РЕгРНК для двухлоскутного праймированного редактирования содержат 3'-удлинение с последовательностью связывания праймера (PBS, зеленый), которая отжигается с цепью геномной ДНК с разрывом, чтобы инициировать обратную транскрипцию, и матрицу обратной транскрипции (матрица RT, голубой), которая является матрицей для синтеза новой ДНК с помощью фермента обратной транскриптазы. В отличие от РЕгРНК, используемых для классического праймированного редактирования, для которых требуется вновь синтезированный отредактированный 3'-лоскут, чтобы конкурировать с эндогенным 5'-лоскутом, нет необходимости кодировать гомологию с целевым участком в матрице RT. Вместо этого 3'-концы двух синтезированных 3'-лоскутов должны содержать только комплементарность друг другу (т.е. 3'-концы 3'-лоскутов являются обратно комплементарными друг другу). Эта комплементарность позволяет двум 3'-лоскутам отжигаться и способствует образованию желаемой отредактированной

последовательности ДНК.

[0204] На **фиг. 93A-93B** показывают результаты Примера 7 с использованием двухлоскутного праймированного редактирования для установки сайтов Bxb1 attB и attP с двухлоскутным праймированным редактированием. На **фиг. 93A** показана установка сайта attB Bxb1 размером 38 п.о. в HEK3. Было сконструировано шесть РЕгРНК, три из которых нацелены на (+) цепь (A1, A2 и A3) и три нацелены на (-) цепь (B1, B2 и B3). Они различаются количеством последовательностей attB, закодированных в матрице RT, что приводит к различному количеству комплементарных нуклеотидов между двумя лоскутами. Матрицу РЕгРНК3x3 оценивали на предмет установки последовательности attB в целевом положении генома в клетках HEK293T. На фиг. 93A и фиг. 93B показана установка сайта attP Bxb1 размером 50 п.о. в HEK3. Было сконструировано шесть РЕгРНК, три из которых нацелены на (+) цепь (A1, A2 и A3) и три нацелены на (-) цепь (B1, B2 и B3). Они различаются количеством последовательностей attP, закодированных в матрице RT, что приводит к различному количеству комплементарных нуклеотидов между двумя лоскутами. Матрицу РЕгРНК 3x3 оценивали на предмет установки последовательности attP в целевом геномном положении в клетках HEK293T. Для обоих вариантов редактирования на фиг. 93A и фиг. 93B установка сайта attB или attP происходит с сопутствующей делецией последовательности геномной ДНК длиной 90 п.о., расположенной между двумя участками разрыва.

[0205] На **фиг. 94** показаны результаты установки сайтов Bxb1 attB и attP в локусы safe-harbor человека с двухлоскутным праймированным редактированием. Установка сайтов attP Bxb1 в локусе AAVS1 (слева) или сайтов attB Bxb1 в локусе CCR5 (справа) в клетках HEK293T. Правильные редактирования показаны синим цветом (AAVS1) и красным цветом (CCR5), а побочный продукт с инделями показан серым цветом.

[0206] На **фиг. 95** представлена схема инверсии геномной последовательности с четырехлоскутным праймированным редактированием. Область геномной ДНК является целевой для инверсии (зеленый и оранжевый сегменты). Четыре РЕгРНК доставляются в клетки с помощью праймирующего редактора PE2. Одна пара РЕгРНК нацелена на одну цепь геномной ДНК и служит матрицей для синтеза двух комплементарных лоскутов ДНК (A и A', синие), в то время как вторая пара нацелена на другую цепь геномной ДНК и является матрицей для синтеза двух комплементарных лоскутов ДНК с ортогональной последовательностью ДНК (B и B', розовые). Комплементарные лоскуты отжигают, образуя дуплексы с выступом 3'. Дуплексы 5'-выступа вырезаются эндогенными ферментами репарации клеток. Разрывы лигируют с получением аллеля продукта, содержащего инвертированную последовательность ДНК (зеленые и оранжевые сегменты) и

последовательности матрицы РЕгРНК в местах соединений инверсии (синий и розовый сегменты).

[0207] На **фиг. 96** показаны результаты секвенирования ампликонов инверсионных соединений AAVS1. Результаты анализа CRISPResso2 инверсии размером 2,7 т.о. в локусе AAVS1 в клетках HEK293T с использованием стратегии четырехлоскутного праймированного редактирования. ПЦР-амплификация и секвенирование ожидаемых инверсионных соединений показали желаемые продукты с последовательностями Bxb1 attP или attB, вставленными в места соединения инверсий.

[0208] На **фиг. 97A-97B** показаны результаты направленной интеграции кольцевой плазмиды ДНК в геном с использованием стратегии четырехлоскутного праймированного редактирования. (Фиг. 97A) Область геномной ДНК и область плазмидной ДНК являются мишенями для интеграции четырехлоскутного праймированного редактирования. Четыре РЕгРНК доставляются в клетки с помощью праймирующего редактора PE2. Две РЕгРНК являются матрицей для комплементарных последовательностей, одна нацелена на одну цепь геномной ДНК, а другая нацелена на одну цепь плазмидной ДНК (генерируя синие лоскуты). Две другие РЕгРНК нацелены на противоположные цепи геномной ДНК и плазмидной ДНК, отличны от первых двух РЕгРНК, и они являются матрицей для синтеза двух комплементарных лоскутов ДНК (розовые), которые ортогональны первой паре. Комплементарные лоскуты отжигаются, образуя дуплексы с выступом 3'. Дуплексы 5'-выступа вырезаются эндогенными ферментами репарации клеток.

[0209] На **фиг. 98A-98B** показаны результаты направленной хромосомной транслокации с четырехлоскутным праймированным редактированием. (Фиг. 98A) РЕгРНК нацелены на две области на разных хромосомах. Комплементарные 3'-лоскуты ДНК соединяют две последовательности хромосом и определяют ориентацию транслокации. (Фиг. 98B) Направленная транслокация между локусами MYC и TIMM44 в клетках HEK239T. Результат анализа CRISPResso2 в результате секвенирования ампликона ожидаемых соединений в результате транслокации между локусом MYC на хромосоме 8 и локусом TIMM44 на хромосоме 19. Большинство прочтений секвенирования соответствуют желаемой последовательности аллеля.

[0210] На **фиг. 99** показана установка сайтов Bxb1 attB и attP с двухлоскутным праймированным редактированием в локусе IDS. Клетки HEK293T трансфицировали PE2 и различными парами РЕгРНК (например, в первой колонке РЕгРНК A1_a и РЕгРНК B2_a с матрицами для установки сайта attP в прямом направлении). Эффективность измерялась HTS. Эти данные показывают, что двухлоскутное редактирование может успешно вставлять представляющую интерес последовательность в локус IDS с эффективностью до

~ 80%.

[0211] На **фиг. 100A-100B** описана дупликация, опосредованная двумя лоскутами. На фиг. 100A приведена схема, показывающая опосредованную двумя лоскутами дупликацию в локусе AAVS1 в клетках 293T. **ИНЖИР**. 100B показаны результаты использования PEгРНК с двумя лоскутами с PE2 для индукции дупликации генетических последовательностей в AAVS1.

[0212] На **фиг. 101** показана новая транслокация MYC-CCR5, индуцированная несколькими лоскутами. Транслокация MYC-CCR5 индуцировалась четырехлоскутными PEгРНК и PE2. События транслокации MYC-CCR5 индуцировались четырехлоскутными PEгРНК. В клетках HEK293T тестировали четыре различных набора PEгРНК. Продукты соединения транслокации между производными chr8 и chr3 амплифицировали с помощью праймеров соединения. На графике показан процент прочтений, выровненных с ожидаемыми аллелями соединения. Результат показывает, что четырехлоскутное редактирование может опосредовать транслокацию генов MYC и CCR5 с чистотой продукта около 100 % на соединении 1 и ~50 % на соединении 2. Репрезентативный график аллелей показывает последовательности, выровненные с ожидаемыми последовательностями аллелей на соединении 1.

[0213] На **фиг. 102A-102B** показано двухлоскутное и мультилоскутное редактирование в других линиях клеток человека. На фиг. 102A показано двухлоскутное редактирование с двумя лоскутами в четырех различных линиях клеток человека. Клетки HEK293T и HeLa трансфицировали двойными PEгРНК и PE2 для редактирования трех разных геномных локусов (IDS, MYC и TIMM44). Клетки U2OS и K562 нуклеофицировали теми же компонентами. Двухлоскутное редактирование показало надежное редактирование во всех четырех клетках человека в целевых локусах, особенно в клетках HEK293T и K562. Клеточный механизм, обеспечивающий двухлоскутное редактирование, сохраняется во многих типах клеток человека. На фиг. 102B показана инверсия 2,7 т.о. на AAVS1 в клетках HeLa, управляемая мультилоскутным редактированием (четырехлоскутное редактирование-PEгРНК).

[0214] На **фиг. 103A-103B** показаны результаты измерения эффективности инверсии с помощью HTS в локусе CCR5. Процент ожидаемого инверсионного аллеля редактирования измеряли с помощью HTS. Четыре набора четыре лоскута- PEгРНК PE2 трансфицировали в клетки HEK293T соответственно. На фиг. 103A показана опосредованная двухлоскутным редактированием дупликация последовательности (~100 нуклеотидов) в локусе CCR5 в клетках HEK293T. Эффективность редактирования достигает ~ 1,5% с помощью анализа секвенирования парных концов HTS 300 циклов. На

фиг. 103B показана опосредованная четырехлоскутным редактированием инверсия последовательности (~95-117 нуклеотидов) в локусе CCR5 в клетках HeLa. Эффективность редактирования достигает ~ 1,2% с помощью анализа секвенирования парных концов HTS 300 циклов. Этот результат показывает, что мультилоскутное редактирование может успешно обеспечивать дупликацию и инверсию точно в локусе CCR5. Специфичность редактирования высока, когда целевая последовательность дублируется (процент вставок <2%).

[0215] На **фиг. 104A-104E** показан целевой путь клеточной репарации для двухлоскутного редактирования. На фиг. 104A-104D, клетки HEK293T трансфицировали плазмидами, которые экспрессируют Exo1, Fen1, белок красной флуоресценции (контроль), ДНК2, Mlh1 neg и ингибитор P53 с РЕгРНК и PE2. Эффективность редактирования измеряли посредством HTS. Эффективность редактирования сравнивали между кандидатом и контролем RFP. На фиг. 104E, клетки HEK293T трансфицировали плазмидами siPHK и РЕгРНК и PE2 для каждого локуса-мишени. В качестве контроля использовали нецелевую siPHK (siNT). Эффективность редактирования измеряли посредством HTS. Эффективность редактирования сравнивали между каждым нокдауном siPHK и siNT ctrl в каждом локусе-мишени. Клетки HEK293T трансфицировали двойной РЕгРНК, PE2 и плазмидами, которые экспрессируют Exo1, Fen1, белок красной флуоресценции (ctrl), DNA2, Mlh1 neg и ингибитор P53 соответственно. Эффективность редактирования измеряли посредством HTS. Эффективность редактирования сравнивали между кандидатом и управляющим RFP. Двусторонний парный t-критерий Стьюдента использовали для измерения статистической разницы между каждым лечением и контролем RFP (P <0,05, *, P <0,01, **, P <0,001, ***). Сверхэкспрессия FEN1 улучшает эффективность двухлоскутного редактирования во всех четырех целевых локусах (MYC, TIMM44, IDS, CCR5).

[0216] На **фиг. 105A-105B** показана опосредованная двумя лоскутами дупликация последовательности в локусе *AAVS1*. На фиг. 105A представлена схематическая диаграмма опосредованной двумя лоскутами дупликации последовательности в локусе AAVS1. На фиг. 105B показано, что при использовании двойных РЕгРНК, которые генерируют две уникальные структуры 3'-лоскута, в локусе AAVS1 в клетках 293T была индуцирована дупликация последовательности ~300 п.о. Ожидаемые аллели амплифицируют специфическими праймерами и подвергают HTS. ~ 94% прочтений соответствуют ожидаемым аллелям с дупликацией. В необработанных образцах продукты дубликации не наблюдаются.

[0217] На **фиг. 106 показана целевая** инверсия геномной последовательности IDS

с четырехлоскутным праймированным редактированием. Было показано, что приблизительно у 13% пациентов с синдромом Хантера имеется инверсия последовательностей гена IDS (Bondeson et al., Human Molecular Genetics, 1995). Для индуцирования этой патогенной инверсии геномной последовательности IDS ~ 40 т.о. в клетках HEK293T применяли четырехлоскутное редактирование. Шесть наборов четвертных РЕгРНК тестировали путем трансфекции клеток HEK293T РЕгРНК и PE2. Праймеры использовали для специфической амплификации инвертированных последовательностей в местах соединения «ab» и соединения «cd». ~ 95% ожидаемых последовательностей инвертированных аллелей наблюдали на обоих соединениях с IDS_QF1. Другие наборы РЕгРНК также дают высокий процент ожидаемых аллельных последовательностей на обоих соединениях. В необработанных образцах продукты инвертированной трансверсии не наблюдаются.

[0218] На **фиг. 107** показано, что модификация 3'-мотива РЕгРНК улучшает эффективность двухлоскутного редактирования в локусе IDS. Для дальнейшего повышения эффективности двухлоскутного редактирования был введен мотив evoPreQ1 для защиты 3'-конца РЕгРНК. При сравнении эффективности редактирования, генерируемой немодифицированными и модифицированными evoPreQ1 двойными РЕгРНК, наблюдается общее увеличение эффективности редактирования модифицированными РЕгРНК в целевом локусе IDS. Повышение эффективности двухлоскутного редактирования может достигать 5,3 раза.

[0219] На **фиг. 108A-108C** показан обзор TwinPE и замены последовательности, опосредованной twinPE. На фиг. 108A показано, что системы twinPE нацелены на последовательности геномной ДНК, которые содержат две последовательности протоспейсеров на противоположных цепях ДНК. Комплексы PE2• РЕгРНК нацеливаются на каждый протоспейсер, генерируют одноцепочечный разрыв и обратно транскрибируют кодированную РЕгРНК матрицу, содержащую желаемую последовательность вставки. После синтеза и высвобождения 3'-лоскутов ДНК существует гипотетический промежуточный продукт, содержащий отожженные 3'-лоскуты, содержащие отредактированную последовательность ДНК, и отожженные 5'-лоскуты, содержащие исходную последовательность ДНК. Вырезание исходной последовательности ДНК, содержащейся в 5'-лоскуте, с последующим лигированием 3'-лоскутов в соответствующий сайт вырезания дает желаемый отредактированный продукт. На фиг. 108B приведен пример опосредованной twinPE замены последовательности из 90 п.о. в *сайте 3 HEK* на последовательность из 38 п.о. Bxb1 *attB*. На фиг. 108C показана оценка twinPE в клетках HEK293T для установки сайта attB *Bxb1* размером *38 п.о.*, как показано на фиг. 108B, или

сайта attP *Bxb1* размером *50 п.о. в* сайте *HEK 3* с использованием PEгРНК, которые являются матрицей для последовательности вставки различной длины. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов.

[0220] На **фиг. 109A-109E** показаны вставка, делеция и перекодирование целевой последовательности с помощью twinPE в клетках человека. На фиг. 109A показана вставка фрагментов FKBP кодирующей последовательности с PE3 (12 п.о., 36 п.о., 108 п.о. или 321 п.о.) или twinPE (108 п.о.) в *сайте 3 HEK* в клетках HEK293T. На фиг. 109B показано прочтение секвенирования последовательности экзонов 4 и 7 в *PAH* в клетках HEK293T с использованием twinPE. Целевую последовательность размером 64 п.о. в экзоне 4 редактировали с использованием перекрывающихся лоскутов длиной 24, 36 или 59 п.о., целевую последовательность размером 46 п.о. в экзоне 7 редактировали с использованием перекрывающихся лоскутов длиной 22 или 42 п.о. или последовательность из 64 п.о. в экзоне 7 редактировали с использованием перекрывающихся лоскутов длиной 24 или 47 п.о. Редактирующую активность сравнивали с использованием стандартных PEгРНК или PEгРНК, содержащих 3'-мотивы evoPreQ1. На фиг. 109C приведена схематическая диаграмма трех различных двухлоскутных стратегий делеции, которые были исследованы для проведения целенаправленных делеций. Стратегия twinPE «Basic-Anchor (BA)» допускает гибкую делецию, начиная с произвольной позиции 3' участка разрыва и заканчивая другим участком разрыва. Стратегия twinPE «Hybrid-Anchor (HA)» позволяет гибко удалять последовательность в произвольно выбранных положениях между двумя участками разрыва. Исследуемая в данном случае стратегия «PrimeDel (PD)» позволяет удалить последовательность, начинающуюся с одного участка разрыва и заканчивающуюся другим участком разрыва. На фиг. 109D показано удаление последовательностей в *сайте 3 HEK* в клетках HEK293T с использованием стратегий BA-twinPE, HA-twinPE или PD, нацеленных на одну и ту же пару протоспейсеров. Редактирующую активность сравнивали с использованием стандартных PEгРНК или PEгРНК, содержащих 3'-мотивы evoPreQ1. На фиг. 109E показана делеция последовательности экзона 51 в локусе *DMD* в клетках HEK293T с использованием BA-twinPE, PD, парной нуклеазы Cas9 или замены последовательности attB, опосредованной *twinPE*. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов. В эксперименте с пропуском экзона 51 DMD выполняли по меньшей мере два независимых биологических повтора.

[0221] На **фиг. 110A-110E** показана сайт-специфическая геномная интеграция ДНК с рекомбиназой twinPE и Bxb1 в клетках человека. На фиг. 110A показан скрининг пар

РЕгРНК twinPE на предмет установки последовательности attP *Bxb1 в* локусе AAVS1 в клетках HEK293T. На фиг. 110B показан скрининг пар РЕгРНК twinPE на предмет установки последовательности Bxb1 *attB* в локусе *CCR5* в клетках HEK293T. На фиг. 110C показана однократная трансфекция доноров ДНК размером 5,6 т.о. с использованием пар РЕгРНК twinPE, нацеленных на *CCR5* (четыре крайних левых столбца) или *AAVS1* (три крайних правых столбца). РЕгРНК twinPE устанавливают *attB* в *CCR5* или *attP* в *AAVS1*. Затем Bxb1 интегрирует донора, несущего соответствующий сайт присоединения, в геномный сайт присоединения. На фиг. 110D показана оптимизация нокаута одиночной трансфекции в *CCR5* с использованием пары РЕгРНК twinPE 531/584. Идентичность матричного редактирования (*attB* против *attP*), идентичность центрального динуклеотида (GT дикого типа против ортогонального мутанта GA) и длину перекрытия между лоскутами варьировали, чтобы определить комбинацию, которая поддерживала самую высокую эффективность нокаута. На фиг. 110E показана вставка последовательности Bxb1 attB в интрон 1 ALB в клеточных линиях HEK293T и Huh7. На фиг. 110F показано сравнение эффективности нокаута одиночной трансфекции в *CCR5* и *ALB* в клеточных линиях HEK293T и Huh7.

[0222] На **фиг. 111A-111E** показана сайт-специфическая инверсия геномной последовательности с помощью TwinPE и рекомбиназы Bxb1 в клетках человека. На фиг. 111A приведена схематическая диаграмма горячих точек рекомбинации в *IDS* и *IDS2*, которые приводят к патогенным инверсиям 39 т.о., и комбинированной стратегии twinPE-Bxb1 для установки или коррекции мутации инверсии *IDS*. На фиг. 111B показан скрининг пар РЕгРНК в *IDS* и *IDS2* для установки вставки сайта рекомбинации *attP* или *attB в локусах IDS* и *IDS2* со специфическими ДНК-мишенями. На фиг. 111C показан анализ секвенирования ДНК вставки attP или attB с последовательной трансфекцией ДНК. На фиг. 111D показана чистота продукта инверсии на инвертированном соединении 1 и соединении 2 (последовательная трансфекция), что указывает на успешную инверсию на двух соединениях. На фиг. 111E показана количественная оценка эффективности инверсии на соединениях (последовательная трансфекция и РНК-нуклеофекция в одном сосуде).

[0223] На **фиг. 112** показано перекодирование последовательностей экзонов 10, 11 и 12 в *PAH* в клетках *HEK293T* посредством twinPE. Целевую последовательность размером 64 п.о. в экзоне 10 редактировали с использованием перекрывающихся лоскутов длинной 28 п.о., целевую последовательность размером 61 и 55 п.о. в экзоне 11 редактировали с использованием перекрывающихся лоскутов длинной 25 п.о., или последовательности 68 и 58 п.о. в экзоне 12 редактировали с использованием перекрывающихся лоскутов длиной 27 и 24 п.о., соответственно. Значения и планки погрешностей отражают среднее значение и

стандартное отклонение для трех независимых биологических репликатов.

[0224] На **фиг. 113** показана трансфекция клональной клеточной линии HEK293T, содержащей гомозиготную вставку сайта attB, плазмидами BxBI и *attP*-содержащими донорными ДНК-плазмидами. Эффективность нокаута составляет от 12 до 17% в целевом сайте по данным ddPCR.

[0225] На **фиг. 114** показано измерение HTS ожидаемых последовательностей соединения, содержащих продукты рекомбинации *attL* и attR, после опосредованного twinPE и BxBI однократного нокаута. Чистота продукта находится в диапазоне от 71 до 95%. Значения и планки погрешностей отражают среднее значение и стандартное отклонение для трех независимых биологических репликатов

[0226] На **фиг. 115A** показана опосредованная *twinPE* эффективность вставки attB с уменьшенной длиной перекрытия лоскутов в двойных РЕгРНК.

[0227] На **фиг. 115B** показаны продукты ПЦР, амплифицированные специфическими праймерами для улавливания рекомбинации между донорной ДНК и плазмидами РЕгРНК, показанными на агарозном геле. Рекомбинация между донорской ДНК и плазмидой РЕгРНК снижалась при меньшем перекрытии лоскутов.

[0228] На **фиг. 116** приведена схематическая диаграмма разработанных стратегий ПЦР для количественной оценки эффективности инверсии *IDS*.

Определения

[0229] Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют значение, обычно известное специалисту в области, к которой относится настоящее изобретение. Следующие ссылки дают специалисту общее определение многих терминов, используемых в настоящем 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), and Hale & Marham, The Harper Collins Dictionary of Biology (1991). В контексте настоящего изобретения следующие термины имеют приписанные им значения, если не указано иное.

Антисмысловая цепь

[0230] В области генетики «антисмысловая» цепь сегмента двухцепочечной ДНК является матричной цепью, и считается, что она имеет ориентацию от 3' к 5'. Напротив, «смысловая» цепь представляет собой сегмент двухцепочечной ДНК, который проходит от 5' до 3' и комплементарен антисмысловой цепи ДНК или матричной цепи, которая проходит от 3' до 5'. В случае сегмента ДНК, кодирующего белок, смысловая цепь представляет собой

цепь ДНК, имеющую ту же последовательность, что и иРНК, которая использует антисмысловую цепь в качестве своей матрицы во время транскрипции и в конечном итоге подвергается (как правило, не всегда) трансляции в белок. Таким образом, антисмысловая цепь отвечает за РНК, которая позже транслируется в белок, в то время как смысловая цепь имеет почти такой же состав, что и иРНК. Необходимо отметить, что для каждого сегмента двухцепочечной ДНК, возможно, будет два набора смысловой и антисмысловой цепи, в зависимости от того, в каком направлении читать (поскольку смысловая цепь и антисмысловая цепь относятся к перспективе). В конечном итоге именно продукт гена или иРНК, определяет, какая нить одного сегмента двухцепочечной ДНК называется смысловой или антисмысловой.

Биспецифический лиганд

[0231] В контексте настоящего изобретения термин «биспецифический лиганд» или «биспецифический фрагмент» относится к лиганду, который связывается с двумя разными лиганд-связывающими доменами. Согласно определенным вариантам осуществления лиганд представляет собой малую молекулу, или пептид, или полипептид. Согласно другим вариантам осуществления лиганд-связывающий домен представляет собой «домен димеризации», который можно вводить в качестве пептидной метки в белок. Согласно различным вариантам осуществления можно индуцировать димеризацию двух белков, каждый из которых содержит одинаковые или разные домены димеризации, посредством связывания каждого домена димеризации с биспецифическим лигандом. В контексте настоящего изобретения «биспецифические лиганды» могут эквивалентно относиться к «химическим индукторам димеризации» или «CID».

Cas9

[0232] Термин «Cas9» или «нуклеаза Cas9» относится к РНК-управляемой нуклеазе, содержащей домен Cas9 или его фрагмент (*например*, белок, содержащий активный или неактивный домен расщепления ДНК Cas9 и/или домен связывания гРНК Cas9). «Домен Cas9» в контексте настоящего изобретения представляет собой белковый фрагмент, содержащий активный или неактивный домен расщепления Cas9 и/или связывающий гРНК домен Cas9. «Белок Cas9» представляет собой полноразмерный белок Cas9. Нуклеаза Cas9 иногда также упоминается как нуклеаза casn1 или CRISPR (Кластеризованные регулярные промежуточные короткие палиндромные повторы)-ассоциированная нуклеаза. CRISPR представляет собой адаптивную иммунную систему, обеспечивающую защиту от мобильных генетических элементов (вирусов, мобильных элементов и конъюгативных плазмид). Кластеры CRISPR содержат спейсеры, последовательности, комплементарные предшествующим мобильным элементам, и нацелены на инвазивные нуклеиновые

кислоты. Кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В системах CRISPR типа II для правильного процессинга пре-crРНК требуется транс-кодируемая малая РНК (tracrRNA), эндогенная рибонуклеаза 3 (rnc) и домен Cas9. tracrRNA служит ориентиром для процессинга пре-crРНК с помощью рибонуклеазы 3. Впоследствии Cas9/crРНК/tracrРНК эндонуклеолитически расщепляет линейную или кольцевую дцДНК-мишень, комплементарную спейсеру. Целевая цепь, некомплементарная crРНК, сначала разрезается эндонуклеолитически, а затем обрезается 3'-5' экзонуклеолитически. В природе для связывания и расщепления ДНК обычно требуется белок и обе РНК. Однако можно сконструировать одиночные гидовые РНК («crРНК» или просто «gNRA») для включения аспектов как crРНК, так и tracrRNA в один вид РНК. См., например, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полные содержания каждой из которых включены в настоящий документ посредством ссылки. Cas9 распознает короткий мотив в повторяющихся последовательностях CRISPR (мотив PAM или смежный с протоспейсером мотив), чтобы помочь отличить свой от чужого. Последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной области техники (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), полные содержания каждой из которых включены в настоящий документ посредством ссылки). Ортологи Cas9 были описаны у различных видов, включая без ограничения *S. pyogenes* и *S. thermophilus*. Дополнительные подходящие нуклеазы и последовательности Cas9 будут очевидны специалистам в данной области техники на основе настоящего раскрытия, и такие нуклеазы и последовательности Cas9 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, полные содержания каждой из которых включены в настоящий документ посредством ссылки. Согласно некоторым вариантам осуществления нуклеаза Cas9 содержит одну или несколько мутаций, которые частично нарушают или инактивируют домен расщепления ДНК.

[0233] Нуклеаза-инактивированный домен Cas9 может взаимозаменяемо называться

белком «dCas9» («мертвая нуклеаза» Cas9). Известны способы создания домена Cas9 (или его фрагмента), имеющего неактивный домен расщепления ДНК (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, полные содержания каждой из которых включены в настоящий документ посредством ссылки). Например, домен расщепления ДНК Cas9, как известно, включает два субдомена, субдомен нуклеазы HNH и субдомен RuvC1. Субдомен HNH расщепляет цепь, комплементарную гРНК, тогда как субдомен RuvC1 расщепляет некомплементарную цепь. Мутации внутри этих субдоменов могут подавлять нуклеазную активность Cas9. Например, мутации D10A и H840A полностью инактивируют нуклеазную активность Cas9 *S. pyogenes* (Jinek *et al.*, *Science*. 337:816-821(2012), Qi *et al.*, *Cell*. 28,152(5):1173-83 (2013)). Согласно некоторым вариантам осуществления предусмотрены белки, содержащие фрагменты Cas9. Например, согласно некоторым вариантам осуществления белок содержит один из двух доменов Cas9: (1) домен связывания гРНК Cas9 или (2) домен расщепления ДНК Cas9. Согласно некоторым вариантам осуществления белки, содержащие Cas9 или их фрагменты, называются «вариантами Cas9». Вариант Cas9 имеет гомологию с Cas9 или его фрагментом. Например, вариант Cas9 на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере приблизительно 99.5% идентичен, по меньшей мере приблизительно 99.8% идентичен или по меньшей мере приблизительно 99.9% идентичен Cas9 дикого типа (например, SpCas9 последовательности SEQ ID NO: 18). Согласно некоторым вариантам 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 замен по сравнению с Cas9 дикого типа (например, SpCas9 последовательности SEQ ID NO: 18). Согласно некоторым вариантам осуществления вариант Cas9 содержит фрагмент последовательности SEQ ID NO: 18 Cas9 (*например*, домен связывания гРНК или домен расщепления ДНК), так что фрагмент на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере

приблизительно 99.5% идентичен или по меньшей мере приблизительно 99.9% идентичен соответствующему фрагменту Cas9 дикого типа (например, SpCas9 последовательности SEQ ID NO: 18). Согласно некоторым вариантам осуществления фрагмент составляет по меньшей мере 30%, по меньшей мере 35%, по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, 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% идентичен, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или по меньшей мере 99.5% от длины аминокислот соответствующего Cas9 дикого типа (например, SpCas9 последовательности SEQ ID NO: 18).

кДНК

[0234] Термин «кДНК» относится к цепи ДНК, скопированной с матрицы РНК. кДНК комплементарна матрице РНК.

Круговой пермутант

[0235] В контексте настоящего изобретения термин «круговой пермутант» относится к белку или полипептиду (например, Cas9), содержащему круговую перестановку, которая представляет собой изменение структурной конфигурации белка, включающее изменение порядка аминокислот, содержащегося в аминокислотной последовательности белка. Другими словами, круговые пермутанты представляют собой белки, которые имеют измененные N- и C-концы по сравнению с аналогом дикого типа, например, C-концевая половина белка дикого типа становится новой N-концевой половиной. Круговая перестановка (или CP) представляет собой, по существу, топологическую перестройку первичной последовательности белка, соединяющую его N- и C-концы, часто с пептидным линкером, при одновременном расщеплении его последовательности в другом положении для создания новых, смежных N- и C-концов. В результате создается белковая структура с другой связностью, но которая часто может иметь такую же в целом подобную трехмерную (3D) форму и, возможно, иметь улучшенные или измененные характеристики, в том числе сниженную протеолитическую чувствительность, улучшенную каталитическую активность, измененное связывание субстрата или лиганда и/или улучшенную термостабильность. Круговые пермутантные белки могут встречаться в природе (например, конканавалин A и лектин). Кроме того, круговая перестановка может происходить в результате посттрансляционных модификаций или может быть создана с использованием рекомбинантных методов.

Cas9 с круговыми перестановками

[0236] Термин «Cas9 с круговыми перестановками» относится к любому белку Cas9

или его варианту, который встречается в виде кругового пермутанта, в результате чего его N- и С-концы были местно перестановлены. Такие белки Cas9 с круговыми перестановками («CP-Cas9») или их варианты сохраняют способность связываться с ДНК при образовании комплекса с гидовой РНК (гРНК). См., Oakes et al., «Protein Engineering of Cas9 for enhanced function,» *Methods Enzymol*, 2014, 546: 491–511 and Oakes et al., «CRISPR-Cas9 CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification,» *Cell*, January 10, 2019, 176: 254-267, каждая из которых включена в настоящий документ посредством ссылки. Настоящее изобретение охватывает любой ранее известный CP-Cas9 или применение нового CP-Cas9 при условии, что полученный белок с круговой перестановкой сохраняет способность связываться с ДНК при образовании комплекса с гидовой РНК (гРНК). Примерами белков CP-Cas9 являются SEQ ID NO: 77-86.

CRISPR

[0237] CRISPR представляет собой семейство последовательностей ДНК (т.е. кластеров CRISPR) у бактерий и архей, которые представляют собой фрагменты предшествующих инфекций, вызванных вирусом, проникшим в прокариоты. Фрагменты ДНК используются прокариотической клеткой для обнаружения и уничтожения ДНК при последующих атаках подобных вирусов и эффективно составляют вместе с набором CRISPR-ассоциированных белков (включая Cas9 и его гомологи) и CRISPR-ассоциированной РНК прокариотическую иммунную систему защиты. В природе кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В некоторых типах систем CRISPR (например, в системах CRISPR типа II) для правильного процессинга пре-crРНК требуется транс-кодируемая малая РНК (tracrRNA), эндогенная рибонуклеаза 3 (rnc) и белок Cas9. tracrRNA служит направляющей для процессинга пре-crРНК с помощью рибонуклеазы 3. Затем Cas9/crРНК/tracrРНК эндонуклеолитически расщепляет линейную или кольцевую двухцепочечную ДНК-мишень, комплементарную РНК. В частности, целевая цепь, некомплементарная crРНК, сначала разрезается эндонуклеолитически, а затем обрезается 3'-5' экзонуклеолитически. В природе для связывания и расщепления ДНК обычно требуется белок и обе РНК. Однако одиночные гидовые РНК («сгРНК» или просто «gNRA») могут быть сконструированы так, чтобы включать аспекты как crРНК, так и tracrРНК в один вид РНК — гидовую РНК. См., *например*, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), полные содержания которых включены в настоящий документ посредством ссылки. Cas9 распознает короткий мотив в повторяющихся последовательностях CRISPR (мотив PAM или смежный с протоспейсером мотив), чтобы помочь отличить свой от чужого. Биология CRISPR, а также последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной

области техники (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), полные содержания каждой из которых включены в настоящий документ посредством ссылки). Ортологи Cas9 были описаны у различных видов, включая без ограничения *S. pyogenes* и *S. thermophilus*. Дополнительные подходящие нуклеазы и последовательности Cas9 будут очевидны специалистам в данной области техники на основе настоящего раскрытия, и такие нуклеазы и последовательности Cas9 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, полные содержания каждой из которых включены в настоящий документ посредством ссылки.

[0238] В некоторых типах систем CRISPR (например, в системах CRISPR типа II) для правильного процессинга пре-crРНК требуется транс-кодируемая малая РНК (tracrРНК), эндогенная рибонуклеаза 3 (rnc) и белок Cas9. tracrRNA служит направляющей для процессинга пре-crРНК с помощью рибонуклеазы 3. Затем Cas9/crРНК/tracrРНК эндонуклеолитически расщепляет линейную или кольцевую нуклеиновую кислоту-мишень, комплементарную РНК. В частности, целевая цепь, некомплементарная crРНК, сначала разрезается эндонуклеолитически, а затем обрезается 3'-5' экзонуклеолитически. В природе для связывания и расщепления ДНК обычно требуется белок и обе РНК. Тем не менее, одиночные гидовые РНК («сгРНК» или просто «gRNA») могут быть сконструированы для включения вариантов как crРНК, так и tracrRNA в один вид РНК - гидовую РНК.

[0239] В общем, «система CRISPR» относится к транскриптам и другим элементам, участвующим в экспрессии или управлении активностью CRISPR-ассоциированных («Cas») генов, включая последовательности, кодирующие ген Cas, tracr (трансактивирующий CRISPR) последовательность (например, tracrRNA или активная частичная tracrРНК), сопряженную последовательность tracr (включающую «прямой повтор» и частичный прямой повтор, процессированный tracrРНК, в контексте эндогенной системы CRISPR), гидовую последовательность (также называемую «спейсер» в контексте

эндогенной системы CRISPR) или другие последовательности и транскрипты из локуса CRISPR. tracrРНК системы комплементарна (полностью или частично) сопряженной последовательности tracr, присутствующей на гидовой РНК.

Матрица для синтеза ДНК

[0240] В контексте настоящего изобретения термин «матрица для синтеза ДНК» относится к участку или части удлиняющего плеча РЕгРНК, которые используются в качестве матричной цепи полимеразой праймирующего редактора для кодирования 3'-лоскута одноцепочечной ДНК, который содержит желаемое редактирование, а затем с помощью механизма праймированного редактирования, заменяет соответствующую эндогенную цепь ДНК в целевом участке. Согласно различным вариантам осуществления матрица для синтеза ДНК показана на фиг. 3A (в контексте РЕгРНК, содержащей 5'-удлиняющее плечо), фиг. 3B (в контексте РЕгРНК, содержащей 3'- удлиняющее плечо), фиг. 3C (в контексте внутреннего удлиняющего плеча), фиг. 3D (в контексте 3′-удлиняющего плеча) и на фиг. 3E (в контексте 5'-удлиняющего плеча). Удлиняющее плечо, включающее матрицу для синтеза ДНК, может состоять из ДНК или РНК. В случае РНК полимеразой праймирующего редактора может быть РНК-зависимая ДНК-полимераза (например, обратная транскриптаза). В случае ДНК полимераза праймирующего редактора может быть ДНК-зависимой ДНК-полимеразой. Согласно различным вариантам осуществления (например, как представлено на фиг. 3D-3E), матрица для синтеза ДНК (4) может содержать «матрицу редактирования» и «плечо гомологии», а также всю или часть необязательного 5'-концевого участка модификатора e2. То есть, в зависимости от природы участка e2 (например, включает ли он шпильку, петлю или вторичную структуру «петля-на-стебле»), полимераза может также не кодировать ни одной части, некоторую часть или весь участок e2. Другими словами, в случае 3'-удлиняющего плеча матрица для синтеза ДНК (3) может включать часть удлиняющего плеча (3), которая простирается от 5'-конца участка связывания праймера (PBS) до 3'-конца ядро гРНК, которая может действовать как матрица для синтеза одноцепочечной ДНК полимеразой (например, обратной транскриптазой). В случае 5'-удлиняющего плеча матрица для синтеза ДНК (3) может включать часть удлиняющего плеча (3), которая простирается от 5'-конца молекулы РЕгРНК до 3'-конца матрицы редактирования. Предпочтительно, матрица для синтеза ДНК исключает участок связывания праймера (PBS) РЕгРНК, имеющий либо 3′-удлиняющее плечо, либо 5′-удлиняющее плечо. Определенные варианты осуществления, описанные в настоящем документе, (например, фиг. 71A) относится к «матрице RT», которая включает матрицу редактирования и плечо гомологии, т.е. последовательность плеча удлинения РЕгРНК, которая фактически используется в качестве матрицы в ходе синтеза ДНК. Термин

«матрица RT» эквивалентен термину «матрица для синтеза ДНК».

[0241] В случае *транс*-праймированного редактирования (например, фиг. 3G и фиг. 3H), участок связывания праймера (PBS) и матрица для синтеза ДНК могут быть сконструированы в виде отдельной молекулы, называемой РНК матрицы транс-праймирующего редактора (tPERT).

Домен димеризации

[0242] Термин «домен димеризации» относится к лиганд-связывающему домену, который связывается со связывающим фрагментом биспецифического лиганда. «Первый» домен димеризации связывается с первым фрагментом связывания биспецифического лиганда, а «второй» домен димеризации связывается со вторым фрагментом связывания того же биспецифического лиганда. Когда первый домен димеризации сливается с первым белком (например, посредством PE, как обсуждается в настоящем документе), а второй домен димеризации (например, посредством PE, как обсуждается в настоящем документе) сливается со вторым белком, первый и второй белок димеризуются в присутствие 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.

По ходу транскрипции

[0243] В контексте настоящего изобретения термины «против хода транскрипции» и «по ходу транскрипции» являются терминами относительности, которые определяют линейное положение по меньшей мере двух элементов, расположенных в молекуле нуклеиновой кислоты (будь то одноцепочечная или двухцепочечная), которая ориентирована в от 5'- к 3'- направление. В частности, первый элемент находится против хода транскрипции от второго элемента в молекуле нуклеиновой кислоты, где первый элемент расположен где-то на 5' от второго элемента. Например, SNP находится против хода транскрипции от Cas9-индуцированного участка разрыва, если SNP находится на 5'-стороне участка разрыва. И наоборот, первый элемент находится по ходу транскрипции от второго элемента в молекуле нуклеиновой кислоты, где первый элемент расположен где-то на 3' от второго элемента. Например, SNP находится по ходу транскрипции от Cas9-индуцированного участка разрыва, если SNP находится на 3'-стороне участка разрыва. Молекула нуклеиновой кислоты может представлять собой ДНК (двухцепочечную или одноцепочечную), РНК (двухцепочечную или одноцепочечную) или гибрид ДНК и РНК. Анализ одинаков для одноцепочечной молекулы нуклеиновой кислоты и для двухцепочечной молекулы, поскольку термины «против хода транскрипции» и «по ходу транскрипции» относятся только к одной цепи молекулы нуклеиновой кислоты, за

исключением того, что необходимо выбрать, какая цепь двухцепочечной молекулы рассматривается. Часто цепь двухцепочечной ДНК, которую можно использовать для определения относительности положения по меньшей мере двух элементов, является «смысловой» или «кодирующей» цепью. В области генетики «смысловая» цепь представляет собой сегмент двухцепочечной ДНК, который проходит от 5' к 3' и комплементарен антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' к 5'. Так, например, нуклеотидное основание SNP находится «по ходу транскрипции» последовательности промотора в геномной ДНК (которая является двухцепочечной), если нуклеотидное основание SNP находится на 3'-стороне промотора на смысловой или кодирующей цепи.

Матрица редактирования

[0244] Термин «матрица редактирования» относится к части плеча удлинения, которая кодирует желаемое изменение в одноцепочечном 3'-лоскуте ДНК, который синтезируется полимеразой, например, ДНК-зависимой ДНК-полимеразой, РНК-зависимой ДНК-полимеразой (например, обратной транскриптазой). Определенные варианты осуществления, описанные в настоящем документе, (например, фиг. 71A) относятся к «матрице RT», которая относится как к матрице редактирования, так и к плечу гомологии, т.е. последовательности плеча удлинения РЕгРНК, которая фактически используется в качестве матрицы при синтезе ДНК. Термин «матрица RT редактирования» также эквивалентен термину «матрица для синтеза ДНК», однако, «матрица RT для редактирования» означает использование праймирующего редактора, имеющего полимеразу, которая представляет собой обратную транскриптазу, а матрица для синтеза ДНК более широким образом обозначает использование праймирующего редактора, имеющего любую полимеразу.

Эффективное количество

[0245] В контексте настоящего изобретения термин «эффективное количество» означает количество биологически активного агента, которое является достаточным для реализации требуемого биологического ответа. Например, согласно некоторым вариантам осуществления эффективное количество праймирующего редактора (PE) может означать количество редактора, которое является достаточным для редактирования нуклеотидной последовательности целевого участка, например, генома. Согласно некоторым вариантам осуществления указанное в настоящем документе эффективное количество праймирующего редактора (PE), например, слитого белка, содержащего домен никазы Cas9 и обратную транскриптазу, может означать количество слитого белка, которое является достаточным для индуцирования редактирования целевого участка, специфично

связанного и редактируемого слитым белком. Как очевидно специалисту в данной области техники, 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, подлежащего редактированию определенного аллеля, генома, или целевого участка, являющихся мишенью клеток или тканей, и используемого агента.

Обратная транскриптаза сниженной точности

[0246] В контексте настоящего изобретения термин обратная транскриптаза «сниженной точности» (или, более широко, любая полимераза) означает 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, которая включена в настоящий документ посредством ссылки. Таким образом, для целей настоящего изобретения термин «сниженной точности» означает RT, которые имеют частоту ошибок в виде более чем одной ошибки на 15000 выстраиваний нуклеотидных оснований (6,7 x 10^{-5} или более), например, 1 ошибку на 14000 нуклеотидных оснований (7,14 x 10^{-5} или более), 1 ошибку на 13000 нуклеотидных оснований или менее (7,7 x 10^{-5} или более), 1 ошибку на 12000 нуклеотидных оснований или менее (7,7 x 10^{-5} или более), 1 ошибку на 11000 нуклеотидных оснований или менее (9,1 x 10^{-5} или более), 1 ошибку на 10000 нуклеотидных оснований или менее (1 x 10^{-4} или 0,0001 или более), 1 ошибку на 9000 нуклеотидных оснований или менее (0,00011 или более), 1 ошибку на 8000 нуклеотидных оснований или менее (0,00013 или более) 1 ошибку на 7000 нуклеотидных оснований или менее (0,00014 или более), 1 ошибку на 6000 нуклеотидных оснований или менее (0,00016 или более), 1 ошибку на 5000 нуклеотидных оснований или менее (0,0002 или более), 1 ошибку на 4000 нуклеотидных оснований или менее (0,00025 или более), 1 ошибку на 3000 нуклеотидных оснований или менее (0,00033 или более), 1 ошибку на 2000 нуклеотидных оснований или менее (0,00050 или более), или 1 ошибку на 1000 нуклеотидных оснований или менее (0,001 или более), или 1 ошибку на 500 нуклеотидных оснований или менее (0,002

или более), или 1 ошибку на 250 нуклеотидных оснований или менее (0,004 или более).

Экстеин

[0247] В контексте настоящего изобретения термин «экстеин» означает полипептидную последовательность, которая фланкирована интеином и лигируется с другим экстеином в процессе сплайсинга белка с образованием зрелого подвергнутого сплайсингу белка. Как правило, интеин фланкирован двумя последовательностями экстеина, которые лигируется вместе, когда интеин катализирует свою собственную эксцизию. Экстеины, соответственно, представляют собой белковый аналог экзонов, обнаруживаемых в гРНК. Например, полипептид, содержащий интеин, может иметь структуру экстеин (1M) - интеин - экстеин(C). После эксцизии интеина и сплайсинга двух экстеинов, полученные структуры представляют собой экстеин (M) - 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 расщепленным интеином, при этом эксцизия расщепленных интеинов вызывает совместный сплайсинг последовательностей экстеина.

Плечо удлинения

[0248] Термин «плечо удлинения» означает компонент нуклеотидной последовательности PEгРНК, который обеспечивает несколько функций, включая участок связывания праймера и матрицу редактирования для обратной транскриптазы. Согласно некоторым вариантам осуществления, например, на фиг. 3D, плечо удлинения расположено на 3'-конце гидовой РНК Согласно другим вариантам осуществления, например, на Фиг. 3H, плечо удлинения расположено на 5'-конце гидовой РНК. Согласно некоторым вариантам осуществления плечо удлинения также включает плечо гомологии. Согласно различным вариантам осуществления плечо удлинения содержит следующие компоненты в направлении от 5' к 3': плечо удлинения, матрицу редактирования, и участок связывания праймера. Поскольку полимеризационная активность обратной транскриптазы осуществляется в направлении от 5' к 3', предпочтительная группировка плеча удлинения, матрицы редактирования и участка связывания праймера соответствует направлению от 5' к 3', так что обратная транскриптаза после праймирования отожженной праймирующей последовательностью обеспечивает полимеризацию одиночной цепи ДНК с использованием матрицы редактирования в качестве цепи комплементарной матрицы. Дополнительные подробности, такие как длина плеча удлинения, описаны в других разделах описания.

[0249] 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 праймирующей последовательностью, которая образована из цепи эндогенной ДНК целевого участка, когда он становится разорванным комплексом праймированного редактирования, тем самым высвобождая 3'-конец на эндогенной цепи с разрывом. Как описано в настоящем документе, связывание праймирующей последовательности с участком связывания праймера на плече удлинения РЕгРНК создает дуплексный участок со свободным 3′ концом (т.е. 3' праймирующей 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 одноцепочечной 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 кодирована в продукте в виде одноцепочечной ДНК (т.е. 3' одноцепочечного лоскута ДНК, содержащего информацию в виде желаемого генетического редактирования) полимеразой комплекса праймированного редактирования, и которая в конечном итоге замещает соответствующую эндогенную цепь ДНК целевого участка, который располагается непосредственно по ходу транскрипции от РЕ-индуцированного участка разрыва. Без ограничения теорией, полимеризация матрицы для синтеза ДНК продолжается по направлению к 5'-концу плеча удлинения до события терминации. Полимеризация может терминироваться рядом способов, включая, но без ограничения (a) достижение 5'- конца РЕгРНК (например, в случае 5' плеча удлинения, где у ДНК-полимеразы просто заканчивается матрица), (b) достижение непроходимой вторичной структуры РНК (например, шпильки или петли на стебле), 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.

Эндонуклеаза лоскутов (например, FEN1)

[0250] В контексте настоящего изобретения термин «эндонуклеаза лоскутов» означает фермент, который катализирует удаление 5'-одноцепочечных лоскутов ДНК. Они представляют собой встречающиеся в природе ферменты, которые процессируют удаление 5' лоскутов, образованных в ходе клеточных процессов, включая репликацию ДНК. В описанных в настоящем документе способах праймированного редактирования могут применяться эндогенно поступающие или представленные в транс эндонуклеазы лоскутов для удаления 5'- лоскута эндогенной ДНК, образованного в целевом участке в ходе праймированное редактировании. Эндонуклеазы лоскутов известны в данной области

техники 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 (каждая из которых включена в настоящий документ посредством ссылки). Иллюстративная эндонуклеаза лоскутов представляет собой FEN1, которая может быть представлена следующей аминокислотной последовательностью:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ	SEQ ID NO:
FEN1 ДИКОГО ТИПА	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIY QFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENG IKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQA AGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLD APSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMR HLTASEAKKLPIQEFHLSRILQELGLNQEQFVDLCILLGS DYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPE NWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFM CGEKQFSEERIRSGVKRLSKSRQGSTQGRLDDFFKVTGS LSSAKRKEPEPKGSTKKKAKTGAAGKFKRGK	SEQ ID NO: 7

Функциональный эквивалент

[0251] Термин «функциональный эквивалент» означает вторую биомолекулу, которая является эквивалентной в плане функции, но не обязательно эквивалентной в плане структуры относительно первой биомолекулы. Например, «эквивалент Cas9» означает белок, который имеет такие же или по существу такие же функции, что и Cas9, но не обязательно такую же аминокислотную последовательность. В контексте настоящего изобретения в описании указан «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

[0252] В контексте настоящего изобретения термин «слитый белок» означает гибридный полипептид, который содержит 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 содержать различные домены, например, домен связывания нуклеиновой кислоты (например, домен связывания гРНК Cas9, который направляет связывание белка с целевым участком) 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 в настоящем документе белков может быть получен с помощью любого способа, известного в данной области техники. Например, указанные в настоящем документе белки могут быть получены с помощью экспрессии рекомбинантного белка и очистки, которые являются особенно подходящими для слитых белков, включающих содержащих линкер. Способы экспрессии рекомбинантных белков и очистки хорошо известны и включают способы, описанные в Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), полное содержание которой включено в настоящий документ посредством ссылки.

Представляющий интерес ген (GOI)

[0253] Термин «представляющий интерес ген» или «GOI» означает 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, ядерный мембранный транспортер, 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 без ограничения, информационную РНК (иРНК), транспортную РНК (тРНК), рибосомную РНК (рРНК), малую ядерную РНК (мяРНК), антисмысловую РНК, гидовую РНК, микроРНК (микроРНК), малую интерферирующую РНК (siРНК) и бесклеточную РНК (бкРНК).

Гидовая РНК («гРНК[11], gRNA»)

[0254] В контексте настоящего изобретения термин «гидовая РНК» означает определенный тип гидовой 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 гидовой РНК. Однако данный термин также охватывает эквивалентные молекулы гидовой нуклеиновой кислоты, которые ассоциируется с эквивалентами, гомологами, ортологами, или паралогами Cas9, которые

встречаются или не встречаются в природе (например, являются сконструированными или рекомбинантными) и которые иным образом программируют эквивалент Cas9 для локализации в определенной целевой нуклеотидной последовательности. Эквиваленты Cas9 могут включать другой napDNAbp из системы CRISPR любого типа (например, типа II, V, VI), включая Cpfl (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 которой включено в настоящий документ посредством ссылки. В настоящем документе указаны иллюстративные последовательности и структуры гидовой РНК. Кроме того, в настоящем документе указаны способы конструирования подходящих последовательностей гидовых РНК. В контексте настоящего изобретения «гидовая РНК» может также означать «традиционную гидовую РНК» для ее противопоставления модифицированным формам гидовой РНК, обозначаемым как «гидовые РНК праймированного редактирования» (или «РЕгРНК, PEgRNA»), которые были изобретены для использования в описанных в настоящем документе способах и композициях для праймированного редактирования.

[0255] Гидовые РНК или РЕгРНК могут содержать различные структурные элементы, которые включают без ограничения:

[0256] Спейсерную последовательность – последовательность в гидовой РНК или РЕгРНК (имеющую приблизительно 20 нуклеотидов в длину), которая связывается с протоспейсером в ДНК-мишени.

[0257] Ядро гРНК (или последовательность каркаса или основной цепи гРНК) - означает последовательность в гРНК, которая отвечает за связывание Cas9, она не включает спейсер/нацеливающую последовательность 20 п.о., которая используется для направления Cas9 к ДНК-мишени.

[0258] Плечо удлинения представляет собой одноцепочечное удлинение на 3'-конце или 5'-конце РЕгРНК, которое содержит последовательность участка связывания праймера и матрицы для синтеза ДНК, которая кодирует посредством полимеразы (например, обратной транскриптазы) одноцепочечный лоскут ДНК, содержащий представляющее интерес изменение, которое затем встраивается в эндогенную ДНК посредством замещения соответствующей эндогенной цепи, тем самым обеспечивая внесение требуемого генетического изменения.

[0259] Терминатор транскрипции – гидовая РНК или РЕгРНК может содержать последовательность терминации транскрипции на 3'-конце молекулы.

Плечо гомологии

[0260] Термин «плечо гомологии» обозначает часть плеча удлинения, которая кодирует часть получаемого кодируемого обратной транскриптазой одноцепочечного лоскута ДНК, который подвергается встраиванию в целевой участок ДНК за счет замещения эндогенной цепи. Часть одноцепочечного локуса ДНК, кодируемая плечом гомологии, является комплементарной не подвергаемой редактированию цепи последовательности-мишени ДНК, которая обеспечивает вытеснение эндогенной цепи и отжиг одноцепочечного локуса ДНК в его месте, тем самым обеспечивая внесение изменения. Данный компонент дополнительно описан в настоящем документе. Плечо гомологии представляет собой часть матрицы для синтеза ДНК, поскольку оно по определению кодируется полимеразой описанных в настоящем документе праймирующих редакторов.

Клетка-хозяин

[0261] В контексте настоящего изобретения термин «клетка-хозяин» означает клетку, которая может являться хозяином, обеспечивать репликацию и экспрессию описанного в настоящем документе вектора, например, вектора, содержащего молекулу нуклеиновой кислоты, кодирующей слитый белок, содержащий Cas9 или эквивалент Cas9 и обратную транскриптазу.

Интеины

[0262] В контексте настоящего изобретения термин «интеин» означает автопроцессирующие полипептидные домены, располагающиеся в организмах всех живых видов. Интеин (промежуточный белок) обеспечивает реализацию уникального автопроцессирующего события, известного как сплайсинг белка, при котором он вырезает себя из более крупного полипептида-предшественника посредством расщепления двух пептидных связей и, в данном процессе, обеспечивает лигирование фланкирующих последовательностей экстеина (внешнего белка) посредством образования новой пептидной связи. Данное перестроение осуществляется после трансляции (или возможно одновременно с трансляцией), поскольку гены интеина являются встроенными в каркас в других генах, кодирующих белок. 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 «расщепленных 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 белка-предшественника с одновременным слиянием

фланкирующих последовательностей белка, известных как экстеины (см. обзор 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)).

[0263] В контексте настоящего изобретения термин «сплайсинг белка» означает процесс, при котором внутренний участок белка-предшественника (интеин) удаляется и фланкирующие участки белка (экстеины) лигируются с образованием 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 катализа 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, но не экспрессируются в виде отдельных белков. Сплайсинг белка может быть также осуществлен *в транс* с использованием расщепленных интеинов, экспрессируемых на отдельных полипептидах, которые спонтанным образом комбинируются с образованием единого интеина, который затем подвергается процессу сплайсинга белка для соединения с отдельными белками.

[0264] Выяснение механизма сплайсинга белков привело к ряду применений 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)). Каждый ссылочный источник включен в

настоящий документ посредством ссылки.

Лиганд-зависимый интеин

[0265] В контексте настоящего изобретения термин «лиганд-зависимый интеин» относится к интеину, который содержит лиганд-связывающий домен. Как правило, лиганд-связывающий домен встраивается в аминокислотную последовательность интеина, в результате чего получается структура интеин (N) – лиганд-связывающий домен – интеин (C). Как правило, лиганд-зависимые интеины не проявляют или проявляют только минимальную активность сплайсинга белков в отсутствие соответствующего лиганда и заметное увеличение активности сплайсинга белков в присутствии лиганда. Согласно некоторым вариантам осуществления лиганд-зависимый интеин не проявляет наблюдаемой активности сплайсинга в отсутствие лиганда, но проявляет активность сплайсинга в присутствии лиганда. Согласно некоторым вариантам осуществления лиганд-зависимый интеин проявляет наблюдаемую активность сплайсинга белков в отсутствие лиганда и активность сплайсинга белков в присутствии соответствующего лиганда, которая по меньшей мере в по меньшей мере 5 раз, по меньшей мере 10 раз, по меньшей мере 50 раз, по меньшей мере 100 раз, по меньшей мере 150 раз, по меньшей мере 200 раз, по меньшей мере 250 раз, по меньшей мере 500 раз, по меньшей мере 1000 раз, по меньшей мере 1500 раз, по меньшей мере 2000 раз, по меньшей мере 2500 раз, по меньшей мере 5000 раз, по меньшей мере 10000 раз, по меньшей мере 20000 раз, по меньшей мере 25000 раз, по меньшей мере 50000 раз, по меньшей мере 100000 раз, по меньшей мере 500000 раз или по меньшей мере 1000000 раз больше, чем активность, наблюдаемая в отсутствие лиганда. Согласно некоторым вариантам осуществления увеличение активности зависит от дозы с по меньшей мере 1 порядком увеличения, по меньшей мере 2 порядками увеличения, по меньшей мере 3 порядками увеличения, по меньшей мере 4 порядками увеличения или по меньшей мере 5 порядками увеличения, что позволяет точно регулировать активности интеина путем регулирования концентрации лиганда. Подходящие лиганд-зависимые интеины известны в данной области техники и включают интеины, представленные 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 каждой из которых

включено в настоящий документ посредством ссылки. Иллюстративные последовательности являются следующими:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
2-4 ИНТЕИН:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 8)
3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYTNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 9)
30R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 10)
30R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE

	LHTLVAEGVVVHNC (SEQ ID NO: 11)
30R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 12)
37R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC ((SEQ ID NO: 13)
37R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 14)
37R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 15)

Линкер

[0266] Термин «линкер» в контексте настоящего изобретения относится к молекуле, связывающей две другие молекулы или фрагменты. Линкер может представлять собой аминокислотную последовательность в случае линкера, соединяющего два слитых белка. Например, Cas9 может быть слит с обратной транскриптазой аминокислотной линкерной последовательностью. Линкер также может представлять собой нуклеотидную последовательность в случае соединения двух нуклеотидных последовательностей вместе. Например, в данном случае традиционная гидовая РНК связана через спейсерную или линкерную нуклеотидную последовательность с РНК-удлинением гидовой РНК праймированного редактирования, которая может содержать последовательность матрицы RT и участок связывания праймера RT. Согласно другим вариантам осуществления линкер представляет собой органическую молекулу, группу, полимер или химическую часть. Согласно некоторым вариантам осуществления линкер составляет 5-100 аминокислот в длину, например, 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 в длину. Также охватываются более длинные или короткие линкеры.

Выделенный

[0267] «Выделенный» означает измененный или изъятый из естественного состояния. Например, нуклеиновая кислота или пептид, естественным образом присутствующие в живом организме, не являются «выделенными», но те же самые нуклеиновые кислоты или пептиды, частично или полностью отделенные от сосуществующих материалов в их естественном состоянии, являются «выделенными». Выделенная нуклеиновая кислота или белок могут существовать в существенно очищенной форме или могут существовать в ненативной среде, такой как, например, клетка-хозяин.

[0268] Согласно некоторым вариантам осуществления представляющий интерес ген кодируется выделенной нуклеиновой кислотой. В контексте настоящего изобретения термин «выделенный» относится к характеристике материала, как представлено в настоящем документе, удаленного из его исходной или естественной среды (например, естественной среды, если он встречается в природе). Следовательно, встречающийся в природе полинуклеотид, или белок, или полипептид, присутствующий в живом животном, не является выделенным, но тот же самый полинуклеотид или полипептид, отделенный вмешательством человека от некоторых или всех сосуществующих материалов в природной системе, является выделенным. Искусственный или сконструированный материал, например, не встречающаяся в природе конструкция нуклеиновой кислоты, такая как конструкции экспрессии и векторы, описанные в настоящем документе,

соответственно, также называются выделенными. Материал не обязательно очищать, чтобы его можно было выделить. Соответственно, материал может быть частью вектора и/или частью композиции, и все же быть выделенным в том смысле, что такой вектор или композиция не является частью окружающей среды, в которой материал встречается в природе.

Методика мечения MS2

[0269] Согласно различным вариантам осуществления (например, как проиллюстрировано в вариантах осуществления на фиг. 72-73 и в примере 19), термин «методика мечения MS2» относится к комбинации «домена взаимодействия РНК-белок» (также известного как «домен рекрутинга РНК-белок или белок») в паре с РНК-связывающим белком, который специфически распознает и связывается с доменом взаимодействия РНК-белок, например, со специфической структурой шпильки. Эти типы систем можно использовать для привлечения различных функций к комплексу праймированного редактирования, который связан с целевым участком. Методика мечения MS2 основана на естественном взаимодействии белка оболочки бактериофага MS2 («MCP» или «MS2cp») со структурой петля на стебле или шпилька, присутствующей в геноме фага, т.е. «шпилькой MS2». В случае праймированного редактирования методика мечения MS2 включает введение шпильки MS2 в желаемую молекулу РНК, участвующую в праймированном редактировании (например, PEгРНК или tPERT), которая затем представляет собой специфическую взаимодействующую связывающую мишень для РНК-связывающего белка, который распознает и связывается с этой структурой. В случае шпильки MS2 она распознается и связывается белком оболочки бактериофага MS2 (MCP). И, если MCP слит с другим белком (например, обратной транскриптазой или другой ДНК-полимеразой), то шпилька MS2 может быть использована для «вербовки» этого другого белка *в транс* по отношению к целевому сайту, занятому комплексом праймированного редактирования.

[0270] Праймирующие редакторы, описанные в настоящем документе, могут включать в качестве аспекта любой известный домен взаимодействия РНК-белок для рекрутирования или «совместной локализации» конкретных представляющих интерес функций в комплексе праймированного редактирования. Обзор других модульных доменов взаимодействия РНК-белок описан в данной области техники, 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, and Zalatan et al., «Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds,» *Cell*, 2015, Vol.160: 339-350, каждая из которых включена в настоящий документ посредством ссылки во всей своей полноте. Другие системы включают шпильку PP7, которая специфически рекрутирует белок PCP, и шпильку «com», которая специфически рекрутирует белок Com. См. Zalatan et al.

[0271] Нуклеотидная последовательность шпильки MS2 (или эквивалентно называемой «аптамером MS2») представляет собой: GCCAACATGAGGATCACCCATGTCTGCAGGGCC (SEQ ID NO: 763).

[0272] Аминокислотная последовательность MCP или MS2cp представляет собой: GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDG NPIPSAIAANSGIY (SEQ ID NO: 764).

[0273] 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 PEгРНК или tPERT, которая эффективно рекрутирует другие эффекторные функции (например, РНК-связывающие белки, имеющие различные функции, например ДНК-полимеразы или другие ДНК-модифицирующие ферменты) для PEгРНК или tPERT, которые модифицированы соответствующим образом, и, таким образом, обеспечивает солокализацию эффекторных функций *в транс* для механизма праймированного редактирования. Это применение 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 праймирующим редактором, который содержит белок MS2cp, слитый с PE2 для обеспечения солокализации комплекса праймирующего редактора (комплекса MS2cp-PE2:crРНК), связанного с целевым участком ДНК и доменом синтеза ДНК молекулы rPERT.

napDNAbp

[0274] В контексте настоящего изобретения термин «программируемый нуклеиновой кислотой ДНК-связывающий белок» или «napDNAbp», примером которого является Cas9, обозначает белки, которые используют РНК:ДНК гибридизацию для нацеливания и связывания определенных последовательностей в молекуле ДНК. Каждый napDNAbp связан по меньшей мере с одной гидовой нуклеиновой кислотой (например,

гидовой РНК), которая локализует napDNAbp в последовательности ДНК, которая содержит цепь ДНК (то есть цепь-мишень), которая комплементарна гидовой нуклеиновой кислоте или ее части (например, протоспейсеру гидовой РНК). Другими словами, направляющая нуклеиновая кислота «программирует» napDNAbp (например, Cas9 или его эквивалент) для локализации и связывания с комплементарной последовательностью.

[0275] Без ограничения теорией, механизм связывания комплекса napDNAbp-гидовая РНК в целом включает стадию образования R-петли, при котором napDNAbp индуцирует раскручивание двухцепочечной ДНК-мишени, тем самым разделяя цепи в участке, связываемом napDNAbp. Затем протоспейсер гидовой РНК гибридизируется с «целевой цепью». Это обеспечивает вытеснение «нецелевой цепи», которая комплементарна целевой цепи, которая образует одноцепочечный участок R-петли. Согласно некоторым вариантам 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 нецелевой цепи в первом месторасположении и/или разрезание целевой цепи во втором месторасположении. В зависимости от нуклеазной активности, ДНК-мишень может быть разрезана с образованием «двухцепочечного разрыва», при котором обе цепи являются разрезанными. Согласно другим вариантам осуществления ДНК-мишень может быть разрезана только в одном участке, то есть ДНК «подвергается разрыву» на одной цепи. Пример napDNAbp с различными видами нуклеазной активности включает «никазу Cas9» («nCas9») и деактивированный Cas9 без нуклеазной активности («мертвый Cas9» или «dCas9»). В настоящем документе приведены иллюстративные последовательности для данных и других napDNAbp.

Никаза

[0276] Термин «никаза» означает Cas9 с деактивированным одним из двух нуклеазных доменов. Данный фермент способен расщеплять только одну цепь ДНК-мишени.

Последовательность локализации в ядре (NLS)

[0277] Термин «последовательность локализации в ядре» или «NLS» означает аминокислотную последовательность, которая промотирует импорт белка в ядро клетки, например, посредством ядерного транспорта. Последовательности ядерной локализации известны в данной области и очевидны специалисту в данной области техники. Например, последовательности NLS описаны Plank *et al.*, в международной PCT заявке PCT/EP2000/011690, поданной 23 ноября 2000 г., опубликованной как WO/2001/038547 31 мая 2001 г., содержание которой включено в настоящий документ посредством ссылки в

части примеров последовательностей ядерной локализации. Согласно некоторым вариантам осуществления NLS содержит аминокислотную последовательность PKKKRKV (SEQ ID NO: 16) или MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 17).

Молекула нуклеиновой кислоты

[0278] В контексте настоящего изобретения термин «нуклеиновая кислота» означает 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).

РЕгРНК

[0279] В контексте настоящего изобретения термины «гидовая РНК праймированного редактирования», или «РЕгРНК», или «удлиненная гидовая РНК» означает специализированную форму гидовой РНК, которая была модифицирована для включения одной или нескольких дополнительных последовательностей для реализации описанных в настоящем документе способов и композиций праймированного редактирования. Как описано в настоящем документе, гидовая РНК для праймированного редактирования содержит один или несколько «удлиненных участков» последовательности нуклеиновой кислоты. Удлиненные участки могут содержать без ограничения одноцепочечную РНК или ДНК. Дополнительно, удлиненные участки могут присутствовать на 3-конце традиционной гидовой РНК. Согласно другим вариантам расположения удлиненные участки могут присутствовать на 5'-конце традиционной гидовой РНК. В других вариантах расположения удлиненный участок может располагаться во внутримолекулярном участке традиционной гидовой РНК, например, в участке в виде ядра гРНК, который ассоциируется и/или связывается с napDNAbp. Удлиненный участок содержит «матрицу для синтеза ДНК», которая кодирует (посредством полимеразы праймирующего редактора) одноцепочечную ДНК, которая, в свою очередь, была

сконструирована для того, чтобы являться (a) гомологичной подвергаемой редактированию эндогенной ДНК-мишени и (b) которая содержит 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, без ограничения аптамеры, структуры типа «петля на стебле», шпильки, петли (например, 3' петлю) или домен рекрутирования РНК-белка (например, шпильку MS2). В контексте настоящего изобретения «участок связывания праймера» содержит последовательность, которая гибридизируется с одноцепочечной последовательностью ДНК, имеющей 3′-конец, образованный из ДНК с разрывом R-петли.

[0280] Согласно определенным вариантам осуществления РЕгРНК соответствуют приведенным на фиг. 3A, на котором показана РЕгРНК, имеющая 5'-плечо удлинения, спейсер и ядро гРНК. 5'-удлинение дополнительно содержит в направлении от 5' к 3' матрицу обратной транскриптазы, участок связывания праймера и линкер. Как показано, матрица обратной транскриптазы также может быть обозначена более широко как «матрица для синтеза ДНК», где полимераза описанного в настоящем документе праймирующего редактора является не RT, а другим типом полимеразы.

[0281] Согласно определенным другим вариантам осуществления РЕгРНК соответствуют приведенным на фиг. 3B, на котором показана РЕгРНК, имеющая 5'-плечо удлинения, спейсер и ядро гРНК. 5'-удлинение дополнительно содержит в направлении от 5' к 3' матрицу обратной транскриптазы, участок связывания праймера и линкер. Как показано, матрица обратной транскриптазы также может быть обозначена более широко как «матрица для синтеза ДНК», где полимераза описанного в настоящем документе праймирующего редактора является не RT, а другим типом полимеразы.

[0282] Согласно другим вариантам осуществления, РЕгРНК соответствуют приведенным на фиг. 3D, на котором показана РЕгРНК, имеющая в направлении от 5' к 3' спейсер (1), ядро гРНК (2) и плечо удлинения (3). Плечо удлинения (3) расположено на 3'-конце РЕгРНК. Плечо удлинения (3) дополнительно содержит в направлении от 5' к 3' «участок связывания праймера» (A), «матрицу редактирования» (B) и «плечо гомологии» (C). Плечо удлинения (3) также может содержать необязательный модификаторный участок на 3' и 5'-концах, который может представлять собой одинаковые или различающиеся последовательности. Кроме того, 3'-конец РЕгРНК может содержать последовательность терминатора транскрипции. Данные элементы последовательности РЕгРНК дополнительно

описаны и определены в настоящем документе.

[0283] Согласно другим вариантам осуществления РЕгРНК соответствуют приведенным на фиг. 3E, на котором показана РЕгРНК, имеющая в направлении от 5' к 3' спейсер (1), ядро гРНК (2) и плечо удлинения (3). Плечо удлинения (3) расположено на 5'-конце РЕгРНК. Плечо удлинения (3) дополнительно содержит в направлении от 5' к 3' «участок связывания праймера» (A), «матрицу редактирования» (B) и «плечо гомологии» (C). Плечо удлинения (3) также может содержать необязательный модификаторный участок на 3' и 5'-концах, который может представлять собой одинаковые или различающиеся последовательности. Кроме того, 3'-конец РЕгРНК может содержать последовательность терминатора транскрипции. Данные элементы последовательности РЕгРНК дополнительно описаны и определены в настоящем документе.

PE1

[0284] В контексте настоящего изобретения «PE1» означает комплекс PE, содержащий слитый белок, содержащий Cas9(H840A) и RT MMLV дикого типа, имеющий следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(wt)] + желаемая РЕгРНК, где слияние PE имеет аминокислотную последовательность SEQ ID NO: 123, которая показана следующим образом

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK**
KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL
VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF
LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG
EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA
AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS
KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG
ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKD
KDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN
LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD
FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV
NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK
SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA

SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK EVLDATLIHQSITGLYETRIDLSQLGGD<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u>*TLNI EDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGI KPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSH QWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADF RIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQR WLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQ EIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMV AAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPAT LLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAK ALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP**</u>SGGSKRTADGSEF<u>EPKKKRKV</u> (SEQ ID NO: 123)

ОБОЗНАЧЕНЯ:

<u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS)</u> ВВЕРХУ:(SEQ ID NO: 124), СНИЗУ: (SEQ ID NO: 133)

CAS9(H840A) (SEQ ID NO: 126)

<u>ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ</u>(SEQ ID NO: 127)

M-MLV обратная транскриптаза (SEQ ID NO: 128).

<u>PE2</u>

[0285] В контексте настоящего изобретения «PE2» относится к комплексу PE, содержащему слитый белок, содержащий Cas9(H840A) и вариант RT MMLV, имеющий следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)] + желаемая PEгРНК, где слияние PE имеет аминокислотную последовательность SEQ ID NO: 134, которая показана далее:

<u>MKRTADGSEFESPKKKRKV</u>**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKD KDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR**

**KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFL KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK EVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TLNI EDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGI KPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSH QWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADF RIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQR WLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQ EIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMV AAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPAT LLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAK ALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP*SGGSKRTADGSEF EPKKKRKV (SEQ ID NO: 134)

ОБОЗНАЧЕНИЯ:

ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS) ВВЕРХУ:(SEQ ID NO: 124), СНИЗУ: (SEQ ID NO: 133)

CAS9(H840A) (SEQ ID NO: 137)

ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ (SEQ ID NO: 127)

M-MLV обратная транскриптаза (SEQ ID NO: 139).

PE3

[0286] В контексте настоящего изобретения «PE3» относится к PE2 плюс гидовая РНК разрыва второй цепи, которая образует комплекс с PE2 и вводит разрыв в нередактируемую цепь ДНК для индуцирования предпочтительно замещения подвергнутой редактированию цепи.

PE3b

[0287] В контексте настоящего изобретения «PE3b» относится к PE3, но где гидовая РНК разрыва второй цепи сконструирована для временного контроля, так что разрыв

второй цепи не осуществляется до реализации внесения требуемого изменения. Это достигается посредством конструирования гРНК со спейсерной последовательностью, которая соответствует только редактируемой цепи, но не исходному аллелю. Используя эту стратегию, упоминаемую далее как PE3b, несовпадения между протоспейсером и неподвергаемым редактированию аллелем должны способствовать отсутствию разрыва посредством сгРНК до реализации редактирования на включающей PAM цепи.

PE-короткая

[0288] В контексте настоящего изобретения «PE-короткая» относится к конструкции PE, которая слита с усеченной на C-конце обратной транскриптазой, и имеет следующую аминокислотную последовательность:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH FLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH LGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIV PQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG RKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD**<u>SGGSSGGSSGSETPGTSESATPESSGGSS GGSS</u>***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPY NLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNE ALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYL*

*GYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVA AGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRV QFGPVVALNPATLLPLPEEGLQHNCLDNSRLIN*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 765)

ОБОЗНАЧЕНИЕ:

ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS) СВЕРХУ:(SEQ ID NO: 124), СНИЗУ: (SEQ ID NO: 133)

CAS9(H840A) (SEQ ID NO: 157)

ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ 1 (SEQ ID NO: 127)

M-MLV TRUNCATED ОБРАТНАЯ ТРАНСКРИПТАЗА

(SEQ ID NO: 766)

Пептидная метка

[0289] Термин «пептидная метка» означает 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 т.д. 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

[0290] В контексте настоящего изобретения термин «полимераза» означает фермент, который синтезирует нуклеотидную цепь и который может применяться совместно с описанными в настоящем документе системами праймирующего редактора. 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». Согласно различным вариантам осуществления система праймирующего редактора содержит ДНК-полимеразу. Согласно различным вариантам 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 РЕгРНК, где плечо удлинения включает цепь ДНК. В таких случаях РЕгРНК может быть обозначена как химерная или гибридная РЕгРНК, которая содержит часть РНК (т.е. компоненты гидовой РНК, включая спейсер и ядро гРНК) и часть ДНК (т.е. плечо удлинения). Согласно различным другим вариантам осуществления ДНК-полимераза может представлять собой «РНК-зависимую ДНК-полимеразу» (где молекула матрицы представляет собой цепь РНК). В таких случаях РЕгРНК представляет собой РНК, то есть включает удлинение РНК. Термин «полимераза» также может обозначать фермент, который катализирует полимеразу нуклеотида (т.е. 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 РЕгРНК) по направлению к 5'-концу цепи матрицы. «ДНК-полимераза» катализирует полимеризацию дезоксинуклеотидов. В контексте настоящего изобретения в отношении ДНК-полимеразы термин «ДНК-полимераза» включает «ее функциональный фрагмент». Термин «ее функциональный фрагмент» означает 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.

Праймированное редактирование и мультилоскутное праймированное редактирование

[0291] В контексте настоящего изобретения термин «праймированное редактирование» или «классическое праймированное редактирование» означает подход для редактирования генов с использованием napDNAbp, полимеразы (например, обратной транскриптазы) и специализированных гидовых 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 праймированного редактирования описаны в вариантах осуществления на фиг. 1A-1H и фиг.72(a)-72(c), среди других фигур. Классическое праймированное редактирование описано в публикации авторов настоящего изобретения Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019), которая включена в настоящий документ

посредством ссылки во всей своей полноте.

[0292] Праймированное редактирование представляет собой платформу для редактирования генома, 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 представленного *в транс* с napDNAbp), где система праймированного редактирования программируется гидовой («PEгРНК») праймированного редактирования (PE), которая как определяет целевой участок, так и служит матрицей для синтеза желаемого редактирования в форме замещающей цепи ДНК в виде удлинения (либо ДНК, либо РНК), сконструированного на гидовой РНК (например, на 5' или 3'-конце, или во внутренней части гидовой РНК). Замещающая цепь, содержащая желаемое редактирование (например, замещение одного нуклеотидного основания), имеет (или является гомологичной) той же последовательности, что и эндогенная цепь (непосредственно по ходу транскрипции от участка разрыва) подвергаемого редактированию целевого участка (за исключением того, что она включает желаемое редактирование). С помощью механизма репарации и/или репликации ДНК эндогенная цепь по ходу транскрипции от участка разрыва замещается новой синтезированной замещающей цепью, содержащей желаемое редактирование. В некоторых случаях праймированное редактирование может быть описано как технология геномного редактирования типа «поиск-и-замена», поскольку праймирующие редакторы, как описано в настоящем документе, не только обеспечивают поиск и нахождение требуемого целевого участка, подлежащего редактированию, но в то же время кодируют замещающую цепь, содержащую желаемое редактирование, которая встраивается вместо соответствующей эндогенной цепи ДНК целевого участка. Праймирующие редакторы согласно настоящему изобретению частично связаны с открытием того, что механизм нацеленной-праймированной обратной транскрипции (TPRT) или «праймированного редактирования» может быть использован или адаптирован для осуществления точного геномного редактирования на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью (например, как показано в различных вариантах осуществления на фиг.1A-1E). TPRT естественным образом используется мобильными элементами ДНК, например, ретротранспозонами млекопитающих без длинных концевых повторов (LTR) и бактериальными интронами Группы II[28,29]. Согласно настоящему изобретению авторы настоящего изобретения использовали продукты слияния белок Cas-обратная транскриптаза или родственные системы для нацеливания на определенную

последовательность ДНК с помощью гидовой РНК, осуществления одноцепочечного разрыва в целевом участке, и применения ДНК с разрывом в качестве праймера для обратной транскрипции сконструированной матрицы обратной транскриптазы, которая встраивается с гидовой РНК. Однако в то время как данная концепция начинается с праймирующих редакторов, которые используют обратную транскриптазу в качестве компонента ДНК-полимеразы, описанные в настоящем документе праймирующие редакторы не ограничиваются обратными транскриптазами, но могут включать использование практически любой ДНК-полимеразы. Действительно, в то время как в настоящем документе могут указываться праймирующие редакторы с «обратными транскриптазами», в настоящем документе устанавливается, что обратные транскриптазы являются лишь одним из типов ДНК-полимераз, которые могут работать в рамках праймированного 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, согласно одному аспекту праймирующие редакторы могут включать Cas9 (или эквивалентный napDNAbp), который программируется для нацеливания на ДНК-последовательность путем его связывания со специализированной гидовой РНК (т.е. РЕгРНК), содержащей спейсерную последовательность, которая отжигается на комплементе протоспейсера в ДНК-мишени. Специализированная гидовая РНК также содержит новую генетическую информацию в форме удлинения, которое кодирует замещающую цепь ДНК, содержащую желаемое генетическое изменение, которая используется для замены соответствующей эндогенной цепи ДНК в целевом участке. Для переноса информации из РЕгРНК в ДНК-мишень механизм праймированного редактирования включает разрыв целевого участка в одной цепи ДНК для высвобождения 3 '-гидроксильной группы. Затем свободная 3 '-гидроксильная группа может использоваться для праймирования полимеризации ДНК кодирующего изменение удлинения на РЕгРНК напрямую в целевой участок. Согласно различным вариантам осуществления удлинение, которое обеспечивает матрицу для полимеризации замещающей содержащей редактирование цепи, может быть образовано из РНК или ДНК. В случае РНК удлинения полимераза праймирующего редактора может представлять собой РНК-зависимую ДНК-полимеразу (например, обратную транскриптазу). В случае ДНК удлинения полимераза праймирующего редактора может представлять собой ДНК-зависимую ДНК-полимеразу. Новая синтезированная цепь (т.е. замещающая цепь ДНК, содержащая требуемое редактирование), которая образовывается описанными в настоящем документе праймирующими редакторами, будет являться гомологичной относительно целевой

геномной последовательности (т.е. 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 лоскут ДНК, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. Согласно определенным вариантам осуществления система может комбинироваться с использованием фермента обратной транскриптазы сниженной точности (например, представленной в виде слитого с Cas9 доменом белка или представленной с Cas9 доменом *в транс)*. Фермент обратной транскриптазы сниженной точности может вносить изменения при синтезе одноцепочечного лоскута ДНК. Таким образом, согласно определенным вариантам осуществления обратная транскриптаза сниженной точности может использоваться для внесения нуклеотидных изменений в ДНК-мишень. В зависимости от обратная транскриптазы сниженной точности, которая используется в системе, изменения могут являться случайными или неслучайными. Разрешение гибридизированного промежуточного продукта (содержащего одноцепочечный лоскут ДНК, синтезированный обратной транскриптазой, гибридизированной с эндогенной цепью ДНК) может включать удаление образованного вытесненного лоскута эндогенной ДНК (например, с помощью эндонуклеазы 5'-концевого лоскута ДНК, FEN1), лигирование синтезированного одноцепочечного лоскута ДНК 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 применений в фундаментальной науке и терапии.

[0293] Согласно различным вариантам осуществления праймированное 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 гидовой РНК для праймированного редактирования (PEгРНК). Со ссылкой на Фиг. 1G, гидовая РНК для праймированного редактирования (PEгРНК) содержит удлинение на 3'- или 5'-конце гидовой РНК или во внутримолекулярном положении в гидовой РНК и кодирует требуемое нуклеотидное изменение (например, изменение одного нуклеотида, вставку или делецию). На стадии (a) комплекс в виде napDNAbp/удлиненной гРНК контактирует с молекулой ДНК, при этом удлиненная гРНК направляет napDNAbp

для связывания с локусом-мишенью. На стадии (b) осуществляется разрыв в одной из цепей локуса-мишени ДНК (например, с помощью нуклеазы или химического агента), тем самым обеспечивая создание свободного 3 '-конца в одной из цепей локуса-мишени. Согласно определенным вариантам осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, т.е. цепи, которая не гибридизируется с последовательностью гидовой РНК, т.е. «нецелевой мишенью цепи». Однако разрыв должен быть осуществлен в любой одной из цепей. То есть разрыв должен быть введен в «целевую цепь» R-петли (т.е. цепь, гибридизированную с протоспейсером удлиненной гРНК) или «нецелевую цепь» (т.е. цепь, образующую одноцепочечную часть R- петли, которая является комплементарной целевой цепи). На стадии (c) 3'-конец цепи ДНК (образованный разрывом) взаимодействует с удлиненной частью гидовой РНК для праймирования обратной транскрипции (т.е. «нацеленной-праймированной мишенью RT»). Согласно определенным вариантам осуществления 3-конец цепи ДНК гибридизируется с определенной праймирующей последовательностью RT на удлиненной части гидовой РНК, т.е. «праймирующей последовательностью обратной транскриптазы» или «участком связывания праймера» на РЕгРНК. На стадии (d) вводится обратная транскриптаза (или другая подходящая ДНК-полимераза), которая синтезирует одиночную цепь ДНК от 3′-конца праймированного участка к 5′-концу гидовой РНК для праймированного редактирования. ДНК-полимераза (например, обратная транскриптаза) может быть слита с napDNAbp или, в качестве альтернативы, быть предоставлена в транс для napDNAbp. Это образует одноцепочечный лоскут ДНК, содержащий требуемое нуклеотидное изменение (например, изменение одного основания, вставку, или делецию, или их комбинацию) и являющийся иным образом гомологичным эндогенной ДНК в участке разрыва или рядом с ним. На стадии (e) napDNAbp и гидовая РНК высвобождаются. Стадии (f) и (g) связаны с разрешением одноцепочечного лоскута ДНК, 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 лоскута ДНК, который образуется, когда 3' одноцепочечный лоскут ДНК вносится и гибридизируется с последовательностью эндогенной ДНК. Без ограничения теорией, процессы репарации и репликации клеточной эндогенной ДНК обеспечивают разрешение несовпадающей ДНК для встраивания нуклеотидного изменения (изменений) с образованием желаемого измененного продукта. Процесс также может быть смещен в сторону образования продукта с «разрывом 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.

[0294] Термин «система праймирующего редактора (PE)», или «праймирующий редактор (PE)», или «система PE», или «PE система редактирования» означает композиции, используемые в способе геномного редактирования с применением описанной в настоящем документе нацеленной-праймированной обратной транскрипции (TPRT), включая без ограничения napDNAbp, обратные транскриптазы, слитые белки (например, содержащие napDNAbp и обратные транскриптазы), гидовые РНК для праймированного редактирования, и комплексы, включающие слитые белки и гидовые РНК праймированного редактирования, а также вспомогательные элементы, например, компоненты для разрыва второй цепи (например, сгРНК второй цепи) и эндонуклеазы для удаления 5'-эндогенного лоскута ДНК (например, FEN1) для обеспечения сдвига процесса праймированного редактирования в сторону образования подвергнутого редактированию продукта.

[0295] Несмотря на то, что в описанных вариантах осуществления PEгРНК представляет собой одиночную молекулу, содержащую гидовую РНК (которая сама содержит спейсерную последовательность и ядро или каркас гРНК) и 5' или 3 ' плечо удлинения, содержащее участок связывания праймера и матрицу для синтеза ДНК (например см. фиг. 3D), PEгРНК может также принимать форму двух отдельных молекул, состоящих из гидовой РНК и матрицы РНК *транс*-праймирующего 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 с модифицированным комплексом праймированного редактирования, который содержит 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 праймированного редактирования.

[0296] В «системе двухлоскутного праймированного редактирования» две PEгРНК используются для нацеливания на противоположные цепи геномного сайта и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 91**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная цепь. Поскольку обе цепи дуплекса синтезируются как отредактированная ДНК, система двухлоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК

нужно только вырезать парные 5'-лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные цепи ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое содержит минимальную геномную гомологию. Нуклеазоактивные варианты праймирующих редакторов, которые разрезают обе цепи ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК.

[0297] Как и классическое праймированное редактирование, мультилоскутное праймированное редактирование (включая двухлоскутное и четырехлоскутное праймированное редактирование) представляет собой универсальный и точный способ редактирования генома, который напрямую записывает новую генетическую информацию в указанный участок ДНК с использованием программируемого нуклеиновой кислотой ДНК-связывающего белка («napDNAbp»), работающего в сочетании с полимеразой (т.е. в форме слитого белка или иным образом представленной *в транс* с napDNAbp), где система праймированного редактирования запрограммирована с помощью гидовой РНК праймированного редактирования (PE) («PEгРНК»), которая одновременно указывает целевой участок и является матрицей для синтеза желаемого редактирования в форме замещающей цепи ДНК посредством удлинения (либо ДНК, либо РНК), сконструированного на гидовой РНК (например, на 5'- или 3'-конце, или во внутренней части гидовой РНК). Цепь замещения, содержащая желаемое редактирование (например, замену одного нуклеотидного основания), имеет ту же последовательность, что и эндогенная цепь целевого участка, подлежащего редактированию (за исключением того, что она включает желаемое редактирование). С помощью механизмов репарации и/или репликации ДНК эндогенная цепь целевого участка заменяется вновь синтезированной замещающей цепью, содержащей желаемое редактирование. В некоторых случаях праймированное редактирование можно рассматривать как технологию редактирования генома «поиск и замена», поскольку праймирующие редакторы, как описано в настоящем документе, не только ищут и находят желаемый целевой сайт для редактирования, но в то же время, кодируют замещающую цепь, содержащую желаемое редактирование, которое устанавливается вместо соответствующей цепи эндогенной ДНК целевого участка.

Праймирующий редактор

[0298] Двойная система праймированного редактирования, описанная в настоящем документе, содержит пару праймирующих редакторов. Термин «праймирующий редактор» относится к описываемым в настоящем документе слитым конструкциям, содержащим napDNAbp (например, никазу Cas9) и обратную транскриптазу и способным осуществлять

праймированное редактирование целевой нуклеотидной последовательности в присутствии РЕгРНК (или «удлиненной гидовой РНК»). Термин «праймирующий редактор» может относиться к слитому белку или к слитому белку, образующему комплекс с РЕгРНК и/или дополнительно образующему комплекс с сгРНК с разрывом второй цепи. Согласно некоторым вариантам осуществления праймирующий редактор может также относиться к комплексу, содержащему слитый белок (обратную транскриптазу, слитую с napDNAbp), РЕгРНК и обычную гидовую РНК, способную управлять стадией разрыва второго участка неотредактированной цепи, как описано в настоящем документе. Согласно другим вариантам осуществления компонент обратной транскриптазы «праймирующего редактора» может быть представлен *в транс*.

[0299] Система двухлоскутного праймированного редактирования, описанная в настоящем документе, содержит пару праймирующих редакторов. Система четырехлоскутного праймированного редактирования, описанная в настоящем документе, содержит четыре праймирующих редактора.

Участок связывания праймера

[0300] Термин «участок связывания праймера» или «PBS» относится к нуклеотидной последовательности, расположенной на РЕгРНК в качестве компонента плеча удлинения (обычно на 3'-конце плеча удлинения) и служит для связывания с праймирующей последовательностью, которая образуется после того, как Cas9 разрывает целевую последовательность праймирующим редактором. Как подробно описано в другом месте, когда никазаный компонент Cas9 праймирующего редактора надрезает одну цепь последовательности ДНК-мишени, формируется 3'-концевой лоскут одноцепочечной ДНК, который служит праймирующей последовательностью, которая отжигается с участком связывания праймера на РЕгРНК для праймирования обратной транскрипции.

Промотор

[0301] 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 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

[0302] В контексте настоящего изобретения термин «протоспейсер» означает последовательность (~20 п.о.) в ДНК, являющуюся смежной относительно последовательности PAM (смежного с протоспейсером мотива). Протоспейсер имеет ту же последовательность, что и спейсерная последовательность гидовой РНК. Гидовая РНК отжигается на комплементе протоспейсерной последовательности на ДНК-мишени (в частности, одной ее цепи, то есть «целевой цепи» против «нецелевой цепи» целевой последовательности 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 по ходу транскрипции от последовательности-мишени в геномной ДНК на нецелевой цепи. Специалисту в данной области техники известно, что «протоспейсер» в литературе в данной области иногда указывается в виде мишень-специфической гидовой последовательности длиной ~20 нуклеотидов на самой гидовой РНК вместо указания его как «спейсера». Таким образом, в некоторых случаях термин «протоспейсер» в контексте настоящего изобретения может использоваться взаимозаменяемым образом с термином «спейсер». Контекст описания, связанный с указанием «протоспейсера» или «спейсера», поможет читателю определить, указывает ли данный термин на гРНК или на ДНК-мишень.

Смежный с протоспейсером мотив (PAM)

[0303] В контексте настоящего изобретения термин «смежная с протоспейсером последовательность» или «PAM» означает 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 нуклеотидное основание с последующими двумя

нуклеотидными 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.

[0304] Например, с учетом того, что аминокислотная последовательность стандартной SpCas9 представляет собой SEQ ГО 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.

[0305] Также будет очевидно, что ферменты Cas9 из бактерий различных видов (то есть ортологи 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 килоосновaние 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 (которая включена в настоящий документ посредством ссылки).

Рекомбиназа

[0306] Термин «рекомбиназа» в контексте настоящего изобретения означает сайт-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, содержание каждой из которых включено в настоящий документ посредством ссылки во всей своей полноте. Указанные в настоящем документе рекомбиназы не являются исключающими примерами рекомбиназ, которые могут применяться в рамках вариантов осуществления изобретения. Способы и композиции согласно настоящему 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, которая включена в настоящий документ посредством ссылки во всей своей полноте). Другие примеры рекомбиназ, которые могут применяться в рамках описанных в настоящем документе 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. Согласно некоторым вариантам 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 связывания определенной ДНК). Рекомбиназы без активности в виде связывания ДНК и способы их конструирования известны и включают описанные в 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, and Proudfoot *et al.*, «Zinc finger recombinases with adaptable DNA sequence specificity.» *PLoS One.* 2011,6(4):e19537, полное содержание каждой из которых включено в настоящий документ посредством 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 которой включено в настоящий документ посредством ссылки). Каталитические домены данных рекомбиназ, таким образом, являются поддающимися рекомбинации с использованием РНК-программируемых нуклеаз с деактивированной нуклеазной активностью (например, dCas9 или ее фрагмента), как описано в настоящем документе, 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 сайт-специфических рекомбиназ, как описано в настоящем документе (см., например, Smith *et al.*, «Diversity in the serine recombinases.» *Mol Microbiol.* 2002,44: 299–307, полное содержание которой включено в настоящий документ посредством ссылки). Подобным образом, известны каталитические домены ядра тирозиновых рекомбиназ (например, Cre, интегразы X), при этом они могут быть также кооптированы для конструирования программируемых сайт-специфических рекомбиназ, как описано в настоящем документе (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, and 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, полные содержания которых включены в настоящий документ посредством ссылки).

Последовательность распознавания рекомбиназой

[0307] Термин «последовательность распознавания рекомбиназой», или

эквивалентно «RRS», или «последовательность-мишень рекомбиназы» в контексте настоящего изобретения означает 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. Согласно различным вариантам осуществления праймирующие редакторы нескольких цепей могут устанавливать один или несколько сайтов рекомбиназы в целевой последовательности или в более чем одной целевой последовательности. Когда с помощью праймирующего редактора нескольких цепей устанавливается более одного сайта рекомбиназы, сайты рекомбиназы могут быть установлены в смежных сайтах-мишенях или несмежных сайтах-мишенях (*например,* в отдельных хромосомах). Согласно различным вариантам осуществления одиночные установленные рекомбиназные сайты можно использовать в качестве «сайтов посадки» для опосредованной рекомбиназой реакции между геномным рекомбиназным сайтом и вторым рекомбиназным сайтом внутри экзогенно поставляемой молекулы нуклеиновой кислоты, *например*, плазмиды. Это обеспечивает целенаправленную интеграцию желаемой молекулы нуклеиновой кислоты. Согласно другим вариантам осуществления, где два сайта рекомбиназы встроены в соседние участки ДНК (*например*, разделенные посредством 25-50 п.о., 50-100 п.о., 100-200 п.о., 200-300 п.о., 300-400 п.о., 400-500 п.о., 500-600 п.о., 600-700 п.о., 700-800 п.о., 800-900 п.о., 900-1000 п.о., 1000-2000 п.о., 2000-3000 п.о., 3000-4000 п.о., 4000-5000 п.о., или более), сайты рекомбиназы можно использовать для опосредованного рекомбиназой вырезания или инверсии промежуточной последовательности или для опосредованного рекомбиназой обмена кассет с экзогенной ДНК, имеющей те же сайты рекомбиназы. Когда два или более сайта рекомбиназы устанавливаются с помощью мультилоскутных праймирующих редакторов на двух разных хромосомах, может происходить транслокация промежуточной последовательности из первого хромосомного местоположения во второе.

Рекомбинировать и рекомбинация

[0308] 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 посредством действия белка рекомбиназы (например, указанного в настоящем документе слитого белка рекомбиназы согласно настоящему изобретению). Рекомбинация может приводить, *среди прочего,* к вставке, инверсии, эксцизии или транслокации нуклеиновых кислот, например, в или между одной или несколькими молекулами нуклеиновой кислоты.

Обратная транскриптаза

[0309] Термин «обратная транскриптаза» означает 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. Настоящее изобретение охватывает применение любых подобных обратных транскриптаз или их вариантов или мутантов.

[0310] Кроме того, настоящее изобретение охватывает применение обратных транскриптаз, которые подвержены ошибкам, то есть которые могут быть обозначены как обратные транскриптазы сниженной точности или обратные транскриптазы, которые не поддерживают встраивание нуклеотидов с высокой точностью при полимеризации. При синтезе одноцепочечного ДНК-лоскута 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 посредством ошибочной полимеризации одноцепочечного лоскута ДНК. Данные ошибки, внесенные при синтезе одноцепочечного лоскута ДНК, затем становятся встроенными в двухцепочечную

молекулу посредством гибридизации с соответствующей эндогенной целевой цепью, удаления эндогенной вытесненной цепи, лигирования, и затем посредством еще одного раунда процессов репарации и/или секвенирования эндогенной ДНК.

Обратная транскрипция

[0311] В контексте настоящего изобретения термин «обратная транскрипция» означает способность фермента синтезировать цепь ДНК (то есть комплементарную ДНК или кДНК) с помощью РНК в качестве матрицы. Согласно некоторым вариантам осуществления обратная транскрипция может представлять собой «обратную транскрипцию сниженной точности», что обозначает свойства определенных ферментов в виде обратной транскриптазы, которые являются подверженными ошибкам в плане их активности, связанной с полимеризацией ДНК.

PACE

[0312] Термин «непрерывная эволюция с помощью фагов (PACE)» в контексте настоящего изобретения означает 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, опубликованной как WO 2010/028347, поданной 11 марта 2010, международной PCT заявке, PCT/US2011/066747, поданной 22 декабря 2011, опубликованной как WO 2012/088381, поданной 28 июня 2012, заявке на патент США, патенте США № 9023594, поданном 5 мая 2015, международной PCT заявке, PCT/US2015/012022, поданной 20 января 2015, опубликованной как WO 2015/134121, поданной 11 сентября 2015, и международной PCT заявке, PCT/US2016/027795, поданной 15 апреля 2016, опубликованной как WO 2016/168631, поданной 20 октября 2016, полные содержания каждой из которых включены в настоящий документ посредством ссылки.

Фаг

[0313] В контексте настоящего изобретения термин «фаг» используется взаимозаменяемым образом с термином «бактериофаг», означает 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 круговой форме. Фаги и фаговые векторы хорошо известны специалистам в данной области, при этом неограничивающие примеры фагов, которые могут применяться для реализации способов PACE, приведены в настоящем документе и представляют собой λ (лизоген), T2, T4, T7, T12, R17, M13, MS2, G4, P1, P2, P4, Phi X174, N4, Φ6 и Φ29. Согласно определенным вариантам 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 включены в настоящий документ посредством ссылки в объеме описания подходящих фагов и клеток-хозяев, а также способов и протоколов выделения, культивирования и манипулирования такими фагами).

Белок, пептид и полипептид

[0314] Термин «белок», «пептид» и «полипептид» используются в настоящем документе 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 в настоящем документе белков может быть получен с помощью любого способа, известного в данной области. Например, указанные в настоящем документе 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 которой включено в настоящий документ посредством ссылки.

Сплайсинг белков

[0315] Термин «сплайсинг белков в контексте настоящего изобретения означает процесс, при котором последовательность, интеин (или расщепленные 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» означает определенный случай, при котором интеины представляют собой расщепленные интеины, при этом они локализованы на различных белках.

Разрыв второй цепи

[0316] Коррекция гетеродуплексной ДНК (т.е. содержащей одну подвергнутую редактированию и одну не подвергнутую редактированию цепь), образованной в результате праймированного редактирования, определяет долгосрочные результаты редактирования. Простыми словами, целью праймированного редактирования является коррекция гетеродуплексной ДНК (подвергнутой редактированию цепи, спаренной с эндогенной не подвергнутой редактированию цепью), образованной в качестве промежуточного продукта PE посредством перманентного встраивания подвергнутой редактированию цепи в комплементарную эндогенную цепь. Подход в виде «разрыва второй цепи» может применяться в настоящем документе для смещения коррекции гетеродуплексной ДНК в сторону перманентной интеграции подвергнутой редактированию цепи в молекулу ДНК. В контексте настоящего изобретения концепция «разрыв второй цепи» означает внесение второго разрыва в участке по ходу транскрипции от первого разрыва (т.е. первоначального участка разрыва, который обеспечивает свободный 3-конец для использования при праймировании обратной транскриптазы на удлиненной части гидовой РНК), предпочтительно на неотредактированной цепи. Согласно определенным вариантам осуществления первый разрыв и второй разрыв находятся на противоположных цепях. Согласно другим вариантам осуществления первый разрыв и второй разрыв находятся на противоположных цепях. Согласно другому варианту осуществления первый разрыв располагается на не подвергаемой редактированию цепи (то есть цепи, которая образует одноцепочечную часть R-петли), а второй разрыв располагается на целевой цепи. Согласно другим вариантам осуществления первый разрыв располагается на подвергаемой редактированию цепи, а второй разрыв располагается на не подвергаемой редактированию

цепи. Второй разрыв может быть расположен на расстоянии по меньшей мере 5 нуклеотидов или по меньшей мере 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 по ходу транскрипции от первого разрыва. Второй разрыв согласно определенным вариантам осуществления может быть введен между приблизительно 5-150 нуклеотидами на не подвергаемой редактированию цепи от участка разрыва, индуцируемого РЕгРНК, или между приблизительно 5-140, или между приблизительно 5-130, или между приблизительно 5-120, или между приблизительно 5-110, или между приблизительно 5-100, или между приблизительно 5-90, или между приблизительно 5-80, или между приблизительно 5-70, или между приблизительно 5-60, или между приблизительно 5-50, или между приблизительно 5-40, или между приблизительно 5-30, или между приблизительно 5-20, или между приблизительно 5-10. Согласно одному варианту осуществления второй разрыв может быть введен между приблизительно 14-116 нуклеотидов от разрыва, индуцируемого РЕгРНК. Без ограничения 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 РЕ. Согласно некоторым вариантам осуществления подвергаемая редактированию цепь представляет собой нецелевую цепь, а не подвергаемая редактированию цепь представляет собой целевую цепь. Согласно другим вариантам осуществления подвергаемая редактированию цепь представляет собой целевую цепь, а не подвергаемая редактированию цепь представляет собой нецелевую цепь.

Смысловая цепь

[0317] В генетике «смысловая» цепь представляет собой сегмент в двухцепочечной 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 смысловая и антисмысловая зависят от перспективы). В конечном итоге то, какая цепь одного сегмента дцДНК считается смысловой или

антисмысловой, определяется генным продуктом или иРНК.

[0318] В контексте РЕгРНК первая стадия представляет собой синтез одноцепочечной комплементарной ДНК (то есть 3'-лоскута одноцепочечной ДНК, который становится встроенным), ориентированной в направлении от 5 ' к 3 ', которая использует в качестве плеча удлинения РЕгРНК без матрицы. Следует ли считать 3'-лоскут одноцепочечной ДНК 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, согласно некоторым вариантам осуществления 3'-лоскут одноцепочечной ДНК (который в целом проходит в направлении от 5' к 3') будет служить в качестве смысловой цепи, поскольку он представляет собой кодирующую цепь. Согласно другим вариантам осуществления 3'-лоскут одноцепочечной ДНК (который в целом проходит в направлении от 5' к 3') будет служить в качестве антисмысловой цепи и, таким образом, матрицы для транскрипции.

Гомология последовательности

[0319] «Гомологическая последовательность» или последовательность, проявляющая «гомологию» к другой последовательности, означает последовательность молекулы нуклеиновой кислоты, проявляющую по меньшей мере приблизительно 65%, 70%, 75%, 80%, 85% или 90% идентичность последовательности с другой молекулой нуклеиновой кислоты. Согласно другим вариантам осуществления «гомологическая последовательность» нуклеиновых кислот может проявлять 93%, 95% или 98% идентичность последовательности с эталонной нуклеиновой кислотой.

[0320] Когда указан процент гомологии или идентичности последовательностей в контексте двух последовательностей нуклеиновых кислот или двух полипептидных последовательностей, процент гомологии или идентичности обычно относится к выравниванию двух или более последовательностей по части их длины при сравнении и выровнены для максимального соответствия. Если не указано иное, гомологию или идентичность последовательностей оценивают по указанной длине нуклеиновой кислоты, полипептида или их части. Согласно некоторым вариантам осуществления гомологию или идентичность оценивают по функциональной части или определенной части длины.

[0321] Выравнивание последовательностей для оценки гомологии последовательностей можно проводить с помощью алгоритмов, известных в данной области техники, таких как алгоритм Basic Local Alignment Search Tool (BLAST), который описан в Altschul et al, J. Mol. Biol. 215:403- 410, 1990. Доступ к общедоступному интернет-интерфейсу для проведения анализов BLAST можно получить через Национальный центр биотехнологической информации. Дополнительные известные алгоритмы включают

алгоритмы, опубликованные в: Smith & Waterman, «Comparison of Biosequences», Adv. Appl. Math. 2:482, 1981, Needleman & Wunsch, «A general method applicable to the search for similarities in the amino acid sequence of two proteins» J. Mol. Biol. 48:443, 1970, Pearson & Lipman «Improved tools for biological sequence comparison», Proc. Natl. Acad. Sci. USA 85:2444, 1988, или путем автоматической реализации этих или подобных алгоритмов. Программы глобального выравнивания также могут использоваться для выравнивания похожих последовательностей приблизительно одинакового размера. Примеры программ глобального выравнивания включают NEEDLE (доступна на по ссылочному адресу www.ebi.ac.uk/Tools/psa/emboss_needle/), которая является частью программного обеспечения EMBOSS (Rice P et al., Trends Genet., 2000, 16: 276-277), программы GGSEARCH fasta.bioch.Virginia.edu/fasta_www2/fasta_www.cgi?rm=compare&pgm=gnw), которая является частью программного обеспечения FASTA (Pearson W and Lipman D, 1988, Proc. Natl. Acad. Sci. USA, 85: 2444-2448). Обе эти программы основаны на алгоритме Нидлмана-Вунша, который используется для нахождения оптимального выравнивания (включая промежутки) двух последовательностей по всей их длине. Подробное обсуждение анализа последовательности также можно найти в Unit 19.3 of Ausubel et al («Current Protocols in Molecular Biology» John Wiley & Sons Inc, 1994-1998, Chapter 15, 1998).

Спейсерная последовательность

[0322] В контексте настоящего изобретения термин «спейсерная последовательность» в отношении гидовой РНК или РЕгРНК относится к части гидовой РНК или РЕгРНК из приблизительно 20 нуклеотидов, которая содержит нуклеотидную последовательность, имеющую ту же последовательность, что и протоспейсерная последовательность в целевой последовательности ДНК. Последовательность спейсера отжигается с комплементом последовательности протоспейсера с образованием гибридной структуры оцРНК/оцДНК в целевом участке и соответствующей структуры оцДНК с R-петлей эндогенной нити ДНК.

Субъект

[0323] Термин «субъект» в контексте настоящего изобретения означает отдельный организм, например, отдельное млекопитающее. Согласно некоторым вариантам осуществления субъект представляет собой человека. Согласно некоторым вариантам осуществления субъект представляет собой отличное от человека млекопитающее. Согласно некоторым вариантам осуществления субъект представляет собой отличного от человека примата. Согласно некоторым вариантам осуществления субъект представляет собой грызуна. Согласно некоторым вариантам осуществления субъект представляет собой овцу, козу, крупный рогатый скот, кошку или собаку. Согласно некоторым вариантам

осуществления субъект представляет собой позвоночное, амфибию, рептилию, рыбу, насекомое, муху, или нематоду. Согласно некоторым вариантам осуществления субъект представляет собой лабораторное животное. Согласно некоторым вариантам осуществления субъект является генетически модифицированным, например, генетически модифицированным отличным от человека субъектом. Субъект может быть любого пола и на любой стадии развития.

Расщепленный интеин

[0324] Несмотря на то, что интеины наиболее часто представлены в виде непрерывного домена, некоторые из них существуют в естественной расщепленной форме. В этом случае два фрагмента экспрессируются в виде отдельных полипептидов и должны быть связаны перед осуществлением сплайсинга, так называемого транссплайсинга белка.

[0325] Примером расщепленного 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 расщепленный 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.

[0326] Дополнительные встречающиеся в природе или сконструированные последовательности расщепленного интеина известны или могут быть реализованы на основе описанных в настоящем документе или доступных в данной области последовательностей цельного интеина. Примеры последовательностей расщепленного 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, который включен в настоящий документ посредством ссылки. Дополнительные последовательности расщепленного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774, и EP2877490, содержание каждой из которых включено в настоящий документ посредством ссылки.

[0327] Кроме того, транс-сплайсинг белка 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 дает возможность экспрессировать белок в виде двух неактивных фрагментов, которые затем подвергают лигированию с образованием функционального продукта, например, как показано на фиг. 66 и 67 в отношении образования полного слитого белка PE из двух отдельно экспрессируемых половин.

Целевой участок

[0328] Термин «целевой участок» относится к последовательности в молекуле нуклеиновой кислоты, которая редактируется описанным в настоящем документе праймирующим редактором (PE). Целевой участок далее относится к последовательности в молекуле нуклеиновой кислоты, с которой связывается комплекс праймирующего редактора (PE) и гРНК.

tPERT

[0329] См. определение «РНК матрица *транс*-праймирующего редактора (tPERT)».

Временной разрыв второй цепи

[0330] В контексте настоящего изобретения термин «временной разрыв второй цепи» означает вариант разрыва 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 двухцепочечным разрывам ДНК. Разрывающая вторую цепь гидовая РНК сконструирована для временного контроля, так что разрыв второй цепи не осуществляется до реализации внесения требуемого изменения. Это достигается посредством конструирования гРНК со спейсерной последовательностью, которая соответствует только подвергаемой редактированию цепи, но не исходному аллелю. Посредством использования данной стратегии несовпадения между протоспейсером и не подвергаемым редактированию аллелем должны способствовать отсутствию разрыва посредством сгРНК до реализации редактирования на включающей PAM цепи.

***Транс*-праймированное редактирование**

[0331] В контексте настоящего изобретения термин «*транс*-праймированное редактирование» или, подобным образом, «двойное транс-праймированное редактирование» относится к модифицированной форме праймированного редактирования или двойному праймированному редактированию, в котором используется расщепленная РЕгРНК, т.е. когда РЕгРНК разделена на две отдельные молекулы: сгРНК и РНК-матрица *транс*-праймированного редактирования (tPERT). сгРНК служит для нацеливания праймирующего редактора (или, в более общем смысле, для нацеливания компонента napDNAbp праймирующего редактора) на желаемый геномный целевой участок, в то время

как tPERT используется полимеразой (например, обратной транскриптазой) для записи новой последовательности ДНК в локус-мишень, как только tPERT рекрутируется *в транс* к редактирующему редактору посредством взаимодействия связывающих доменов, расположенных на праймирующем редакторе и на tPERT. Согласно одному варианту осуществления связывающие домены могут включать фрагменты рекрутирования РНК-белок, такие как аптамер MS2, расположенный на tPERT, и белок MS2cp, слитый с праймирующим редактором. Преимущество *транс*-праймированного редактирования заключается в том, что, отделив матрицу синтеза ДНК от гидовой РНК, можно потенциально использовать матрицы большей длины.

[0332] Вариант осуществления *транс*-праймированного редактирования показан на фиг. 3G и 3H. На фиг. 3G слева показан состав комплекса *транс*-праймированного редактирования («комплекс RP-PE:гРНК»), который содержит napDNAbp, слитый с полимеразой (например, обратной транскриптазой) и рекрутирующим белком rPERT (например, MS2sc), и который находится в комплексе с гидовой РНК. На фиг. 3G дополнительно показана отдельная молекула tPERT, которая содержит элементы плеча удлинения РЕгРНК, включая матрицу для синтеза ДНК и последовательность связывания праймера. Молекула tPERT также включает рекрутинговый домен РНК-белок (который в данном случае представляет собой структуру «петля на стебле» и может представлять собой, например, аптамер MS2). Как показано в процессе, описанном на фиг. 3H, комплекс RP-PE:гРНК связывается с целевой последовательностью ДНК и разрывает ее. Затем рекрутирующий белок (RP) рекрутирует tPERT для совместной локализации в комплексе праймированного редактирования, связанном с целевым участком ДНК, тем самым позволяя участку связывания праймера связываться с последовательностью праймера на цепи с разрывом, а затем позволяя полимеразе (например, RT) синтезировать одну цепь ДНК против матрицы для синтеза ДНК до 5' tPERT.

[0333] Хотя tPERT показан на фиг. 3G и фиг. 3H, как содержащий PBS и матрицу для синтеза ДНК на 5'-конце домена рекрутирования РНК-белка, tPERT в других конфигурациях может быть разработан с PBS и матрицей для синтеза ДНК, расположенной на 3'-конце домена рекрутирования РНК-белка. Однако tPERT с 5'-удлинением имеет то преимущество, что синтез одноцепочечной ДНК естественным образом заканчивается на 5'-конце tPERT и, таким образом, не возникает риска использования какой-либо части домена рекрутирования РНК-белок в качестве матрицы на стадии синтеза ДНК праймированного редактирования.

Матрица РНК *транс*-праймирующего редактора (tPERT)

[0334] В контексте настоящего изобретения а «матрица РНК транс-праймирующего

редактора (tPERT)» относится к компоненту, используемому в транс-праймированном редактировании или двойном транс-праймированном редактировании, модифицированной версии праймированного редактирования, которая работает путем разделения РЕгРНК на две отдельные молекулы: гидовая РНК и молекула tPERT. Молекула tPERT запрограммирована на совместную локализацию с комплексом праймированного редактора в целевом участке ДНК, перенося участок связывания праймера и матрицу для синтеза ДНК в главный редактор к праймирующему редактору в транс. Например, см. фиг. 3G для варианта осуществления транс-праймирующего редактора (tPE), который показывает двухкомпонентную систему, содержащую (1) комплекс RP-PE:гРНК и (2) tPERT, который включает участок связывания праймера и матрицу для синтеза ДНК, присоединенную к домену рекрутирования РНК-белок, в котором компонент RP (рекрутирующий белок) комплекса RP-PE:гРНК рекрутирует tPERT к целевому участку, подлежащему редактированию, тем самым связывая PBS и матрицу для синтеза ДНК с праймирующим редактором в транс. Другими словами, tPERT сконструирован так, чтобы содержать (полностью или частично) плечо удлинения РЕгРНК, которое включает участок связывания праймера и матрицу для синтеза ДНК.

Трансзиции

[0335] В контексте настоящего изобретения термин «транзиции» означает замену пуриновых нуклеотидных оснований (A ↔ G) или замену пиримидиновых нуклеотидных оснований (C ↔ T). Данный класс замен затрагивает нуклеотидные основания одинаковой формы. Описанные в настоящем документе композиции и способы способны обеспечивать индуцирование одной или нескольких транзиций в молекуле-мишени ДНК. Описанные в настоящем документе 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 двухцепочечной ДНК с парными нуклеотидными 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 в настоящем документе композиции и способы способны обеспечивать индуцирование одной или нескольких транзиций в молекуле-мишени ДНК. Описанные в настоящем документе 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

[0336] В контексте настоящего изобретения термин «трансверсии» означает замену пуриновых нуклеотидных оснований пиримидиновыми нуклеотидными основаниями или

наоборот и, таким образом, включает взаимную замену нуклеотидных 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 двухцепочечной ДНК с парными нуклеотидными 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 в настоящем документе композиции и способы способны обеспечивать индуцирование одной или нескольких трансверсий в молекуле-мишени ДНК. Описанные в настоящем документе 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

[0337] 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 в настоящем документе. В контексте настоящего изобретения 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 в настоящем документе. Согласно некоторым вариантам осуществления лечение может осуществляться после развития одного или нескольких симптомов и/или после диагностирования заболевания. Согласно другим вариантам 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.

Нарушение, связанное с тринуклеотидными повторами

[0338] В контексте настоящего изобретения термин «нарушение, связанное с тринуклеотидными повторами» (или, альтернативно, «связанное экспансией повторов заболевание» или «связанное с повторной экспансией заболевание») означает набор генетических нарушений, 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 интеграциями 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.

[0339] 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.

[0340] Нарушения, связанные с нуклеотидными повторами, могут включать заболевания, при которых триплетный повтор присутствует в некодирующем участке (то есть некодирующее связанное с тринуклеотидными повторами заболевание) или в кодирующем участке.

[0341] Система праймирующего редактора (PE), описанная в настоящем документе, может использоваться для лечения нарушений, связанных с нуклеотидными повторами, которые могут включать синдром ломкой Х-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.

Против хода транскрипции

[0342] В контексте настоящего изобретения термины «против хода транскрипции» и «по ходу транскрипции» являются терминами относительности, которые определяют линейное положение по меньшей мере двух элементов, расположенных в молекуле нуклеиновой кислоты (будь то одноцепочечной или двухцепочечной), которая ориентирована 5'-к-3' направление. В частности, первый элемент находится против хода транскрипции от второго элемента в молекуле нуклеиновой кислоты, где первый элемент расположен где-то на расстоянии 5' от второго элемента. Например, SNP находится против хода транскрипции от Cas9-индуцированного участка разрыва, если SNP находится на 5'-стороне участка разрыва. И наоборот, первый элемент находится по ходу транскрипции от второго элемента в молекуле нуклеиновой кислоты, где первый элемент расположен где-то на расстоянии 3' от второго элемента. SNP находится по ходу транскрипции от Cas9-индуцированного участка разрыва, если SNP находится на 3'-стороне участка разрыва. Молекула нуклеиновой кислоты может представлять собой ДНК (двухцепочечную или одноцепочечную), РНК (двухцепочечная или одноцепочечная) или гибрид ДНК и РНК. Анализ одинаков для одноцепочечной молекулы нуклеиновой кислоты и для двухцепочечной молекулы, поскольку термины против хода транскрипции и по ходу транскрипции относятся только к одной цепи молекулы нуклеиновой кислоты, за исключением того, что необходимо выбрать, какая цепь двухцепочечной молекулы рассматривается. Часто цепь двухцепочечной ДНК, которую можно использовать для определения относительности положения по меньшей мере двух элементов, является «смысловой» или «кодирующей» цепью. В генетике «смысловая» цепь представляет собой сегмент двухцепочечной ДНК, который проходит от 5' до 3' и комплементарен антисмысловой цепи ДНК, или матричной цепи, которая проходит от 3' до 5'. Так, например, нуклеотидное основание SNP находится «по ходу транскрипции» последовательности промотора в геномной ДНК (которая является двухцепочечной), если нуклеотидное основание SNP находится на 3'-стороне промотора на смысловой или кодирующей цепи.

Вариант

[0343] В контексте настоящего изобретения термин «вариант» означает проявление качеств, которые имеют паттерн, являющийся отличным от находящегося в природе, например, вариант Cas9 представляет собой Cas9, включающую одно или несколько изменений 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

[0344] Термин «вектор» в контексте настоящего изобретения означает 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 конъюгатные плазмиды. Дополнительные подходящие векторы будут очевидны специалистам в данной области на основе настоящего описания.

Дикого типа

[0345] В контексте настоящего изобретения термин «дикого типа» представляет собой термин в данной области техники, который понимается специалистами в данной области и означает 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 лоскут ДНК

[0346] В контексте настоящего изобретения термин «5′- эндогенный лоскут ДНК» означает цепь ДНК, расположенную непосредственно по ходу транскрипции от индуцированного РЕ участка разрыва в ДНК-мишени. Разрыв цепи ДНК-мишени с помощью РЕ высвобождает 3'-гидроксильную группу на стороне против хода транскрипции от участка разрыва и 5'-гидроксильную группу на стороне по ходу транскрипции от участка разрыва. Эндогенная цепь, оканчивающаяся 3'-гидроксильной группой, используется для праймирования ДНК-полимеразы праймирующего редактора (например, когда ДНК-полимераза представляет собой обратную транскриптазу). Эндогенная цепь на стороне по ходу транскрипции от участка разрыва, которая начинается со свободной 5'-гидроксильной группы, обозначается как «5'- эндогенный лоскут ДНК» и в конечном итоге удаляется и замещается новой синтезированной замещающей цепью (т.е. «3'- замещающим лоскутом ДНК»), кодируемой удлинением РЕгРНК.

Удаление 5′- лоскута эндогенной ДНК

[0347] В контексте настоящего изобретения термин «удаление 5’-лоскута эндогенной ДНК» или «удаление 5’-лоскута» относится к удалению 5’-лоскута эндогенной ДНК, который образуется, когда лоскут одноцепочечной ДНК, синтезированный с помощью RT, конкурентно внедряется и гибридизуется с эндогенной ДНК, вытесняя при этом эндогенную цепь. Удаление этой эндогенно смещенной цепи может привести к образованию желаемого продукта, содержащего желаемую замену нуклеотида. Собственные ферменты репарации ДНК клетки могут катализировать удаление или вырезание 5'-эндогенного лоскута (например, эндонуклеаза лоскута, такая как EXO1 или FEN1). Кроме того, клетки-хозяева могут быть трансформированы для экспрессии одного или нескольких ферментов, катализирующих удаление указанных 5'-эндогенных лоскутов, тем самым направляя процесс к образованию продукта (например, лоскутной эндонуклеазы). Эндонуклеазы лоскутов известны в данной области техники, и их можно найти в описании 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 (каждая из которых включена в настоящий документ посредством ссылки).

3′-замещающий лоскут ДНК

[0348] В контексте настоящего изобретения термин «3′-замещающий лоскут ДНК» или просто «замещающий лоскут ДНК» относится к цепи ДНК, которая синтезируется праймирующим редактором и кодируется плечом удлинения PEгРНК праймирующего редактора. В частности, 3'-замещающий лоскут ДНК кодируется полимеразной матрицей PEгРНК. 3'-замещающий лоскут ДНК содержит ту же последовательность, что и 5'-эндогенный лоскут ДНК, за исключением того, что он также содержит отредактированную последовательность (например, замену одного нуклеотида). 3’-замещающий лоскут ДНК отжигается с ДНК-мишенью, вытесняя или заменяя 5’-эндогенный лоскут ДНК (который может быть вырезан, например, эндонуклеазой 5’-лоскута, такой как FEN1 или EXO1), а затем лигируется, чтобы присоединиться к ДНК-мишени с соединением 3’-конца 3’-замещающего лоскута ДНК с открытым 5’-гидроксильным концом эндогенной ДНК (обнажается после вырезания 5’-эндогенного лоскута ДНК), тем самым, тем самым реформируя фосфодисложноэфирную связь и устанавливая 3’- замещающий лоскут ДНК с образованием гетеродуплексной ДНК, содержащей одну отредактированную цепь и одну неотредактированную цепь. Процессы репарации ДНК разрешают гетеродуплекс путем копирования информации в отредактированной цепи в комплементарную цепь, навсегда

встраивая редактирование в ДНК. Этот процесс разрешения можно довести до завершения, надрезав неотредактированную цепи, т.е. посредством «разрыва второй цепи», как описано в настоящем документе.

Подробное описание некоторых вариантов осуществления

[0349] Адоптированные системы кластеризованных регулярных промежуточных коротких палиндромных повторов (CRISPR) для редактирования генома произвели революцию в науках о жизни [1–3]. Хотя разрушение генов с помощью CRISPR в настоящее время является стандартным подходом, точная установка редактирования отдельных нуклеотидов остается серьезной проблемой, несмотря на то, что это необходимо для изучения или исправления большого количества мутаций, вызывающих заболевание. Репарация, направляемая гемологией, (HDR) способна обеспечить такие редактирования, но имеет недостатки низкой эффективностью (часто <5%), необходимости использования донорных матриц репарации ДНК и вредных эффектов образования двухцепочечных разрывов ДНК (DSB). Недавно в лаборатории Prof. David Liu *et al.* разработали редактирование оснований, которое обеспечивает эффективное редактирование отдельных нуклеотидов без DSB. Редакторы оснований (BE) объединяют систему CRISPR с ферментами дезаминазы, модифицирующими основания, для преобразования целевых пар оснований C•G или A•T в A•T или G•C, соответственно[4–6]. Хотя современные BE уже широко используются исследователями во всем мире, они обеспечивают только четыре из двенадцати возможных преобразований пар оснований и не могут исправлять небольшие вставки или делеции. Кроме того, целевая область редактирования оснований ограничена редактированием нецелевых оснований C или A, соседних с целевым основанием («неспецифическое редактирование»), и требованием, чтобы последовательность PAM существовала на расстоянии 15 ± 2 п.о. от целевого основания. Таким образом, преодоление этих ограничений значительно расширит фундаментальные исследования и терапевтические применения редактирования генома.

[0350] Были разработаны новые подходы к точному редактированию (*например*, классическое праймированное редактирование), которые обладают многими преимуществами по сравнению с редактированием оснований, а именно, избегают двухцепочечных разрывов и донорных матриц репарации ДНК, преодолевая при этом его основные ограничения. Предлагаемый подход, описанный в настоящем документе, обеспечивает прямую вставку отредактированных цепей ДНК в целевых участках генома с использованием нацеленной-праймированной обратной транскрипции (TPRT). При конструировании, раскрытом в настоящем документе, гидовая РНК (гРНК) CRISPR

сконструирована таким образом, чтобы нести матричную последовательность обратной транскриптазы (RT), кодирующую одноцепочечную ДНК, содержащую желаемое нуклеотидное изменение. Целевой участок ДНК с разрывом нуклеазой CRISPR (Cas9) будет служить в качестве праймера для обратной транскрипции матричной последовательности на модифицированной гРНК, что позволяет прямое включение любого желаемого нуклеотидоного редактирования.

[0351] Механизм нацеленной-праймированной обратной транскрипции (TPRT) можно использовать или адаптировать для проведения точного редактирования генома на основе CRISPR/Cas с высокой эффективностью и генетической гибкостью (например, как показано в различных вариантах осуществления на фиг. 1A-1F). Слияния белка Cas с обратной транскриптазой используют для нацеливания на определенную последовательность ДНК с помощью модифицированной гидовой РНК («удлиненная гидовая РНК»), создания одноцепочечного разрыва в целевом участке и использования ДНК с разрывом в качестве праймера для обратной транскрипции сконструированной матрицы обратной транскриптазы, которую вставляют в удлиненную гидовую РНК. Вновь синтезированная цепь будет гомологична геномной целевой последовательности, за исключением включения желаемого нуклеотидного изменения (например, замены одного нуклеотида, делеции или вставки или их комбинации). Вновь синтезированную цепь ДНК можно назвать лоскутом одноцепочечной ДНК, который будет конкурировать за гибридизацию с комплементарной гомологичной эндогенной цепью ДНК, тем самым вытесняя соответствующую эндогенную цепь. Результат этого гибридизованного промежуточного продукта может включать удаление полученного смещенного лоскута эндогенной ДНК (например, с помощью эндонуклеазы 5'-концевого лоскута ДНК, FEN1), лигирование синтезированного лоскута одноцепочечной ДНК с ДНК-мишенью и ассимиляцию желаемого нуклеотидного изменения в результате процессов репарации и/или репликации клеточной ДНК. Поскольку матричный синтез ДНК обеспечивает точность до одного нуклеотида, область применения этого подхода очень широка и, как ожидается, он может быть использован для множества применений в фундаментальной науке и терапии.

[0352] Настоящее изобретение относится к новой платформе для редактирования генома, называемой «мультилоскутное праймированное редактирование» (включая, например, «двухлоскутное праймированное редактирование» и «четырехлоскутное праймированное редактирование») и представляющей собой инновационное усовершенствование «праймированного редактирования» или «классического праймированного редактирования», как описано авторами настоящего изобретения в

Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019), которая включена в настоящий документ посредством ссылки. В то время, как классическое праймированное редактирование согласно различным вариантам осуществления полимеризует в участке разрыва одиночный 3'-лоскут, который интегрируется в нуклеиновую кислоту-мишень на той же цепи, описанные в настоящем документе системы мультилоскутного праймированного редактирования включают различные конструкции, системы и методики, которые, согласно различным вариантам осуществления, создают пары или множество пар 3'-лоскутов на разных цепях, которые образуют дуплексы, содержащие желаемые редактирования, и которые включаются в молекулы-мишени нуклеиновой кислоты, например, в определенных локусах или участках редактирования в геноме. Согласно различным аспектам пары или множество пар 3'-лоскутов образуют дуплексы, поскольку они содержат обратные комплементарные последовательности, которые отжигаются друг с другом после того, как были сгенерированы праймирующими редакторами, описанными в настоящем документе. Дуплексы включаются в целевой участок с помощью клеточных механизмов, которые естественным образом замещают эндогенные последовательности дуплексов, расположенные между соседними участками разрыва. Согласно определенным вариантам осуществления новые дуплексные последовательности могут быть введены в одно или несколько мест (например, в соседние геномные локусы или в два разных участка хромосомы) и могут содержать одну или множество представляющих интерес последовательностей, например, последовательность, кодирующую белок, пептид-кодирующую последовательность или РНК-кодирующую последовательность. Согласно одному варианту осуществления новые дуплексные последовательности, установленные системами мультилоскутного праймированного редактирования, могут содержать сайт рекомбиназы, *например,* сайт *attB* (38 п.о.) и/или *attP* (50 п.о.) сайт рекомбинзы Bxb1, или сайт рекомбиназы, распознаваемый рекомбиназой Hin, рекомбиназой Gin, рекомбиназой Tn3, рекомбиназой β-six, рекомбиназой CinH, рекомбиназой ParA, рекомбиназой γδ, рекомбиназой ϕC31, рекомбиназой TP901, рекомбиназой TG1, рекомбиназой φBT1, рекомбиназой R4, рекомбиназой φRV1, рекомбиназой φFC1, рекомбиназой MR11, рекомбиназой A118, рекомбиназой U153 и рекомбиназой gp29, рекомбиназой Cre, рекомбиназой FLP, рекомбиназой R, рекомбиназой Lambda, рекомбиназой HK101, рекомбиназой HK022 и рекомбиназой pSAM2.

[0353] Согласно различным аспектам в настоящем описании раскрыта система мультилоскутного праймированного редактирования (включая, например, систему двойного праймированного редактирования и систему четвертного праймированного

редактирования), которая решает проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в неотредактированную комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. В системе двухлоскутного праймированного редактирования, например, две РЕгРНК две PEgRNA используются для нацеливания на противоположные цепи участка генома и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 90**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная нить. Поскольку обе цепи дуплекса синтезируются в виде отредактированной ДНК, система двухлоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК нужно только вырезать парные 5'- лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные цепь ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое содержит минимальную геномную гомологию. Варианты с нуклеазной активностью праймирующих редакторов, которые разрезают обе цепи ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК. Система четырехлоскутного праймированного редактирования, использующая четыре РЕгРНК, обеспечивает подобные преимущества.

[1] napDNAbp

[0354] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут содержать программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp).

[0355] Согласно одному аспекту napDNAbp может быть связан или в комплексе с по меньшей мере одной гидовой нуклеиновой кислотой (например, гидовая РНК или РЕгРНК), которая локализует napDNAbp в последовательности ДНК, которая содержит цепь ДНК (т.е. целевая цепь), которая комплементарна гидовой нуклеиновой кислоте или ее части (например, спейсер гидовой РНК, который гибридизуется с протоспейсером ДНК-мишени). Другими словами, гидовая нуклеиновая кислота «программирует» napDNAbp (например, Cas9 или его эквивалент) на локализацию и связывание с комплементарной

последовательностью протоспейсера в ДНК.

[0356] Любой подходящий napDNAbp может быть использован в мультилоскутных праймирующих редакторах, описанных в настоящем документе. Согласно различным вариантам осуществления napDNAbp может представлять собой любую систему CRISPR-Cas класса 2, включая любой фермент CRISPR-Cas типа II, типа V или типа VI. Учитывая быстрое развитие CRISPR-Cas как инструмента редактирования генома, номенклатура, используемая для описания и/или идентификации ферментов CRISPR-Cas, таких как Cas9 и ортологи Cas9, постоянно развивалась. В настоящем документе упоминаются ферменты CRISPR-Cas с номенклатурой, которая может быть устаревшей и/или новой. Квалифицированный специалист сможет идентифицировать конкретный фермент CRISPR-Cas, упоминаемый в настоящем документе, на основе используемой номенклатуры, будь то устаревшая (т.е. «унаследованная») или новая номенклатура. Номенклатура CRISPR-Cas широко обсуждается в Makarova et al., «Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?,» *The CRISPR Journal*, Vol. 1. No. 5, 2018, полное содержание которой включено в настоящий документ посредством ссылки. Конкретная номенклатура CRISPR-Cas, используемая в любом конкретном случае в настоящем документе, никоим образом не является ограничивающей, и специалист в данной области сможет определить о каком ферменте CRISPR-Cas идет речь.

[0357] Например, следующие ферменты CRISPR-Cas класса 2 типа II, типа V и типа VI имеют следующие признанные в данной области техники устаревшие (т.е. унаследованные) и новые названия. Каждый из этих ферментов и/или их варианты можно использовать с мультилоскутными праймирующими редакторами, описанными в настоящем документе:

Устаревшая номенклатура	**Существующая в настоящее время номенклатура***
ферменты CRISPR-Cas типа II	
Cas9	такое же
ферменты CRISPR-Cas типа V	
Cpf1	Cas12a
CasX	Cas12e
C2c1	Cas12b1
Cas12b2	такое же
C2c3	Cas12c
CasY	Cas12d

C2c4	такое же
C2c8	такое же
C2c5	такое же
C2c10	такое же
C2c9	такое же
ферменты CRISPR-Cas типа VI	
C2c2	Cas13a
Cas13d	такое же
C2c7	Cas13c
C2c6	Cas13b

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

[0358] Без ограничения какой-либо теорией, механизм действия определенных napDNAbp, рассматриваемых в настоящем документе, предусматривает стадию формирования R-петли, посредством которой napDNAbp индуцирует раскручивание двухцепочечной ДНК-мишени, тем самым разделяя цепи в области, связанной с napDNAbp. Затем спейсер гидовой РНК гибридизуется с «целевой цепью» в последовательности протоспейсера. Это вытесняет «нецелевую цепь», комплементарную целевой цепи, которая образует одноцепочечную область R-петли. Согласно некоторым вариантам осуществления napDNAbp включает одну или несколько нуклеазных активностей, которые затем разрезают ДНК, оставляя различные типы повреждений. Например, napDNAbp может обладать нуклеазной активностью, которая разрезает нецелевую цепь в первом месте и/или разрезает целевую цепь во втором месте. В зависимости от нуклеазной активности ДНК-мишень может быть разрезана с образованием «двухцепочечного разрыва», при котором разрезаются обе цепи. Согласно другим вариантам осуществления ДНК-мишень может быть разрезана только в одном участке, т.е. ДНК «разрезана» в одной цепи. Примеры napDNAbp с различной нуклеазной активностью включают «никазу Cas9» («nCas9») и деактивированную Cas9, не имеющую нуклеазной активности («мертвая Cas9» или «dCas9»).

[0359] Приведенное ниже описание различных napDNAbp, которые можно использовать в связи с раскрытыми в настоящем документе мультилоскутными праймирующими редакторами, не является ограничением каким-либо образом. Мультилоскутные праймирующие редакторы могут содержать канонический белок SpCas9, или любой ортолог белка Cas9, или любой вариант белка Cas9, включая любой встречающийся в природе вариант, мутант или иным образом сконструированный вариант Cas9, который известен или может быть получен или эволюционировал посредством

направленного эволюционного процесса или иного мутагенного процесса. Согласно различным вариантам осуществления Cas9 или варианты Cas9 обладают никазной активностью, т.е., расщепляют только одну цепь целевой последовательности ДНК. Согласно другим вариантам осуществления Cas9 или варианты Cas9 не обладают нуклеазной активностью, т.е. представляют собой «мертвые» белки Cas9. Другими вариантами белков Cas9, которые могут быть использованы, являются те, которые имеют меньшую молекулярную массу, чем канонический SpCas9 (например, для облегчения доставки), или имеют модифицированную или подвергнутую перестановке структуру первичных аминокислот (например, форматы круговых пермутантов).

[0360] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, также могут содержать эквиваленты Cas9, включая белки Cas12a (Cpf1) и Cas12b1, которые являются результатом конвергентной эволюции. Согласно настоящему изобретению napDNAbp (например, SpCas9, вариант Cas9 или эквиваленты Cas9) могут также содержать различные модификации, которые изменяют/усиливают их PAM-специфичность. Наконец, настоящее изобретение охватывает любой Cas9, вариант Cas9 или эквивалент Cas9, последовательность которого на 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% идентична эталонной последовательности Cas9, такой как эталонная каноническая последовательность SpCas9 или эталонный эквивалент Cas9 (например, Cas12a (Cpf1)).

[0361] NapDNAbp может быть нуклеазой, ассоциированной с CRISPR (короткие палиндромные повторы с регулярно расположенными кластерами). Как указано выше, CRISPR представляет собой адаптивную иммунную систему, обеспечивающую защиту от мобильных генетических элементов (вирусов, мобильных элементов и конъюгативных плазмид). Кластеры CRISPR содержат спейсеры, последовательности, комплементарные предшествующим мобильным элементам, и нацелены на инвазивные нуклеиновые кислоты. Кластеры CRISPR транскрибируются и процессируются в РНК CRISPR (crРНК). В системах CRISPR типа II для правильного процессинга пре-crРНК требуется транс-кодируемая малая РНК (tracrRNA), эндогенная рибонуклеаза 3 (rnc) и белок Cas9. tracrRNA служит направляющей для процессинга пре-crРНК с помощью рибонуклеазы 3. Затем Cas9/crРНК/tracrРНК эндонуклеолитически расщепляет линейную или кольцевую дцДНК-мишень, комплементарную спейсеру. Целевая цепь, некомплементарная crРНК, сначала разрезается эндонуклеолитически, а затем обрезается 3'-5' экзонуклеолитически. В природе

для связывания и расщепления ДНК обычно требуется белок и обе РНК. Однако одиночные гидовые РНК («crРНК» или просто «gRNA») могут быть сконструированы таким образом, чтобы включать аспекты как crРНК, так и tracrRNA в один вид РНК. См., например, Jinek M. et al., Science 337:816-821 (2012), полное содержание которой включено в настоящий документ посредством ссылки.

[0362] Согласно некоторым вариантам осуществления napDNAbp направляет расщепление одной или обеих цепей в месте расположения целевой последовательности, например, в пределах целевой последовательности и/или в комплементе целевой последовательности. Согласно некоторым вариантам осуществления napDNAbp направляет расщепление одной или обеих цепей в пределах приблизительно 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 целевой последовательности. Согласно некоторым вариантам осуществления вектор кодирует napDNAbp, который мутирован по отношению к соответствующему ферменту дикого типа, так что мутантная napDNAbp не способна расщеплять одну или обе цепи целевого полинуклеотида, содержащего целевую последовательность. Например, замена аспартата на аланин (D10A) в каталитическом домене RuvC I Cas9 из *S. pyogenes* превращает 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 в отношении канонической последовательности SpCas9 или эквивалентных положений аминокислот в других вариантах Cas9 или эквивалентах Cas9.

[0363] В контексте настоящего изобретения термин «белок Cas» относится к полноразмерному белку Cas, полученному из природного источника, рекомбинантному белку Cas, имеющему последовательности, которые отличаются от встречающегося в природе белка Cas, или любому фрагменту белка Cas, который, тем не менее, сохраняет все или значительное количество необходимых основных функций, необходимых для раскрытых способов, т.е. (i) наличие программируемого нуклеиновой кислотой связывания белка Cas с ДНК-мишенью и (ii) способность разрывать последовательность ДНК-мишени на одной цепи. Белки Cas, рассматриваемые в настоящем документе, охватывают белки CRISPR Cas9, а также эквиваленты, варианты (например, никаза Cas9 (nCas9) или Cas9 без нуклеазной активности (dCas9)) гомологи, ортологи или паралоги Cas9, встречающиеся в природе 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, содержание которых включено в настоящий документ посредством ссылки.

[0364] Термины «Cas9» или «нуклеаза Cas9», или «фрагмент Cas9», или «домен Cas9» охватывают любой встречающийся в природе Cas9 из любого организма, любой природный эквивалент Cas9 или его функциональный фрагмент, любой гомолог, ортолог или паралог Cas9 из любого организма, а также любой мутант или вариант Cas9, встречающийся в природе или сконструированный. Термин Cas9 не предназначен для особого ограничения и может упоминаться как «Cas9 или эквивалент». Примеры белков Cas9 дополнительно описаны в настоящем документе и/или описаны в данной области техники и включены в настоящий документ в качестве ссылки. Настоящее раскрытие не ограничено в отношении конкретного Cas9, который используется в мультилоскутных праймирующих редакторах, описанных в настоящем документе.

[0365] Как отмечено в настоящем документе, последовательности и структуры нуклеазы Cas9 хорошо известны специалистам в данной области техники (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), полные содержания каждой из которых включены в настоящий документ посредством ссылки).

[0366] Примеры Cas9 и эквивалентов Cas9 представлены ниже, однако, эти конкретные примеры не являются ограничением. Мультилоскутные праймирующие редакторы согласно настоящему изобретению могут использовать любой подходящий napDNAbp, включая любой подходящий Cas9 или эквивалент Cas9.

A. <u>Канонический дикого типа SpCas9</u>

[0367] Согласно одному варианту осуществления конструкции мультилоскутного праймирующего редактора, описанные в настоящем документе, могут содержать

«каноническую SpCas9» нуклеазу из *S. pyogenes*, которая широко используется в качестве инструмента для геномной инженерии и относится к подгруппе ферментов типа II класса 2 системы CRISPR-Cas. Этот белок Cas9 представляет собой большой многодоменный белок, содержащий два различных нуклеазных домена. Точечные мутации могут быть введены в Cas9 для отмены одной или обеих нуклеазных активностей, что приводит к образованию никазы Cas9 (nCas9) или мертвого Cas9 (dCas9), соответственно, которые все еще сохраняют свою способность связываться с ДНК запрограммированным сгРНК способом. В принципе, при слиянии с другим белком или доменом Cas9 или его вариант (например, nCas9) может нацеливать этот белок практически на любую последовательность ДНК просто путем совместной экспрессии с соответствующей сгРНК. Используемый в настоящем документе канонический белок SpCas9 относится к белку дикого типа из *Streptococcus pyogenes*, имеющему следующую аминокислотную последовательность:

Описание	Последовательность	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* M1 SwissProt № доступа Q99ZW2 Дикого типа	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD AKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKL INGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVV	SEQ ID NO: 18

	GTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLK SVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADA NLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
SpCas9 Обратная трансляция SwissProt № доступа Q99ZW2 *Streptococcus pyogenes*	ATGGATAAAAAATATAGCATTGGCCTGGATATTGGCACC AACAGCGTGGGCTGGGCGGTGATTACCGATGAATATAAA GTGCCGAGCAAAAAATTTAAAGTGCTGGGCAACACCGAT CGCCATAGCATTAAAAAAAACCTGATTGGCGCGCTGCTG TTTGATAGCGGCGAAACCGCGGAAGCGACCCGCCTGAAA CGCACCGCGCGCCGCCGCTATACCCGCCGCAAAAACCGC ATTTGCTATCTGCAGGAAATTTTTAGCAACGAAATGGCGA AAGTGGATGATAGCTTTTTTCATCGCCTGGAAGAAAGCTT TCTGGTGGAAGAAGATAAAAAACATGAACGCCATCCGAT TTTTGGCAACATTGTGGATGAAGTGGCGTATCATGAAAA ATATCCGACCATTTATCATCTGCGCAAAAAACTGGTGGAT AGCACCGATAAAGCGGATCTGCGCCTGATTTATCTGGCG CTGGCGCATATGATTAAATTTCGCGGCCATTTTCTGATTG AAGGCGATCTGAACCCGGATAACAGCGATGTGGATAAAC TGTTTATTCAGCTGGTGCAGACCTATAACCAGCTGTTTGA AGAAAACCCGATTAACGCGAGCGGCGTGGATGCGAAAGC GATTCTGAGCGCGCGCCTGAGCAAAAGCCGCCGCCTGGA AAACCTGATTGCGCAGCTGCCGGGCGAAAAAAAAAACGG CCTGTTTGGCAACCTGATTGCGCTGAGCCTGGGCCTGACC CCGAACTTTAAAAGCAACTTTGATCTGGCGGAAGATGCG AAACTGCAGCTGAGCAAAGATACCTATGATGATGATCTG GATAACCTGCTGGCGCAGATTGGCGATCAGTATGCGGAT CTGTTTCTGGCGGCGAAAAACCTGAGCGATGCGATTCTGC TGAGCGATATTCTGCGCGTGAACACCGAAATTACCAAAG CGCCGCTGAGCGCGAGCATGATTAAACGCTATGATGAAC ATCATCAGGATCTGACCCTGCTGAAAGCGCTGGTGCGCC AGCAGCTGCCGGAAAAATATAAAGAAATTTTTTTTGATC AGAGCAAAAACGGCTATGCGGGCTATATTGATGGCGGCG CGAGCCAGGAAGAATTTTATAAATTTATTAAACCGATTCT	SEQ ID NO: 19

	GGAAAAAATGGATGGCACCGAAGAACTGCTGGTGAAACT GAACCGCGAAGATCTGCTGCGCAAACAGCGCACCTTTGA TAACGGCAGCATTCCGCATCAGATTCATCTGGGCGAACT GCATGCGATTCTGCGCCGCCAGGAAGATTTTTATCCGTTT CTGAAAGATAACCGCGAAAAAATTGAAAAAATTCTGACC TTTCGCATTCCGTATTATGTGGGCCCGCTGGCGCGCGGCA ACAGCCGCTTTGCGTGGATGACCCGCAAAAGCGAAGAAA CCATTACCCCGTGGAACTTTGAAGAAGTGGTGGATAAAG GCGCGAGCGCGCAGAGCTTTATTGAACGCATGACCAACT TTGATAAAAACCTGCCGAACGAAAAAGTGCTGCCGAAAC ATAGCCTGCTGTATGAATATTTTACCGTGTATAACGAACT GACCAAAGTGAAATATGTGACCGAAGGCATGCGCAAACC GGCGTTTCTGAGCGGCGAACAGAAAAAAGCGATTGTGGA TCTGCTGTTTAAAACCAACCGCAAAGTGACCGTGAAACA GCTGAAAGAAGATTATTTTAAAAAAATTGAATGCTTTGAT AGCGTGGAAATTAGCGGCGTGGAAGATCGCTTTAACGCG AGCCTGGGCACCTATCATGATCTGCTGAAAATTATTAAAG ATAAAGATTTTCTGGATAACGAAGAAAACGAAGATATTC TGGAAGATATTGTGCTGACCCTGACCCTGTTTGAAGATCG CGAAATGATTGAAGAACGCCTGAAAACCTATGCGCATCT GTTTGATGATAAAGTGATGAAACAGCTGAAACGCCGCCG CTATACCGGCTGGGGCCGCCTGAGCCGCAAACTGATTAA CGGCATTCGCGATAAACAGAGCGGCAAAACCATTCTGGA TTTTCTGAAAAGCGATGGCTTTGCGAACCGCAACTTTATG CAGCTGATTCATGATGATAGCCTGACCTTTAAAGAAGAT ATTCAGAAAGCGCAGGTGAGCGGCCAGGGCGATAGCCTG CATGAACATATTGCGAACCTGGCGGGCAGCCCGGCGATT AAAAAAGGCATTCTGCAGACCGTGAAAGTGGTGGATGAA CTGGTGAAAGTGATGGGCCGCCATAAACCGGAAAACATT GTGATTGAAATGGCGCGCGAAAACCAGACCACCCAGAAA GGCCAGAAAAACAGCCGCGAACGCATGAAACGCATTGA AGAAGGCATTAAAGAACTGGGCAGCCAGATTCTGAAAGA ACATCCGGTGGAAAACACCCAGCTGCAGAACGAAAAACT GTATCTGTATTATCTGCAGAACGGCCGCGATATGTATGTG GATCAGGAACTGGATATTAACCGCCTGAGCGATTATGAT GTGGATCATATTGTGCCGCAGAGCTTTCTGAAAGATGATA GCATTGATAACAAAGTGCTGACCCGCAGCGATAAAAACC GCGGCAAAAGCGATAACGTGCCGAGCGAAGAAGTGGTG	

	AAAAAAATGAAAAACTATTGGCGCCAGCTGCTGAACGCG AAACTGATTACCCAGCGCAAATTTGATAACCTGACCAAA GCGGAACGCGGCGGCCTGAGCGAACTGGATAAAGCGGG CTTTATTAAACGCCAGCTGGTGGAAACCCGCCAGATTACC AAACATGTGGCGCAGATTCTGGATAGCCGCATGAACACC AAATATGATGAAAACGATAAACTGATTCGCGAAGTGAAA GTGATTACCCTGAAAAGCAAACTGGTGAGCGATTTTCGC AAAGATTTTCAGTTTTATAAAGTGCGCGAAATTAACAACT ATCATCATGCGCATGATGCGTATCTGAACGCGGTGGTGG GCACCGCGCTGATTAAAAAATATCCGAAACTGGAAAGCG AATTTGTGTATGGCGATTATAAAGTGTATGATGTGCGCAA AATGATTGCGAAAAGCGAACAGGAAATTGGCAAAGCGA CCGCGAAATATTTTTTTTATAGCAACATTATGAACTTTTTT AAAACCGAAATTACCCTGGCGAACGGCGAAATTCGCAAA CGCCCGCTGATTGAAACCAACGGCGAAACCGGCGAAATT GTGTGGGATAAAGGCCGCGATTTTGCGACCGTGCGCAAA GTGCTGAGCATGCCGCAGGTGAACATTGTGAAAAAAACC GAAGTGCAGACCGGCGGCTTTAGCAAAGAAAGCATTCTG CCGAAACGCAACAGCGATAAACTGATTGCGCGCAAAAAA GATTGGGATCCGAAAAAATATGGCGGCTTTGATAGCCCG ACCGTGGCGTATAGCGTGCTGGTGGTGGCGAAAGTGGAA AAAGGCAAAAGCAAAAAACTGAAAAGCGTGAAAGAACT GCTGGGCATTACCATTATGGAACGCAGCAGCTTTGAAAA AAACCCGATTGATTTTCTGGAAGCGAAAGGCTATAAAGA AGTGAAAAAAGATCTGATTATTAAACTGCCGAAATATAG CCTGTTTGAACTGGAAAACGGCCGCAAACGCATGCTGGC GAGCGCGGGCGAACTGCAGAAAGGCAACGAACTGGCGC TGCCGAGCAAATATGTGAACTTTCTGTATCTGGCGAGCCA TTATGAAAAACTGAAAGGCAGCCCGGAAGATAACGAACA GAAACAGCTGTTTGTGGAACAGCATAAACATTATCTGGA TGAAATTATTGAACAGATTAGCGAATTTAGCAAACGCGT GATTCTGGCGGATGCGAACCTGGATAAAGTGCTGAGCGC GTATAACAAACATCGCGATAAACCGATTCGCGAACAGGC GGAAAACATTATTCATCTGTTTACCCTGACCAACCTGGGC GCGCCGGCGGCGTTTAAATATTTTGATACCACCATTGATC GCAAACGCTATACCAGCACCAAAGAAGTGCTGGATGCGA CCCTGATTCATCAGAGCATTACCGGCCTGTATGAAACCCG CATTGATCTGAGCCAGCTGGGCGGCGAT	

[0368] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут включать канонический SpCas9 или любой его вариант, последовательность которого на по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентична последовательности Cas9 дикого типа, представленной выше. Эти варианты могут включать варианты SpCas9, содержащие одну или несколько мутаций, включая любую известную мутацию, о которой сообщается в SwissProt № доступа Q99ZW2 (SEQ ID NO: 18), которые включают:

Мутация SpCas9 (относительно аминокислотной последовательности канонической последовательности SpCas9, SEQ ID NO: 18)	Функция/характеристика (как сообщалось) (см. UniProtKB —Q99ZW2 (CAS9_STRPT1) — включена в настоящий документ посредством ссылки)
D10A	Мутант никазы, который расщепляет протоспейсерную цепь (но не расщепляет непротоспейсерную цепь)
S15A	Снижение активности расщепления ДНК
R66A	Снижение активности расщепления ДНК
R70A	Без расщепления ДНК
R74A	Снижение расщепления ДНК
R78A	Снижение расщепления ДНК
делеция 97-150	Нет нуклеазной активности
R165A	Снижение расщепления ДНК
делеция 175-307	Приблизительно на 50 % уменьшилось расщепление ДНК
делеция 312-409	Нет нуклеазной активности
E762A	Никаза
H840A	Мутант никазы, который расщепляет непротоспейсерную цепь, но не расщепляет протоспейсерную цепь
N854A	Никаза
N863A	Никаза
H982A	Снижение расщепления ДНК
D986A	Никаза
делеция 1099-1368	Нет нуклеазной активности
R1333A	Снижение связывания ДНК

[0369] Другие последовательности SpCas9 дикого типа, которые можно

использовать согласно настоящему изобретению, включают:

Описание	Последовательности	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* MGAS1882 дикого типа NC_017053.1	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAA TAGCGTCGGATGGGCGGTGATCACTGATGATTATAAGGTTCC GTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACA GTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTG GAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGT AGAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAG GAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTC TTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAG AAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAA GTTGCTTATCATGAGAAATATCCAACTATCTATCATCTGCGA AAAAAATTGGCAGATTCTACTGATAAAGCGGATTTGCGCTTA ATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATT TTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGG ACAAACTATTTATCCAGTTGGTACAAATCTACAATCAATTAT TTGAAGAAAACCCTATTAACGCAAGTAGAGTAGATGCTAAA GCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAA AATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGCTTG TTTGGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATT TTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGC TTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGG CGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTA AGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAG TAAATAGTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGA TTAAGCGCTACGATGAACATCATCAAGACTTGACTCTTTTAA AAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAA ATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATT GATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAA ACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGT GAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCT TTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGC TGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTT TAAAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTC GAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTC GTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCC CATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTC AATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTC	SEQ ID NO: 20

	CAAATGAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGT ATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTA CTGAGGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAG AAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAA GTAACCGTTAAGCAATTAAAAGAAGATTATTTCAAAAAAAT AGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAG ATTTAATGCTTCATTAGGCGCCTACCATGATTTGCTAAAAAT TATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAG ATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAG ATAGGGGGATGATTGAGGAAAGACTTAAAACATATGCTCAC CTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGT TATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTG AAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATC CATGATGATAGTTTGACATTTAAAGAAGATATTCAAAAAGC ACAGGTGTCTGGACAAGGCCATAGTTTACATGAACAGATTG CTAACTTAGCTGGCAGTCCTGCTATTAAAAAAGGTATTTTAC AGACTGTAAAAATTGTTGATGAACTGGTCAAAGTAATGGGG CATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAA TCAGACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTA TGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAG ATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAAT GAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATG TATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTAT GATGTCGATCACATTGTTCCACAAAGTTTCATTAAAGACGAT TCAATAGACAATAAGGTACTAACGCGTTCTGATAAAAATCGT GGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAA GATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAA TCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTG GAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCC AATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACAA ATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGAT AAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAA TTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTA CGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTA AATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAA CTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGAT GTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAA	

	AGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTT CTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGCA AACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATT GTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCGCAAAGT ATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGT ACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAA GAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGAT CCAAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTAT TCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAA GAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTAT GGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGA AGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTA AACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTA AACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAAT GAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTA GCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAA CGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCATTATTT AGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGT TATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATA TAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAA ATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGC TGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGATA TACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCA ATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCA GCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* MGAS1882 дикого типа NC_017053.1	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSI KKNLIGALLFGSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEGDL NPDNSDVDKLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKS RRLENLIAQLPGEKRNGLFGNLIALSLGLTPNFKSNFDLAEDAK LQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR VNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIF FDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF	SEQ ID NO: 21

	KKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDEL VKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLS DYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEF VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITL ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTV AYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGD	
SpCas9 *Streptococcus pyogenes* дикого типа SWBC2D7W014	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAAT TCCGTTGGATGGGCTGTCATAACCGATGAATACAAAGTACCT TCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTC GATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGG CGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGA GAAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAA GAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTC TTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAG AAACATGAACGGCACCCCATCTTTGGAAACATAGTAGATGA GGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTCAG AAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGT TAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGC ACTTTCTCATTGAGGGTGATCTAAATCCGGACAACTCGGATG TCGACAAACTGTTCATCCAGTTAGTACAAACCTATAATCAGT TGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCG AAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTA GAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGG GTTGTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACC AAATTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAATT	SEQ ID NO: 22

	GCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATC TACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGG CTGCCAAAAACCTTAGCGATGCAATCCTCCTATCTGACATAC TGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTT CAATGATCAAAAGGTACGATGAACATCACCAAGACTTGACA CTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATAT AAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGG TTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGT TTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAG TTGCTTGTAAAACTCAATCGCGAAGATCTACTGCGAAAGCAG CGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTA GGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTAT CCGTTCCTCAAAGACAATCGTGAAAAGATTGAGAAAATCCT AACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGG GAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAAGAAA CGATTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTG CGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGACA AGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTTA CTTTACGAGTATTTCACAGTGTACAATGAACTCACGAAAGTT AAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCTAAGC GGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACT TTAAGAAAATTGAATGCTTCGATTCTGTCGAGATCTCCGGGG TAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACC TCCTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGAA GAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTTACC CTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAAC ATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAA GAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAAC TTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTATT CTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTT ATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGAT ATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCA CGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTA AGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAG ATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAA ACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAA	

	AGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAA ATACCCAATTGCAGAACGAGAAACTTTACCTCTATTACCTAC AAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATA AACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAA TCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACA CGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAG CGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGC TCCTAAATGCGAAACTGATAACGCAAAGAAAGTTCGATAAC TTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAA GGCCGGATTTATTAAACGTCAGCTCGTGGAAACCCGCCAAAT CACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATA CGAAATACGACGAGAACGATAAGCTGATTCGGGAAGTCAAA GTAATCACTTTAAAGTCAAAATTGGTGTCGGACTTCAGAAAG GATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACCAC CATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCA CTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTAT GGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAA AAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCT TTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCT GGCAAACGGAGAGATACGCAAACGACCTTTAATTGAAACCA ATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGAC TTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAAC ATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAA GGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCG CTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTC GATAGCCCTACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAA GTTGAGAAGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGA ATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAA GAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAG TAAAAAAGGATCTCATAATTAAACTACCAAAGTATAGTCTGT TTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCC GGAGAGCTTCAAAAGGGGAACGAACTCGCACTACCGTCTAA ATACGTGAATTTCCTGTATTTAGCGTCCCATTACGAGAAGTT GAAAGGTTCACCTGAAGATAACGAACAGAAGCAACTTTTTG TTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGCAA ATTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAAT CTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAA ACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTAC	

	TCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGA CACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGG TGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTAT ATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGA TCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGA CCATGACGGTGATTATAAAGATCATGACATCGATTACAAGG ATGACGATGACAAGGCTGCAGGA	
SpCas9 *Streptococcus pyogenes* дикого типа Кодируемый продукт SWBC2D7W0 14	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRE KIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRY TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITG	SEQ ID NO: 23

	LYETRIDLSQLGGDGSPKKKRKVSSDYKDHDGDYKDHDIDYK DDDDKAAG	
SpCas9 *Streptococcus pyogenes* M1GAS дикого типа NC_002737.2	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAA TAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCC GTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACA GTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTG GAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGT AGAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAG GAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTC TTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAG AAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAA GTTGCTTATCATGAGAAATATCCAACTATCTATCATCTGCGA AAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTA ATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATT TTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGG ACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTAT TTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAA GCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAA AATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTA TTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATT TTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGC TTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGG CGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTA AGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAG TAAATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGA TTAAACGCTACGATGAACATCATCAAGACTTGACTCTTTTAA AAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAA ATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATT GATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAA ACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTGGT GAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCT TTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGC TGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTT TAAAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTC GAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTC GTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCC CATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTC AATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTC	SEQ ID NO: 24

	CAAATGAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGT ATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTA CTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAG AAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAA GTAACCGTTAAGCAATTAAAAGAAGATTATTTCAAAAAAAT AGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAG ATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAAT TATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAG ATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAG ATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCAC CTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGT TATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGT ATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTG AAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATC CATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGC ACAAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGC AAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACA GACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGC GGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAA AATCAGACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCG TATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTC AGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAA ATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACA TGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATT ATGATGTCGATCACATTGTTCCACAAAGTTTCCTTAAAGACG ATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATC GTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAA AAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTT AATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACG TGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACG CCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACA AATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATG ATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTA AATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAG TACGTGAGATTAACAATTACCATCATGCCCATGATGCGTATC TAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAA AACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATG ATGTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGC	

	AAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAAC TTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGC AAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAAT TGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCGCAAAG TATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAG TACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAA AGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGA TCCAAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTA TTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGA AGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATT ATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTA GAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCAT TAAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCG TAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAA ATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATT TAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGAT AACGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCATTA TTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCG TGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGC ATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAG AAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTC CCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAAC GATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCC ATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGA GTCAGCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* M1GAS дикого типа Кодируемый продукт NC_002737.2 (100% идентичен каноническом у Q99ZW2 дикого типа)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRE KIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED	SEQ ID NO: 25

	YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRY TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITG LYETRIDLSQLGGD	

[0370] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут включать любую из вышеописанных последовательностей SpCas9, или любой их вариант, имеющий идентичность последовательности по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99%.

B. Ортологи Cas9 дикого типа

[0371] Согласно другим вариантам осуществления белок Cas9 может быть ортологом Cas9 дикого типа из другого вида бактерий, отличным от канонического Cas9 из *S. pyogenes*. Например, следующие ортологи Cas9 можно использовать в связи с конструкциями мультилоскутного праймирующего редактора, описанными в настоящем документе. Кроме того, любые варианты ортологов Cas9, последовательность которых на по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентична последовательности любого из указанных ниже ортологов, также могут быть использованы с мультилоскутными праймирующими редакторами согласно настоящему изобретению.

Описание	Последовательность
LfCas9 *Lactobacillus fermentum* дикого типа GenBank: *SNX31424.1* 1	MKEYHIGLDIGTSSIGWAVTDSQFKLMRIKGKTAIGVRLFEEGKTAAERRT FRTTRRRLKRRKWRLHYLDEIFAPHLQEVDENFLRRLKQSNIHPEDPTKNQ AFIGKLLFPDLLKKNERGYPTLIKMRDELPVEQRAHYPVMNIYKLREAMI NEDRQFDLREVYLAVHHIVKYRGHFLNNASVDKFKVGRIDFDKSFNVLNE AYEELQNGEGSFTIEPSKVEKIGQLLLDTKMRKLDRQKAVAKLLEVKVAD KEETKRNKQIATAMSKLVLGYKADFATVAMANGNEWKIDLSSETSEDEIE KFREELSDAQNDILTEITSLFSQIMLNEIVPNGMSISESMMDRYWTHERQL AEVKEYLATQPASARKEFDQVYNKYIGQAPKERGFDLEKGLKKILSKKEN WKEIDELLKAGDFLPKQRTSANGVIPHQMHQQELDRIIEKQAKYYPWLAT ENPATGERDRHQAKYELDQLVSFRIPYYVGPLVTPEVQKATSGAKFAWA KRKEDGEITPWNLWDKIDRAESAEAFIKRMTVKDTYLLNEDVLPANSLLY QKYNVLNELNNVRVNGRRLSVGIKQDIYTELFKKKKTVKASDVASLVMA KTRGVNKPSVEGLSDPKKFNSNLATYLDLKSIVGDKVDDNRYQTDLENIIE WRSVFEDGEIFADKLTEVEWLTDEQRSALVKKRYKGWGRLSKKLLTGIV DENGQRIIDLMWNTDQNFKEIVDQPVFKEQIDQLNQKAITNDGMTLRERV ESVLDDAYTSPQNKKAIWQVVRVVEDIVKAVGNAPKSISIEFARNEGNKG EITRSRRTQLQKLFEDQAHELVKDTSLTEELEKAPDLSDRYYFYFTQGGKD MYTGDPINFDEISTKYDIDHILPQSFVKDNSLDNRVLTSRKENNKKSDQVP AKLYAAKMKPYWNQLLKQGLITQRKFENLTKDVDQNIKYRSLGFVKRQL VETRQVIKLTANILGSMYQEAGTEIIETRAGLTKQLREEFDLPKVREVNDY HHAVDAYLTTFAGQYLNRRYPKLRSFFVYGEYMKFKHGSDLKLRNFNFF HELMEGDKSQGKVVDQQTGELITTRDEVAKSFDRLLNMKYMLVSKEVH DRSDQLYGATIVTAKESGKLTSPIEIKKNRLVDLYGAYTNGTSAFMTIIKFT GNKPKYKVIGIPTTSAASLKRAGKPGSESYNQELHRIIKSNPKVKKGFEIVV PHVSYGQLIVDGDCKFTLASPTVQHPATQLVLSKKSLETISSGYKILKDKP AIANERLIRVFDEVVGQMNRYFTIFDQRSNRQKVADARDKFLSLPTESKY EGAKKVQVGKTEVITNLLMGLHANATQGDLKVLGLATFGFFQSTTGLSLS EDTMIVYQSPTGLFERRICLKDI (SEQ ID NO: 26)
SaCas9 *Staphylococcus aureus* дикого типа GenBank: AYD60528.1	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINAS GVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS

Описание	Последовательность
	RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVT VKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYG DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPL IETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPK RNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 27)
SaCas9 *Staphylococcus aureus*	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSK RGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKL SEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYV AELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDT YIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYA YNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIA KEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIA KILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDE LWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSI KVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRT TGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSV SFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKG KGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFR VNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIF KEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDF KDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDK LKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNY LTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRF

Описание	Последовательность
	DVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFI ASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPP HIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKK (SEQ ID NO: 28)
StCas9 *Streptococcus thermophilus* UniProtKB/Swiss-Prot: G3ECR1.2 Дикого типа	MLFNKCIIISINLDFSNKEKCMTKPYSIGLDIGTNSVGWAVITDNYKVPSKK MKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILY LQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKVYHDEF PTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQ KNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKN SGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDY SDVFLKAKKLYDAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKE YIRNISLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKNLLAEFEGAD YFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKE RIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEA FINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLD SKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTY HDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVL KKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHD DALSFKKKIQKAQIIGDEDKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVK VMGGRKPESIVVEMARENQYTNQGKSNSQQRLKRLEKSLKELGSKILKEN IPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDRLSNYDIDHIIPQ AFLKDNSIDNKVLVSSASNRGKSDDFPSLEVVKKRKTFWYQLLKSKLISQ RKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKD ENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVIAS ALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLA DGRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNH GLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSFA VLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIE LPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRI SNTINENHRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAFQSW QNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSL LKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO: 29)
LcCas9 *Lactobacillus crispatus* NCBI Эталонная последовательно	MKIKNYNLALTPSTSAVGHVEVDDDLNILEPVHHQKAIGVAKFGEGETAE ARRLARSARRTTKRRANRINHYFNEIMKPEIDKVDPLMFDRIKQAGLSPLD ERKEFRTVIFDRPNIASYYHNQFPTIWHLQKYLMITDEKADIRLIYWALHS LLKHRGHFFNTTPMSQFKPGKLNLKDDMLALDDYNDLEGLSFAVANSPEI EKVIKDRSMHKKEKIAELKKLIVNDVPDKDLAKRNNKIITQIVNAIMGNSF

Описание	Последовательность
сть: WP_133478044.1 Дикого типа	HLNFIFDMDLDKLTSKAWSFKLDDPELDTKFDAISGSMTDNQIGIFETLQKI YSAISLLDILNGSSNVVDAKNALYDKHKRDLNLYFKFLNTLPDEIAKTLKA GYTLYIGNRKKDLLAARKLLKVNVAKNFSQDDFYKLINKELKSIDKQGLQ TRFSEKVGELVAQNNFLPVQRSSDNVFIPYQLNAITFNKILENQGKYYDFL VKPNPAKKDRKNAPYELSQLMQFTIPYYVGPLVTPEEQVKSGIPKTSRFA WMVRKDNGAITPWNFYDKVDIEATADKFIKRSIAKDSYLLSELVLPKHSL LYEKYEVFNELSNVSLDGKKLSGGVKQILFNEVFKKTNKVNTSRILKALA KHNIPGSKITGLSNPEEFTSSLQTYNAWKKYFPNQIDNFAYQQDLEKMIEW STVFEDHKILAKKLDEIEWLDDDQKKFVANTRLRGWGRLSKRLLTGLKD NYGKSIMQRLETTKANFQQIVYKPEFREQIDKISQAAAKNQSLEDILANSY TSPSNRKAIRKTMSVVDEYIKLNHGKEPDKIFLMFQRSEQEKGKQTEARS KQLNRILSQLKADKSANKLFSKQLADEFSNAIKKSKYKLNDKQYFYFQQL GRDALTGEVIDYDELYKYTVLHIIPRSKLTDDSQNNKVLTKYKIVDGSVAL KFGNSYSDALGMPIKAFWTELNRLKLIPKGKLLNLTTDFSTLNKYQRDGYI ARQLVETQQIVKLLATIMQSRFKHTKIIEVRNSQVANIRYQFDYFRIKNLNE YYRGFDAYLAAVVGTYLYKVYPKARRLFVYGQYLKPKKTNQENQDMHL DSEKKSQGFNFLWNLLYGKQDQIFVNGTDVIAFNRKDLITKMNTVYNYK SQKISLAIDYHNGAMFKATLFPRNDRDTAKTRKLIPKKKDYDTDIYGGYT SNVDGYMLLAEIIKRDGNKQYGFYGVPSRLVSELDTLKKTRYTEYEEKLK EIIKPELGVDLKKIKKIKILKNKVPFNQVIIDKGSKFFITSTSYRWNYRQLILS AESQQTLMDLVVDPDFSNHKARKDARKNADERLIKVYEEILYQVKNYMP MFVELHRCYEKLVDAQKTFKSLKISDKAMVLNQILILLHSNATSPVLEKLG YHTRFTLGKKHNLISENAVLVTQSITGLKENHVSIKQML (SEQ ID NO: 30)
PdCas9 *Pedicoccus damnosus* NCBI Эталонная последовательно сть: WP_062913273.1 Дикого типа	MTNEKYSIGLDIGTSSIGFAVVNDNNRVIRVKGKNAIGVRLFDEGKAAAD RRSFRTTRRSFRTTRRRLSRRRWRLKLLREIFDAYITPVDEAFFIRLKESNLS PKDSKKQYSGDILFNDRSDKDFYEKYPTIYHLRNALMTEHRKFDVREIYL AIHHIMKFRGHFLNATPANNFKVGRLNLEEKFEELNDIYQRVFPDESIEFRT DNLEQIKEVLLDNKRSRADRQRTLVSDIYQSSEDKDIEKRNKAVATEILKA SLGNKAKLNVITNVEVDKEAAKEWSITFDSESIDDDLAKIEGQMTDDGHEI IEVLRSLYSGITLSAIVPENHTLSQSMVAKYDLHKDHLKLFKKLINGMTDT KKAKNLRAAYDGYIDGVKGKVLPQEDFYKQVQVNLDDSAEANEIQTYID QDIFMPKQRTKANGSIPHQLQQQELDQIIENQKAYYPWLAELNPNPDKKR QQLAKYKLDELVTFRVPYYVGPMITAKDQKNQSGAEFAWMIRKEPGNIT PWNFDQKVDRMATANQFIKRMTTTDTYLLGEDVLPAQSLLYQKFEVLNE LNKIRIDHKPISIEQKQQIFNDLFKQFKNVTIKHLQDYLVSQGQYSKRPLIE GLADEKRFNSSLSTYSDLCGIFGAKLVEENDRQEDLEKIIEWSTIFEDKKIY RAKLNDLTWLTDDQKEKLATKRYQGWGRLSRKLLVGLKNSEHRNIMDIL

Описание	Последовательность
	WITNENFMQIQAEPDFAKLVTDANKGMLEKTDSQDVINDLYTSPQNKKAI RQILLVVHDIQNAMHGQAPAKIHVEFARGEERNPRRSVQRQRQVEAAYE KVSNELVSAKVRQEFKEAINNKRDFKDRLFLYFMQGGIDIYTGKQLNIDQ LSSYQIDHILPQAFVKDDSLTNRVLTNENQVKADSVPIDIFGKKMLSVWGR MKDQGLISKGKYRNLTMNPENISAHTENGFINRQLVETRQVIKLAVNILAD EYGDSTQIISVKADLSHQMREDFELLKNRDVNDYHHAFDAYLAAFIGNYL LKRYPKLESYFVYGDFKKFTQKETKMRRFNFIYDLKHCDQVVNKETGEIL WTKDEDIKYIRHLFAYKKILVSHEVREKRGALYNQTIYKAKDDKGSGQES KKLIRIKDDKETKIYGGYSGKSLAYMTIVQITKKNKVSYRVIGIPTLALARL NKLENDSTENNGELYKIIKPQFTHYKVDKKNGEIIETTDDFKIVVSKVRFQ QLIDDAGQFFMLASDTYKNNAQQLVISNNALKAINNTNITDCPRDDLERL DNLRLDSAFDEIVKKMDKYFSAYDANNFREKIRNSNLIFYQLPVEDQWEN NKITELGKRTVLTRILQGLHANATTTDMSIFKIKTPFGQLRQRSGISLSENA QLIYQSPTGLFERRVQLNKIK (SEQ ID NO: 31)
FnCas9 *Fusobaterium nucleatum* NCBI Эталонная последовательность: WP_060798984.1	MKKQKFSDYYLGFDIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFEEA KTAAERRVQRNSRRRLKRRKWRLNLLEEIFSNEILKIDSNFFRRLKESSLW LEDKSSKEKFTLFNDDNYKDYDFYKQYPTIFHLRNELIKNPEKKDIRLVYL AIHSIFKSRGHFLFEGQNLKEIKNFETLYNNLIAFLEDNGINKIIDKNNIEKL EKIVCDSKKGLKDKEKEFKEIFNSDKQLVAIFKLSVGSSVSLNDLFDTDEY KKGEVEKEKISFREQIYEDDKPIYYSILGEKIELLDIAKTFYDFMVLNNILA DSQYISEAKVKLYEEHKKDLKNLKYIIRKYNKGNYDKLFKDKNENNYSA YIGLNKEKSKKEVIEKSRLKIDDLIKNIKGYLPKVEEIEEKDKAIFNKILNKI ELKTILPKQRISDNGTLPYQIHEAELEKILENQSKYYDFLNYEENGIITKDKL LMTFKFRIPYYVGPLNSYHKDKGGNSWIVRKEEGKILPWNFEQKVDIEKS AEEFIKRMTNKCTYLNGEDVIPKDTFLYSEYVILNELNKVQVNDEFLNEEN KRKIIDELFKENKKVSEKKFKEYLLVKQIVDGTIELKGVKDSFNSNYISYIR FKDIFGEKLNLDIYKEISEKSILWKCLYGDDKKIFEKKIKNEYGDILTKDEIK KINTFKFNNWGRLSEKLLTGIEFINLETGECYSSVMDALRRTNYNLMELLS SKFTLQESINNENKEMNEASYRDLIEESYVSPSLKRAIFQTLKIYEEIRKITG RVPKKVFIEMARGGDESMKNKKIPARQEQLKKLYDSCGNDIANFSIDIKE MKNSLISYDNNSLRQKKLYLYYLQFGKCMYTGREIDLDRLLQNNDTYDID HIYPRSKVIKDDSFDNLVLVLKNENAEKSNEYPVKKEIQEKMKSFWRFLK EKNFISDEKYKRLTGKDDFELRGFMARQLVNVRQTTKEVGKILQQIEPEIK IVYSKAEIASSFREMFDFIKVRELNDTHHAKDAYLNIVAGNVYNTKFTEKP YRYLQEIKENYDVKKIYNYDIKNAWDKENSLEIVKKNMEKNTVNITRFIK EKKGQLFDLNPIKKGETSNEIISIKPKVYNGKDDKLNEKYGYYKSLNPAYF LYVEHKEKNKRIKSFERVNLVDVNNIKDEKSLVKYLIENKKLVEPRVIKK

Описание	Последовательность
	VYKRQVILINDYPYSIVTLDSNKLMDFENLKPLFLENKYEKILKNVIKFLED NQGKSEENYKFIYLKKKDRYEKNETLESVKDRYNLEFNEMYDKFLEKLD SKDYKNYMNNKKYQELLDVKEKFIKLNLFDKAFTLKSFLDLFNRKTMAD FSKVGLTKYLGKIQKISSNVLSKNELYLLEESVTGLFVKKIKL (SEQ ID NO: 32)
EcCas9 *Enterococcus cecorum* NCBI Эталонная последовательность: WP_047338501.1 Дикого типа	RRKQRIQILQELLGEEVLKTDPGFFHRMKESRYVVEDKRTLDGKQVELPY ALFVDKDYTDKEYYKQFPTINHLIVYLMTTSDTPDIRLVYLALHYYMKNR GNFLHSGDINNVKDINDILEQLDNVLETFLDGWNLKLKSYVEDIKNIYNRD LGRGERKKAFVNTLGAKTKAEKAFCSLISGGSTNLAELFDDSSLKEIETPKI EFASSSLEDKIDGIQEALEDRFAVIEAAKRLYDWKTLTDILGDSSSLAEARV NSYQMHHEQLLELKSLVKEYLDRKVFQEVFVSLNVANNYPAYIGHTKIN GKKKELEVKRTKRNDFYSYVKKQVIEPIKKKVSDEAVLTKLSEIESLIEVD KYLPLQVNSDNGVIPYQVKLNELTRIFDNLENRIPVLRENRDKIIKTFKFRIP YYVGSLNGVVKNGKCTNWMVRKEEGKIYPWNFEDKVDLEASAEQFIRR MTNKCTYLVNEDVLPKYSLLYSKYLVLSELNNLRIDGRPLDVKIKQDIYE NVFKKNRKVTLKKIKKYLLKEGIITDDDELSGLADDVKSSLTAYRDFKEK LGHLDLSEAQMENIILNITLFGDDKKLLKKRLAALYPFIDDKSLNRIATLNY RDWGRLSERFLSGITSVDQETGELRTIIQCMYETQANLMQLLAEPYHFVEA IEKENPKVDLESISYRIVNDLYVSPAVKRQIWQTLLVIKDIKQVMKHDPERI FIEMAREKQESKKTKSRKQVLSEVYKKAKEYEHLFEKLNSLTEEQLRSKKI YLYFTQLGKCMYSGEPIDFENLVSANSNYDIDHIYPQSKTIDDSFNNIVLVK KSLNAYKSNHYPIDKNIRDNEKVKTLWNTLVSKGLITKEKYERLIRSTPFS DEELAGFIARQLVETRQSTKAVAEILSNWFPESEIVYSKAKNVSNFRQDFEI LKVRELNDCHHAHDAYLNIVVGNAYHTKFTNSPYRFIKNKANQEYNLRK LLQKVNKIESNGVVAWVGQSENNPGTIATVKKVIRRNTVLISRMVKEVDG QLFDLTLMKKGKGQVPIKSSDERLTDISKYGGYNKATGAYFTFVKSKKRG KVVRSFEYVPLHLSKQFENNNELLKEYIEKDRGLTDVEILIPKVLINSLFRY NGSLVRITGRGDTRLLLVHEQPLYVSNSFVQQLKSVSSYKLKKSENDNAK LTKTATEKLSNIDELYDGLLRKLDLPIYSYWFSSIKEYLVESRTKYIKLSIEE KALVIFEILHLFQSDAQVPNLKILGLSTKPSRIRIQKNLKDTDKMSIIHQSPS GIFEHEIELTSL (SEQ ID NO: 33)
AhCas9 *Anaerostipes hadrus* NCBI Эталонная последовательность:	MQNGFLGITVSSEQVGWAVTNPKYELERASRKDLWGVRLFDKAETAEDR RMFRTNRRLNQRKKNRIHYLRDIFHEEVNQKDPNFFQQLDESNFCEDDRT VEFNFDTNLYKNQFPTVYHLRKYLMETKDKPDIRLVYLAFSKFMKNRGH FLYKGNLGEVMDFENSMKGFCESLEKFNIDFPTLSDEQVKEVRDILCDHKI AKTVKKKNIITITKVKSKTAKAWIGLFCGCSVPVKVLFQDIDEEIVTDPEKI SFEDASYDDYIANIEKGVGIYYEAIVSAKMLFDWSILNEILGDHQLLSDAM

Описание	Последовательность
WP_044924278.1 Дикого типа	IAEYNKHHDDLKRLQKIIKGTGSRELYQDIFINDVSGNYVCYVGHAKTMS SADQKQFYTFLKNRLKNVNGISSEDAEWIDTEIKNGTLLPKQTKRDNSVIP HQLQLREFELILDNMQEMYPFLKENREKLLKIFNFVIPYYVGPLKGVVRK GESTNWMVPKKDGVIHPWNFDEMVDKEASAECFISRMTGNCSYLFNEKV LPKNSLLYETFEVLNELNPLKINGEPISVELKQRIYEQLFLTGKKVTKKSLT KYLIKNGYDKDIELSGIDNEFHSNLKSHIDFEDYDNLSDEEVEQIILRITVFE DKQLLKDYLNREFVKLSEDERKQICSLSYKGWGNLSEMLLNGITVTDSNG VEVSVMDMLWNTNLNLMQILSKKYGYKAEIEHYNKEHEKTIYNREDLM DYLNIPPAQRRKVNQLITIVKSLKKTYGVPNKIFFKISREHQDDPKRTSSRK EQLKYLYKSLKSEDEKHLMKELDELNDHELSNDKVYLYFLQKGRCIYSG KKLNLSRLRKSNYQNDIDYIYPLSAVNDRSMNNKVLTGIQENRADKYTYF PVDSEIQKKMKGFWMELVLQGFMTKEKYFRLSRENDFSKSELVSFIEREIS DNQQSGRMIASVLQYYFPESKIVFVKEKLISSFKRDFHLISSYGHNHLQAA KDAYITIVVGNVYHTKFTMDPAIYFKNHKRKDYDLNRLFLENISRDGQIA WESGPYGSIQTVRKEYAQNHIAVTKRVVEVKGGLFKQMPLKKGHGEYPL KTNDPRFGNIAQYGGYTNVTGSYFVLVESMEKGKKRISLEYVPVYLHERL EDDPGHKLLKEYLVDHRKLNHPKILLAKVRKNSLLKIDGFYYRLNGRSGN ALILTNAVELIMDDWQTKTANKISGYMKRRAIDKKARVYQNEFHIQELEQ LYDFYLDKLKNGVYKNRKNNQAELIHNEKEQFMELKTEDQCVLLTEIKK LFVCSPMQADLTLIGGSKHTGMIAMSSNVTKADFAVIAEDPLGLRNKVIY SHKGEK (SEQ ID NO: 34)
KvCas9 *Kandleria vitulina* NCBI Эталонная последовательность: WP_031589969.1 Дикого типа	MSQNNNKIYNIGLDIGDASVGWAVVDEHYNLLKRHGKHMWGSRLFTQA NTAVERRSSRSTRRRYNKRRERIRLLREIMEDMVLDVDPTFFIRLANVSFL DQEDKKDYLKENYHSNYNLFIDKDFNDKTYYDKYPTIYHLRKHLCESKE KEDPRLIYLALHHIVKYRGNFLYEGQKFSMDVSNIEDKMIDVLRQFNEINL FEYVEDRKKIDEVLNVLKEPLSKKHKAEKAFALFDTTKDNKAAYKELCA ALAGNKFNVTKMLKEAELHDEDEKDISFKFSDATFDDAFVEKQPLLGDCV EFIDLLHDIYSWVELQNILGSAHTSEPSISAAMIQRYEDHKNDLKLLKDVIR KYLPKKYFEVFRDEKSKKNNYCNYINHPSKTPVDEFYKYIKKLIEKIDDPD VKTILNKIELESFMLKQNSRTNGAVPYQMQLDELNKILENQSVYYSDLKD NEDKIRSILTFRIPYYFGPLNITKDRQFDWIIKKEGKENERILPWNANEIVDV DKTADEFIKRMRNFCTYFPDEPVMAKNSLTVSKYEVLNEINKLRINDHLIK RDMKDKMLHTLFMDHKSISANAMKKWLVKNQYFSNTDDIKIEGFQKEN ACSTSLTPWIDFTKIFGKINESNYDFIEKIIYDVTVFEDKKILRRRLKKEYDL DEEKIKKILKLKYSGWSRLSKKLLSGIKTKYKDSTRTPETVLEVMERTNM NLMQVINDEKLGFKKTIDDANSTSVSGKFSYAEVQELAGSPAIKRGIWQA LLIVDEIKKIMKHEPAHVYIEFARNEDEKERKDSFVNQMLKLYKDYDFED

Описание	Последовательность
	ETEKEANKHLKGEDAKSKIRSERLKLYYTQMGKCMYTGKSLDIDRLDTY QVDHIVPQSLLKDDSIDNKVLVLSSENQRKLDDLVIPSSIRNKMYGFWEKL FNNKIISPKKFYSLIKTEFNEKDQERFINRQIVETRQITKHVAQIIDNHYENT KVVTVRADLSHQFRERYHIYKNRDINDFHHAHDAYIATILGTYIGHRFESL DAKYIYGEYKRIFRNQKNKGKEMKKNNDGFILNSMRNIYADKDTGEIVW DPNYIDRIKKCFYYKDCFVTKKLEENNGTFFNVTVLPNDTNSDKDNTLAT VPVNKYRSNVNKYGGFSGVNSFIVAIKGKKKKGKKVIEVNKLTGIPLMYK NADEEIKINYLKQAEDLEEVQIGKEILKNQLIEKDGGLYYIVAPTEIINAKQ LILNESQTKLVCEIYKAMKYKNYDNLDSEKIIDLYRLLINKMELYYPEYRK QLVKKFEDRYEQLKVISIEEKCNIIKQILATLHCNSSIGKIMYSDFKISTTIGR LNGRTISLDDISFIAESPTGMYSKKYKL (SEQ ID NO: 35)
EfCas9 *Enterococcus faecalis* NCBI Эталонная последовательность: WP_016631044.1 Дикого типа	MRLFEEGHTAEDRRLKRTARRRISRRRNRLRYLQAFFEEAMTDLDENFFA RLQESFLVPEDKKWHRHPIFAKLEDEVAYHETYPTIYHLRKKLADSSEQA DLRLIYLALAHIVKYRGHFLIEGKLSTENTSVKDQFQQFMVIYNQTFVNGE SRLVSAPLPESVLIEEELTEKASRTKKSEKVLQQFPQEKANGLFGQFLKLM VGNKADFKKVFGLEEEAKITYASESYEEDLEGILAKVGDEYSDVFLAAKN VYDAVELSTILADSDKKSHAKLSSSMIVRFTEHQEDLKKFKRFIRENCPDE YDNLFKNEQKDGYAGYIAHAGKVSQLKFYQYVKKIIQDIAGAEYFLEKIA QENFLRKQRTFDNGVIPHQIHLAELQAIIHRQAAYYPFLKENQEKIEQLVTF RIPYYVGPLSKGDASTFAWLKRQSEEPIRPWNLQETVDLDQSATAFIERMT NFDTYLPSEKVLPKHSLLYEKFMVFNELTKISYTDDRGIKANFSGKEKEKI FDYLFKTRRKVKKKDIIQFYRNEYNTEIVTLSGLEEDQFNASFSTYQDLLK CGLTRAELDHPDNAEKLEDIIKILTIFEDRQRIRTQLSTFKGQFSAEVLKKLE RKHYTGWGRLSKKLINGIYDKESGKTILDYLVKDDGVSKHYNRNFMQLI NDSQLSFKNAIQKAQSSEHEETLSETVNELAGSPAIKKGIYQSLKIVDELVA IMGYAPKRIVVEMARENQTTSTGKRRSIQRLKIVEKAMAEIGSNLLKEQPT TNEQLRDTRLFLYYMQNGKDMYTGDELSLHRLSHYDIDHIIPQSFMKDDS LDNLVLVGSTENRGKSDDVPSKEVVKDMKAYWEKLYAAGLISQRKFQRL TKGEQGGLTLEDKAHFIQRQLVETRQITKNVAGILDQRYNAKSKEKKVQII TLKASLTSQFRSIFGLYKVREVNDYHHGQDAYLNCVVATTLLKVYPNLAP EFVYGEYPKFQTFKENKATAKAIIYTNLLRFFTEDEPRFTKDGEILWSNSYL KTIKKELNYHQMNIVKKVEVQKGGFSKESIKPKGPSNKLIPVKNGLDPQK YGGFDSPVVAYTVLFTHEKGKKPLIKQEILGITIMEKTRFEQNPILFLEEKG FLRPRVLMKLPKYTLYEFPEGRRRLLASAKEAQKGNQMVLPEHLLTLLYH AKQCLLPNQSESLAYVEQHQPEFQEILERVVDFAEVHTLAKSKVQQIVKLF EANQTADVKEIAASFIQLMQFNAMGAPSTFKFFQKDIERARYTSIKEIFDA

Описание	Последовательность
	TIIYQSPTGLYETRRKVVD (SEQ ID NO: 36)
Staphylococcus aureus Cas9	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRG ARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEE EFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYID LLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEI LVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILT IYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDN SFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRI SKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNN LDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEW KKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDY KYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKK LINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTK YSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVY LDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFY NNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKT IASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 37)
Geobacillus thermodenitrificans Cas9	MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLA RSARRRLRRRKHRLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEA LDRKLNNDELARILLHLAKRRGFRSNRKSERTNKENSTMLKHIEENQSILS SYRTVAEMVVKDPKFSLHKRNKEDNYTNTVARDDLEREIKLIFAKQREYG NIVCTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFEPKEKRAPKATYTFQ SFTVWEHINKLRLVSPGGIRALTDDERRLIYKQAFHKNKITFHDVRTLLNL PDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVYGKGAAKSF RPIDFDTFGYALTMFKDDTDIRSYLRNEYEQNGKRMENLADKVYDEELIE ELLNLSFSKFGHLSLKALRNILPYMEQGEVYSTACERAGYTFTGPKKKQK TVLLPNIPPIANPVVMRALTQARKVVNAIIKKYGSPVSIHIELARELSQSFDE RRKMQKEQEGNRKKNETAIRQLVEYGLTLNPTGLDIVKFKLWSEQNGKC AYSLQPIEIERLLEPGYTEVDHVIPYSRSLDDSYTNKVLVLTKENREKGNR TPAEYLGLGSERWQQFETFVLTNKQFSKKKRDRLLRLHYDENEENEFKNR NLNDTRYISRFLANFIREHLKFADSDDKQKVYTVNGRITAHLRSRWNFNK

Описание	Последовательность
	NREESNLHHAVDAAIVACTTPSDIARVTAFYQRREQNKELSKKTDPQFPQ PWPHFADELQARLSKNPKESIKALNLGNYDNEKLESLQPVFVSRMPKRSIT GAAHQETLRRYIGIDERSGKIQTVVKKKLSEIQLDKTGHFPMYGKESDPRT YEAIRQRLLEHNNDPKKAFQEPLYKPKKNGELGPIIRTIKIIDTTNQVIPLND GKTVAYNSNIVRVDVFEKDGKYYCVPIYTIDMMKGILPNKAIEPNKPYSE WKEMTEDYTFRFSLYPNDLIRIEFPREKTIKTAVGEEIKIKDLFAYYQTIDSS NGGLSLVSHDNNFSLRSIGSRTLKRFEKYQVDVLGNIYKVRGEKRVGVAS SSHSKAGETIRPL (SEQ ID NO: 38)
ScCas9 *S. canis* 1375 AA 159.2 кДа	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLMG ALLFDSGETAEATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFFQ RLEESFLVEEDKKNERHPIFGNLADEVAYHRNYPTIYHLRKKLADSPEKA DLRLIYLALAHIIKFRGHFLIEGKLNAENSDVAKLFYQLIQTYNQLFEESPL DEIEVDAKGILSARLSKSKRLEKLIAVFPNEKKNGLFGNIIALALGLTPNFK SNFDLTEDAKLQLSKDTYDDDLDELLGQIGDQYADLFSAAKNLSDAILLS DILRSNSEVTKAPLSASMVKRYDEHHQDLALLKTLVRQQFPEKYAEIFKD DTKNGYAGYVGIGIKHRKRTTKLATQEEFYKFIKPILEKMDGAEELLAKL NRDDLLRKQRTFDNGSIPHQIHLKELHAILRRQEEFYPFLKENREKIEKILTF RIPYYVGPLARGNSRFAWLTRKSEEAITPWNFEEVVDKGASAQSFIERMT NFDEQLPNKKVLPKHSLLYEYFTVYNELTKVKYVTERMRKPEFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIIGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM KQLKRRHYTGWGRLSRKMINGIRDKQSGKTILDFLKSDGFSNRNFMQLIH DDSLTFKEEIEKAQVSGQGDSLHEQIADLAGSPAIKKGILQTVKIVDELVK VMGHKPENIVIEMARENQTTTKGLQQSRERKKRIEEGIKELESQILKENPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDS IDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSEADKAGFIKRQLVETRQITKHVARILDSRMNTKRDKNDKPI REVKVITLKSKLVSDFRKDFQLYKVRDINNYHHAHDAYLNAVVGTALIK KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKRFFYSNIMNFFKTE VKLANGEIRKRPLIETNGETGEVVWNKEKDFATVRKVLAMPQVNIVKKT EVQTGGFSKESILSKRESAKLIPRKKGWDTRKYGGFGSPTVAYSILVVAKV EKGKAKKLKSVKVLVGITIMEKGSYEKDPIGFLEAKGYKDIKKELIFKLPK YSLFELENGRRRMLASATELQKANELVLPQHLVRLLYYTQNISATTGSNN LGYIEQHREEFKEIFEKIIDFSEKYILKNKVNSNLKSSFDEQFAVSDSILLSNS FVSLLKYTSFGASGGFTFLDLDVKQGRLRYQTVTEVLDATLIYQSITGLYE TRTDLSQLGGD (SEQ ID NO: 39)

[0372] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут включать любую из вышеуказанных последовательностей ортологов Cas9 или любые их варианты, имеющие по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичность последовательности.

[0373] NapDNAbp может включать любые подходящие гомологи и/или ортологи или встречающиеся в природе ферменты, такие как Cas9. Гомологи и/или ортологи Cas9 были описаны у различных видов, включая без ограничения *S. pyogenes* и *S. thermophilus.* Предпочтительно фрагмент Cas сконфигурирован (например, мутагенизирован, рекомбинантно сконструирован или иным образом получен из природного источника) как никаза, т.е. способен расщеплять только одну цепь двухцепочечной ДНК-мишени. Дополнительные подходящие нуклеазы и последовательности Cas9 будут очевидны специалистам в данной области техники на основе настоящего раскрытия, и такие нуклеазы и последовательности Cas9 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, полные содержания которой включено в настоящий документ посредством ссылки. Согласно некоторым вариантам осуществления нуклеаза Cas9 имеет неактивный (например, инактивированный) домен расщепления ДНК, то есть Cas9 является никазой. Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 80% идентична аминокислотной последовательности белка Cas9, как представлено любым из вариантов, приведенных в Таблице 3. Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 92%, 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% идентична аминокислотной последовательности белка Cas9, как представлено любым из ортологов Cas, приведенных в таблицах выше.

C. Мертвый вариант Cas9

[0374] Согласно определенным вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут включать мертвый Cas9, например, мертвый SpCas9, который не обладает нуклеазной активностью из-за одной или нескольких мутаций, которые инактивируют оба нуклеазных домена Cas9, а именно домен RuvC (который расщепляет непротоспейсерную цепь ДНК) и домен HNH (который расщепляет протоспейсерную цепь ДНК). Инактивация нуклеазы может быть обусловлена

одной или несколькими мутациями, которые приводят к одной или нескольким заменам и/или делециям в аминокислотной последовательности кодируемого белка или любых его вариантов, имеющих по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичность последовательности.

[0375] В контексте настоящего изобретения термин «dCas9» относится к Cas9 без нуклеазной активности или нуклеаза-мертвому Cas9, или его функциональному фрагменту, и охватывает любой встречающийся в природе dCas9 из любого организма, любой встречающийся в природе эквивалент dCas9 или его функциональный фрагмент, любой гомолог, ортолог или паралог dCas9 из любого организма, а также любой мутант или вариант dCas9, встречающийся в природе или сконструированный. Термин dCas9 не является ограничивающим и может упоминаться как «dCas9 или эквивалент». Иллюстративные белки dCas9 и способ получения белков dCas9 дополнительно описаны в настоящем документе и/или описаны в данной области техники и включены в настоящий документ посредством ссылки.

[0376] Согласно другим вариантам осуществления dCas9 соответствует, или содержит частично или полностью, аминокислотной последовательности Cas9, имеющей одну или несколько мутаций, которые инактивируют нуклеазную активность Cas9. Согласно другим вариантам осуществления предусмотрены варианты Cas9, имеющие мутации, отличные от D10A и H840A, которые могут приводить к полной или частичной инактивации эндогенной нуклеазной активности Cas9 (например, nCas9 или dCas9, соответственно). Такие мутации, например, включают другие аминокислотные замены в D10 и H840 или другие замены в нуклеазных доменах Cas9 (например, замены в нуклеазном субдомене HNH и/или субдомене RuvC1) со ссылкой на последовательность дикого типа, такую как Cas9 из *Streptococcus pyogenes* (эталонная последовательность NCBI: NC_017053.1). Согласно некоторым вариантам осуществления предусмотрены варианты или гомологи Cas9 (например, варианты Cas9 из *Streptococcus pyogenes* (NCBI эталонная последовательность: NC_017053.1 (SEQ ID NO: 20))), которые на по меньшей мере приблизительно 70% идентичны, по меньшей мере приблизительно 80% идентичны, по меньшей мере приблизительно 90% идентичны, по меньшей мере приблизительно 95% идентичны, по меньшей мере приблизительно 98% идентичны, по меньшей мере приблизительно 99% идентичны, по меньшей мере приблизительно 99.5% идентичны или по меньшей мере приблизительно 99.9% идентичны NCBI эталонной последовательности NC_017053.1. Согласно некоторым вариантам осуществления предусмотрены варианты dCas9 (например, варианты NCBI эталонной последовательности: NC_017053.1 (SEQ ID NO: 20)), имеющие аминокислотные последовательности, которые короче или длиннее, чем

NC_017053.1 (SEQ ID NO: 20) на приблизительно 5 аминокислот, на приблизительно 10 аминокислот, на приблизительно 15 аминокислот, на приблизительно 20 аминокислот, на приблизительно 25 аминокислот, на приблизительно 30 аминокислот, на приблизительно 40 аминокислот, на приблизительно 50 аминокислот, на приблизительно 75 аминокислот, на приблизительно 100 аминокислот или более.

[0377] Согласно одному варианту осуществления мертвый Cas9 может быть основан на канонической последовательности SpCas9 Q99ZW2 и может иметь следующую последовательность, которая содержит D10X и H810X, где X может представлять собой любую аминокислоту (замены подчеркнуты и выделены жирным шрифтом), или вариант может быть 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.

[0378] Согласно одному варианту осуществления мертвый Cas9 может быть основан на канонической последовательности SpCas9 Q99ZW2 и может иметь следующую последовательность, которая содержит замены D10A и H810A (подчеркнуты и выделены жирным шрифтом), или может представлять собой вариант SEQ ID NO: 41, имеющий по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичность последовательности.

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**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIP YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFE DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI	SEQ ID NO: 40

Описание	Последовательность	SEQ ID NO:
	VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**X**IVPQSFL KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI LDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREIN NYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
мертвый Cas9 или dCas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**A** и H810**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIP YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFE DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**A**IVPQSFL KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI LDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREIN NYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI	SEQ ID NO: 41

Описание	Последовательность	SEQ ID NO:
	AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	

D. Никазный вариант Cas9

[0379] Согласно одному варианту осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, содержат никазу Cas9. Термин «никаза Cas9» или «nCas9» относится к варианту Cas9, который способен вносить одноцепочечный разрыв в двухцепочечную молекулу ДНК-мишень. Согласно некоторым вариантам осуществления никаза Cas9 содержит только один функционирующий нуклеазный домен. Cas9 дикого типа (например, канонический SpCas9) содержит два отдельных нуклеазных домена, а именно, домен RuvC (который расщепляет непротоспейсерную цепь ДНК) и домен HNH (который расщепляет протоспейсерную цепь ДНК). Согласно одному варианту осуществления никаза Cas9 содержит мутацию в домене RuvC, которая инактивирует нуклеазную активность RuvC. Например, сообщалось о мутациях в аспартате (D) 10, гистидине (H) 983, аспартате (D) 986 или глутамате (E) 762 как мутациях с потерей функции нуклеазного домена RuvC и создании функциональной никазы Cas9 (например, Nishimasu et al., «Crystal structure of Cas9 in complex with guide РНК and target DNA,» *Cell* 156(5), 935–949, которая включена в настоящий документ посредством ссылки). Таким образом, мутации никазы в домене RuvC могут включать D10X, H983X, D986X или E762X, где X представляет собой любую аминокислоту, отличную от аминокислоты дикого типа. Согласно определенным вариантам осуществления никаза может представлять собой D10A, H983A, D986A, или E762A, или их комбинацию.

[0380] Согласно различным вариантам осуществления никаза Cas9 может иметь мутацию в нуклеазном домене RuvC и иметь одну из следующих аминокислотных последовательностей или ее 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.

Описание	Последовательность	SEQ ID NO:
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D10**X**, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGL**X**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 42
Никаза Cas9 *Streptococcus pyogenes*	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL	SEQ ID NO: 43

Описание	Последовательность	SEQ ID NO:
Q99ZW2 Cas9 с E762X, где X представляет собой любую альтернативную аминокислоту	IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIXMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с H983X, где X представляет собой любую	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV	SEQ ID NO: 44

Описание	Последовательность	SEQ ID NO:
альтернативную аминокислоту	RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHXAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с D986X, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL	SEQ ID NO: 45

Описание	Последовательность	SEQ ID NO:
	PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAH**X**AYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c D10**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS	SEQ ID NO: 46

Описание	Последовательность	SEQ ID NO:
	GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c E762A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VI**A**MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI	SEQ ID NO: 47

Описание	Последовательность	SEQ ID NO:
	VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c H983A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYH**A**AHDAYLNAVVGTALIKKYPKLES	SEQ ID NO: 48

Описание	Последовательность	SEQ ID NO:
	EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c D986A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFL IEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEER LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAH**A**AYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS	SEQ ID NO: 49

Описание	Последовательность	SEQ ID NO:
	FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	

[0381] Согласно другому варианту осуществления никаза Cas9 содержит мутацию в домене HNH, которая инактивирует нуклеазную активность HNH. Например, сообщалось о мутациях в гистидине (H) 840 или аспарагине (R) 863 как о мутациях с потерей функции нуклеазного домена HNH и создании функциональной никазы Cas9 (например, Nishimasu et al., «Crystal structure of Cas9 in complex with guide РНК and target DNA,» *Cell* 156(5), 935–949, которая включена в настоящий документ посредством ссылки). Таким образом, мутации никазы в домене HNH могут включать H840X и R863X, где X представляет собой любую аминокислоту, отличную от аминокислоты дикого типа. Согласно определенным вариантам осуществления никаза может представлять собой H840A, или R863A или их комбинацию.

[0382] Согласно различным вариантам осуществления никаза Cas9 может иметь мутацию в нуклеазном домене HNH и иметь одну из следующих аминокислотных последовательностей или ее 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% идентичности последовательности.

Описание	Последовательность	SEQ ID NO:
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**X**, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF	SEQ ID NO: 50

Описание	Последовательность	SEQ ID NO:
	LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNS RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR DMYVDQELDINRLSDYDVD**X**IVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c H840**A**	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNS RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR	SEQ ID NO: 51

Описание	Последовательность	SEQ ID NO:
	DMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 с R863X, где X представляет собой любую альтернативную аминокислоту	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNS RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NXGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	SEQ ID NO: 52

Описание	Последовательность	SEQ ID NO:
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Никаза Cas9 *Streptococcus pyogenes* Q99ZW2 Cas9 c R863**A**	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNS RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK N**A**GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK	SEQ ID NO: 53

Описание	Последовательность	SEQ ID NO:
	GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	

[0383] Согласно некоторым вариантам осуществления N-концевой метионин удален из никазы Cas9 или из любого варианта, ортолога или эквивалента, раскрытого или охватываемого в настоящем документе. Например, никазы Cas9 без метионина включает следующие последовательности или 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.

Описание	Последовательность
Никаза Cas9 (Met минус) *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**X**, где X представляет собой любую альтернативную аминокислоту	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVD**X**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYW RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

	RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQS ITGLYETRIDLSQLGGD (SEQ ID NO: 54)
Никаза Cas9 (Met минус) *Streptococcus pyogenes* Q99ZW2 Cas9 с H840**A**	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKN**R**GKSDNVPSEEVVKKMKNYW RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQS ITGLYETRIDLSQLGGD (SEQ ID NO: 55)
Никаза Cas9 (Met минус) *Streptococcus pyogenes* Q99ZW2 Cas9 с R863X, где X представляет собой любую	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE

альтернативную аминокислоту	TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**X**GKSDNVPSEEVVKKMKNYW RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQS ITGLYETRIDLSQLGGD(SEQ ID NO: 56)
Никаза Cas9 (Met минус) *Streptococcus pyogenes* Q99ZW2 Cas9 c R863**A**	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVDHIVPQSFLKDDSIDNKVLTRSDKN**A**GKSDNVPSEEVVKKMKNYW RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF

	FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQS ITGLYETRIDLSQLGGD (SEQ ID NO: 57)

E. Другие варианты Cas9

[0384] Помимо мертвого Cas9 и никазных вариантов Cas9, белки Cas9 согласно настоящему изобретению могут также включать другие «варианты Cas9», на по меньшей мере приблизительно 70% идентичные, по меньшей мере приблизительно 80% идентичные, по меньшей мере приблизительно 90% идентичные, по меньшей мере приблизительно 95% идентичные, по меньшей мере приблизительно 96% идентичные, по меньшей мере приблизительно 97% идентичные, по меньшей мере приблизительно 98% идентичные, по меньшей мере приблизительно 99% идентичные, по меньшей мере приблизительно 99.5% идентичные или по меньшей мере приблизительно 99.9% идентичные любому эталонному белку Cas9, включая любой Cas9 дикого типа, или мутант Cas9 (например, мертвый Cas9 или никаза Cas9), или фрагмент Cas9, или круговой пермутант Cas9, или другой вариант Cas9 раскрытые в настоящем документе или известные в данной области техники. Согласно некоторым вариантам 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 эталонным Cas9. Согласно некоторым вариантам осуществления вариант Cas9 содержит фрагмент эталонного Cas9 (*например*, домен связывания гРНК или домен расщепления ДНК), так что фрагмент, на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере приблизительно 99.5% идентичен или по меньшей мере приблизительно 99.9% идентичен соответствующему фрагменту Cas9 дикого типа. Согласно некоторым вариантам осуществления фрагмент составляет по меньшей мере 30%, по меньшей мере 35%, по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, 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% идентичен, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или по меньшей мере 99.5% от длины аминокислот соответствующего Cas9 дикого типа (*например*, SEQ ID NO: 18).

[0385] Согласно некоторым вариантам осуществления согласно настоящему изобретению также могут использоваться фрагменты Cas9, которые сохраняют их функциональность, и которые представляют собой фрагменты любого из раскрытого в настоящем документе белка Cas9. Согласно некоторым вариантам осуществления фрагмент Cas9 составляет по меньшей мере 100 аминокислот в длину. Согласно некоторым вариантам осуществления фрагмент составляет 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 в длину.

[0386] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, раскрытые в настоящем документе, могут содержать один из вариантов Cas9, описанных далее, или их вариант Cas9, который на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере приблизительно 99.5% идентичен или по меньшей мере приблизительно 99.9% идентичен любому из эталонных вариантов Cas9.

F. Варианты Cas9 небольшого размера

[0387] Согласно некоторым вариантам осуществления мультилоскутные праймирующие редакторы, охватываемые настоящим изобретением, могут включать белок Cas9 с меньшей молекулярной массой, чем каноническая последовательность SpCas9. Согласно некоторым вариантам осуществления варианты Cas9 меньшего размера могут способствовать доставке в клетки, например, с помощью вектора экспрессии, наночастиц или других средств доставки. Согласно определенным вариантам осуществления варианты Cas9 меньшего размера могут включать ферменты, отнесенные к категории ферментов типа II систем CRISPR-Cas класса 2. Согласно некоторым вариантам осуществления варианты Cas9 меньшего размера могут включать ферменты, отнесенные к категории ферментов типа V систем CRISPR-Cas класса 2. Согласно другим вариантам осуществления Cas9 меньшего размера могут включать ферменты, отнесенные к категории ферментов типа VI систем CRISPR-Cas класса 2.

[0388] Канонический белок SpCas9 имеет длину 1368 аминокислот и расчетную молекулярную массу 158 килодальтон. В контексте настоящего изобретения термин «вариант Cas9 небольшого размера» относится к любому варианту Cas9, встречающемуся в природе, сконструированному или полученному иным образом, который составляет менее по меньшей мере 1300 аминокислот, или по меньшей мере менее 1290 аминокислот, или менее 1280 аминокислот, или менее 1270 аминокислот, или менее 1260 аминокислот, или менее 1250 аминокислот, или менее 1240 аминокислот, или менее 1230 аминокислот, или менее 1220 аминокислот, или менее 1210 аминокислот, или менее 1200 аминокислот, или менее 1190 аминокислот, или менее 1180 аминокислот, или менее 1170 аминокислот, или менее 1160 аминокислот, или менее 1150 аминокислот, или менее 1140 аминокислот, или менее 1130 аминокислот, или менее 1120 аминокислот, или менее 1110 аминокислот, или менее 1100 аминокислот, или менее 1050 аминокислот, или менее 1000 аминокислот, или менее 950 аминокислот, или менее 900 аминокислот, или менее 850 аминокислот, или менее 800 аминокислот, или менее 750 аминокислот, или менее 700 аминокислот, или менее 650 аминокислот, или менее 600 аминокислот, или менее 550 аминокислот, или менее 500 аминокислот, но по меньшей мере более приблизительно 400 аминокислот, и сохраняет требуемые функции белка Cas9. Варианты Cas9 могут включать те, которые классифицируются как ферменты типа II, типа V или типа VI системы CRISPR-Cas класса 2.

[0389] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, раскрытые в настоящем документе, могут содержать один из вариантов Cas9 небольшого размера, описанных ниже, или их вариант Cas9, на по меньшей мере приблизительно 70% идентичный, по меньшей мере приблизительно 80% идентичный, по меньшей мере приблизительно 90% идентичный, по меньшей мере приблизительно 95% идентичный, по меньшей мере приблизительно 96% идентичный, по меньшей мере приблизительно 97% идентичный, по меньшей мере приблизительно 98% идентичный, по меньшей мере приблизительно 99% идентичный, по меньшей мере приблизительно 99.5% идентичный или по меньшей мере приблизительно 99.9% идентичный любому эталонному белку Cas9 небольшого размера.

Описание	Последовательность	SEQ ID NO:
SaCas9 *Staphylococcus aureus*	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANV ENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHS ELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVN EVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGE VRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDL	SEQ ID NO: 58

Описание	Последовательность	SEQ ID NO:
1053 AA 123 кДа	LETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEEL RSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIE NVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNL KVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTN LNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTND NQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFI QSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKR NRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLE AIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENS KKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKK EYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYF RVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAE DALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPE IETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLIN DTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPE KLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNY LTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKV VKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNS KCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNN DLLNRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKK YSTDILGNLYEVKSKKHPQIIKK	
NmeCas9 *N. meningitidis* 1083 AA 124.5 кДа	MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLG VRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRT RRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAALDRKL TPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGV AGNAHALQTGDFRTPAELALNKFEKESGHIRNQRSDYSHTF SRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRP ALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLN NLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLG LEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGL KDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEIL EALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYG DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGV VRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREK AAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEI NLGRLNEKGYVEIDAALPFSRTWDDSFNNKVLVLGSENQN KGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRIL	SEQ ID NO: 59

Описание	Последовательность	SEQ ID NO:
	LQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGK KRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVA CSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTH FPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAE KLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSAKR LDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARL EAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKT GVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQV AKGILPDRAVVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVI TKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGV KTALSFQKYQIDELGKEIRPCRLKKRPPVR	
CjCas9 *C. jejuni* 984 AA 114.9 кДа	MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGE SLALPRRLARSARKRLARRKARLNHLKHLIANEFKLNYED YQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFARVILH IAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGE YLYKEYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELK LIFKKQREFGFSFSKKFEEEVLSVAFYKRALKDFSHLVGNCS FFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGILYTKD DLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYF IEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKAL AKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPLMLEGKK YDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNPVVL RAIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKI EKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKE FCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMNKVL VFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQ KRILDKNYKDKEQKNFKDRNLNDTRYIARLVLNYTKDYLD FLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSALRHTWGF SAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSA ELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEIFVSKPE RKKPSGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKV NGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDFALKVLP NKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDM QEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNAN EKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK	SEQ ID NO: 60
GeoCas9 *G.*	MRYKIGLDIGITSVGWAVMNLDIPRIEDLGVRIFDRAENPQT GESLALPRRLARSARRRLRRRKHRLERIRRLVIREGILTKEE	SEQ ID NO: 61

Описание	Последовательность	SEQ ID NO:
stearothermophilus 1087 AA 127 кДа	LDKLFEEKHEIDVWQLRVEALDRKLNNDELARVLLHLAKR RGFKSNRKSERSNKENSTMLKHIEENRAILSSYRTVGEMIV KDPKFALHKRNKGENYTNTIARDDLEREIRLIFSKQREFGN MSCTEEFENEYITIWASQRPVASKDDIEKKVGFCTFEPKEKR APKATYTFQSFIAWEHINKLRLISPSGARGLTDEERRLLYEQ AFQKNKITYHDIRTLLHLPDDTYFKGIVYDRGESRKQNENI RFLELDAYHQIRKAVDKVYGKGKSSSFLPIDFDTFGYALTL FKDDADIHSYLRNEYEQNGKRMPNLANKVYDNELIEELLN LSFTKFGHLSLKALRSILPYMEQGEVYSSACERAGYTFTGP KKKQKTMLLPNIPPIANPVVMRALTQARKVVNAIIKKYGSP VSIHIELARDLSQTFDERRKTKKEQDENRKKNETAIRQLME YGLTLNPTGHDIVKFKLWSEQNGRCAYSLQPIEIERLLEPGY VEVDHVIPYSRSLDDSYTNKVLVLTRENREKGNRIPAEYLG VGTERWQQFETFVLTNKQFSKKKRDRLLRLHYDENEETEF KNRNLNDTRYISRFFANFIREHLKFAESDDKQKVYTVNGRV TAHLRSRWEFNKNREESDLHHAVDAVIVACTTPSDIAKVT AFYQRREQNKELAKKTEPHFPQPWPHFADELRARLSKHPK ESIKALNLGNYDDQKLESLQPVFVSRMPKRSVTGAAHQET LRRYVGIDERSGKIQTVVKTKLSEIKLDASGHFPMYGKESD PRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGEPGPVIRT VKIIDTKNQVIPLNDGKTVAYNSNIVRVDVFEKDGKYYCVP VYTMDIMKGILPNKAIEPNKPYSEWKEMTEDYTFRFSLYPN DLIRIELPREKTVKTAAGEEINVKDVFVYYKTIDSANGGLEL ISHDHRFSLRGVGSRTLKRFEKYQVDVLGNIYKVRGEKRV GLASSAHSKPGKTIRPLQSTRD	
LbaCas12a *L. bacterium* 1228 AA 143.9 кДа	MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDE KRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISLFR KKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFKKDII ETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENMFSEEA KSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEKI LNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEK IKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLSFYGE GYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSA GIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKA VVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKE IIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMK DLLDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILL	SEQ ID NO: 62

Описание	Последовательность	SEQ ID NO:
	KVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKD KETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGN YEKINYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYK NGTFKKGDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNF SETEKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVEEG KLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDENNHGQIR LSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTT LSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLL KHDDNPYVIGIDRGERNLLYIVVVDGKGNIVEQYSLNEIINN FNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGYI SQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQ KFEKMLIDKLNYMVDKKSNPCATGGALKGYQITNKFESFK SMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKK FISSFDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYS YGNRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQ QGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDV DFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIA RKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKH	
BhCas12b *B. hisashii* 1108 AA 130.4 кДа	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLI RQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQKCNS FTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLY PLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKW EEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWM EKSRNQSVRRLDKDMFIQALERFLSWESWNLKVKEEYEKV EKEYKTLEERIKEDIQALKALEQYEKERQEQLLRDTLNTNE YRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQRK HPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDK KKKDAKQQATFTLADPINHPLWVRFEERSGSNLNKYRILTE QLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSR QFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFD RDHLRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIH RDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVM SIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHR ASFNIKLPGETLVKSREVLRKAREDNLKLMNQKLNFLRNV LHFQQFEDITEREKRVTKWISRQENSDVPLVYQDELIQIREL MYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKSLSDGRK GLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRF	SEQ ID NO: 63

Описание	Последовательность	SEQ ID NO:
	AIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKW QAKNPACQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPR QVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKE KLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFIS LSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAY QVDGQTVYIPESKDQKQKIIEEFGEGYFILKDGVYEWVNAG KLKIKKGSSKQSSSELVDSDILKDSFDLASELKGEKLMLYR DPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDD SSKQSM	

G. Эквиваленты Cas9

[0390] Согласно некоторым вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут включать любой эквивалент Cas9. В контексте настоящего изобретения термин «эквивалент Cas9» является широким термином, который охватывает любой белок napDNAbp, который выполняет ту же функцию, что и Cas9 в мультилоскутных праймирующих редакторах согласно настоящему изобретению, несмотря на то, что его первичная аминокислотная последовательность и/или его трехмерная структура могут быть разными и/или не связанными с эволюционной точки зрения. Таким образом, в то время как эквиваленты Cas9 включают любой ортолог, гомолог, мутант или вариант Cas9, описанный или включенный в настоящее описание, которые являются эволюционно родственными, эквиваленты Cas9 также охватывают белки, которые могли эволюционировать в результате конвергентных эволюционных процессов с достижением той же или подобной функции, что и Cas9, но которые не обязательно имеют какое-либо сходство в отношении аминокислотной последовательности и/или трехмерной структуры. Мультилоскутные праймирующие редакторы, описанные в настоящем документе, охватывают любой эквивалент Cas9, который обеспечивает ту же или подобную функцию, что и Cas9, несмотря на то, что эквивалент Cas9 может быть основан на белке, возникшем в результате конвергентной эволюции. Например, если Cas9 относится к ферменту типа II системы CRISPR-Cas, эквивалент Cas9 может относиться к ферменту типа V или типа VI системы CRISPR-Cas.

[0391] Например, Cas12e (CasX) является эквивалентом Cas9, который, как сообщается, имеет ту же функцию, что и Cas9, но эволюционировал в результате конвергентной эволюции. Таким образом, 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, предполагается использовать с мультилоскутными праймирующими

редакторами, описанными в настоящем документе. Кроме того, любой вариант или модификация Cas12e (CasX) возможны и входят в объем настоящего изобретения.

[0392] Cas9 представляет собой бактериальный фермент, который эволюционировал у самых разных видов. Однако охватываемые настоящим изобретением эквиваленты Cas9 также могут быть получены из архей, которые составляют домен и царство одноклеточных прокариотических микробов, отличных от бактерий.

[0393] Согласно некоторым вариантам осуществления эквиваленты Cas9 могут относиться к Cas12e (CasX) или Cas12d (CasY), которые описаны, например, в Burstein et al., «New CRISPR–Cas systems from uncultivated microbes.» *Cell Res.* 2017 Feb 21. doi: 10.1038/cr.2017.21, полные содержания которой включено в настоящий документ посредством ссылки. Используя метагеномику с разрешением генома, был идентифицирован ряд систем CRISPR-Cas, в том числе первый сообщенный Cas9 в простейшей области жизни. Этот дивергентный белок Cas9 был обнаружен у малоизученных наноархей как часть активной системы CRISPR-Cas. У бактерий были обнаружены две ранее неизвестные системы, CRISPR-Cas12e и CRISPR-Cas12d, которые являются одними из самых компактных обнаруженных систем. Согласно некоторым вариантам осуществления Cas9 относится к Cas12e или варианту Cas12e. Согласно некоторым вариантам осуществления Cas9 относится к Cas12d или варианту Cas12d. Очевидно, что другие направляемые РНК ДНК-связывающие белки могут использоваться в качестве программируемого нуклеиновой кислотой ДНК-связывающего белка (napDNAbp) и входят в объем настоящего изобретения. Также см. Liu et al., «CasX enzymes comprises a distinct family of RNA-guided genome editors,» *Nature*, 2019, Vol.566: 218-223. Охватывается любой из этих эквивалентов Cas9.

[0394] Согласно некоторым вариантам осуществления эквивалент Cas9 содержит аминокислотную последовательность, которая на 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 99.5% идентична встречающемуся в природе белку Cas12e (CasX) или Cas12d (CasY). Согласно некоторым вариантам осуществления napDNAbp представляет собой встречающийся в природе белок Cas12e (CasX) или Cas12d (CasY). Согласно некоторым вариантам осуществления napDNAbp содержит аминокислотную последовательность, которая на 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 99.5% идентична

фрагменту Cas дикого типа или любому фрагменту Cas, раскрытому в настоящем документе.

[0395] Согласно различным вариантам 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. Одним примером программируемого нуклеиновой кислотой ДНК-связывающего белка, который имеет специфичность PAM, отличную от Cas9, являются кластеризованные регулярные промежуточные короткие палиндромные повторы из Prevotella и Francisella 1 (т.е., Cas12a (Cpf1)). Подобно Cas9, Cas12a (Cpf1) также является эффектором CRISPR класса 2, но он является членом подгруппы ферментов типа V, а не подгруппы ферментов типа II. Было показано, что Cas12a (Cpf1) опосредует надежную интерференцию ДНК с характеристиками, отличными от Cas9. Cas12a (Cpf1) представляет собой направляемую одной РНК эндонуклеазу, в которой отсутствует tracrRNA, и использует T-богатый мотив, прилегающий к протоспейсеру (TTN, TTTN или YTN). Более того, Cpf1 расщепляет ДНК посредством ступенчатого двухцепочечного разрыва ДНК. Показано, что из 16 белков семейства Cpf1 два фермента из *Acidaminococcus* и *Lachnospiraceae* обладают эффективной активностью редактирования генома в клетках человека. Белки Cpf1 известны в данной области техники и были описаны ранее, например, Yamano *et al.*, «Crystal structure of Cpf1 in complex with гидовая РНК and target DNA.» *Cell* (165) 2016, p. 949-962, полное содержание которой включено в настоящий документ посредством ссылки.

[0396] Согласно другим вариантам осуществления белок Cas может включать любой CRISPR-ассоциированный белок, включая без ограничения Cas12a, Cas12b1, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (также известный как Csn1 и 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 их модифицированные варианты, и предпочтительно содержащий никазную мутацию (например, мутацию, соответствующую мутации D10A последовательности полипептида Cas9 дикого типа SEQ ID NO: 18).

[0397] Согласно различным другим вариантам осуществления napDNAbp может представлять собой любой из следующих белков: Cas9, Cas12a (Cpf1), Cas12e (CasX), Cas12d (CasY), Cas12b1 (C2c1), Cas13a (C2c2), Cas12c (C2c3), GeoCas9, CjCas9, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, Cas9 с круговыми перестановками, или домен Argonaute (Ago), или их вариант.

[0398] Иллюстративные последовательности эквивалентного белка Cas9 могут

включать следующие:

Описание	Последовательность
AsCas12a (ранее известный как Cpf1) *Acidaminococcus sp. (штамм BV3L6)* UniProtKB U2UMQ6	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPI IDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQATYRNA IHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENA LLRSFDKFTTYFSGFYENRKNVFSAEDISTAIPHRIVQDNFPKFKENCHIFTRL ITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGIS REAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILE EFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISS ALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIISAAG KELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLL DWFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFK LNFQMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEP TEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIE PLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSK YTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGK LYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRP KSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDE ARALLPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNA YLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREK ERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFGFK SKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTS FAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDF LHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFI AGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDD SHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWP MDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRN (SEQ ID NO: 64)
AsCas12a никаза (например, R1226A)	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPI IDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQATYRNA IHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENA LLRSFDKFTTYFSGFYENRKNVFSAEDISTAIPHRIVQDNFPKFKENCHIFTRL ITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGIS REAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILE EFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISS ALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIISAAG KELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLL DWFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFK

	LNFQMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEP TEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIE PLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSK YTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGK LYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRP KSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDE ARALLPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNA YLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREK ERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFGFK SKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTS FAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDF LHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFI AGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDD SHAIDTMVALIRSVLQMANSNAATGEDYINSPVRDLNGVCFDSRFQNPEWP MDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRN (SEQ ID NO: 65)
LbCas12a (ранее известный как Cpf1) *Lachnospiraceae bacterium GAM79* Эталонная последовательность WP_119623382.1	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQQE LKEIMDDYYRTFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKIQNEK RKEICCYFTSDKRFKDLFNAKLITDILPNFIKDNKEYTEEEKAEKEQTRVLFQ RFATAFTNYFNQRRNNFSEDNISTAISFRIVNENSEIHLQNMRAFQRIEQQYP EEVCGMEEEYKDMLQEWQMKHIYSVDFYDRELTQPGIEYYNGICGKINEH MNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFESDQEVYDALNEFI KTMKKKEIIRRCVHLGQECDDYDLGKIYISSNKYEQISNALYGSWDTIRKCI KEEYMDALPGKGEKKEEKAEAAAKKEEYRSIADIDKIISLYGSEMDRTISAK KCITEICDMAGQISIDPLVCNSDIKLLQNKEKTTEIKTILDSFLHVYQWGQTFI VSDIIEKDSYFYSELEDVLEDFEGITTLYNHVRSYVTQKPYSTVKFKLHFGSP TLANGWSQSKEYDNNAILLMRDQKFYLGIFNVRNKPDKQIIKGHEKEEKGD YKKMIYNLLPGPSKMLPKVFITSRSGQETYKPSKHILDGYNEKRHIKSSPKFD LGYCWDLIDYYKECIHKHPDWKNYDFHFSDTKDYEDISGFYREVEMQGYQI KWTYISADEIQKLDEKGQIFLFQIYNKDFSVHSTGKDNLHTMYLKNLFSEEN LKDIVLKLNGEAELFFRKASIKTPIVHKKGSVLVNRSYTQTVGNKEIRVSIPE EYYTEIYNYLNHIGKGKLSSEAQRYLDEGKIKSFTATKDIVKNYRYCCDHYF LHLPITINFKAKSDVAVNERTLAYIAKKEDIHIIGIDRGERNLLYISVVDVHGN IREQRSFNIVNGYDYQQKLKDREKSRDAARKNWEEIEKIKELKEGYLSMVIH YIAQLVVKYNAVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLHYLVF KDREVCEEGGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFV NLFSFKNLTNRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTILAST KWKVYTNGTRLKRIVVNGKYTSQSMEVELTDAMEKMLQRAGIEYHDGKD

	LKGQIVEKGIEAEIIDIFRLTVQMRNSRSESEDREYDRLISPVLNDKGEFFDTA TADKTLPQDADANGAYCIALKGLYEVKQIKENWKENEQFPRNKLVQDNKT WFDFMQKKRYL (SEQ ID NO: 66)
PcCas12a – ранее известный как Cpf1 *Prevotella copri* Эталонная последовательность WP_119227726.1	MAKNFEDFKRLYSLSKTLRFEAKPIGATLDNIVKSGLLDEDEHRAASYVKV KKLIDEYHKVFIDRVLDDGCLPLENKGNNNSLAEYYESYVSRAQDEDAKKK FKEIQQNLRSVIAKKLTEDKAYANLFGNKLIESYKDKEDKKKIIDSDLIQFINT AESTQLDSMSQDEAKELVKEFWGFVTYFYGFFDNRKNMYTAEEKSTGIAY RLVNENLPKFIDNIEAFNRAITRPEIQENMGVLYSDFSEYLNVESIQEMFQLD YYNMLLTQKQIDVYNAIIGGKTDDEHDVKIKGINEYINLYNQQHKDDKLPK LKALFKQILSDRNAISWLPEEFNSDQEVLNAIKDCYERLAENVLGDKVLKSL LGSLADYSLDGIFIRNDLQLTDISQKMFGNWGVIQNAIMQNIKRVAPARKHK ESEEDYEKRIAGIFKKADSFSISYINDCLNEADPNNAYFVENYFATFGAVNTP TMQRENLFALVQNAYTEVAALLHSDYPTVKHLAQDKANVSKIKALLDAIKS LQHFVKPLLGKGDESDKDERFYGELASLWAELDTVTPLYNMIRNYMTRKP YSQKKIKLNFENPQLLGGWDANKEKDYATIILRRNGLYYLAIMDKDSRKLL GKAMPSDGECYEKMVYKFFKDVTTMIPKCSTQLKDVQAYFKVNTDDYVL NSKAFNKPLTITKEVFDLNNVLYGKYKKFQKGYLTATGDNVGYTHAVNVW IKFCMDFLNSYDSTCIYDFSSLKPESYLSLDAFYQDANLLLYKLSFARASVSY INQLVEEGKMYLFQIYNKDFSEYSKGTPNMHTLYWKALFDERNLADVVYK LNGQAEMFYRKKSIENTHPTHPANHPILNKNKDNKKKESLFDYDLIKDRRY TVDKFMFHVPITMNFKSVGSENINQDVKAYLRHADDMHIIGIDRGERHLLY LVVIDLQGNIKEQYSLNEIVNEYNGNTYHTNYHDLLDVREEERLKARQSWQ TIENIKELKEGYLSQVIHKITQLMVRYHAIVVLEDLSKGFMRSRQKVEKQVY QKFEKMLIDKLNYLVDKKTDVSTPGGLLNAYQLTCKSDSSQKLGKQSGFLF YIPAWNTSKIDPVTGFVNLLDTHSLNSKEKIKAFFSKFDAIRYNKDKKWFEF NLDYDKFGKKAEDTRTKWTLCTRGMRIDTFRNKEKNSQWDNQEVDLTTE MKSLLEHYYIDIHGNLKDAISAQTDKAFFTGLLHILKLTLQMRNSITGTETDY LVSPVADENGIFYDSRSCGNQLPENADANGAYNIARKGLMLIEQIKNAEDLN NVKFDISNKAWLNFAQQKPYKNG (SEQ ID NO: 67)
ErCas12a – ранее известный как Cpf1 *Eubacterium rectale* Эталонная последовательно	MFSAKLISDILPEFVIHNNNYSASEKEEKTQVIKLFSRFATSFKDYFKNRANC FSANDISSSSCHRIVNDNAEIFFSNALVYRRIVKNLSNDDINKISGDMKDSLK EMSLEEIYSYEKYGEFITQEGISFYNDICGKVNLFMNLYCQKNKENKNLYKL RKLHKQILCIADTSYEVPYKFESDEEVYQSVNGFLDNISSKHIVERLRKIGEN YNGYNLDKIYIVSKFYESVSQKTYRDWETINTALEIHYNNILPGNGKSKADK VKKAVKNDLQKSITEINELVSNYKLCPDDNIKAETYIHEISHILNNFEAQELK YNPEIHLVESELKASELKNVLDVIMNAFHWCSVFMTEELVDKDNNFYAELE EIYDEIYPVISLYNLVRNYVTQKPYSTKKIKLNFGIPTLADGWSKSKEYSNNA IILMRDNLYYLGIFNAKNKPDKKIIEGNTSENKGDYKKMIYNLLPGPNKMIP

сть WP_119223642.1	KVFLSSKTGVETYKPSAYILEGYKQNKHLKSSKDFDITFCHDLIDYFKNCIAI HPEWKNFGFDFSDTSTYEDISGFYREVELQGYKIDWTYISEKDIDLLQEKGQ LYLFQIYNKDFSKKSSGNDNLHTMYLKNLFSEENLKDIVLKLNGEAEIFFRK SSIKNPIIHKKGSILVNRTYEAEEKDQFGNIQIVRKTIPENIYQELYKYFNDKS DKELSDEAAKLKNVVGHHEAATNIVKDYRYTYDKYFLHMPITINFKANKTS FINDRILQYIAKEKDLHVIGIDRGERNLIYVSVIDTCGNIVEQKSFNIVNGYDY QIKLKQQEGARQIARKEWKEIGKIKEIKEGYLSLVIHEISKMVIKYNAIIAME DLSYGFKKGRFKVERQVYQKFETMLINKLNYLVFKDISITENGGLLKGYQLT YIPDKLKNVGHQCGCIFYVPAAYTSKIDPTTGFVNIFKFKDLTVDAKREFIKK FDSIRYDSDKNLFCFTFDYNNFITQNTVMSKSSWSVYTYGVRIKRRFVNGRF SNESDTIDITKDMEKTLEMTDINWRDGHDLRQDIIDYEIVQHIFEIFKLTVQM RNSLSELEDRDYDRLISPVLNENNIFYDSAKAGDALPKDADANGAYCIALKG LYEIKQITENWKEDGKFSRDKLKISNKDWFDFIQNKRYL (SEQ ID NO: 68)
CsCas12a – ранее известный как Cpf1 *Clostridium sp. AF34-10BH* Эталонная последовательность WP_118538418.1	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQQE LKEIMDDYYRAFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKIQNEK RKEICCYFTSDKRFKDLFNAKLITDILPNFIKDNKEYTEEEKAEKEQTRVLFQ RFATAFTNYFNQRRNNFSEDNISTAISFRIVNENSEIHLQNMRAFQRIEQQYP EEVCGMEEEYKDMLQEWQMKHIYLVDFYDRVLTQPGIEYYNGICGKINEH MNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFESDQEVYDALNEFI KTMKEKEIICRCVHLGQKCDDYDLGKIYISSNKYEQISNALYGSWDTIRKCI KEEYMDALPGKGEKKEEKAEAAAKKEEYRSIADIDKIISLYGSEMDRTISAK KCITEICDMAGQISTDPLVCNSDIKLLQNKEKTTEIKTILDSFLHVYQWGQTFI VSDIIEKDSYFYSELEDVLEDFEGITTLYNHVRSYVTQKPYSTVKFKLHFGSP TLANGWSQSKEYDNNAILLMRDQKFYLGIFNVRNKPDKQIIKGHEKEEKGD YKKMIYNLLPGPSKMLPKVFITSRSGQETYKPSKHILDGYNEKRHIKSSPKFD LGYCWDLIDYYKECIHKHPDWKNYDFHFSDTKDYEDISGFYREVEMQGYQI KWTYISADEIQKLDEKGQIFLFQIYNKDFSVHSTGKDNLHTMYLKNLFSEEN LKDIVLKLNGEAELFFRKASIKTPVVHKKGSVLVNRSYTQTVGDKEIRVSIPE EYYTEIYNYLNHIGRGKLSTEAQRYLEERKIKSFTATKDIVKNYRYCCDHYF LHLPITINFKAKSDIAVNERTLAYIAKKEDIHIIGIDRGERNLLYISVVDVHGNI REQRSFNIVNGYDYQQKLKDREKSRDAARKNWEEIEKIKELKEGYLSMVIH YIAQLVVKYNAVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLHYLVF KDREVCEEGGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFV NLFSFKNLTNRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTMLAS TKWKVYTNGTRLKRIVVNGKYTSQSMEVELTDAMEKMLQRAGIEYHDGK DLKGQIVEKGIEAEIIDIFRLTVQMRNSRSESEDREYDRLISPVLNDKGEFFDT ATADKTLPQDADANGAYCIALKGLYEVKQIKENWKENEQFPRNKLVQDNK TWFDFMQKKRYL (SEQ ID NO: 69)

BhCas12b *Bacillus hisashii* Эталонная последовательность WP_095142515.1	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYEHHEQ DPKNPKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRELYEELV PSSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPS WEEEKKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEK SRNQSVRRLDKDMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKE DIQALKALEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDE NEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYA TFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLNKYRILTEQLHT EKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK HAFTYKDESIKFPLKGTLGGARVQFDRDHLRRYPHKVESGNVGRIYFNMTV NIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGL RVMSIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLP GETLVKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTK WISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEV KHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQ RFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQI ILFEDLSNYNPYEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQ FSSRFHAKTGSPGIRCSVVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDL YPDKGGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAY QVDGQTVYIPESKDQKQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQ SSSELVDSDILKDSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGK LERILISKLTNQYSISTIEDDSSKQSM (SEQ ID NO: 70)
ThCas12b *Thermomonas hydrothermalis* Эталонная последовательность WP_072754838	MSEKTTQRAYTLRLNRASGECAVCQNNSCDCWHDALWATHKAVNRGAK AFGDWLLTLRGGLCHTLVEMEVPAKGNNPPQRPTDQERRDRRVLLALSWL SVEDEHGAPKEFIVATGRDSADDRAKKVEEKLREILEKRDFQEHEIDAWLQ DCGPSLKAHIREDAVWVNRRALFDAAVERIKTLTWEEAWDFLEPFFGTQYF AGIGDGKDKDDAEGPARQGEKAKDLVQKAGQWLSARFGIGTGADFMSMA EAYEKIAKWASQAQNGDNGKATIEKLACALRPSEPPTLDTVLKCISGPGHKS ATREYLKTLDKKSTVTQEDLNQLRKLADEDARNCRKKVGKKGKKPWADE VLKDVENSCELTYLQDNSPARHREFSVMLDHAARRVSMAHSWIKKAEQRR RQFESDAQKLKNLQERAPSAVEWLDRFCESRSMTTGANTGSGYRIRKRAIE GWSYVVQAWAEASCDTEDKRIAAARKVQADPEIEKFGDIQLFEALAADEAI CVWRDQEGTQNPSILIDYVTGKTAEHNQKRFKVPAYRHPDELRHPVFCDFG NSRWSIQFAIHKEIRDRDKGAKQDTRQLQNRHGLKMRLWNGRSMTDVNLH WSSKRLTADLALDQNPNPNPTEVTRADRLGRAASSAFDHVKIKNVFNEKE WNGRLQAPRAELDRIAKLEEQGKTEQAEKLRKRLRWYVSFSPCLSPSGPFIV YAGQHNIQPKRSGQYAPHAQANKGRARLAQLILSRLPDLRILSVDLGHRFA AACAVWETLSSDAFRREIQGLNVLAGGSGEGDLFLHVEMTGDDGKRRTVV

	YRRIGPDQLLDNTPHPAPWARLDRQFLIKLQGEDEGVREASNEELWTVHKL EVEVGRTVPLIDRMVRSGFGKTEKQKERLKKLRELGWISAMPNEPSAETDE KEGEIRSISRSVDELMSSALGTLRLALKRHGNRARIAFAMTADYKPMPGGQ KYYFHEAKEASKNDDETKRRDNQIEFLQDALSLWHDLFSSPDWEDNEAKK LWQNHIATLPNYQTPEEISAELKRVERNKKRKENRDKLRTAAKALAENDQL RQHLHDTWKERWESDDQQWKERLRSLKDWIFPRGKAEDNPSIRHVGGLSIT RINTISGLYQILKAFKMRPEPDDLRKNIPQKGDDELENFNRRLLEARDRLREQ RVKQLASRIIEAALGVGRIKIPKNGKLPKRPRTTVDTPCHAVVIESLKTYRPD DLRTRRENRQLMQWSSAKVRKYLKEGCELYGLHFLEVPANYTSRQCSRTG LPGIRCDDVPTGDFLKAPWWRRAINTAREKNGGDAKDRFLVDLYDHLNNL QSKGEALPATVRVPRQGGNLFIAGAQLDDTNKERRAIQADLNAAANIGLRA LLDPDWRGRWWYVPCKDGTSEPALDRIEGSTAFNDVRSLPTGDNSSRRAPR EIENLWRDPSGDSLESGTWSPTRAYWDTVQSRVIELLRRHAGLPTS (SEQ ID NO: 71)
LsCas12b *Laceyella sacchari* WP_132221894.1	MSIRSFKLKLKTKSGVNAEQLRRGLWRTHQLINDGIAYYMNWLVLLRQED LFIRNKETNEIEKRSKEEIQAVLLERVHKQQQRNQWSGEVDEQTLLQALRQL YEEIVPSVIGKSGNASLKARFFLGPLVDPNNKTTKDVSKSGPTPKWKKMKD AGDPNWVQEYEKYMAERQTLVRLEEMGLIPLFPMYTDEVGDIHWLPQASG YTRTWDRDMFQQAIERLLSWESWNRRVRERRAQFEKKTHDFASRFSESDV QWMNKLREYEAQQEKSLEENAFAPNEPYALTKKALRGWERVYHSWMRLD SAASEEAYWQEVATCQTAMRGEFGDPAIYQFLAQKENHDIWRGYPERVIDF AELNHLQRELRRAKEDATFTLPDSVDHPLWVRYEAPGGTNIHGYDLVQDT KRNLTLILDKFILPDENGSWHEVKKVPFSLAKSKQFHRQVWLQEEQKQKKR EVVFYDYSTNLPHLGTLAGAKLQWDRNFLNKRTQQQIEETGEIGKVFFNISV DVRPAVEVKNGRLQNGLGKALTVLTHPDGTKIVTGWKAEQLEKWVGESG RVSSLGLDSLSEGLRVMSIDLGQRTSATVSVFEITKEAPDNPYKFFYQLEGTE MFAVHQRSFLLALPGENPPQKIKQMREIRWKERNRIKQQVDQLSAILRLHK KVNEDERIQAIDKLLQKVASWQLNEEIATAWNQALSQLYSKAKENDLQWN QAIKNAHHQLEPVVGKQISLWRKDLSTGRQGIAGLSLWSIEELEATKKLLTR WSKRSREPGVVKRIERFETFAKQIQHHINQVKENRLKQLANLIVMTALGYK YDQEQKKWIEVYPACQVVLFENLRSYRFSFERSRRENKKLMEWSHRSIPKL VQMQGELFGLQVADVYAAYSSRYHGRTGAPGIRCHALTEADLRNETNIIHE LIEAGFIKEEHRPYLQQGDLVPWSGGELFATLQKPYDNPRILTLHADINAAQ NIQKRFWHPSMWFRVNCESVMEGEIVTYVPKNKTVHKKQGKTFRFVKVEG SDVYEWAKWSKNRNKNTFSSITERKPPSSMILFRDPSGTFFKEQEWVEQKTF WGKVQSMIQAYMKKTIVQRMEE (SEQ ID NO: 72)
DtCas12b	MVLGRKDDTAELRRALWTTHEHVNLAVAEVERVLLRCRGRSYWTLDRRG DPVHVPESQVAEDALAMAREAQRRNGWPVVGEDEEILLALRYLYEQIVPSC

Dsulfonatronum thiodismutans *WP_031386437*	LLDDLGKPLKGDAQKIGTNYAGPLFDSDTCRRDEGKDVACCGPFHEVAGK YLGALPEWATPISKQEFDGKDASHLRFKATGGDDAFFRVSIEKANAWYEDP ANQDALKNKAYNKDDWKKEKDKGISSWAVKYIQKQLQLGQDPRTEVRRK LWLELGLLPLFIPVFDKTMVGNLWNRLAVRLALAHLLSWESWNHRAVQDQ ALARAKRDELAALFLGMEDGFAGLREYELRRNESIKQHAFEPVDRPYVVSG RALRSWTRVREEWLRHGDTQESRKNICNRLQDRLRGKFGDPDVFHWLAED GQEALWKERDCVTSFSLLNDADGLLEKRKGYALMTFADARLHPRWAMYE APGGSNLRTYQIRKTENGLWADVVLLSPRNESAAVEEKTFNVRLAPSGQLS NVSFDQIQKGSKMVGRCRYQSANQQFEGLLGGAEILFDRKRIANEQHGATD LASKPGHVWFKLTLDVRPQAPQGWLDGKGRPALPPEAKHFKTALSNKSKF ADQVRPGLRVLSVDLGVRSFAACSVFELVRGGPDQGTYFPAADGRTVDDPE KLWAKHERSFKITLPGENPSRKEEIARRAAMEELRSLNGDIRRLKAILRLSVL QEDDPRTEHLRLFMEAIVDDPAKSALNAELFKGFGDDRFRSTPDLWKQHCH FFHDKAEKVVAERFSRWRTETRPKSSSWQDWRERRGYAGGKSYWAVTYL EAVRGLILRWNMRGRTYGEVNRQDKKQFGTVASALLHHINQLKEDRIKTG ADMIIQAARGFVPRKNGAGWVQVHEPCRLILFEDLARYRFRTDRSRRENSR LMRWSHREIVNEVGMQGELYGLHVDTTEAGFSSRYLASSGAPGVRCRHLV EEDFHDGLPGMHLVGELDWLLPKDKDRTANEARRLLGGMVRPGMLVPWD GGELFATLNAASQLHVIHADINAAQNLQRRFWGRCGEAIRIVCNQLSVDGS TRYEMAKAPKARLLGALQQLKNGDAPFHLTSIPNSQKPENSYVMTPTNAGK KYRAGPGEKSSGEEDELALDIVEQAEELAQGRKTFFRDPSGVFFAPDRWLPS EIYWSRIRRRIWQVTLERNSSGRQERAEMDEMPY (SEQ ID NO: 73)

[0399] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут также содержать варианты Cas12a (Cpf1) (dCpf1), которые можно использовать в качестве домена программируемого гидовой нуклеотидной последовательностью ДНК-связывающего белка. Белок Cas12a (Cpf1) имеет RuvC-подобный эндонуклеазный домен, который подобен домену RuvC Cas9, но не содержит эндонуклеазный домен HNH, и N-конец Cas12a (Cpf1) не имеет альфа-спирального блока узнавания Cas9. В Zetsche *et al.*, *Cell*, 163, 759–771, 2015 (которая включена в настоящий документ посредством ссылки) показано, что RuvC-подобный домен Cas12a (Cpf1) отвечает за расщепление обеих цепей ДНК, а инактивация RuvC-подобного домена инактивирует нуклеазную активность Cas12a (Cpf1).

[0400] Согласно некоторым вариантам осуществления napDNAbp является отдельным эффектором микробной системы CRISPR-Cas. Отдельные эффекторы микробных систем CRISPR-Cas включают без ограничения Cas9, Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2) и Cas12c (C2c3). Как правило, микробные системы CRISPR-Cas делятся на системы класса 1 и класса 2. Системы класса 1 имеют мультисубъединичные

эффекторные комплексы, в то время как системы класса 2 имеют один эффекторный белок. Например, Cas9 и Cas12a (Cpf1) являются эффекторами класса 2. Помимо Cas9 и Cas12a (Cpf1), три различные системы CRISPR-Cas класса 2 (Cas12b1, Cas13a и Cas12c) были описаны 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 которой включено в настоящий документ посредством ссылки.

[0401] Эффекторы двух систем, Cas12b1 и Cas12c, содержат RuvC-подобные эндонуклеазные домены, родственные Cas12a. Третья система, Cas13a, содержит эффектор с двумя предсказанными доменами РНКазы HEPN. Продукция зрелой РНК CRISPR не зависит от tracrРНК, в отличие от продукции РНК CRISPR с помощью Cas12b1. Cas12b1 зависит как от РНК CRISPR, так и от tracrРНК для расщепления ДНК. Было показано, что бактериальный Cas13a обладает уникальной РНКазной активностью для созревания РНК CRISPR, отличной от его РНК-активируемой активности расщепления одноцепочечной РНК. Эти функции РНКазы отличаются друг от друга и от поведения Cas12a в отношении процессинга РНК CRISPR. 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 которой включено в настоящий документ посредством ссылки. *In vitro* биохимический анализ Cas13a в *Leptotrichia shahii* показал, что Cas13a направляется одной РНК CRISPR и может быть запрограммирован на расщепление мишеней оцРНК, несущих комплементарные протоспейсеры. Каталитические остатки в двух консервативных доменах HEPN опосредуют расщепление. Мутации в каталитических остатках генерируют каталитически неактивные 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 которой включено в настоящий документ посредством ссылки.

[0402] О кристаллической структуре Cas12b1 (AacC2c1) *Alicyclobaccillus acidoterrastris* сообщалось в комплексе с химерной гидовой РНК в виде одной молекулы (сгРНК). 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 которой включено в настоящий документ посредством ссылки. Также сообщалось о кристаллической структуре C2c1 *Alicyclobacillus acidoterrestris*, связанного с ДНК-мишенями, в виде тройных 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 которой включено в настоящий документ посредством ссылки. Каталитически компетентные конформации AacC2c1, как с цепями

ДНК-мишени, так и с цепями ДНК, не являющейся мишенью, были захвачены независимо друг от друга в одном каталитическом кармане RuvC, при этом C2c1-опосредованное расщепление приводит к ступенчатому семинуклеотидному разрыву ДНК-мишени. Структурные сравнения тройных комплексов C2c1 и ранее идентифицированных аналогов Cas9 и Cpf1 демонстрируют разнообразие механизмов, используемых системами CRISPR-Cas9.

[0403] Согласно некоторым вариантам осуществления napDNAbp может представлять собой белок C2c1, C2c2 или C2c3. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок C2c1. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок Cas13a. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок Cas12c. Согласно некоторым вариантам осуществления napDNAbp содержит аминокислотную последовательность, которая на 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 99.5% идентична белку Cas12b1 (C2c1), Cas13a (C2c2) или Cas12c (C2c3) природного происхождения. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок Cas12b1 (C2c1), Cas13a (C2c2) или Cas12c (C2c3) природного происхождения.

H. Cas9 круговые пермутанты

[0404] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, раскрытые в настоящем документе, могут содержать круговой пермутант Cas9.

[0405] Термин «Cas9 с круговыми перестановками» или «круговой пермутант» Cas9 или «CP-Cas9» относится к любому белку Cas9 или его варианту, который встречается или был модифицирован при конструировании в виде варианта кругового пермутанта, что означает, что N-конец и C-конец белка Cas9 (например, белка Cas9 дикого типа) были локально перестановлены. Такие белки Cas9 с круговыми перестановками или их варианты сохраняют способность связываться с ДНК при образовании комплекса с гидовой РНК (гРНК). См., Oakes et al., «Protein Engineering of Cas9 for enhanced function,» *Methods Enzymol*, 2014, 546: 491–511 и Oakes et al., «CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification,» *Cell*, January 10, 2019, 176: 254-267, каждая из которых включена в настоящий документ посредством ссылки. Настоящее раскрытие охватывает любой ранее известный CP-Cas9 или применение нового CP-Cas9

при условии, что полученный белок с круговыми перестановками сохраняет способность связываться с ДНК при образовании комплекса с гидовой РНК (гРНК).

[0406] Любой из белков Cas9, описанных в настоящем документе, включая любой вариант, ортолог или встречающийся в природе Cas9 или его эквивалент, может быть реконфигурирован в виде варианта кругового пермутанта.

[0407] Согласно различным вариантам осуществления круговые пермутанты Cas9 может иметь следующую структуру:

N-конец-[исходный C-конец] – [необязательный линкер] – [исходный N-конец]-C-конец.

[0408] В качестве примера настоящее изобретение охватывает следующие круговые пермутанты канонического Cas9 *S. pyogenes* (1368 аминокислот UniProtKB - Q99ZW2 (CAS9_STRP1) (нумерация основана на положении аминокислот 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 круговые пермутанты других белков Cas9 (включая другие ортологи, варианты Cas9 и т.д.).

[0409] Согласно конкретным вариантам осуществления круговой пермутант Cas9 имеет следующую структуру (на основе Cas9 *S. pyogenes* (1368 аминокислот UniProtKB - Q99ZW2 (CAS9_STRP1) (нумерация на основе положения аминокислот 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 круговые пермутанты других белков Cas9 (включая другие ортологи, варианты Cas9 и т.д.).

[0410] Согласно другим вариантам осуществления, круговой пермутант Cas9 имеет следующую структуру (на основе Cas9 *S. pyogenes* (1368 аминокислот UniProtKB - Q99ZW2 (CAS9_STRP1) (нумерация на основе положения аминокислот в SEQ ID NO: 18):

N-конец-[103-1368]-[необязательный линкер]-[1-102]-С-конец,

N-конец-[1029-1368]-[необязательный линкер]-[1-1028]-С-конец,

N-конец-[1042-1368]-[необязательный линкер]-[1-1041]-С-конец,

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

N-конец-[1301-1368]-[необязательный линкер]-[1-1300]-С-конец, или соответствующие круговые пермутанты других белков Cas9 (включая другие ортологи, варианты Cas9 и т.д.).

[0411] Согласно некоторым вариантам осуществления круговой пермутант может быть образован путем связывания С-концевого фрагмента Cas9 с N-концевым фрагментом Cas9 либо непосредственно, либо с использованием линкера, такого как аминокислотный линкер. Согласно некоторым вариантам осуществления С-концевой фрагмент может соответствовать С-концевым аминокислотам Cas9 на 95% или более (например, аминокислотам приблизительно 1300-1368), или С-концевым аминокислотам Cas9 на 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10% или 5% или более (например, любой последовательности SEQ ID NO: 77-86). N-концевая часть может соответствовать N-концевым аминокислотам Cas9 на 95% или более (например, аминокислотам приблизительно 1-1300), или N-концевым аминокислотам Cas9 на 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10% или 5% или более (например, последовательности SEQ ID NO: 18).

[0412] Согласно некоторым вариантам осуществления круговой пермутант может быть образован путем связывания С-концевого фрагмента Cas9 с N-концевым фрагментом Cas9 либо непосредственно, либо с использованием линкера, такого как аминокислотный линкер. Согласно некоторым вариантам осуществления С-концевой фрагмент, перестановленный на N-конец, включает или соответствует С-концевым аминокислотам Cas9 на 30% или менее (*например,* аминокислотам 1012-1368 последовательности SEQ ID NO: 18). Согласно некоторым вариантам осуществления концевой фрагмент, перестановленный на N-конец, включает или соответствует С-концевым аминокислотам Cas9 на 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% (*например,* последовательности Cas9 SEQ ID NO: 18). Согласно некоторым вариантам осуществления

С-концевой фрагмент, перестановленный на N-конец, включает или соответствует С-концевым 410 остаткам Cas9 или менее (*например,* последовательности Cas9 SEQ ID NO: 18). Согласно некоторым вариантам осуществления С-концевой фрагмент, перестановленный на N-конец, включает или соответствует С-концевым 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 остаткам Cas9 (*например,* последовательности Cas9 SEQ ID NO: 18). Согласно некоторым вариантам осуществления С-концевой фрагмент, перестановленный на N-конец, включает или соответствует С-концевым 357, 341, 328, 120 или 69 остаткам Cas9 (например, последовательности Cas9 SEQ ID NO: 18).

[0413] Согласно другим вариантам осуществления варианты кругового пермутанта Cas9 могут быть определены как топологическая перестройка первичной структуры Cas9 на основе следующего способа, основанного на Cas9 *S. pyogenes* последовательности SEQ ID NO: 18: (a) выбор участка круговой перестановки (CP), соответствующего внутреннему аминокислотному остатку первичной структуры Cas9, который рассекает исходный белок на две половины: N-концевой участок и С-концевой участок, (b) модификация последовательности белка Cas9 (например, с помощью методов генной инженерии) путем перемещения исходной С-концевой области (содержащей аминокислоту участка CP) перед исходной N-концевой областью, тем самым образуя новый N-конец белка Cas9, который теперь начинается с аминокислотного остатка участка CP. Участок CP может быть расположен в любом домене белка Cas9, включая, например, домен helical-II, домен RuvCIII или домен CTD. Например, участок CP может быть расположен (относительно Cas9 *S. pyogenes* SEQ ID NO: 18) при исходных аминокислотных остатках 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249 или 1282. Таким образом, после перемещения на N-конец исходные аминокислоты 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249 или 1282 станут новыми N-концевыми аминокислотами. Номенклатура этих белков CP-Cas9 может обозначаться 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. Это описание не предназначено для ограничения получения вариантов CP из SEQ ID NO: 18, но может быть реализовано для получения вариантов CP в любой последовательности Cas9 либо в участках CP, которые соответствуют этим положениям, либо в других участках CP полностью. Настоящее описание никоим образом не предназначено для ограничения конкретных участков CP. Практически любой участок CP может быть использован для образования варианта CP-Cas9.

[0414] Иллюстративные аминокислотные последовательности CP-Cas9, основанные

на последовательности Cas9 SEQ ID NO: 18, представлены ниже, в которых линкерные последовательности указаны подчеркиванием, а необязательные остатки метионина (M) выделены жирным шрифтом. Следует понимать, что в описании представлены последовательности CP-Cas9, которые не включают линкерную последовательность или включают другие линкерные последовательности. Следует понимать, что последовательности CP-Cas9 могут быть основаны на последовательностях Cas9, отличных от последовательностей SEQ ID NO: 18, и любые приведенные в настоящем документе примеры не предназначены для ограничения. Иллюстративные последовательности CP-Cas9 являются следующими:

Название CP	Последовательность	SEQ ID NO:
CP1012	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLA NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIV KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATL IHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGD KKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKP	SEQ ID NO: 77

Название CP	Последовательность	SEQ ID NO:
	ENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYG	
CP1028	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL SQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNS VGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL ENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL QLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE LLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIL DFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYW	SEQ ID NO: 78

Название CP	Последовательность	SEQ ID NO:
	RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQ	
CP1041	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELE NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTI DRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGG SGGSGGSGGSGGSGGDKKYSIGLAIGTNSVGWAVITDEYKV PSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDK KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFR IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK	SEQ ID NO: 79

Название CP	Последовательность	SEQ ID NO:
	MIAKSEQEIGKATAKYFFYS	
CP1249	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSG GSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEYKV PSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDK KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFR IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELA LPSKYVNFLYLASHYEKLKGS	SEQ ID NO: 80
CP1300	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSG	SEQ ID NO: 81

Название CP	Последовательность	SEQ ID NO:
	GSGGDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNT DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC YLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIK FRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHH QDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE EFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPL ARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRI EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKS DNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDK LIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIK LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVLSAYNKHRD	

[0415] Циклические пермутанты Cas9 могут быть полезны в описанных в настоящем документе конструкциях для мультилоскутного праймированного редактирования, описанного в настоящем документе. Иллюстративные С-концевые фрагменты Cas9 на основе последовательности Cas9 SEQ ID NO: 18, которые могут быть перестановлены на N-конец Cas9, представлены ниже. Следует понимать, что такие С-концевые фрагменты

Cas9 являются иллюстративными и не предназначены для ограничения. Эти иллюстративные фрагменты CP-Cas9 имеют следующие последовательности:

Название CP	Последовательность	SEQ ID NO:
CP1012 C-концевой фрагмент	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITL ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVN IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAG ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFV EQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 82
CP1028 C-концевой фрагмент	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL SQLGGD	SEQ ID NO: 83
CP1041 C-концевой фрагмент	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELE NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTI DRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 84
CP1249 C-концевой фрагмент	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 85
CP1300 C-концевой фрагмент	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGD	SEQ ID NO: 86

I. Варианты Cas9 с модифицированными специфичностями PAM

[0416] Мультилоскутные праймирующие согласно настоящему раскрытию также могут содержать варианты Cas9 с модифицированными специфичностями PAM. Согласно некоторым аспектам настоящее изобретение относится к белкам Cas9, которые проявляют активность в отношении целевой последовательности, которая не содержит канонический PAM (5'-NGG-3', где N представляет собой A, C, G или T) на своем 3'-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5'-NGG-3' PAM на своем 3'-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NNG-3′ PAM на ее 3′-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NNA-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NNC-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NNT-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NGT-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NGA-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NGC-3′ PAM на своем 3’-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NAA-3′ PAM на своем 3'-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NAC-3′ PAM на своем 3'-конце. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NAT-3′ PAM на своем 3'-конце. Согласно другим вариантам осуществления, белок Cas9 проявляет активность в отношении целевой последовательности, содержащей последовательность 5′-NAG-3′ PAM на своем 3'-конце.

[0417] Следует понимать, что любая из аминокислотных мутаций, описанных в настоящем документе (например, A262T) от первого аминокислотного остатка (например,

A) до второго аминокислотного остатка (например, T), также может включать мутации первого аминокислотного остатка на аминокислотный остаток, который подобен (например, консервативен) второму аминокислотному остатку. Например, мутация аминокислоты с гидрофобной боковой цепью (например, аланин, валин, изолейцин, лейцин, метионин, фенилаланин, тирозин или триптофан) может представлять собой мутацию на вторую аминокислоту с другой гидрофобной боковой цепью (например, аланин, валин, изолейцин, лейцин, метионин, фенилаланин, тирозин или триптофан). Например, мутация аланина на треонин (например, мутация A262T) также может быть мутацией аланина на аминокислоту, которая по размеру и химическим свойствам аналогична треонину, например, серин. В качестве другого примера, мутация аминокислоты с положительно заряженной боковой цепью (например, аргинин, гистидин или лизин) может представлять собой мутацию на вторую аминокислоту с другой положительно заряженной боковой цепью (например, аргинин, гистидин или лизин). В качестве другого примера, мутация аминокислоты с полярной боковой цепью (например, серин, треонин, аспарагин или глутамин) может представлять собой мутацию на вторую аминокислоту с другой полярной боковой цепью (например, серин, треонин, аспарагин, или глютамин). Дополнительные аналогичные пары аминокислот включают без ограничения следующие: фенилаланин и тирозин, аспарагин и глютамин, метионин и цистеин, аспарагиновая кислота и глутаминовая кислота, и аргинин и лизин. Специалисту в данной области техники понятно, что такие консервативные аминокислотные замены, вероятно, окажут незначительное влияние на структуру белка и, вероятно, будут хорошо переноситься без ущерба для функции. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на треонин может представлять собой аминокислотную мутацию на серин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на аргинин может представлять собой аминокислотную мутацию на лизин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на изолейцин, может представлять собой аминокислотную мутацию на аланин, валин, метионин или лейцин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на лизин может представлять собой аминокислотную мутацию на аргинин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на аспарагиновую кислоту может представлять собой аминокислотную мутацию на глутаминовую кислоту или

аспарагин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на валин может представлять собой аминокислотную мутацию на аланин, изолейцин, метионин или лейцин. Согласно некоторым вариантам осуществления любая из раскрытых в настоящем документе аминокислотных мутаций одной аминокислоты на глицин может представлять собой аминокислотную мутацию на аланин. Однако следует понимать, что дополнительные консервативные аминокислотные остатки могут быть очевидны специалистом в данной области техники, и любые аминокислотные мутации в другие консервативные аминокислотные остатки также входят в объем настоящего изобретения.

[0418] Согласно некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций, которые проявляют активность в отношении целевой последовательности, содержащей последовательность 5'-NAA-3' PAM на своем 3'-конце. Согласно некоторым вариантам осуществления комбинация мутаций присутствует в любом из клонов, перечисленных в таблице 1. Согласно некоторым вариантам осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в таблице 1. Согласно некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций любого из клонов Cas9, перечисленных в таблице 1.

[0419] Таблица 1. Клоны PAM NAA

Мутации 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

[0420] Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 80% идентична аминокислотной последовательности белка Cas9, представленной в любом из вариантов в Таблице 1. Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 92%, 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% идентична аминокислотной последовательности белка Cas9, представленной в любом из вариантов в Таблице 1.

[0421] Согласно некоторым вариантам осуществления белок Cas9 проявляет повышенную активность в отношении целевой последовательности, не содержащей канонической PAM (5'-NGG-3') на своем 3'-конце, по сравнению с Cas9 *Streptococcus pyogenes*, представленным SEQ ID NO: 18. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, имеющей 3'-конец, не примыкающий непосредственно к канонической последовательности PAM (5'-NGG-3'), которая по меньшей мере в 5 раз повышена по сравнению с активностью Cas9 *Streptococcus pyogenes*, представленным SEQ ID NO: 18 в отношении той же целевой последовательности. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, не примыкающей непосредственно к канонической последовательности PAM (5´-NGG-3´), которая по меньшей мере в 10 раз, по меньшей мере в 50 раз, по меньшей мере в 100 раз, по меньшей мере 500 раз, по меньшей мере 1000 раз, по меньшей мере 5000 раз, по меньшей мере 10000 раз, по меньшей мере 50000 раз, по меньшей мере 100000 раз по меньшей мере 500000 раз или по меньшей мере 1000000 раз повышена по сравнению с активностью *Streptococcus pyogenes*, как представлено SEQ ID NO: 18, в отношении той же целевой последовательности. Согласно некоторым вариантам осуществления 3'-конец целевой последовательности непосредственно примыкает к последовательности AAA, GAA, CAA или TAA. Согласно некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций, которые проявляют активность в отношении целевой последовательности, содержащей последовательность 5'-NAC-3' PAM на своем 3'-конце. Согласно некоторым вариантам осуществления комбинация мутаций представлена в любом из клонов, перечисленных в таблице 2. Согласно некоторым вариантам осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 2. Согласно

некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций любого из клонов Cas9, перечисленных в Таблице 2.

[0422] Таблица 2. Клоны PAM NAC

МУТАЦИИ 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

[0423] Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 80% идентична аминокислотной последовательности белка Cas9, представленной в любом из вариантов в

Таблице 2. Согласно некоторым вариантам осуществления белок Cas9 содержит аминокислотную последовательность, которая на по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 92%, 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% идентична аминокислотной последовательности белка Cas9, представленной в любом из вариантов в Таблице 2.

[0424] Согласно некоторым вариантам осуществления белок Cas9 проявляет повышенную активность в отношении целевой последовательности, не содержащей канонической PAM (5'-NGG-3') на своем 3'-конце, по сравнению с Cas9 *Streptococcus pyogenes*, представленным SEQ ID NO: 18. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, имеющей 3'-конец, не примыкающий непосредственно к канонической последовательности PAM (5'-NGG-3'), которая по меньшей мере в 5 раз повышена по сравнению с активностью Cas9 *Streptococcus pyogenes*, представленным SEQ ID NO: 18 в отношении той же целевой последовательности. Согласно некоторым вариантам осуществления белок Cas9 проявляет активность в отношении целевой последовательности, не примыкающей непосредственно к канонической последовательности PAM (5´-NGG-3´), которая по меньшей мере в 10 раз, по меньшей мере в 50 раз, по меньшей мере в 100 раз, по меньшей мере 500 раз, по меньшей мере 1000 раз, по меньшей мере 5000 раз, по меньшей мере 10000 раз, по меньшей мере 50000 раз, по меньшей мере 100000 раз по меньшей мере 500000 раз или по меньшей мере 1000000 раз повышена по сравнению с активностью *Streptococcus pyogenes*, как представлено SEQ ID NO: 18, в отношении той же целевой последовательности. Согласно некоторым вариантам осуществления 3'-конец целевой последовательности непосредственно примыкает к последовательности AAC, GAC, CAC или TAC.

[0425] Согласно некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций, которые проявляют активность в отношении целевой последовательности, содержащей последовательность 5'-NAT-3' PAM на своем 3'-конце. Согласно некоторым вариантам осуществления комбинация мутаций представлена в любом из клонов, перечисленных в таблице 3. Согласно некоторым вариантам осуществления комбинация мутаций представляет собой консервативные мутации клонов, перечисленных в Таблице 3. Согласно некоторым вариантам осуществления белок Cas9 содержит комбинацию мутаций любого из клонов Cas9, перечисленных в Таблице 3.

[0426] Таблица 3. Клоны PAM NAT

МУТАЦИИ 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

[0427] Приведенное выше описание различных napDNAbp, которые можно использовать в связи с раскрытыми в настоящем документе мультилоскутными праймирующими редакторами никоим образом не предназначено для ограничения. Мультилоскутные праймирующие редакторы могут содержать канонический белок SpCas9, или любой ортолог белка Cas9, или любой вариант белка Cas9, включая любой встречающийся в природе вариант, мутант или иным образом сконструированный вариант

Cas9, который известен или может быть получен или эволюционировал посредством направленного эволюционного процесса или иного мутагенного процесса. Согласно различным вариантам осуществления Cas9 или варианты Cas9 обладают никазной активностью, т.е., расщепляют только одну цепь целевой последовательности ДНК. Согласно другим вариантам осуществления Cas9 или варианты Cas9 не обладают нуклеазной активностью, т.е. представляют собой «мертвые» белки Cas9. Другими вариантами белков Cas9, которые могут быть использованы, являются те, которые имеют меньшую молекулярную массу, чем канонический SpCas9 (например, для облегчения доставки), или имеют модифицированную или подвергнутую перестановке структуру первичных аминокислот (например, форматы круговых пермутантов). Мультилоскутные праймирующие редакторы, описанные в настоящем документе, также могут содержать эквиваленты Cas9, включая белки Cas12a (Cpf1) и Cas12b1, которые являются результатом конвергентной эволюции. Согласно настоящему изобретению napDNAbp (например, SpCas9, вариант Cas9 или эквиваленты Cas9) могут также содержать различные модификации, которые изменяют/усиливают их PAM-специфичность. Наконец, настоящее изобретение охватывает любой Cas9, вариант Cas9 или эквивалент Cas9, последовательность которого на 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% идентична эталонной последовательности Cas9, такой как эталонная каноническая последовательность SpCas9 или эталонный эквивалент Cas9 (например, Cas12a/Cpf1).

[0428] Согласно конкретному варианту осуществления вариант Cas9 с расширенными возможностями PAM представляет собой SpCas9 (H840A) VRQR (SEQ ID NO: 87), который имеет следующую аминокислотную последовательность (с заменами V, R, Q, R относительно SpCas9 (H840A) последовательности SEQ ID NO: 51, показанными жирным шрифтом с подчеркиванием). Кроме того, остаток метионина в SpCas9 (H840) был удален для SpCas9 (H840A) VRQR):

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDE VAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK

MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY
AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDS
LTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE
VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSV
LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN
GRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIE
QISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK**Q**Y
RSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 87)

[0429] Согласно другому конкретному варианту осуществления вариант Cas9 с расширенными возможностями PAM представляет собой SpCas9 (H840A) VRER, который имеет следующую аминокислотную последовательность (с заменами V, R, E, R относительно SpCas9 (H840A) последовательности SEQ ID NO: 51, показанными жирным шрифтом с подчеркиванием) Кроме того, остаток метионина в SpCas9 (H840) был удален для SpCas9 (H840A) VRER):

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDE
VAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV
QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS
NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY
AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDS
LTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE

VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSV
LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN
GRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIE
QISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK**E**Y
RSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 88)

[0430] Согласно некоторым вариантам осуществления napDNAbp, который функционирует с неканонической последовательностью PAM, представляет собой белок Argonaute. Одним из примеров такого программируемого нуклеиновой кислотой ДНК-связывающего белка является белое Argonaute из *Natronobacterium gregoryi* (NgAgo). NgAgo представляет собой эндонуклеазу, направляемую одноцепочечной ДНК. NgAgo связывается с 5′-фосфорилированной одноцепочечной ДНК из ~24 нуклеотидов (гДНК) для направления на ее целевой участок и делает двухцепочечные разрывы ДНК в сайте гДНК. В отличие от Cas9, система NgAgo-гДНК не требует мотива, прилегающего к протоспейсеру (PAM). Использование NgAgo без нуклеазной активности (dNgAgo) может значительно расширить количество оснований-мишеней. Характеристика и использование NgAgo описаны 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, каждая из которых включена в настоящий документ посредством ссылки.

[0431] Согласно некоторым вариантам осуществления napDNAbp представляет собой прокариотический гомолог белка Argonaute. Прокариотические гомологи белков Argonaute известны и 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 которой включено в настоящий документ посредством ссылки. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок Argunaute (MpAgo) Marinitoga piezophila. CRISPR-ассоциированный белок Marinitoga piezophila Argunaute (MpAgo) расщепляет одноцепочечные целевые последовательности с использованием 5'-фосфорилированных направляющих. 5′-направляющие используются всеми известными белками Argonaute. Кристаллическая структура комплекса MpAgo-РНК показывает участок связывания направляющей цепи, содержащий остатки, которые блокируют 5'-фосфатные взаимодействия. Эти данные предполагают эволюцию подкласса Argonaute с неканонической специфичностью в отношении 5'-гидроксилированной направляющей. См., например, Kaya et al., «A bacterial Argonaute with noncanonical guide РНК specificity», *Proc Natl Acad Sci U S A*. 2016 Apr 12,113(15):4057-62, полное содержание

которой включено в настоящий документ посредством ссылки). Следует понимать, что могут быть использованы другие белки Argonaute, и они входят в объем настоящего раскрытия.

[0432] Согласно некоторым аспектам настоящее изобретение относится к доменам Cas9, обладающим различной специфичностью PAM. Как правило, белки Cas9, такие как Cas9 из *S. pyogenes* (spCas9), требуют канонической последовательности PAM NGG для связывания определенной области нуклеиновой кислоты. Это может ограничить возможность редактирования нужных оснований в геноме. Согласно некоторым вариантам осуществления раскрытые в настоящем документе слитые белки для редактирования оснований могут нуждаться в размещении в точном месте, например, когда целевое основание помещается в область из 4 оснований (*например,* «окно редактирования»), которая находится на расстоянии приблизительно 15 оснований против хода транскрипции от 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 которой включено в настоящий документ посредством ссылки. Соответственно, согласно некоторым вариантам осуществления любой из слитых белков, раскрытых в настоящем документе, может содержать домен Cas9, способный связывать нуклеотидную последовательность, не содержащую каноническую (*например,* NGG) последовательность PAM. Домены Cas9, которые связываются с неканоническими последовательностями PAM, были описаны в данной области техники и должны быть очевидны для специалистов в данной области техники. Например, домены Cas9, которые связывают неканонические последовательности PAM, были описаны в Kleinstiver, B. P., *et al.*, «Engineered CRISPR-Нуклеаза Cas9s 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), полные содержания каждой из которых включены в настоящий документ посредством ссылки.

[0433] Например, домен napDNAbp с измененной специфичностью PAM, такой как домен с по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичностью последовательности с Cpf1 *Francisella* novicida дикого типа (D917, E1006 и D1255) (SEQ ID NO: 74), имеет следующую аминокислотную последовательность:

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEI
LSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLID
AKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSN
DIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFS

LDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSV
LFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDL
SKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSL
ETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQAS
AEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIR
NYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAI
KENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKF
EFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDS
VVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIP
KKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLL
LKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDS
ARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLE
KMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFV
NQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRN
SDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRN
SKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKK
LNLVIKNEEYFEFVQNRNN (SEQ ID NO: 74)

[0434] Дополнительный домен napDNAbp с измененной специфичностью PAM, такой как домен, имеющий по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичности последовательности с Cas9 *Geobacillus thermodenitrificans* дикого типа (SEQ ID №: 75), имеет следующую аминокислотную последовательность:

MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLARSARRRLRRRKHR
LERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEALDRKLNNDELARILLHLAKRRGFRSN
RKSERTNKENSTMLKHIEENQSILSSYRTVAEMVVKDPKFSLHKRNKEDNYTNTVARDDLEREI
KLIFAKQREYGNIVCTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFEPKEKRAPKATYTFQSF
TVWEHINKLRLVSPGGIRALTDDERRLIYKQAFHKNKITFHDVRTLLNLPDDTRFKGLLYDRNT
TLKENEKVRFLELGAYHKIRKAIDSVYGKGAAKSFRPIDFDTFGYALTMFKDDTDIRSYLRNEY
EQNGKRMENLADKVYDEELIEELLNLSFSKFGHLSLKALRNILPYMEQGEVYSTACERAGYTF
TGPKKKQKTVLLPNIPPIANPVVMRALTQARKVVNAIIKKYGSPVSIHIELARELSQSFDERRK
MQKEQEGNRKKNETAIRQLVEYGLTLNPTGLDIVKFKLWSEQNGKCAYSLQPIEIERLLEPGYT
EVDHVIPYSRSLDDSYTNKVLVLTKENREKGNRTPAEYLGLGSERWQQFETFVLTNKQFSKKK
RDRLLRLHYDENEENEFKNRNLNDTRYISRFLANFIREHLKFADSDDKQKVYTVNGRITAHLR
SRWNFNKNREESNLHHAVDAAIVACTTPSDIARVTAFYQRREQNKELSKKTDPQFPQPWPHFA
DELQARLSKNPKESIKALNLGNYDNEKLESLQPVFVSRMPKRSITGAAHQETLRRYIGIDERSG
KIQTVVKKKLSEIQLDKTGHFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGE
LGPIIRTIKIIDTTNQVIPLNDGKTVAYNSNIVRVDVFEKDGKYYCVPIYTIDMMKGILPNKAIEP

NKPYSEWKEMTEDYTFRFSLYPNDLIRIEFPREKTIKTAVGEEIKIKDLFAYYQTIDSSNGGLSLV
SHDNNFSLRSIGSRTLKRFEKYQVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 75)

[0435] Согласно некоторым вариантам осуществления программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) представляет собой программируемый нуклеиновой кислотой ДНК-связывающий, для которого не требуется каноническая (NGG) последовательность PAM. Согласно некоторым вариантам осуществления napDNAbp представляет собой белок Argonaute. Одним примером такого программируемого нуклеиновой кислотой ДНК-связывающего белка является белок Argonaute из *Natronobacterium gregoryi* (NgAgo). NgAgo представляет собой эндонуклеазу, направляемую одноцепочечной ДНК. NgAgo связывается с 5′-фосфорилированной одноцепочечной ДНК из ~24 нуклеотидов (гДНК) для направления на ее целевой участок и делает двухцепочечные разрывы ДНК в сайте гДНК. В отличие от Cas9, система NgAgo-гДНК не требует мотива, прилегающего к протоспейсеру (PAM). Использование NgAgo без нуклеазной активности (dNgAgo) может значительно расширить количество оснований-мишеней. Характеристика и использование NgAgo описаны 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, каждая из которых включена в настоящий документ посредством ссылки. Последовательность Argonaute *Natronobacterium gregoryi* представлена в SEQ ID NO: 76.

[0436] Описанные слитые белки могут содержать домен napDNAbp, имеющий по меньшей мере 80 %, по меньшей мере 85 %, по меньшей мере 90 %, по меньшей мере 95 % или по меньшей мере 99 % идентичности последовательности с Argonaute *Natronobacterium gregoryi* дикого типа (SEQ ID NO: 76), который имеет следующую аминокислотную последовательность:

MTVIDLDSTTTADELTSGHTYDISVTLTGVYDNTDEQHPRMSLAFEQDNGERRYITLWKNTTP
KDVFTYDYATGSTYIFTNIDYEVKDGYENLTATYQTTVENATAQEVGTTDEDETFAGGEPLDH
HLDDALNETPDDAETESDSGHVMTSFASRDQLPEWTLHTYTLTATDGAKTDTEYARRTLAYTV
RQELYTDHDAAPVATDGLMLLTPEPLGETPLDLDCGVRVEADETRTLDYTTAKDRLLARELVE
EGLKRSLWDDYLVRGIDEVLSKEPVLTCDEFDLHERYDLSVEVGHSGRAYLHINFRHRFVPKLT
LADIDDDNIYPGLRVKTTYRPRRGHIVWGLRDECATDSLNTLGNQSVVAYHRNNQTPINTDLL
DAIEAADRRVVETRRQGHGDDAVSFPQELLAVEPNTHQIKQFASDGFHQQARSKTRLSASRCS
EKAQAFAERLDPVRLNGSTVEFSSEFFTGNNEQQLRLLYENGESVLTFRDGARGAHPDETFSK
GIVNPPESFEVAVVLPEQQADTCKAQWDTMADLLNQAGAPPTRSETVQYDAFSSPESISLNVA
GAIDPSEVDAAFVVLPPDQEGFADLASPTETYDELKKALANMGIYSQMAYFDRFRDAKIFYTR
NVALGLLAAAGGVAFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINIAATATAVYKDGTILGH

SSTRPQLGEKLQSTDVRDIMKNAILGYQQVTGESPTHIVIHRDGFMNEDLDPATEFLNEQGVEY
DIVEIRKQPQTRLLAVSDVQYDTPVKSIAAINQNEPRATVATFGAPEYLATRDGGGLPRPIQIERV
AGETDIETLTRQVYLLSQSHIQVHNSTARLPITTAYADQASTHATKGYLVQTGAFESNVGFL
(SEQ ID NO: 76)

[0437] Кроме того, можно использовать любые доступные способы для получения или конструирования варианта или мутантного белка Cas9. В контексте настоящего изобретения термин «мутация» относится к замене остатка в последовательности, например, последовательности нуклеиновой кислоты или аминокислотной последовательности, другим остатком, или делеции, или вставке одного или нескольких остатков в последовательности. Мутации обычно описываются в данном документе путем идентификации исходного остатка, за которым следует положение остатка в последовательности и идентификации нового замещающего остатка. Различные способы получения аминокислотных замен (мутаций), предусмотренных в настоящем документе, хорошо известны в данной области и представлены, 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, такие как полиморфизмы с одним основанием, области микродупликации, вставки/делеции и инверсии, и никоим образом не предназначены для ограничения. Мутации могут включать мутации с «потерей функции», которые являются нормальным результатом мутации, снижающей или устраняющей активность белка. Большинство мутаций с потерей функции являются рецессивными, поскольку у гетерозигот вторая копия хромосомы несет немутированный вариант гена, кодирующего полностью функциональный белок, присутствие которой компенсирует эффект мутации. Мутации также охватывают мутации «приобретения функции», которые придают аномальную активность белку или клетке, которые иначе не присутствуют в нормальном состоянии. Многие мутации с приобретением функции находятся в регуляторных последовательностях, а не в кодирующих областях, и поэтому могут иметь ряд последствий. Например, мутация может привести к тому, что один или несколько генов будут экспрессироваться в неправильных тканях, и эти ткани приобретут функции, которых им обычно не хватает. Из-за своей природы мутации с приобретением функции обычно являются доминирующими.

[0438] Мутации могут быть введены в эталонный белок Cas9 с помощью сайт-направленного мутагенеза. Более старые способы сайт-направленного мутагенеза, известные в данной области техники, основаны на субклонировании последовательности, подлежащей мутации, в вектор, такой как вектор бактериофага M13, который позволяет выделить одноцепочечную ДНК-матрицу. В этих способах мутагенный праймер (т.е.

праймер, способный к отжигу с мутируемым сайтом, но несущий один или несколько несовпадающих нуклеотидов в мутируемом сайте) отжигают с одноцепочечной матрицей, а затем полимеризуют комплемент матрицы, начиная с 3'-конца мутагенного праймера. Полученные дуплексы затем трансформируют в бактерии-хозяева, а бляшки проверяют на наличие желаемой мутации. Совсем недавно в сайт-направленном мутагенезе использовали методологии ПЦР, преимущество которых состоит в том, что не требуется одноцепочечная матрица. Кроме того, были разработаны способы, не требующие субклонирования. При проведении сайт-направленного мутагенеза на основе ПЦР необходимо учитывать несколько вопросов. Во-первых, в этих способах желательно уменьшить количество циклов ПЦР, чтобы предотвратить распространение нежелательных мутаций, вносимых полимеразой. Во-вторых, необходимо использовать отбор, чтобы уменьшить количество немутировавших родительских молекул, сохраняющихся в реакции. В-третьих, предпочтительным является способ ПЦР с увеличенной длиной, чтобы можно было использовать один набор праймеров для ПЦР. И, в-четвертых, из-за активности концевого удлинения некоторых термостабильных полимераз, не зависящей от матрицы, часто необходимо включать в процедуру стадию шлифовки концов перед лигированием тупых концов мутантного продукта, полученного в результате ПЦР.

[0439] Мутации также могут быть вызваны процессами направленной эволюции, такими как непрерывная эволюция с помощью фага (PACE) или прерывистая эволюция с помощью фага (PANCE). В контексте настоящего изобретения термин «непрерывная эволюция с помощью фагов (PACE)» относится к непрерывной эволюции, в которой используют фаги в качестве вирусных векторов. Общая концепция технологии PACE была описана, например, в международной PCT заявке, PCT/US2009/056194, поданной 8 сентября 2009, опубликованной как WO 2010/028347 11 марта 2010, международной PCT заявке, PCT/US2011/066747, поданной 22 декабря 2011, опубликованной как WO 2012/088381 28 июня 2012, заявке США, патенте США № 9023594, поданном 5 мая 2015, международной PCT заявке, PCT/US2015/012022, поданной 20 января 2015, опубликованной как WO 2015/134121 11 сентября 2015, и международной PCT заявке, PCT/US2016/027795, поданной 15 апреля 2016, опубликованной как WO 2016/168631 20 октября 2016, полные содержания каждой из которых включены в настоящий документ посредством ссылки. Варианты Cas9 также могут быть получены с помощью прерывистой эволюции с помощью фага (PANCE)», которая в контексте настоящего изобретения относится к прерывистой эволюции, в которой используются фаги в качестве вирусных векторов. PANCE представляет собой упрощенный способ быстрой направленной эволюции in vivo с использованием серийных переносов в колбах эволюционирующего

«фага отбора» (SP), который содержит представляющий интерес ген для эволюции, с использованием свежих клеток-хозяев *E. coli*, тем самым позволяя генам внутри хозяина *E. .coli* оставаться постоянными, в то время как гены, содержащиеся в SP, непрерывно эволюционируют. Серийные переносы в колбах долгое время служили широко доступным подходом к лабораторной эволюции микробов, а совсем недавно аналогичные подходы были разработаны для эволюции бактериофагов. Система PANCE отличается меньшей строгостью, чем система PACE.

[0440] Любой из упомянутых выше ссылочных источников, которые относятся к Cas9 или эквивалентам Cas9, настоящим полностью включен посредством ссылки, если это еще не указано.

J. Разделенные домены napDNAbp для разделенной доставки PE

[0441] Согласно различным вариантам осуществления праймирующие редакторы, описанные в настоящем документе, могут быть доставлены в клетки в виде двух или более фрагментов, которые собираются внутри клетки (путем либо пассивной сборки, либо активной сборки, такой как использование последовательностей расщепленного интеина) в реконструированный праймирующий редактор. В некоторых случаях самосборка может быть пассивной, когда два или более фрагмента праймирующего редактора связываются внутри клетки ковалентно или нековалентно, с реконструированием праймирующего редактора. В других случаях самосборка может катализироваться доменами димеризации, установленными на каждом из фрагментов. Примеры доменов димеризации описаны в настоящем документе. В других случаях самосборка может катализироваться последовательностями расщепленного интеина, установленными на каждом из фрагментов праймирующего редактора.

[0442] Разделенная доставка PE может быть выгодна для решения различных ограничений по размеру различных 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 (доставка на основе рибонуклеопротеинов). И каждый из этих способов доставки может быть более эффективным и/или действенным при разделении праймирующего редактора на более мелкие части. Оказавшись внутри клетки, более мелкие части могут собраться в функциональный праймирующий редактор. В зависимости от способа разделения разделенные фрагменты праймирующего редактора могут быть повторно собраны нековалентным или ковалентным способом с повторным образованием праймирующего редактора. Согласно одному варианту осуществления праймирующий редактор может быть

разделен в одном или нескольких сайтах разделения на два или более фрагментов. Фрагменты могут быть немодифицированными (кроме разделенных). Как только фрагменты доставлены в клетку (например, путем прямой доставки рибонуклеопротеинового комплекса, или путем доставки нуклеиновых кислот, например, доставки иРНК, или доставки на основе вирусного вектора), фрагменты могут повторно ассоциировать ковалентно или нековалентно с восстановлением праймирующего редактора. Согласно другому варианту осуществления праймирующий редактор может быть разделен при одном или нескольких сайтах разделения на два или более фрагментов. Каждый из фрагментов может быть модифицирован с включением домена димеризации, при этом каждый образованный фрагмент соединяется с доменом димеризации. После доставки или экспрессии в клетке домены димеризации различных фрагментов ассоциируют и связываются друг с другом, объединяя различные фрагменты праймирующего редактора, с повторным образованием функционального праймирующего редактора. Согласно другому варианту осуществления фрагмент праймирующего редактора может быть модифицирован с включением расщепленного интеина. После доставки или экспрессии в клетке домены расщепленного интеина различных фрагментов ассоциируют и связываются друг с другом, а затем подвергаются транс-сплайсингу, что приводит к вырезанию доменов расщепленного интеина из каждого из фрагментов и сопутствующему образованию пептидной связи между фрагментами, тем самым восстанавливая праймирующий редактор.

[0443] Согласно одному варианту осуществления праймирующий редактор может быть доставлен с использованием подхода расщепленного интеина.

[0444] Место сайта расщепления может быть расположено между любой одной или несколькими парами остатков в праймирующем редакторе и в любых его доменах, включая домен napDNAbp, полимеразный домен (например, домен RT), линкерный домен, который соединяет домен napDNAb. и полимеразный домен.

[0445] Согласно одному варианту осуществления, как представлено на фиг. 66, праймирующий редактор (PE) разделен по сайту расщепления внутри napDNAbp.

[0446] Согласно определенным вариантам осуществления napDNAbp представляет собой следующую каноническую полипептидную последовательность SpCas9 SEQ ID NO: 18:

SpCas9 *Streptococcus pyogenes* M1 SwissProt №	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL	SEQ ID NO: 18

доступа Q99ZW2 Дикого типа 1368 AA	AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRE KIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLS DYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGD	

[0447] Согласно определенным вариантам осуществления SpCas9 расщепляется на два фрагмента в сайте расщепления, 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, расположенных в любом месте 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 канонической последовательности SpCas9 SEQ ID NO: 18.

[0448] Согласно определенным вариантам осуществления napDNAbp расщепляется на два фрагмента в сайте расщепления, расположенном на паре остатков, которая соответствует любым двум парам остатков, расположенным в любом месте 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 канонической последовательности SpCas9 SEQ ID NO: 18.

[0449] Согласно определенным вариантам осуществления SpCas9 расщепляется на два фрагмента в сайте расщепления, 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 в любом месте 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 канонической последовательности SpCas9 SEQ ID NO: 18. Согласно определенным вариантам осуществления сайт расщепления расположен в одном или нескольких сайтах полипептидной связи (т. е. «сайт расщепления или сайт расщепления расщепленного интеина»), слит с расщепленным интеином, а затем доставляется в клетки в виде отдельно кодируемых слитых белков. Как только слитые белки расщепленного интеина (т.е. половины белка) экспрессируются внутри клетки, белки подвергаются транс-сплайсингу с образованием полного или цельного PE с одновременным удалением соединенных последовательностей расщепленного интеина.

[0450] Например, как показано на фиг. 66, N-концевой экстеин может быть слит с первым расщепленным интеином (например, N-интеином), а C-концевой экстеин может быть слит со вторым расщепленным интеином (например, C-интеином). N-концевой экстеин сливается с C-концевым экстеином с повторным образованием цельного слитого белка праймирующего редактора, содержащего домен napDNAbp и полимеразный домен (например, домен RT) после самоассоциации N-интеина и C-интеина внутри клетки с последующей их самоэксцизией и сопутствующим образованием пептидной связи между N-концевой экстеиновой и C-концевой экстеиновой частями целого праймирующего редактора (PE).

[0451] Для достижения преимуществ доставки расщепленного PE с использованием расщепленных интеинов, праймирующий редактор должен быть разделен при одном или нескольких сайтах расщепления с созданием по меньшей мере двух отдельных половин праймирующего редактора, каждая из которых может быть воссоединена внутри клетки, если каждая половина слита с последовательностью расщепленного интеина.

[0452] Согласно определенным вариантам осуществления праймирующий редактор

при одном сайте разделения. Согласно определенным вариантам осуществления праймирующий редактор разделен при двух сайтах разделения, или трех сайтах разделения, или четырех сайтах разделения, или более.

[0453] Согласно предпочтительному варианту осуществления праймирующий редактор разделен при одном сайте разделения с созданием двух отдельных половин праймирующего редактора, каждая из которых может быть слита с последовательностью расщепленного интеина.

[0454] Иллюстративным расщепленным интеином является интеин *Ssp* DnaE, который содержит две субъединицы, а именно DnaE-N и DnaE-C. Две разные субъединицы кодируются отдельными генами, а именно *dnaE-n* и *dnaE-c*, которые кодируют субъединицы DnaE-N и DnaE-C соответственно. DnaE представляет собой природный расщепленный интеин PCC6803 *Synechocytis sp.* и способен управлять транс-сплайсингом двух отдельных белков, каждый содержащий слияние либо с DnaE-N, либо с DnaE-C.

[0455] Дополнительные встречающиеся в природе или сконструированные последовательности расщепленного интеина известны или могут быть получены из последовательностей цельного интеина, описанных в настоящем документе или доступных в данной области техники. Примеры последовательностей расщепленных интеинов 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 protin trans-splicing by a naturally split DnaE intein from Nostc punctiforme, FEBS Lett, 580: 1853-1858, каждая из которых включена в настоящий документ посредством ссылки. Дополнительные последовательности расщепленного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774 и EP 2877490, содержание каждой из которых включено в настоящий документ посредством ссылки.

[0456] Кроме того, транс-сплайсинг белка 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 дает возможность экспрессировать белок в виде двух неактивных фрагментов, которые затем подвергают лигированию с образованием функционального продукта, например, как показано на фиг. 66 и 67 в отношении образования полного слитого белка PE из двух отдельно экспрессируемых половин.

[0457] Согласно различным вариантам осуществления, описанные в настоящем

документе, способы непрерывной эволюции (например, PACE) могут использоваться для эволюции первой части редактора оснований. Первая часть может включать один компонент или домен, например, домен Cas9, домен дезаминазы или домен UGI. Затем отдельно эволюционировавший компонент или домен можно слить с оставшимися частями редактора оснований в клетке путем раздельной экспрессии как эволюционировавшей части, так и оставшихся неэволюционировавших частей с полипептидными доменами расщепленного интеина. В более широком смысле первая часть может включать любую первую аминокислотную часть редактора оснований, которая желательно должна эволюционировать с использованием способа непрерывной эволюции, описанного в настоящем документе. Вторая часть согласно этому варианту осуществления будет относиться к оставшейся аминокислотной части редактора оснований, которая не эволюционировала с использованием способов, описанных в настоящем документе. Каждая из эволюционировавшей первой части и второй части редактора оснований может быть экспрессирована с полипептидными доменами расщепленного интеина в клетке. Естественные механизмы сплайсинга белка в клетке будут повторно собирать эволюционировавшую первую часть и неэволюционировавшую вторую часть с образованием эволюционировавшего редактора оснований в виде единственного слитого белка. Эволюционировавшая первая часть может содержать либо N-, либо C-концевую часть одного слитого белка. Аналогичным образом, использование второй пары ортогональных транс-сплайсинговых интеинов может позволить эволюционировавшей первой части содержать внутреннюю часть единого слитого белка.

[0458] Таким образом, любой из эволюционировавших и неэволюционированных компонентов редакторов оснований, описанных в настоящем документе, может быть экспрессирован с помощью меток расщепленного интеина, чтобы облегчить образование полного редактора оснований, содержащего эволюционировавший и неэволюционировавший компонент внутри клетки.

[0459] Механизм процесса сплайсинга белка изучен очень подробно (Chong, et al., J. Biol. Chem. 1996, 271, 22159-22168, Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153) и консервативные аминокислоты были обнаружены в интеиновых и экстеиновых точках сплайсинга (Xu, et al., EMBO Journal, 1994, 13 5517-522). Конструкции, описанные в настоящем документе, содержат последовательность интеина, слитую с 5'-концом первого гена (например, эволюционировавшая часть редактора оснований). Подходящие последовательности интеина могут быть выбраны из любых белков, о которых известно, что они содержат элементы сплайсинга белков. 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 этого гена на определенную органеллу пептидный сигнал может быть слит с кодирующей последовательностью гена. После второго гена последовательность интеин-ген может повторяться сколь угодно часто для экспрессии множества белков в одной и той же клетке. Для мультиинтеинсодержащих конструкций может быть полезно использовать интеиновые элементы из разных источников. После последовательности последнего экспрессируемого гена должна быть вставлена последовательность терминации транскрипции. Согласно одному варианту осуществления модифицированная единица интеинового сплайсинга конструируется таким образом, чтобы она могла как катализировать удаление экстеинов из интеинов, так и предотвращать лигирование экстеинов. Было обнаружено, что мутагенез С-концевого экстеинового соединения в ДНК-полимеразе GB-D Pyrococcus species приводит к измененному элементу 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 либо в аланин, либо в глицин индуцировала расщепление, но предотвращала лигирование. Мутация эквивалентных остатков в других единицах интеинового сплайсинга также должна предотвращать лигирование экстеина из-за сохранения аминокислот на С-концевом соединении экстеина с интеином. Предпочтительным интеином, не содержащим эндонуклеазный домен, является белок GyrA Mycobacterium xenopi (Telenti, et al. J. Bacteriol. 1997, 179, 6378-6382). Другие обнаруживаются в природе или были созданы искусственно путем удаления эндонуклеазных доменов из интеинов, содержащих эндонуклеазу (Chong, et al. J. Biol. Chem. 1997, 272, 15587-15590). Согласно предпочтительному варианту осуществления интеин выбирают таким образом, чтобы он состоял из минимального количества аминокислот, необходимых для выполнения функции сплайсинга, например, интеин из белка GyrA Mycobacterium xenopi (Telenti, A., et al., J. Bacteriol. 1997, 179, 6378-6382). Согласно альтернативному варианту осуществления выбирают интеин без эндонуклеазной активности, такой как интеин из белка GyrA Mycobacterium xenopi или интеин VMA Saccharaomyces cerevisiae, который был модифицирован с удалением эндонуклеазных доменов (Chong, 1997). Дальнейшая модификация единицы интеинового сплайсинга может позволить изменить скорость реакции расщепления, позволяя контролировать дозу белка путем простой модификации последовательности гена сплайсинговой единицы.

[0460] Интеины также могут существовать в виде двух фрагментов, кодируемых двумя отдельно транскрибируемыми и транслируемыми генами. Эти так называемые расщепленные интеины самоассоциируются и катализируют активность транс-сплайсинга белков. Расщепленные интеины были идентифицированы у различных 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.

[0461] 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 эукариот. Недавно биоинформатический анализ метагеномных данных об окружающей среде выявил 26 различных локусов с новым расположением генома. В каждом локусе консервативная область, кодирующая фермент, прерывается расщепленным интеином с автономным геном эндонуклеазы, вставленным между секциями, кодирующими субдомены интеина. Среди них полностью собраны пять локусов: ДНК-хеликазы (gp41-1, gp41-8), инозин-5'-монофосфатдегидрогеназы (IMPDH-1), и каталитических субъединиц рибонуклеотидредуктазы (NrdA-2 и NrdJ-1). Эта разорванная генная организация, по-видимому, присутствует в основном у фагов (Dassa et al, Nucleic Acids Research. 57:2560-2573 (2009)).

[0462] Расщепленный интеин Npu DnaE был охарактеризован как имеющий самую высокую скорость реакции транс-сплайсинга белка. Кроме того, реакция сплайсинга белка Npu DnaE считается надежной и высокоэффективной в отношении различных экстеиновых последовательностей, температура от 6 до 37°C и присутствие до 6 M мочевины (Zettler J. et al, FEBS Letters. 553:909-914 (2009), Iwai I. et al, FEBS Letters 550: 1853-1858 (2006)). Как и ожидали, когда была введена мутация Cysl Ala в N-домен этих интеинов, начальный сдвиг N в S-ацил и, следовательно, сплайсинг белка блокировался. К сожалению, реакция C-концевого расщепления также была практически полностью ингибирована. Зависимость циклизации аспарагина на C-концевом соединении сплайсинга от ацильного сдвига на N-концевой разрывающейся пептидной связи, по-видимому, является уникальным свойством, общим для аллелей естественно расщепленного интеина DnaE (Zettler J. et al. FEBS Letters. 555:909-914 (2009)).

[0463] Механизм сплайсинга белков обычно состоит из четырех стадий [29-30]: 1) ацильный сдвиг NS или NO на N-конце интеина, который разрывает пептидную связь против хода транскрипции и образует сложноэфирную связь между N-экстеином и боковой цепью первой аминокислоты интеина (Cys или Ser), 2) транс-эстерификация, перемещающая N-экстеин к C-концу интеина с образованием новой сложноэфирной связи, связывающей N-экстеин с боковой цепью первой аминокислоты C-экстеина (Cys, Ser или Thr), 3) циклизация Asn с разрывом пептидной связи между интеином и C-экстеином, и 4) ацильный сдвиг SN или ON, который заменяет сложноэфирную связь на пептидную связь между N-экстеином и C-экстеином.

[0464] Транс-сплайсинг белков, катализируемый расщепленными интеинами, представляет собой полностью ферментативный метод лигирования белков [31]. Расщепленный 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 расщепленного интеина могут нековалентно связываться с образованием активного интеина и катализировать реакцию сплайсинга, по существу, таким же образом, как это делает непрерывный интеин. Расщепленные интеины обнаружены в природе, а также созданы в лабораториях [31-35]. В контексте настоящего изобретения термин термин «расщепленный интеин» относится к любому интеину, в котором существует один или несколько разрывов пептидной связи между N-концевой и C-концевой аминокислотными последовательностями, так что N-концевая и C-концевая последовательности становятся отдельными молекулами, которые могут нековалентно повторно ассоциироваться или восстанавливаться в интеин, который является функциональным для реакций транс-сплайсинга. Любой каталитически активный интеин или его фрагмент можно использовать для получения расщепленного интеина для применения в способах согласно настоящему изобретению. Например, согласно одному аспекту расщепленный интеин может быть получен из эукариотического интеина. Согласно другому аспекту расщепленный интеин может быть получен из бактериального интеина. Согласно другому аспекту расщепленный интеин может быть получен из интеина архей. Предпочтительно полученный таким образом расщепленный интеин будет иметь только те аминокислотные последовательности, которые необходимы для катализа реакций транс-сплайсинга.

[0465] В контексте настоящего изобретения «N-концевой расщепленный интеин (In)» относится к любой последовательности интеина, которая содержит N-концевую аминокислотную последовательность, функциональную для реакций транс-сплайсинга. Таким образом, In также содержит последовательность, которая подвергается сплайсингу, когда происходит транс-сплайсинг. In может содержать последовательность, которая является модификацией N-концевой части последовательности интеина, встречающейся в природе. Например, In может содержать дополнительные аминокислотные остатки и/или мутированные остатки при условии, что включение таких дополнительных и/или мутантных остатков не делает In нефункциональным при транс-сплайсинге. Предпочтительно включение дополнительных и/или мутантных остатков улучшает или усиливает активность транс-сплайсинга In.

[0466] В контексте настоящего изобретения «C-концевой расщепленный интеин (Ic)» относится к любой последовательности интеина, которая содержит C-концевую

аминокислотную последовательность, функциональную для реакций транс-сплайсинга. Согласно одному аспекту Ic содержит от 4 до 7 смежных аминокислотных остатков, по меньшей мере 4 аминокислоты из которых происходят из последней β-цепи интеина, из которого он был получен. Таким образом, Ic также содержит последовательность, которая подвергается сплайсингу, когда происходит транс-сплайсинг. Ic может содержать последовательность, которая является модификацией С-концевой части встречающейся в природе последовательности интеина. Например, Ic может содержать дополнительные аминокислотные остатки и/или мутированные остатки при условии, что включение таких дополнительных и/или мутантных остатков не делает In нефункциональным при транс-сплайсинге. Предпочтительно включение дополнительных и/или мутантных остатков улучшает или усиливает транс-сплайсинговую активность Ic.

[0467] Согласно некоторым вариантам осуществления настоящего изобретения пептид, связанный с Ic или In, может содержать дополнительный химический фрагмент, включая без ограничения флуоресцентные группы, биотин, полиэтиленгликоль (PEG), аналоги аминокислот, неприродные аминокислоты, фосфатные группы, гликозильные группы, радиоизотопные метки и фармацевтические молекулы. Согласно другим вариантам осуществления пептид, связанный с Ic, может содержать одну или несколько химически реакционноспособных групп, включая без ограничения кетон, альдегид, остатки Cys и остатки Lys. N-интеин и С-интеин расщепленного интеина могут нековалентно связываться с образованием активного интеина и катализировать реакцию сплайсинга, когда присутствует «интеин-сплайсинговый полипептид (ISP)». В контексте настоящего изобретения «интеин-сплайсинговый полипептид (ISP)» относится к части аминокислотной последовательности расщепленного интеина, которая сохраняется, когда Ic, In или оба удаляются из расщепленного интеина. Согласно определенным вариантам осуществления In содержит ISP. Согласно другому варианту осуществления Ic содержит ISP. Согласно другому варианту осуществления ISP представляет собой отдельный пептид, который не связан ковалентно ни с In, ни с Ic.

[0468] Расщепленные интеины могут быть созданы из непрерывных интеинов путем создания одного или нескольких сайтов расщепления в неструктурированной петле или промежуточной аминокислотной последовательности между -12 консервативными бета-цепями, обнаруживаемыми в структуре мини-интеинов [25-28]. Может существовать некоторая гибкость в расположении сайта расщепления в областях между бета-цепями при условии, что создание расщепления не нарушит структуру интеина, в частности структурированными бета-цепями, в достаточной степени, чтобы сплайсинговая активность белка не была потеряна.

[0469] При транс-сплайсинге белка один белок-предшественник состоит из части N-экстеина, за которой следует N-интеин, другой белок-предшественник состоит из C-интеина, за которым следует часть C-экстеина, и реакция транс-сплайсинга (катализируемая N- и C-интеинами вместе) вырезает две последовательности интеина и связывает две последовательности экстеина пептидной связью. Транс-сплайсинг белков, являясь ферментативной реакцией, может работать с очень низкими (например, микромолярными) концентрациями белков и может осуществляться в физиологических условиях.

[2] Другие программируемые нуклеазы

[0470] Согласно различным вариантам осуществления описанные в настоящем документе мультилоскутные праймирующие редакторы содержат napDNAbp, например, белок Cas9. Данные белки «программируются» посредством образования комплексов с гидовой РНК (или РЕгРНК, как может быть в этом случае), которая направляет белок Cas9 в целевой участок на ДНК, которая обладает последовательностью, являющейся комплементарной спейсерной части гРНК (или РЕгРНК), и которая также обладает требуемой последовательностью PAM. Однако согласно определенному варианту осуществления, раскрытому в настоящем документе, napDNAbp может быть замещен другим типом программируемого белка, например, нуклеазой цинкового пальца или подобной активатору транскрипции эффекторной нуклеазой (TALEN).

[0471] На **фиг. 1J** показан такой вариант предполагаемого в настоящем документе праймированного редактирования, с 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, гидовой РНК), но вместо этого могут программироваться путем задания специфичности ДНК-связывающего домена, например, (и в частности) нуклеазы. Как и при праймированном 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

праймированному 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 содержит 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 *в транс* 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) FEN 1 для обеспечения удаления эндогенной цепи на цепи с разрывом для смещения реакции в сторону замещения эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:гРНК для создания второго участка разрыва на противоположной цепи, что может обеспечивать интеграцию репарации синтеза посредством предпочтительной клеточной репарации не подвергнутой редактированию цепи). Аналогично праймированному 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 редактирование, перманентно в целевой участок ДНК.

[0472] Подходящие альтернативно программируемые нуклеазы хорошо известны в данной области техники и могут использоваться вместо комплекса napDNAbp:гРНК для конструирования альтернативной системы праймированного редактирования, которая может быть запрограммирована для избирательного связывания с целевым участком ДНК и которая может быть дополнительно модифицирована описанным выше образом для солокализации полимеразы и РНК или ДНК плеча удлинения, содержащего 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 описанных в настоящем документе способах и композициях праймированного редактирования. 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 контексте настоящего изобретения термин «TALEN» является широким и включает мономерный TALEN, который может расщеплять двухцепочечную ДНК без помощи со стороны другого TALEN. Термин «TALEN» также используется для обозначения одного или обоих членов пары TALEN, которые конструируются для 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 № 12/965590, заявку на патент США №13/426991 (патент США № 8450471), заявку на патент США № 13/427040 (патент США № 8440431), заявку на патент США № 13/427137 (патент США № 8440432) и заявку на патент США № 13/738381, все из которых включены в настоящий документ посредством ссылки. Кроме того, TALEN описаны в WO 2015/027134, US 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), каждая из которых включена в настоящий документ посредством ссылки.

[0473] Как представлено на фиг. 1J, нуклеазы цинкового пальца также могут использоваться в качестве альтернативных программируемых нуклеаз согласно праймированному редактированию вместо napDNAb, например, никаз Cas9. Подобно TALEN, белки ZFN могут быть модифицированы таким образом, что они функционируют в качестве никаз, то есть посредством конструирования ZFN, так что оно расщепляет только одну цепь ДНК-мишени таким же образом, что и napDNAbp, используемый с описанными в настоящем документе мультилоскутными праймирующими редакторами. Белки ZFN были обширно описаны в данной области техники, 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, and Gaj et al., «ZFN, TALEN, and

CRISPR/Cas-based methods for genome engineering,» *Trends Biotechnol.* 2013, Vol.31: 397-405, каждая из которых включена в настоящий документ посредством ссылки во всей своей полноте.

[3] Полимеразы (например, обратные транскриптазы)

[0474] Согласно различным вариантам осуществления система мультилоскутного праймирующего редактора, раскрытые в настоящем документе, включает полимеразу (например, ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, например, обратную транскриптазу) или ее вариант, который может быть представлен в виде слитого белка с napDNAbp или другой программируемой нуклеазой или представлен *в транс*.

[0475] Любая полимераза может быть использована в мультилоскутных праймирующих редакторах, раскрытых в настоящем документе. 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 DEEP VENT®, 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), каждая из которых включена в настоящий документ посредством ссылки). Для синтеза более длинных молекул нуклеиновой кислоты (например, молекул нуклеиновой кислоты длиной более чем около 3-5 т.о.) могут использоваться по меньшей мере две ДНК-полимеразы. Согласно определенным вариантам осуществления одна из полимераз может по существу не иметь 3' экзонуклеазной активности, а другая может иметь 3' экзонуклеазную активность. Такие пары могут включать полимеразы, которые являются одинаковыми или различающимися. Примеры ДНК-полимераз, которые по существу не имеют 3' экзонуклеазной активности, включают без ограничения ДНК-полимеразы Taq, Tne(exo-),

Tma(exo-), Pfu(exo-), Pwo(exo-), exo-KOD и Tth и их мутанты, варианты и производные.

[0476] Предпочтительно полимеразы, которые могут использоваться в описанных в настоящем документе мультилоскутных праймирующих редакторах, представляют собой «зависимую от матрицы» полимеразу (поскольку полимеразы должны полагаться на матрицу для синтеза ДНК для определения последовательности цепи ДНК при синтезе во время праймированного редактирования. В контексте настоящего изобретения термин «молекула матричной ДНК» означает цепь нуклеиновой кислоты, из которой синтезируется комплементарная цепь нуклеиновой кислоты ДНК-полимеразой, например, при реакции праймированного удлинения матрицы для синтеза ДНК РЕгРНК.

[0477] В контексте настоящего изобретения термин «зависимым от матрицы образом» означает процесс, который предусматривает зависимое от матрицы удлинение молекулы праймера (например, синтез ДНК посредством ДНК-полимеразы). 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» означает широкую концепцию комплементарности последовательностей между участками двух полинуклеотидных цепей или между двумя нуклеотидами посредством спаривания оснований. Известно, что нуклеотид аденин способен образовывать определенные водородные связи («спаривание оснований») с нуклеотидом, который представляет собой тимин или урацил. Подобным образом, известно, что нуклеотид цитозин способен обеспечивать спаривание оснований с нуклеотидом гуанином. Таким образом, в случае праймированного редактирования может быть сказано, что одна цепь ДНК, синтезированная полимеразой праймирующего редактора на основе матрицы для синтеза ДНК, является «комплементарной» последовательности матрицы для синтеза ДНК.

A. Иллюстративные полимеразы

[0478] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, содержат полимеразу. Настоящее 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 мультилоскутных праймирующих редакторах согласно настоящему 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 согласно настоящему изобретению, также могут быть сконструированы для включения определенных аминокислотных замещений, например, тех, которые конкретно указаны в настоящем документе. Согласно определенным предпочтительным вариантам осуществления полимеразы, которые могут использоваться в мультилоскутных праймирующих редакторах согласно настоящему изобретению, представляют собой полимеразы на основе матрицы, т.е. они синтезируют нуклеотидные последовательности зависимым от матрицы образом.

[0479] Полимераза представляет собой фермент, который синтезирует нуклеотидную цепь и который может применяться совместно с описанными в настоящем документе системами мультилоскутных праймирующих 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». Согласно различным вариантам осуществления система мультилоскутных праймирующих редакторов содержит ДНК-полимеразу. Согласно различным вариантам 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 РЕгРНК, где плечо удлинения включает цепь ДНК. В таких случаях РЕгРНК может быть обозначена как химерная или гибридная РЕгРНК, которая включает часть в виде РНК (т.е. компоненты гидовой РНК, включая спейсер и ядро гРНК) и часть в виде ДНК (т.е. плечо удлинения). Согласно различным другим вариантам осуществления ДНК-полимераза может представлять собой «РНК-зависимую ДНК-полимеразу» (т.е. молекула матрицы представляет собой цепь РНК). В таких случаях РЕгРНК представляет собой РНК, т.е. 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 РЕгРНК) по направлению к 5'-концу цепи матрицы. «ДНК-полимераза» катализирует полимеризацию дезоксинуклеотидов. В контексте настоящего изобретения в отношении

ДНК-полимеразы термин «ДНК-полимераза» включает «ее функциональный фрагмент». Термин «ее функциональный фрагмент» означает 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.

[0480] Согласно некоторым вариантам 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 от 3' к 5' exo T7 «Sequenase»).

[0481] Согласно другим вариантам 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* DPl/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 желаемым температурам в рамках анализов.

[0482] Термостабильные архейные ДНК-полимеразы выделяют из вида *Pyrococcus* (furiosus, вид GB-D, *woesii, abysii, horikoshii),* вида *Thermococcus* (kodakaraensis KOD1, litoralis, вид 9 degrees North-7, вид JDF-3, gorgonarius), *Pyrodictium occultum* и *Archaeoglobus fulgidus*.

[0483] 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.

[0484] Существует множество коммерчески доступных ДНК-полимераз pol I, некоторые из которых были модифицированы для снижения или устранения экзонуклеазной активности от 5' к 3'.

[0485] 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).

[0486] Дополнительные эубактерии, являющиеся родственными тем, которые перечислены выше, описаны в Thermophilic Bacteria (Kristjansson, J. K., ed.) CRC Press, Inc., Boca Raton, Fla., 1992.

[0487] Настоящее 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 включено в настоящий документ посредством ссылки во всей своей полноте.

[0488] 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. Иллюстративные обратные транскриптазы

[0489] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, содержат обратную транскриптазу в качестве полимеразы. Настоящее 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 мультилоскутных праймирующих редакторах согласно настоящему 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 применяться согласно настоящему изобретению, также могут быть сконструированы для включения определенных аминокислотных замещений, например, тех, которые

конкретно указаны в настоящем документе.

[0490] 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 гРНК. Данные ферменты, которые синтезируют комплементарную ДНК (кДНК) с использованием гРНК 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 and Le Grice et al., 1991, J. Virol. 65: 7004-07, каждая из которых включена посредством ссылки). Недавно были созданы мутанты и слитые белки в попытке улучшить свойства, такие как термостабильность, точность и активность. В настоящем документе 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 техники способов.

[0491] 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 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) (каждая из которых включена в настоящий документ посредством ссылки во всей своей полноте).

RT дикого типа

[0492] Иллюстративные ферменты для применения в описанных в настоящем документе мультилоскутных праймирующих редакторах могут включать без ограничения обратную транскриптазу M-MLV и обратную транскриптазу RSV. Ферменты, имеющие активность обратной транскриптазы, являются коммерчески доступными. Согласно определенным вариантам осуществления обратная транскриптаза представлена в транс относительно других компонентов системы мультилоскутного праймирующего редактора (PE). То есть обратная транскриптаза экспрессируется или иным образом предоставляется в виде отдельного компонента, то есть не в виде слитого белка с napDNAbp.

[0493] Специалисту в данной области техники очевидно, что обратные транскриптазы дикого типа, включая 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 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 вируса саркомы Рауса (RAV) и обратную транскриптазу миелобластоз-ассоциированного вируса (MAV), могут подходящим образом использоваться в описанных в настоящем документе способах и композициях согласно настоящему изобретению.

[0494] Иллюстративные ферменты RT дикого типа являются следующими:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ОБРАТНАЯ ТРАНСКРИПТАЗА (M-MLV RT) ДИКОГО ТИПА ВИРУС ЛЕЙКЕМИИ МЫШЕЙ МОЛОНИ ИСПОЛЬЗУЕТСЯ В PE1 (СЛИТЫЙ БЕЛОК ПРАЙМИРУЮЩЕГО	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA PLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWN TPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPS HQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRL PQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELD CQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRW LTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYP LTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEK QGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALL LDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTD

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
РЕДАКТОРА 1, РАСКРЫТОГО В НАСТОЯЩЕМ ДОКУМЕНТЕ)	QPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLL TSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSP(SEQ ID NO: 700)
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУС ЛЕЙКЕМИИ МЫШЕЙ МОЛОНИ ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ AAA66622.1	AFPLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNLAKVK GITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSA PDIGRKLGRLEDLKSKTLGDLVREAEKIFNKRETPEEREERIRRETEE KEERRRTVDEQKEKERDRRRHREMSKLLATVVIGGQEQDRQEGERK RPQLDKDQCAYCKEKGHWAKDCPKKPRGPRGPRPQTSLLTLGDXG GQGQDPPPEPRITLKVGGQPVTFLVDTGAQHSVLTQNPGPLSDKSA WVQGATGGKRYRWTTDRKVHLATGKVTHSFLHVPDCPYPLLGRDL LTKLKAQIHFEGSGAQVVGPMGQPLQVLTLNIEDEYRLHETSKEPDV SLGFTWLSDFPQAWAESGGMGLAVRQAPLIIPLKATSTPVSIKQYPM SQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQ DLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRD LADFR(SEQ ID NO: 90)
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИРУС ЛЕЙКЕМИИ КОШЕК ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ NP955579.1	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAHCQA PVLIQLKATATPISIRQYPMPHEAYQGIKPHIRRMLDQGILKPCQSPW NTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHPTVPNPYNLLSTLPP SHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPEIGLSGQLTWTRL PQGFKNSPTLFDEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTE CLEGTKALLETLGNKGYRASAKKAQICLQEVTYLGYSLKDGQRWL TKARKEAILSIPVPKNSRQVREFLGTAGYCRLWIPGFAELAAPLYPLT RPGTLFQWGTEQQLAFEDIKKALLSSPALGLPDITKPFELFIDENSGF AKGVLVQKLGPWKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKD AGKLTLGQPLTILTSHPVEALVRQPPNKWLSNARMTHYQAMLLDAE RVHFGPTVSLNPATLLPLPSGGNHHDCLQILAETHGTRPDLTDQPLPD ADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSAQRAEL IALTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRRGLLTSEGKE IKNKNEILALLEALFLPKRLSIIHCPGHQKGDSPQAKGNRLADDTAK KAATETHSSLTVL SEQ ID NO: 91)
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИЧ-1 RT, ЦЕПЬ A	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKIS KIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIP HPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIR YQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYV

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ITL3-A	GSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELH PDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLXK LLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEI QKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQ KITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNT PPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRGRQ KVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPD QSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRK V(SEQ ID NO: 92) SEE MARTINELLI ET AL., VIROLOGY, 1990, 174(1): 135-144, WHICH IS INCORPORATED BY REFERENCE
ОБРАТНАЯ ТРАНСКРИПТАЗА ВИЧ-1 RT, ЦЕПЬ B ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ITL3-B	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKIS KIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIP HPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIR YQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYV GSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELH PDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCK LLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEI QKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQ KITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNT PPLVKLWYQLEKEPIVGAETF(SEQ ID NO: 93) SEE STAMMERS ET AL., J. MOL. BIOL., 1994, 242(4): 586-588, WHICH IS INCORPORATED BY REFERENCE
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИРУСА САРКОМЫ РАУСА ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ACL14945	TVALHLAIPLKWKPNHTPVWIDQWPLPEGKLVALTQLVEKELQLGHI EPSLSCWNTPVFVIRKASGSYRLLHDLRAVNAKLVPFGAVQQGAPV LSALPRGWPLMVLDLKDCFFSIPLAEQDREAFAFTLPSVNNQAPARR FQWKVLPQGMTCSPTICQLIVGQILEPLRLKHPSLRMLHYMDDLLL AASSHDGLEAAGEEVISTLERAGFTISPDKVQKEPGVQYLGYKLGST YAAPVGLVAEPRIATLWDVQKLVGSLQWLRPALGIPPRLRGPFYEQL RGSDPNEAREWNLDMKMAWREIVQLSTTAALERWDPALPLEGAVA RCEQGAIGVLGQGLSTHPRPCLWLFSTQPTKAFTAWLEVLTLLITKL RASAVRTFGKEVDILLLPACFRDELPLPEGILLALRGFAGKIRSSDTPS IFDIARPLHVSLKVRVTDHPVPGPTVFTDASSSTHKGVVVWREGPR WEIKEIADLGASVQQLEARAVAMALLLWPTTPTNVVTDSAFVAKML LKMGQEGVPSTAAAFILEDALSQRSAMAAVLHVRSHSEVPGFFTEG NDVADSQATFQAYPLREAKDLHTALHIGPRALSKACNISMQQAREV

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	VQTCPHCNSAPALEAGVNPRGLGPLQIWQTDFTLEPRMAPRSWLAV TVDTASSAIVVTQHGRVTSVAAQHHWATVIAVLGRPKAIKTDNGSCF TSKSTREWLARWGIAHTTGIPGNSQGQAMVERANRLLKDKIRVLAE GDGFMKRIPTSKQGELLAKAMYALNHFERGENTKTPIQKHWRPTVL TEGPPVKIRIETGEWEKGWNVLVWGRGYAAVKNRDTDKVIWVPSR KVKPDIAQKDEVTKKDEASPLFA(SEQ ID NO: 94) SEE YASUKAWA ET AL., J. BIOCHEM. 2009, 145(3): 315-324, WHICH IS INCORPORATED BY REFERENCE
ОБРАТНАЯ ТРАНСКРИПТАЗА RT ВИРУСА МОЗАИКИ ЦВЕТНОЙ КАПУСТЫ ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ AGT42196	MMDHLLQKTQIQNQTEQVMNITNPNSIYIKGRLYFKGYKKIELHCF VDTGASLCIASKFVIPEEHWINAERPIMVKIADGSSITINKVCRDIDLII AGEIFHIPTVYQQESGIDFIIGNNFCQLYEPFIQFTDRVIFTKDRTYPVH IAKLTRAVRVGTEGFLESMKKRSKTQQPEPVNISTNKIAILSEGRRLS EEKLFITQQRMQKIEELLEKVCSENPLDPNKTKQWMKASIKLSDPSK AIKVKPMKYSPMDREEFDKQIKELLDLKVIKPSKSPHMAPAFLVNNE AEKRRGKKRMVVNYKAMNKATVGDAYNLPNKDELLTLIRGKKIFS SFDCKSGFWQVLLDQDSRPLTAFTCPQGHYEWNVVPFGLKQAPSIF QRHMDEAFRVFRKFCCVYVDDILVFSNNEEDHLLHVAMILQKCNQH GIILSKKKAQLFKKKINFLGLEIDEGTHKPQGHILEHINKFPDTLEDK KQLQRFLGILTYASDYIPKLAQIRKPLQAKLKENVPWKWTKEDTLY MQKVKKNLQGFPPLHHPLPEEKLIIETDASDDYWGGMLKAIKINEG TNTELICRYASGSFKAAEKNYHSNDKETLAVINTIKKFSIYLTPVHFLI RTDNTHFKSFVNLNYKGDSKLGRNIRWQAWLSHYSFDVEHIKGTD NHFADFLSREFNRVNS(SEQ ID NO: 95) SEE FARZADFAR ET AL., VIRUS GENES, 2013, 47(2): 347-356, WHICH IS INCORPORATED BY REFERENCE
ОБРАТНАЯ ТРАНСКРИПТАЗА KLEBSIELLA PNEUMONIA ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ RFF81513.1	MKEKISKIDKNFYTDIFIKTSFQNEFEAGGVIPPIAKNQVSTISNKNKT FYSLAHSSPHYSIQTRIEKFLLKNIPLSASSFAFRKERSYLHYLEPHTQ NVKYCHLDIVSFFHSIDVNIVRDTFSVYFSDEFLVKEKQSLLDAFMA SVTLTAELDGVEKTFIPMGFKSSPSISNIIFRKIDILIQKFCDKNKITYT RYADDLLFSTKKENNILSSTFFINEISSILSINKFKLNKSKYLYKEGTIS LGGYVIENILKDNSSGNIRLSSSKLNPLYKALYEIKKGSSSKHICIKVF NLKLKRFIYKKNKEKFEAKFYSSQLKNKLLGYRSYLLSFVIFHKKYK CINPIFLEKCVFLISEIESIMNRKF(SEQ ID NO: 96)
ОБРАТНАЯ	MKITSNNVTAVINGKGWHSINWKKCHQHVKTIQTRIAKAACNQQW

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
ТРАНСКРИПТАЗА RT ESCHERICHIA COLI ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ TGH57013	RTVGRLQRLLVRSFSARALAVKRVTENSGRKTPGVDGQIWSTPESK WEAIFKLRRKGYKPLPLKRVFIPKSNGKKRPLGIPVMLDRAMQALH LLGLEPVSETNADHNSYGFRPARCTADAIQQVCNMYSSRNASKWVL EGDIKGCFEHISHEWLLENIPMDKQILRNWLKAGIIEKSIFSKTLSGTP QGGIISPVLANMALDGLERLLQNRFGRNRLI(SEQ ID NO: 97)
ОБРАТНАЯ ТРАНСКРИПТАЗА RT BACILLUS SUBTILIS ЭТАЛОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ QBJ66766	MSKIKINYEKYHIKPFPHFDQRIKVNKKVKENLQNPFYIAAHSFYPFI HYKKISYKFKNGTLSSPKERDIFYSGHMDGYIYKHYGEILNHKYNN TCIGKGIDHVSLAYRNNKMGKSNIHFAAEVINFISEQQQAFIFVSDFS SYFDSLDHAILKEKLIEVLEEQDKLSKDWWNVFKHITRYNWVEKEE VISDLECTKEKIARDKKSRERYYTPAEFREFRKRVNIKSNDTGVGIPQ GTAISAVLANVYAIDLDQKLNQYALKYGGIYRRYSDDIIMVLPMTSD GQDPSNDHVSFIKSVVKRNKVTMGDSKTSVLYYANNNIYEDYQRK RESKMDYLGFSFDGMTVKIREKSLFKYYHRTYKKINSINWASVKKE KKVGRKKLYLLYSHLGRNYKGHGNFISYCKKAHAVFEGNKKIESLI NQQIKRHWKKIQKRLVDV(SEQ ID NO: 98)
RT ИНТРОНА ГРУППЫ II EUBACTERIUM RECTALE	DTSNLMEQILSSDNLNRAYLQVVRNKGAEGVDGMKYTELKEHLAK NGETIKGQLRTRKYKPQPARRVEIPKPDGGVRNLGVPTVTDRFIQQA IAQVLTPIYEEQFHDHSYGFRPNRCAQQAILTALNIMNDGNDWIVDI DLEKFFDTVNHDKLMTLIGRTIKDGDVISIVRKYLVSGIMIDDEYEDS IVGTPQGGNLSPLLANIMLNELDKEMEKRGLNFVRYADDCIIMVGSE MSANRVMRNISRFIEEKLGLKVNMTKSKVDRPSGLKYLGFGFYFDP RAHQFKAKPHAKSVAKFKKRMKELTCRSWGVSNSYKVEKLNQLIR GWINYFKIGSMKTLCKELDSRIRYRLRMCIWKQWKTPQNQEKNLV KLGIDRNTARRVAYTGKRIAYVCNKGAVNVAISNKRLASFGLISMLD YYIEKCVTC(SEQ ID NO: 99)
RT ИНТРОНА ГРУППЫ II GEOBACILLUS STEAROTHERMOPHILUS	ALLERILARDNLITALKRVEANQGAPGIDGVSTDQLRDYIRAHWSTI HAQLLAGTYRPAPVRRVEIPKPGGGTRQLGIPTVVDRLIQQAILQELT PIFDPDFSSSSFGFRPGRNAHDAVRQAQGYIQEGYRYVVDMDLEKFF DRVNHDILMSRVARKVKDKRVLKLIRAYLQAGVMIEGVKVQTEEGT PQGGPLSPLLANILLDDLDKELEKRGLKFCRYADDCNIYVKSLRAGQ RVKQSIQRFLEKTLKLKVNEEKSAVDRPWKRAFLGFSFTPERKARIR LAPRSIQRLKQRIRQLTNPNWSISMPERIHRVNQYVMGWIGYFRLVE TPSVLQTIEGWIRRRLRLCQWLQWKRVRTRIRELRALGLKETAVMEI ANTRKGAWRTTKTPQLHQALGKTYWTAQGLKSLTQR(SEQ ID NO:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	100)

Вариант RT и RT сниженной точности

[0495] Обратная 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 признаки, включая их термостабильность, процессирование (скорость встраивания dNTP), и точность (или процент ошибок). Предусмотренные в настоящем документе 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).

[0496] Согласно различным вариантам осуществления обратная транскриптаза может представлять собой вариант обратной транскриптазы. В контексте настоящего изобретения термин «вариант обратной транскриптазы» включает любой встречающийся в природе или генетически сконструированный вариант, содержащий 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 двухцепочечную кДНК. В ретровирусах и ретротранспозонах данная кДНК затем может встраиваться в геном хозяина, из которого могут быть получены новые копии РНК посредством 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.

[0497] Специалисту в данной области очевидно, что варианты обратных транскриптаз, являющиеся производными из других обратных транскриптаз, включая 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 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 MC AV 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 описанных в настоящем документе способов и композиции согласно настоящему изобретению.

[0498] 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®).

[0499] Кроме того, мутантные обратные транскриптазы могут быть получены с

помощью мутации в виде вставки или усечения (N-концевых, внутренних, или С- концевых вставок или усечений) в соответствии со способами, известными специалисту в данной области техники. В контексте настоящего изобретения термин «мутация» означает замещение остатка в последовательности, например, нуклеиновой кислоты или аминокислотной последовательности, другим остатком, или делецию или вставку одного или нескольких остатков в последовательности. Мутации, как правило, описаны в настоящем документе на основе указания исходного остатка и далее положения остатка в последовательности и на основе указания нового замещенного остатка. Различные способы осуществления указанных в настоящем документе аминокислотных замещений (мутаций) хорошо известны в данной области техники 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.

[0500] Старые известные в данной области техники способы сайт-направленного мутагенеза основаны на субклонировании подвергаемой мутированию последовательности в вектор, например, вектор на основе бактериофага М13, который обеспечивает изолирование одноцепочечной матрицы ДНК. В этих способах 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.

[0501] 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 этих способах 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 конечную стадию шлифовки в процедуру перед лигированием тупого конца, полученного с помощью ПЦР мутантного продукта.

[0502] В данной области техники 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.

[0503] Пример способа случайного мутагенеза представляет собой так называемый «способ на основе ПЦР сниженной точности». 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.

[0504] Согласно различным аспектам RT мультилоскутных праймирующих редакторов может представлять собой вариант обратной транскриптазы «сниженной точности». Могут применяться обратные транскриптазы сниженной точности, которые

известны и/или доступны в данной области техники. 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., «DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1,» *Nucleic Acids Res*, 2001, 29: 2217-2222, которая включена посредством ссылки. Таким образом, в контексте настоящего изобретения обратные транскриптазы, считающиеся «сниженной точности» или считающиеся имеющими «подверженную ошибкам точность», соответствуют тем, которые имеют величину ошибок в виде менее чем одной ошибки на 15000 синтезированных нуклеотидов.

[0505] Обратная транскриптаза сниженной точности 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 в настоящем документе значением термин «непрерывная эволюция с помощью фагов (PACE)» означает 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, опубликованной как WO 2010/028347 11 марта 2010, международной PCT заявке, PCT/US2011/066747, поданной 22 декабря 2011, опубликованной как WO 2012/088381, поданной 28 июня 2012, заявке США, патенте США № 9023594, поданном 5 мая, 2015, международной PCT заявке, PCT/US2015/012022, поданной 20 января 2015, опубликованной как WO 2015/134121, поданной 11 сентября 2015, и международной PCT заявке, PCT/US2016/027795, поданной 15 апреля 2016, опубликованной как WO 2016/168631, поданной 20 октября 2016, полные содержания каждой из которых включены в настоящий документ посредством ссылки.

[0506] Обратные транскриптазы сниженной точности также могут быть получены с помощью «прерывистой эволюции с помощью фагов (PANCE)», которая в соответствии с используемым в настоящем документе значением означает 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.

[0507] В литературе были описаны другие обратные транскриптазы сниженной точности, каждая из которых предусмотрена для применения в рамках описанных в настоящем документе способов и композиций. Например, подверженные ошибкам обратные транскриптазы были описаны в Bebenek et al., "Error-prone Polymerization by HIV-1 Reverse Transcriptase," *J Biol ('hem,* 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 в настоящем документе 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, ультрачистую обратную транскриптазу MMLV GPR, все от TAKARA BIO USA, INC. (ранее CLONTECH).

[0508] Настоящее 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 H может являться нежелательной для синтеза длинных кДНК, поскольку РНК-матрица может расщепляться до завершения полноразмерной обратной транскрипции. Активность РНКазы H 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 H.

[0509] Настоящее 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 нуклеотидных 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.

[0510] Также в настоящем документе предусмотрены варианты 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'-лоскута оцДНК в результате процесса праймированного 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.

[0511] Предусмотренные в настоящем документе варианты обратных транскриптаз, включая RT сниженной точности, 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.

[0512] Используемые согласно настоящему изобретению ферменты в виде вариантов RT могут также включать другие «варианты RT», на по меньшей мере приблизительно 70% идентичные, по меньшей мере приблизительно 80% идентичные, по меньшей мере приблизительно 90% идентичные, по меньшей мере приблизительно 95% идентичные, по меньшей мере приблизительно 96% идентичные, по меньшей мере приблизительно 97% идентичные, по меньшей мере приблизительно 98% идентичные, по меньшей мере приблизительно 99% идентичные, по меньшей мере приблизительно 99.5% идентичные или по меньшей мере приблизительно 99.9% идентичные любому исходному белку RT, включая RT дикого типа, или мутантную RT, или фрагмент RT, или другой вариант RT, описанный или охватываемый в настоящем документе или известный в данной области техники.

[0513] Согласно некоторым вариантам 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 до400, или до 500 или более аминокислотных изменений по сравнению с эталонной RT. Согласно некоторым вариантам осуществления вариант RT содержит фрагмент эталонной RT, такой как фрагмент, который на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере приблизительно 99.5% идентичен или по меньшей мере приблизительно 99.9% идентичен соответствующему фрагменту эталонной RT. Согласно некоторым вариантам осуществления фрагмент составляет на по меньшей мере 30%, по меньшей мере 35%, по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, 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% идентичен, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или по меньшей мере 99.5% от длины аминокислот соответствующего дикого типа RT (M-MLV обратная транскриптаза) (*например*, SEQ ID NO: 89) или любых последовательностей обратной транскриптазы SEQ ID NO: 90-100.

[0514] Согласно некоторым вариантам осуществления также могут использоваться фрагменты RT, которые сохраняют свою функциональность и которые представляют собой фрагменты любых описанных в настоящем документе белков RT. Согласно некоторым

вариантам осуществления фрагмент RT составляет по меньшей мере 100 аминокислот в длину. Согласно некоторым вариантам осуществления фрагмент составляет по меньшей мере 100, 150, 200, 250, 300, 350, 400, 450, 500, 550 или до 600 или более аминокислот в длину.

[0515] Согласно другим вариантам осуществления 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. Согласно некоторым вариантам осуществления полимеразную функцию. В некоторых вариантах осуществления усеченный 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. Согласно другим вариантам 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 C-конце белка. Согласно другим вариантам осуществления усеченный вариант RT имеет усечения на N-конце и C-конце, которые имеют одинаковую или различающуюся длину.

[0516] Например, мультилоскутные праймирующие редакторы, раскрытые в настоящем документе, могут включать усеченную версию обратной транскриптазы M-MLV. Согласно этому варианту 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 п.о. 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)	*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIP LKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVK KPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLK DAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEA LHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLR EFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQA LLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLD PVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDN SRLIN* (SEQ ID NO: 766)

[0517] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы раскрытые в настоящем документе могут содержать один из вариантов RT, описанных в настоящем документе, или их вариант RT, который на по меньшей мере приблизительно 70% идентичен, по меньшей мере приблизительно 80% идентичен, по меньшей мере приблизительно 90% идентичен, по меньшей мере приблизительно 95% идентичен, по меньшей мере приблизительно 96% идентичен, по меньшей мере приблизительно 97% идентичен, по меньшей мере приблизительно 98% идентичен, по меньшей мере приблизительно 99% идентичен, по меньшей мере приблизительно 99.5% идентичен или по меньшей мере приблизительно 99.9% идентичен любым эталонным вариантам Cas9.

[0518] Согласно другим вариантам осуществления в способах и композициях согласно настоящему изобретению 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 которой включено в настоящий документ посредством ссылки.

[0519] Согласно определенным другим вариантам осуществления обратная транскриптаза представлена как компонент слитого-белка, также содержащий napDNAbp. Другими словами, согласно некоторым вариантам осуществления обратная транскриптаза слита с napDNAbp в виде слитого-белка.

[0520] Согласно различным вариантам осуществления варианты обратной транскриптазы могут быть сконструированы из обратной транскриптазы M-MLV дикого

типа, как представлено SEQ ID NO: 89.

[0521] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) могут включать вариант RT, содержащий 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 последовательности SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа.

[0522] Некоторые примеры обратных транскриптаз, которые могут быть слиты с белками napDNAbp или быть представлены в виде отдельных белков в соответствии с различными вариантами осуществления настоящего изобретения, приведены ниже. Примеры обратных транскриптаз включают варианты, имеющие по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичности последовательности со следующими ферментами дикого типа или частичными ферментами:

Описание	Последовательность (замещения варианта относительно дикого типа)
Обратная транскриптаза (M-MLV RT) дикого типа вирус лейкоза мышей Молони Используется в PE1 (слитый белок праймирующего редактора 1, раскрытый в настоящем документе)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 89)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR

Описание	Последовательность (замещения варианта относительно дикого типа)
	LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 701)
M-MLV RT D200N T330P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 702)
M-MLV RT D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA

Описание	Последовательность (замещения варианта относительно дикого типа)
	LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 740)
M-MLV RT D200N T330P L603W E69K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQ**K**ARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 703)
M-MLV RT D200N T330P L603W E302R	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLR**R**FLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 704)
M-MLV RT D200N T330P L603W E607K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE

Описание	Последовательность (замещения варианта относительно дикого типа)
	LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 705)
M-MLV RT D200N T330P L603W L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 706)
M-MLV RT D200N T330P L603W L435G	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVI**G**APHAVEALVKQPPDRWLSNARMTHYQALLLDT DRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 707)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN

Описание	Последовательность (замещения варианта относительно дикого типа)
T330P L603W N454K	DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLS**K**ARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 708)
M-MLV RT D200N T330P L603W T306K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLG**K**AGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 709)
M-MLV RT D200N T330P L603W W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL**F**IPG FAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFEL FVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTW YTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA

Описание	Последовательность (замещения варианта относительно дикого типа)
	EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 710)
M-MLV RT D200N T330P L603W D524G E562Q D583N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYT**G**GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRA**Q**LIALTQALK MAEGKKLNVYT**N**SRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 711)
M-MLV RT D200N T330P L603W E302R W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLR**R**FLGTAGFCRL**F**IPG FAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFEL FVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTW YTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 712)
M-MLV RT D200N T330P L603W E607K L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIP

Описание	Последовательность (замещения варианта относительно дикого типа)
	GFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIA VLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTS**K**GKEIKNKDEILALLKA LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 713)
M-MLV RT P51L S67K T197A H204R E302K F309N W313F T330P L435G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII**L**LK ATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT NDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCL RLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP**A**LFDEAL**R**RDLADFR IQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQ KQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLR**K**FLGTAG**N**CRL**F**I PGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPF ELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAI AVLTKDAGKLTMGQPLVI**G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLD TDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYT**G**GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 714)
M-MLV RT D200N P51L S67K T197A H204R E302K F309N W313F T330P L345G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII**L**LK ATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT NDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCL RLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP**A**LF**N**EAL**R**RDLADFR IQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQ KQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLR**K**FLGTAG**N**CRL**F**I PGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPF ELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAI AVLTKDAGKLTMGQPLVI**G**APHAVEALVKQPPDRWLS**K**ARMTHYQALLLD TDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYT**G**GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK MAEGKKLNVYT**N**SRYAFATAHI**Q**GEIYRRRGLLTSEGKEIKNKDEILALLKAL FLPKRLSIIHCPGHQKGHSAEARGNRMA**N**QAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 715)
M-MLV RT	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK

Описание	Последовательность (замещения варианта относительно дикого типа)
D200N T330P L603W T306K W313F в PE2	ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLG**K**AGFCRL**F**IPG FAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFEL FVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTW YTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 716)

[0523] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) могут включать вариант RT, содержащий 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 последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту.

[0524] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий P51X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой L.

[0525] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий S67X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным

вариантам осуществления X представляет собой K.

[0526] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий E69X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой K.

[0527] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий L139X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой P.

[0528] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий T197X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой A.

[0529] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий D200X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой N.

[0530] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий H204X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным

вариантам осуществления X представляет собой R.

[0531] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий F209X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой N.

[0532] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий E302X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой K.

[0533] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий E302X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой R.

[0534] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий T306X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой K.

[0535] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий F309X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным

вариантам осуществления X представляет собой N.

[0536] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий W313X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой F.

[0537] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий T330X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой P.

[0538] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий L345X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой G.

[0539] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий L435X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой G.

[0540] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий N454X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным

вариантам осуществления X представляет собой K.

[0541] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий D524X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой G.

[0542] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий E562X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой Q.

[0543] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий D583X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой N.

[0544] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий H594X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой Q.

[0545] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий L603X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным

вариантам осуществления X представляет собой W.

[0546] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий E607X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой K.

[0547] Согласно различным другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, (с RT либо в качестве партнера слияния, либо *в транс*) может включать вариант RT, содержащий D653X мутацию в последовательности RT M-MLV дикого типа SEQ ID NO: 89 или в соответствующем аминокислотном положении в другой полипептидной последовательности RT дикого типа, где «X» может представлять собой любую аминокислоту. Согласно определенным вариантам осуществления X представляет собой N.

[0548] Некоторые примеры обратных транскриптаз, которые могут быть слиты с белками napDNAbp или быть представлены в виде отдельных белков в соответствии с различными вариантами осуществления данного изобретения, приведены ниже. Примеры обратных транскриптаз включают варианты, имеющие по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или по меньшей мере 99% идентичности последовательности со следующими ферментами дикого типа или частичными ферментами:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
ОБРАТНАЯ ТРАНСКРИПТАЗА (M-MLV RT) ДИКОГО ТИПА ВИРУС ЛЕЙКОЗА МЫШЕЙ МОЛОНИ, ИСПОЛЬЗУЕТСЯ В PE1 (СЛИТЫЙ БЕЛОК	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFD EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
ПРАЙМИРУЮЩЕГО РЕДАКТОРА 1, РАСКРЫТЫЙ В НАСТОЯЩЕМ ДОКУМЕНТЕ)	GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 89)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 106)
M-MLV RT D200N T330P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 107)
M-MLV RT D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 108)
M-MLV RT E69K D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQ**K**ARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 109)
M-MLV RT D200N T330P L603W E302R	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLR**R**FLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVK QPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN CLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHI HGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA EARGNRMADQAARKAAITETPDTSTLLIENSSP(SEQ ID NO: 110)
M-MLV RT D200N T330P L603W E607K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTS**K**GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 111)
M-MLV RT D200N T330P L603W L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 112)
M-MLV RT D200N T330P L603W L435G	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI**G**APHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 113)
M-MLV RT D200N T330P L603W N454K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 114)
M-MLV RT D200N T330P L603W T306K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLG**K**AGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVK QPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN CLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHI HGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	EARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 115)
M-MLV RT D200N T330P L603W W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 116)
M-MLV RT D200N T330P L603W D524G E562Q D583N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRA**Q**LIALTQALKMAEGKKLNVYT**N**SRYAFATAHI HGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA EARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 117)
M-MLV RT D200N T330P L603W E302R W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLR**R**FLGTAGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
	QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 118)
M-MLV RT D200N T330P L603W E607K L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG**P**PPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTS**K**GKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 119)
M-MLV RT P51L S67K T197A H204R E302K F309N W313F T330P L435G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII **L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP**A**LFD EAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI**G**APHAVEALVK QPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN CLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYT**N**SRYAFATAHI **Q**GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA EARGNRMA**N**QAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 120)
M-MLV RT P51L	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (ЗАМЕЩЕНИЯ ВАРИАНТА ОТНОСИТЕЛЬНО ДИКОГО ТИПА)
S67K T197A D200N H204R E302K F309N W313F T330P L345G N454K D524G D583N H594Q D653N	**L**LKATSTPVSIKQYPM**K**QEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSP**A**LF**N** EAL**R**RDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLR**K**FLGTAG**N**CRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI**G**APHAVEALVK QPPDRWLS**K**ARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN CLDILAEAHGTRPDLTDQPLPDADHTWYT**G**GSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYT**N**SRYAFATAHI **Q**GEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA EARGNRMA**N**QAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 121)
M-MLV RT D200N T330P L603W T306K W313F IN PE2	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPV KKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N** EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLG**K**AGFCRL**F**IPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIK QALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIH GEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 122)

[0549] Система мультилоскутного праймирующего редактора, описанного в настоящем документе, охватывает любую общедоступную обратную транскриптазу, описанную и раскрытую в любом из следующих патентов США (каждый из которых включен в настоящий документ посредством ссылки во всей своей полноте): Патент США № 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 техники. Содержание каждой из них полностью

включено в настоящий документ посредством ссылки.

[0550] 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).

[0551] 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).

[0552] Zhao, C., Liu, F. & Pyle, A. M. An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA* **24**, 183–195 (2018).

[0553] Zimmerly, S. & Wu, L. An Unexplored Diversity of Reverse Transcriptases in Bacteria. *Microbiol Spectr* **3**, MDNA3-0058–2014 (2015).

[0554] Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. *Annual Review of Genetics* **35**, 501–538 (2001).

[0555] Perach, M. & Hizi, A. Catalytic Features of the Recombinant Reverse Transcriptase of Bovine Leukemia Virus Expressed in Bacteria. *Virology* **259**, 176–189 (1999).

[0556] Lim, D. *et al.* Crystal structure of the moloney murine leukemia virus RNase H domain. *J. Virol.* **80**, 8379–8389 (2006).

[0557] 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).

[0558] Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biol.* **2**, REVIEWS1017 (2001).

[0559] 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).

[0560] 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).

[0561] 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).

[0562] 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).

[0563] 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).

[0564] 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).

[0565] 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).

[0566] Xiong, Y. & Eickbush, T. H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* **9**, 3353–3362 (1990).

[0567] Herschhorn, A. & Hizi, A. Retroviral reverse transcriptases. *Cell. Mol. Life Sci.* **67**, 2717–2747 (2010).

[0568] 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).

[0569] Liu, M. *et al.* Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science* **295**, 2091–2094 (2002).

[0570] 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).

[0571] 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).

[0572] 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).

[0573] 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).

[0574] 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).

[0575] 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).

[0576] Das, D. & Georgiadis, M. M. The Crystal Structure of the Monomeric Reverse Transcriptase from Moloney Murine Leukemia Virus. *Structure* **12**, 819–829 (2004).

[0577] 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).

[0578] 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).

[0579] Monot, C. *et al.* The Specificity and Flexibility of L1 Reverse Transcription Priming at Imperfect T-Tracts. *PLOS Genetics* **9**, e1003499 (2013).

[0580] 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).

[0581] Любая из приведенных выше ссылок, связанная с обратными транскриптазами, полностью включена в настоящий документ посредством ссылки, если иное еще не было указано.

[4] Праймирующие редакторы

[0582] Настоящее изобретение относится к системе, содержащей праймирующие редакторы. Согласно одному аспекту настоящее изобретение относится к системам одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, содержащим первый комплекс праймированного редактирования и второй комплекс праймированного редактирования, где каждый из первого и второго комплексов праймированного редактирования содержит (1) праймирующий редактор, содержащий (i) программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и (ii) полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и (2) РЕгРНК, содержащую спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера, где матрица для синтеза ДНК РЕгРНК первого комплекса праймированного редактирования кодирует первую последовательность одноцепочечной ДНК, и матрица для синтеза ДНК РЕгРНК второго комплекса праймированного редактирования кодирует вторую последовательность одноцепочечной ДНК, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть по сравнению с последовательностью ДНК в целевом участке, подлежащем редактированию, который вставляется в целевой участок, подлежащий редактированию.

[0583] Согласно другому аспекту настоящее изобретение относится к системам

редактирования одной или нескольких последовательностей двухцепочечной ДНК, причем система содержит:

a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем редактированию,

b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем редактированию,

c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с первым участком связывания на первой цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается со вторым участком связывания на второй цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, третий РЕгРНК содержит третью матрицу для синтеза ДНК, кодирующую третью

последовательность одноцепочечной ДНК, и четвертая РЕгРНК содержит четвертую матрицу для синтеза ДНК, кодирующую четвертую последовательность одноцепочечной ДНК,

где первая и третья последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где вторая и четвертая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

[0584] Согласно некоторым вариантам осуществления праймирующие редакторы, используемые в системах, описанных в настоящем документе (*например*, первый праймирующий редактор, второй праймирующий редактор, третий праймирующий редактор и/или четвертый праймирующий редактор в системах, описанных выше) представлены в виде слитых белков. Согласно некоторым вариантам осуществления слитые белки праймирующего редактора содержат napDNAbp и полимеразу (например, ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, такую как обратная транскриптаза). Согласно некоторым вариантам осуществления napDNAbp и полимераза необязательно соединены линкером с образованием слитого белка. Различные конфигурации слитых белков праймирующего редактора и дополнительных доменов слитых белков описаны далее в настоящем документе. Согласно различным вариантам осуществления системы и способы, представленные в настоящем описании, предусматривают использование любого из слитых белков праймирующего редактора, описанных в настоящем документе.

[0585] Согласно некоторым вариантам осуществления комплексы праймированного редактирования, используемые в системах, описанных в настоящем документе, содержат праймирующий редактор (например, первый праймирующий редактор, второй праймирующий редактор, третий праймирующий редактор и/или четвертый праймирующий редактор в системах, описанных выше), где компоненты одного или нескольких праймирующих редакторов представлены *в транс*, как дополнительно подробно описано в настоящем документе. Согласно некоторым вариантам осуществления праймирующий редактор содержит napDNAbp и полимеразу, экспрессируемую *в транс*. Согласно некоторым вариантам осуществления napDNAbp и полимераза экспрессируются из одного или нескольких векторов (*например*, оба компонента экспрессируются из одного и того же вектора или каждый компонент экспрессируется из другого вектора). Согласно определенным вариантам осуществления праймирующие редакторы содержат дополнительные компоненты, как описано в настоящем документе, экспрессируемые *в транс*. Согласно некоторым вариантам осуществления праймирующие редакторы, применяемые в системах, описанных в настоящем документе, могут содержать как один

или несколько праймирующих редакторов, в виде слитых белков, так и один или несколько праймирующих редакторов, компоненты которых представлены *в транс*.

[0586] Согласно некоторым вариантам осуществления системы мультилоскутного праймирующего редактора, описанные в настоящем документе, охватывают слитые белки, содержащие napDNAbp и полимеразу (например, например, ДНК-зависимую ДНК-полимеразу или РНК-зависимую ДНК-полимеразу, такую как обратная транскриптаза), и необязательно соединенные линкером. Настоящее изобретение также охватывает любой подходящий napDNAbp и полимеразу (например, ДНК-зависимая ДНК-полимераза или РНК-зависимая ДНК-полимераза, такая как обратная транскриптаза), которые должны быть объединены в один слитый белок. Примеры napDNAbp и полимераз (например, ДНК-зависимая ДНК-полимераза или РНК-зависимая ДНК-полимераза, такая как обратная транскриптаза) определены в настоящем документе. Поскольку полимеразы хорошо известны в данной области техники, а аминокислотные последовательности легко доступны, настоящее раскрытие никоим образом не ограничивается конкретными полимеразами, указанными в настоящем документе.

[0587] Согласно различным вариантам осуществления слитые белки могут иметь любую подходящую структурную конфигурацию. Например, слитый белок может включать в направлении от N-конца к С-концу napDNAbp, слитый с полимеразой (например, ДНК-зависимой ДНК-полимеразой или РНК-зависимой ДНК-полимеразой, такой как обратная транскриптаза). Согласно другим вариантам осуществления слитый белок может содержать от N-конца к С-концу полимеразу (например, обратную транскриптазу), слитую с napDNAbp. Слитый домен необязательно может быть соединен линкером, например, аминокислотной последовательностью. Согласно другим вариантам осуществления слитые белки могут иметь структуру NH_2-[napDNAbp]-[полимераза]-COOH, или NH_2-[полимераза]-[napDNAbp]-COOH, где в каждом случае «]-[» указывает на присутствие необязательной последовательности линкера. Согласно вариантам осуществления, в которых полимеразой является обратная транскриптаза, слитые белки могут содержать структуру NH_2-[napDNAbp]-[RT]-COOH, или NH_2-[RT]-[napDNAbp]-COOH, где в каждом случае «]-[» указывает на присутствие необязательной последовательности линкера.

[0588] Иллюстративный слитый белок изображен на фиг. 14, на котором показан слитый белок, содержащий обратную транскриптазу MLV («MLV-RT»), слитую с никазой Cas9 («Cas9(H840A)») через линкерную последовательность. Этот пример не предназначен для ограничения объема слитых белков, которые могут быть использованы для системы праймирующего редактора (PE), описанной в настоящем документе.

[0589] Согласно различным вариантам осуществления слитый белок мультилоскутного праймирующего редактора может иметь следующую аминокислотную последовательность (упоминаемую в настоящем документе как «PE1»), которая включает вариант Cas9, содержащий мутацию H840A (т.е., никаза Cas9) и RT M-MLV дикого типа, а также N-концевую последовательность NLS (19 аминокислот) и аминокислотный линкер (32 аминокислот), который соединяет C-конец никазного домена Cas9 и N-конец домена RT. Слитый белок PE1 имеет следующую структуру: [NLS]-[Cas9(H840A)]-[линкер]-[MMLV_RT(wt)]. Аминокислотная последовательность PE1 и ее отдельные компоненты являются следующими:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
PE1 СЛИТЫЙ БЕЛОК CAS9(H840A)-MMLV_RT(WT)	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVP** **SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT** **RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF** **GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF** **RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK** **AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN** **FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI** **LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK** **YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK** **LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR** **EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD** **KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK** **YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI** **ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDI** **VLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS** **RKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ** **KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR** **HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE** **HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIV** **PQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL** **LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV** **AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVR** **EINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK** **MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN** **GETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL** **PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK** **KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY**

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVLDATLIHQSITGLYETRIDLSQLGGD<u>**SGGSSGGSSGSETPGT SESATPESSGGSSGGSS**</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQ GILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYN LLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSEL DCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEA RKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTL FNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLV ILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATL LPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEG QRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTD SRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCP GHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP*<u>SGGSKRTAD GSEFEPKKKRKV</u> (SEQ ID NO: 123)

ОБОЗНАЧЕНИЯ:
<u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS)</u> TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133)
CAS9(H840A) (SEQ ID NO: 126)
<u>ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ</u> (SEQ ID NO: 127)
M-MLV ОБРАТНАЯ ТРАНСКРИПТАЗА (SEQ ID NO: 128) |
| PE1 – N-КОНЦЕВАЯ NLS | MKRTADGSEFESPKKKRKV (SEQ ID NO: 124) |
| PE1 – CAS9 (H840A) (MET MINUS) | DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIA LSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD |

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFE DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQT TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQ NGRDMYVDQELDINRLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKNR GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLE SEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLA NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKV EKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 130)
PE1 – ЛИНКЕР МЕЖДУ ДОМЕНОМ CAS9 И ДОМЕНОМ RT (33 АМИНОКИСЛОТ)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127)
PE1 – M-MLV RT	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTP LLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQ WYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQG FKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQG TRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARK ETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTL FNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGV LTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLT MGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGP VVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWY

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ
	TDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQAL KMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEIL ALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETP DTSTLLIENSSP (SEQ ID NO: 132)
PE1 – C-КОНЦЕВАЯ NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 133)

[0590] Согласно другому варианту осуществления слитый белок мультилоскутного праймирующего редактора может иметь следующую аминокислотную последовательность (обозначаемую в настоящем документе как «PE2»), которая включает вариант Cas9, содержащий мутацию H840A (т.е. никазу Cas9) и RT M-MLV, содержащую мутации D200N, T330P, L603W, T306K и W313F, а также N-концевую последовательность NLS (19 аминокислот) и аминокислотный линкер (33 аминокислоты), который соединяет C-конец домена никазы Cas9 с N-концом домена RT. Слитый белок PE2 имеет следующую структуру: [NLS]-[Cas9(H840A)]-[линкер] - [MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]. Аминокислотная последовательность PE2 является следующей:

PE2 СЛИТЫЙ БЕЛОК CAS9(H840A) -MMLV_RT D200N T330P L603W T306K W313F	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQ EIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPT IYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKL FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKN GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGT EELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGIL QTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY DVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH**

	DAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFL YLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS TKEVLDATLIHQSITGLYETRIDLSQLGGD<u>**SGGSSGGSSGSETPGTSESATP ESSGGSSGGSS**</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVR QAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVK KPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFC LRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKY LGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAP LYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAK GVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLV ILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEE GLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRG WLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARK AAITETPDTSTLLIENSSP**</u>SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 134) ОБОЗНАЧЕНИЯ: <u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS)</u> TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 137) **<u>ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ</u>** (SEQ ID NO: 127) *M-MLV ОБРАТНАЯ ТРАНСКРИПТАЗА* (SEQ ID NO: 139)
PE2 – N-КОНЦЕВАЯ NLS	MKRTADGSEFESPKKKRKV (SEQ ID NO: 124)
PE2 – CAS9 (H840A) (МЕТ МИНУС)	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARL SKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKR YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK

	PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKG ASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**A**IVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGD YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN GETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 141)
PE2 – ЛИНКЕР МЕЖДУ ДОМЕНОМ CAS9 И ДОМЕНОМ RT (33 АМИНОКИСЛОТ)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 127)
PE2 – MMLV_RT D200N T330P L603W T306K W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLG**K**AGFCRL**F**IPG FAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELF VDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLT KDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQ FGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTD

	GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK KLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRL SIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 143)
PE2 – C-КОНЦЕВАЯ NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 133)

[0591] Согласно другим вариантам осуществления слитый белок праймирующего редактора может иметь следующие аминокислотные последовательности:

PE СЛИТЫЙ БЕЛОК MMLV_RT(WT)-32AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAET GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTV LDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEAL HRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWI PGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFV DEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGK LTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNP ATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRK AGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHI HGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP**SGGSSGGSSGSETPGTSESATPESSGGS SGGSS**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENP INASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPN FKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAIL LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLP NEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKD KDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLT FKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN

	TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSID NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKY PKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITL ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIR EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGL YETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 145) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS) TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 147) **ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ** (SEQ ID NO: 127) *M-MLV ОБРАТНАЯ ТРАНСКРИПТАЗА* (SEQ ID NO: 149)
РЕ СЛИТЫЙ БЕЛОК MMLV_RT(WT)-60AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAET GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTV LDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEAL HRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWI PGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFV DEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGK LTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNP ATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRK AGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHI HGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSPS**GGSSGGSSGSETPGTSESATPESAGSY PYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS**DKKYSIGLDIGTNSVGWA VITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL ENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYD

	DDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK KDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFK YFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTAD GSEFEPKKKRKV (SEQ ID NO: 150) ОБОЗНАЧЕНИЯ: ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS) TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133) **CAS9(H840A)**(SEQ ID NO: 153) **АМИНОКИСЛОТНЫЙ ЛИНКЕР** (SEQ ID NO: 175) *M-MLV ОБРАТНАЯ ТРАНСКРИПТАЗА* (SEQ ID NO: 149)
РЕ СЛИТЫЙ БЕЛОК CAS9(H840A)-FEN1-MMLV_RT D200N T330P	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQ EIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPT IYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKL FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKN GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGT**

L603W T306K W313F	**EELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN** **REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS** **AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK** **PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR** **FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT** **YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG** **FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGIL** **QTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE** **GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY** **DVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ** **LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL** **DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH** **DAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK** **YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV** **LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD** **SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG** **YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFL** **YLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL** **DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS** **TKEVLDATLIHQSITGLYETRIDLSQLGGD<u>SGGSSGGSSGSETPGTSESATP</u>** <u>**ESSGGSSGGSS**</u>GIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIA VRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGE LAKRSERRAEAEKQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSL MGIPYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEA KKLPIQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHK SIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIK FMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKEPEPK GSTKKKAKTGAAGKFKRGK<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u> *TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATST* *PVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQ* *DLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFA* *FEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVD* *DLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRW* *LTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFN* *WGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRR* *PVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ* *PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEA* *HGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS*

	*AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNK DEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL IENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 154) KEY: ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS) TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 157) **33-АМИНОКИСЛОТНЫЙ ДИНКЕР 1** (SEQ ID NO: 127) *M-MLV ОБРАТНАЯ ТРАНСКРИПТАЗА* (SEQ ID NO: 159) **33-АМИНОКИСЛОТНЫЙ ЛИНКЕР 2** (SEQ ID NO: 127) FEN1 (SEQ ID NO: 161)

[0592] Согласно другим вариантам осуществления слитые белки мультилоскутного праймирующего редактора могут быть основаны на SaCas9 или на никаза SpCas9s с измененными специфичностями PAM, как например, следующие иллюстративные последовательности:

SACAS9-M-MLV RT ПРАЙМИРУЮЩИЙ РЕДАКТОР	MKRTADGSEFESPKKKRKVGKRNYILGLDIGITSVGYGIIDYETRDVID AGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLL TDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVE EDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKT SDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPF GWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVI TRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVT STGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEE LTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAI FNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKY GLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENA KYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFD NSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKG KGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRS YFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIA NADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITP HQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNL NGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEK NPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYP NSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNS KCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIE

	VNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVK SKKHPQIIKKGSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTLNIED EYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKA TSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKK PGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLD LKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTL FDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQT LGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQP TPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQ QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPW RRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPAT LLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQ EGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKK LNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLP KRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENS SPSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 162)
SPCAS9(H840A)-VRQR-ВИРУС МЫШИНОГО ЛЕЙКОЗА МОЛОНИ ОБРАТНАЯ ТРАНСКРИПТАЗА ПРАЙМИРУЮЩЕГО РЕДАКТОРА	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSK KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN PDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTY DDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWM TRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVT VKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRY TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDA

	YLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKK YGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGS SGGSSGSETPGTSESATPESSGGSSGGSSTLNIEDEYRLHETSKEPDVSL GSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQE ARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLRE VNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQ PLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQ HPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQ ICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKA GFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAP ALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFAT AHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQ KGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADG SEFEPKKKRKV (SEQ ID NO: 163)
SPCAS9(H840A)-VRER-ВИРУС МЫШИНОГО ЛЕЙКОЗА МОЛОНИ ОБРАТНАЯ ТРАНСКРИПТАЗА ПРАЙМИРУЮЩЕГО РЕДАКТОРА	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSK KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN PDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTY DDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWM TRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVT VKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRY

	TGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDA YLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKK YGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTIDRKEYRSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSS GGSSGSETPGTSESATPESSGGSSGGSSTLNIEDEYRLHETSKEPDVSLG STWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEA RLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREV NKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQP LFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQH PDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQI CQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKA GFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAP ALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPV AAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFAT AHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQ KGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADG SEFEPKKKRKV (SEQ ID NO: 164)

[0593] Согласно другим вариантам осуществления слитые белки мультилоскутного праймирующего редактора, охватываемые настоящим изобретением, могут содержать никазу Cas9 (например, Cas9 (H840A)), слитую с усеченной версией обратной транскриптазы M-MLV. Согласно этому варианту осуществления обратная транскриптаза также содержит 4 мутации (D200N, T306K, W313F, T330P, следует отметить, что мутация L603W, присутствующая в PE2, больше не присутствует из-за усечения). Последовательность ДНК, кодирующая этот усеченный редактор, на 522 п.о. меньше, чем

PE2, что делает ее потенциально полезным для применений, в которых доставка последовательности ДНК затруднена из-за ее размера (например, доставка аденоассоциированного вируса и лентивируса). Этот вариант осуществления обозначается как Cas9(H840A)-MMLV-RT(усеченный) или «PE2-короткий» или «PE2-усеченный» и имеет следующую аминокислотную последовательность:

CAS9(H840A)-MMLV-RT(УСЕЧЕННЫЙ) ИЛИ PE2-КОРОТКИЙ	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTK VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED IQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMG RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQ LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVR KMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD**<u>SGGSSGGSSGSETP GTSESATPESSGGSSGGSS</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQ*</u>

	AWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLL DQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNP YNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQ LTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSE LDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTE ARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGT LFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVL TQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQ PLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNP ATLLPLPEEGLQHNCLDNSRLIN<u>SGGSKRTADGSEFEPKKKRKV</u> (SEQ ID NO: 765) ОБОЗНАЧЕНИЯ: <u>ПОСЛЕДОВАТЕЛЬНОСТЬ ЛОКАЛИЗАЦИИ В ЯДРЕ (NLS)</u> TOP:(SEQ ID NO: 124), BOTTOM: (SEQ ID NO: 133) **CAS9(H840A)** (SEQ ID NO: 157) **<u>ЛИНКЕР ИЗ 33 АМИНОКИСЛОТ 1</u>** (SEQ ID NO: 127) *M-MLV TRUNCATED ОБРАТНАЯ ТРАНСКРИПТАЗА* (SEQ ID NO: 766)

[0594] См. фиг. 75, на которой представлена гистограмма, сравнивающая эффективность (т.е. «% от общего числа прочтений секвенирования со специфическим редактированием или инделями») PE2, PE2-усеченный, PE3 и PE3-усеченный по различным целевым участкам в различных клеточных линиях. Данные показывают, что праймирующие редакторы, содержащие варианты укороченной RT, были примерно так же эффективны, как и праймирующие редакторы, содержащие белки неусеченной RT.

[0595] Согласно различным вариантам осуществления слитые белки мультилоскутного праймирующего редактора, охватываемые настоящим изобретением, также могут включать любые варианты описанные выше последовательностей, имеющих аминокислотную последовательность, которая на по меньшей мере приблизительно 70% идентична, по меньшей мере приблизительно 80% идентична, по меньшей мере приблизительно 90% идентична, по меньшей мере приблизительно 95% идентична, по меньшей мере приблизительно 96% идентична, по меньшей мере приблизительно 97% идентична, по меньшей мере приблизительно 98% идентична, по меньшей мере приблизительно 99% идентична, по меньшей мере приблизительно 99.5% идентична или по меньшей мере приблизительно 99.9% идентична PE1, PE2 или любым указанным выше слитым последовательностям праймирующего редактора.

[0596] Согласно определенным вариантам осуществления линкеры могут

применяться для связывания пептидов, или пептидных доменов, или фрагментов согласно настоящему изобретению (например, napDNAbp, связанный или слитый с обратной транскриптазой).

[5] Линкеры и другие домены

[0597] Слитые белки PE могут содержать различные другие домены помимо napDNAbp (например, домен Cas9) и домен полимеразы (например, домен RT). Например, в случае, когда napDNAbp представляет собой Cas9, а полимераза представляет собой RT, слитые белки PE могут содержать один или несколько линкеров, которые соединяют домен Cas9 с доменом RT. Линкеры также могут присоединять другие функциональные домены, такие как последовательности локализации в ядре (NLS) или FEN1 (или другую эндонуклеазу лоскутов), к слитым белкам PE или их доменам.

[0598] Кроме того, в вариантах осуществления, включающих *транс*-праймированное редактирование, линкеры можно использовать для связывания рекрутирующего белка tPERT с праймирующим редактором, например, между рекрутирующим белком tPERt и napDNAbp. См., например, фиг. 3G, где приведен пример схемы *транс*-праймирующего редактора (tPE), который включает линкеры для раздельного слияния домена полимеразы и домена рекрутирующего белка с napDNAbp.

A. Линкеры

[0599] Как определено выше, в контексте настоящего изобретения термин «линкер» относится к химической группе или молекуле, связывающей две молекулы или группы, например, связывающий домен и домен расщепления нуклеазы. Согласно некоторым вариантам осуществления линкер соединяет домен связывания гРНК РНК-программируемой нуклеазы и каталитический домен полимеразы (например, обратной транскриптазы). Согласно некоторым вариантам осуществления линкер соединяет dCas9 и обратную транскриптазу. Как правило, линкер расположен между двумя группами, молекулами или другими фрагментами или окружен ими и соединен с каждым из них ковалентной связью, таким образом, соединяя их. Согласно некоторым вариантам осуществления линкер представляет собой аминокислоту или множество аминокислот (например, пептид или белок). Согласно некоторым вариантам осуществления линкер представляет собой органическую молекулу, группу, полимер или химический фрагмент. Согласно некоторым вариантам осуществления линкер составляет 5-100 аминокислот в длину, например, 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 в длину. Также рассматриваются более длинные или короткие линкеры.

[0600] Линкер может быть таким же простым, как ковалентная связь, или он может быть полимерным линкером длиной в несколько атомов. Согласно определенным вариантам осуществления линкер представляет собой полипептид или основан на аминокислотах. Согласно другим вариантам осуществления линкер не является пептидоподобным. Согласно определенным вариантам осуществления линкер представляет собой ковалентную связь (например, связь углерод-углерод, дисульфидную связь, связь углерод-гетероатом и т.д.). Согласно определенным вариантам осуществления линкер представляет собой углерод-азотную связь амидной связи. Согласно определенным вариантам осуществления линкер представляет собой циклический или ациклический, замещенный или незамещенный, разветвленный или неразветвленный алифатический или гетероалифатический линкер. Согласно определенным вариантам осуществления линкер является полимерным (например, полиэтиленом, полиэтиленгликолем, полиамидом, сложным полиэфиром и т.д.). Согласно определенным вариантам осуществления линкер содержит мономер, димер или полимер аминоалкановой кислоты. Согласно определенным вариантам осуществления линкер содержит аминоалкановую кислоту (например, глицин, этановую кислоту, аланин, бета-аланин, 3-аминопропановую кислоту, 4-аминобутановую кислоту, 5-пентановую кислоту и т.д.). Согласно определенным вариантам осуществления линкер содержит мономер, димер или полимер аминогексановой кислоты (Ahx). Согласно определенным вариантам осуществления линкер основан на карбоциклическом фрагменте (например, циклопентане, циклогексане). Согласно другим вариантам осуществления линкер содержит фрагмент полиэтиленгликоля (ПЭГ). Согласно другим вариантам осуществления линкер содержит аминокислоты. Согласно другим вариантам осуществления линкер содержит пептид. Согласно определенным вариантам осуществления линкер содержит арильный или гетероарильный фрагмент. Согласно определенным вариантам осуществления линкер основан на фенильном кольце. Линкер может включать функционализированные фрагменты для облегчения присоединения нуклеофила (например, тиола, амино) из пептида к линкеру. В качестве части линкера можно использовать любой электрофил. Примеры электрофилов включают без ограничения активированные сложные эфиры, активированные амиды, акцепторы Михаэля, алкилгалогениды, арилгалогениды, ацилгалогениды и изотиоцианаты.

[0601] Согласно некоторым другим вариантам осуществления линкер содержит 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. Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность (GGS)N (SEQ ID NO: 176), ГДЕ n представляет собой 1, 3 или 7. Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGS (SEQ ID NO: 174). Согласно другим вариантам осуществления линкер содержит аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS (SEQ ID NO: 175, 60AA).

[0602] Согласно определенным вариантам осуществления линкеры могут быть использованы для связывания любых пептидов или пептидных доменов или фрагментов согласно настоящему изобретению (например, napDNAbp, связанный или слитый с обратной транскриптазой).

[0603] Как определено выше, термин «линкер» в контексте настоящего изобретения относится к химической группе или молекуле, связывающей две молекулы или группы, например, связывающий домен и домен расщепления нуклеазы. Согласно некоторым вариантам осуществления линкер соединяет домен связывания гРНК РНК-программируемой нуклеазы и каталитический домен рекомбиназы. Согласно некоторым вариантам осуществления линкер соединяет dCas9 и обратную транскриптазу. Как правило, линкер расположен между двумя группами, молекулами или другими фрагментами или окружен ими и соединен с каждым из них ковалентной связью, таким образом, соединяя их. Согласно некоторым вариантам осуществления линкер представляет собой аминокислоту или множество аминокислот (например, пептид или белок). Согласно некоторым вариантам осуществления линкер представляет собой органическую молекулу, группу, полимер или химический фрагмент. Согласно некоторым вариантам осуществления длина линкера составляет 5-100 аминокислот в длину, например, 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 в длину. Также рассматриваются более длинные или короткие линкеры.

[0604] Линкер может быть таким же простым, как ковалентная связь, или он может быть полимерным линкером длиной в несколько атомов. Согласно определенным

вариантам осуществления линкер представляет собой полипептид или основан на аминокислотах. Согласно другим вариантам осуществления линкер не является пептидоподобным. Согласно определенным вариантам осуществления линкер представляет собой ковалентную связь (например, связь углерод-углерод, дисульфидную связь, связь углерод-гетероатом и т.д.). Согласно определенным вариантам осуществления линкер представляет собой углерод-азотную связь амидной связи. Согласно определенным вариантам осуществления линкер представляет собой циклический или ациклический, замещенный или незамещенный, разветвленный или неразветвленный алифатический или гетероалифатический линкер. Согласно определенным вариантам осуществления линкер является полимерным (например, полиэтиленом, полиэтиленгликолем, полиамидом, сложным полиэфиром и т.д.). Согласно определенным вариантам осуществления линкер содержит мономер, димер или полимер аминоалкановой кислоты. Согласно определенным вариантам осуществления линкер содержит аминоалкановую кислоту (например, глицин, этановую кислоту, аланин, бета-аланин, 3-аминопропановую кислоту, 4-аминобутановую кислоту, 5-пентановую кислоту и т.д.). Согласно определенным вариантам осуществления линкер содержит мономер, димер или полимер аминогексановой кислоты (Ahx). Согласно определенным вариантам осуществления линкер основан на карбоциклическом фрагменте (например, циклопентан, циклогексан). Согласно другим вариантам осуществления линкер содержит фрагмент полиэтиленгликоля (ПЭГ). Согласно другим вариантам осуществления линкер содержит аминокислоты. Согласно определенным вариантам осуществления линкер содержит пептид. Согласно определенным вариантам осуществления линкер содержит арильный или гетероарильный фрагмент. Согласно определенным вариантам осуществления линкер основан на фенильном кольце. Линкер может включать функционализированные фрагменты для облегчения присоединения нуклеофила (например, тиола, амино) из пептида к линкеру. В качестве части линкера можно использовать любой электрофил. Примеры электрофилов включают без ограничения активированные сложные эфиры, активированные амиды, акцепторы Михаэля, алкилгалогениды, арилгалогениды, ацилгалогениды и изотиоцианаты.

[0605] Согласно некоторым другим вариантам осуществления линкер содержит 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. Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность (GGS)N (SEQ ID NO: 176), где n представляет собой 1, 3 или 7. Согласно некоторым вариантам

осуществления линкер содержит аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGS (SEQ ID NO: 174).

[0606] В частности, следующие линкеры могут использоваться в различных вариантах осуществления для соединения доменов праймирующего 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),

SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGG SSGG S (SEQ ID NO: 175).

B. <u>Последовательность локализации в ядре (NLS)</u>

[0607] Согласно различным вариантам осуществления слитые белки PE могут содержать одну или несколько последовательностей локализации в ядре (NLS), которые способствуют транслокации белка в ядро клетки. Такие последовательности хорошо известны в данной области и могут включать следующие примеры:

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ	SEQ ID NO:
NLS SV40 БОЛЬШОЙ T-AG	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 БОЛЬШОЙ ПОЛИОМЫ T-AG	VSRKRPRP	SEQ ID NO: 194
NLS АНТИГЕНА ВИРУСА ГЕПАТИТА D	EGAPPAKRAR	SEQ ID NO: 195
NLS МЫШИНОГО P53	PPQPKKKPLDGE	SEQ ID NO: 196
NLS PE1 И PE2	SGGSKRTADGSEFEPKKKRKV	SEQ ID NO: 133

[0608] Примеры NLS выше являются неограничивающими. Слитые белки PE могут содержать любую известную последовательность NLS, включая описанные в Cokol et al., «Finding nuclear localization signals,» *EMBO Rep.*, 2000, 1(5): 411-415 and Freitas et al., «Mechanisms and Signals for the Nuclear Import of Proteins,» *Current Genomics*, 2009, 10(8): 550-7, каждая из которых включена в настоящий документ посредством ссылки.

[0609] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы и конструкции, кодирующие праймирующие редакторы, раскрытые в настоящем документе, дополнительно содержат один или несколько, предпочтительно по меньшей мере два сигнала локализации в ядре. Согласно определенным вариантам осуществления мультилоскутные праймирующие редакторы содержат по меньшей мере два NLS. Согласно вариантам осуществления по меньшей мере с двумя NLS эти NLS могут быть одними и теми же NLS или могут быть разными NLS. Кроме того, NLS могут быть экспрессированы как часть слитого белка с оставшимися частями мультилоскутных праймирующих редакторов. Согласно некоторым вариантам осуществления одна или несколько NLS представляют собой NLS из двух частей («bpNLS»). Согласно определенным вариантам осуществления раскрытые слитые белки содержат две NLS из двух частей. Согласно некоторым вариантам осуществления раскрытые слитые белки содержат более двух NLS из двух частей.

[0610] Место слияния NLS может быть на N-конце, C-конце или в последовательности праймирующего редактора (например, вставлено между кодируемым компонентом napDNAbp (например, Cas9) и полимеразным доменом (например, доменом обратной транскриптазы).

[0611] NLS могут быть любой известной последовательностью NLS в данной области техники. NLS также могут быть любой NLS, открытой в будущем для локализации

в ядре. NLS также могут быть любыми NLS природного происхождения или любыми NLS неприродного происхождения (например, NLS с одной или несколькими желаемыми мутациями).

[0612] Термин «последовательность локализации в ядре» или «NLS» относится к аминокислотной последовательности, которая способствует импорту белка в ядро клетки, например, посредством ядерного транспорта. Последовательности локализации в ядре известны специалистам в данной области техники. Например, последовательности NLS описаны в Plank et al., международная заявка PCT/EP2000/011690, поданная 23 ноября 2000 г., опубликованная как WO/2001/038547 31 мая 2001 г., содержание которой включено в настоящий документ посредством ссылки. Согласно некоторым вариантам осуществления NLS содержит аминокислотную последовательность PKKKRKV (SEQ ID NO: 16), MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 17), KRTADGSEFESPKKKRKV (SEQ ID NO: 3864), или KRTADGSEFEPKKKRKV (SEQ ID NO: 125). Согласно другим вариантам осуществления NLS содержит аминокислотные последовательности NLSKRPAAIKKAGQAKKKK (SEQ ID NO: 136), PAAKRVKLD (SEQ ID NO: 192), RQRRNELKRSF (SEQ ID NO: 3934), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 3935).

[0613] Согласно одному аспекту мультилоскутный праймирующий редактор может быть модифицирован одним или несколькими сигналами локализации в ядре (NLS), предпочтительно по меньшей мере двумя NLS. Согласно определенным вариантам осуществления мультилоскутные праймирующие редакторы модифицированы двумя или более NLS. Раскрытие охватывает использование любого сигнала локализации в ядре, известного в данной области техники на момент раскрытия, или любого сигнала локализации в ядре, который идентифицирован или иным образом стал доступным в уровне техники после момента подачи настоящей заявки. Иллюстративным сигналом локализации в ядре является пептидная последовательность, которая направляет белок в ядро клетки, в которой эта последовательность экспрессируется. Сигнал локализации в ядре является преимущественно основным, может располагаться почти в любом месте аминокислотной последовательности белка, обычно содержит короткую последовательность из от четырех аминокислот (Autieri & Agrawal, (1998) *J. Biol. Chem* . 273: 14731-37, включено в настоящий документ посредством ссылки) до восьми аминокислот и обычно содержит большое количество остатков лизина и аргинина (Magin *et al* ., (2000) *Virology* 274: 11-16, включено в настоящий документ посредством ссылки). Сигналы локализации в ядре часто включают остатки пролина. Были идентифицированы различные сигналы локализации в ядре, которые использовались для осуществления 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, которая включена посредством ссылки. В настоящее время полагают, что в транслокации участвуют белки ядерных пор.

[0614] Большинство NLS можно разделить на три общие группы: (i) однокомпонентная NLS, примером которой является NLS с большим Т-антигеном SV40 (PKKKRKV (SEQ ID NO: 16)), (ii) мотив из двух частей, состоящий из двух основных доменов, разделенных переменным числом спейсерных аминокислот, представленный NLS *Xenopus nucleoplasmin* (KRXXXXXXXXXXKKKL (SEQ ID NO: 3936))), и (iii) неканонические последовательности, такие как M9 белка hnRNP A1, NLS нуклеопротеина вируса гриппа и NLS белка Gal4 дрожжей (Dingwall and Laskey 1991).

[0615] Сигналы локализации в ядре появляются в различных точках аминокислотной последовательности белков. NLS были идентифицированы на N-конце, C-конце и в центральной области белков. Таким образом, настоящее изобретение относится к мультилоскутным праймирующим редакторам, которые могут быть модифицированы одной или несколькими NLS на C-конце, N-конце, а также во внутренней области мультилоскутного праймирующего редактора. Остатки более длинной последовательности, которые не функционируют как составные остатки NLS, должны быть выбраны так, чтобы не мешать, например, тонически или стерически, самому сигналу локализации в ядре. Следовательно, хотя нет строгих ограничений на состав NLS-содержащей последовательности, на практике такая последовательность может быть функционально ограничена по длине и составу.

[0616] В настоящем раскрытии рассматриваются любые подходящие средства, с помощью которых можно модифицировать мультилоскутные праймирующие редакторы для включения в него одной или нескольких NLS. Согласно одному аспекту мультилоскутные праймирующие редакторы могут быть сконструированы для экспрессии белка праймирующего редактора, который трансляционно слит на его N-конце или его C-конце (или на обоих) с одной или несколькими NLS, т.е. для образования слитой конструкции праймирующего редактора-NLS. Согласно другим вариантам осуществления праймирующий редактор нуклеотидная последовательность, кодирующая праймирующий редактор, может быть генетически модифицирована для включения рамки считывания, которая кодирует одну или несколько NLS во внутренней области кодируемого праймирующего редактора. Кроме того, NLS могут включать различные аминокислотные линкеры или спейсерные области, кодируемые между праймирующим редактором и аминокислотной последовательностью N-концевой, C-концевой или внутренней NLS, например, и в центральной области белков. Таким образом, настоящее изобретение также

относится к нуклеотидным конструкциям, векторам и клеткам-хозяевам для экспрессии слитых белков, которые содержат праймирующий редактор и одну или несколько NLS.

[0617] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут также содержать сигналы локализации в ядре, которые связаны с праймирующим редактором через один или несколько линкеров, например, и полимерный, аминокислотный, на основе нуклеиновой кислоты, полисахаридный, химический или на основе нуклеиновой кислоты линкерный элемент. Линкеры в предполагаемом объеме настоящего изобретения не имеют каких-либо ограничений и могут представлять собой молекулы любого подходящего типа (например, полимер, аминокислоту, полисахарид, нуклеиновую кислоту, липид или любой синтетический химический линкерный домен) и быть присоединены к праймирующему редактору с помощью любой подходящей стратегии, которая осуществляет образование связи (например, ковалентной связи, водородной связи) между праймирующим редактором и одной или несколькими NLS.

C. Эндонуклеазы лоскутов (например, FEN1)

[0618] Согласно различным вариантам осуществления слитые белки PE могут содержать одну или несколько эндонуклеаз лоскута (например, FEN1), которая относится к ферменту, катализирующему удаление 5'-одноцепочечной лоскутов ДНК. Это встречающиеся в природе ферменты, которые удаляют 5'-лоскуты, образующиеся в ходе клеточных процессов, включая репликацию ДНК. Описанные в настоящем документе способы мультилоскутного праймированного редактирования могут использовать эндогенно поставляемые эндонуклеазы лоскутов или эндонуклеазы, обеспечиваемые *в транс*, для удаления 5'-эндогенного лоскута ДНК, образованного в целевом участке в ходе мультилоскутного праймированного редактирования. Эндонуклеазы лоскутов известны в данной области техники, и их можно найти 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 (каждая из которых включена в настоящий документ посредством ссылки). Иллюстративной эндонуклеазой лоскута является FEN1, которая может быть представлена следующей аминокислотной последовательностью:

Описание	последовательность	SEQ ID NO:
FEN1 Дикого типа (wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQF LIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPV YVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQE	SEQ ID NO: 198

	VEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASC AALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLP IQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRA VDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEV LDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSK SRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTG AAGKFKRGK	

[0619] Эндонуклеазы лоскута могут также включать любой вариант, мутант FEN1 или другой ортолог, гомолог или вариант эндонуклеазы лоскута. Неограничивающие примеры варианта FEN1 являются следующими:

Описание	Последовательность	SEQ ID NO:
FEN1 K168R (относитель но FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQF LIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPV YVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQE VEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASC AALV**R**AGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLP IQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRA VDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEV LDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSK SRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTG AAGKFKRGK	SEQ ID NO: 199
FEN1 S187A (относитель но FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQF LIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPV YVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQE VEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASC AALVKAGKVYAAATEDMDCLTFG**A**PVLMRHLTASEAKKL PIQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRA VDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEV LDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSK SRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTG AAGKFKRGK	SEQ ID NO: 200
FEN1 K354R (относитель но FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQF LIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPV YVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQE VEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASC AALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLP IQEFHLSRILQELGLNQEQFVDLCILLGSDYCESIRGIGPKRA	SEQ ID NO: 201

	VDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEV LDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSK SRQGSTQGRLDDFFKVTGSLSSA**R**RKEPEPKGSTKKKAKTG AAGKFKRGK	
GEN1	MGVNDLWQILEPVKQHIPLRNLGGKTIAVDLSLWVCEAQT VKKMMGSVMKPHLRNLFFRISYLTQMDVKLVFVMEGEPPK LKADVISKRNQSRYGSSGKSWSQKTGRSHFKSVLRECLHML ECLGIPWVQAAGEAEAMCAYLNAGGHVDGCLTNDGDTFL YGAQTVYRNFTMNTKDPHVDCYTMSSIKSKLGLDRDALVG LAILLGCDYLPKGVPGVGKEQALKLIQILKGQSLLQRFNRW NETSCNSSPQLLVTKKLAHCSVCSHPGSPKDHERNGCRLCK SDKYCEPHDYEYCCPCEWHRTEHDRQLSEVENNIKKKACC CEGFPFHEVIQEFLLNKDKLVKVIRYQRPDLLLFQRFTLEKM EWPNHYACEKLLVLLTHYDMIERKLGSRNSNQLQPIRIVKT RIRNGVHCFEIEWEKPEHYAMEDKQHGEFALLTIEEESLFEA AYPEIVAVYQKQKLEIKGKKQKRIKPKENNLPEPDEVMSFQ SHMTLKPTCEIFHKQNSKLNSGISPDPTLPQESISASLNSLLLP KNTPCLNAQEQFMSSLRPLAIQQIKAVSKSLISESSQPNTSSH NISVIADLHLSTIDWEGTSFSNSPAIQRNTFSHDLKSEVESELS AIPDGFENIPEQLSCESERYTANIKKVLDEDSDGISPEEHLLS GITDLCLQDLPLKERIFTKLSYPQDNLQPDVNLKTLSILSVKE SCIANSGSDCTSHLSKDLPGIPLQNESRDSKILKGDQLLQEDY KVNTSVPYSVSNTVVKTCNVRPPNTALDHSRKVDMQTTRKI LMKKSVCLDRHSSDEQSAPVFGKAKYTTQRMKHSSQKHNS SHFKESGHNKLSSPKIHIKETEQCVRSYETAENEESCFPDSTK SSLSSLQCHKKENNSGTCLDSPLPLRQRLKLRFQST	SEQ ID NO: 202
ERCC5	MGVQGLWKLLECSGRQVSPEALEGKILAVDISIWLNQALKG VRDRHGNSIENPHLLTLFHRLCKLLFFRIRPIFVFDGDAPLLK KQTLVKRRQRKDLASSDSRKTTEKLLKTFLKRQAIKTAFRS KRDEALPSLTQVRRENDLYVLPPLQEEEKHSSEEEDEKEWQ ERMNQKQALQEEFFHNPQAIDIESEDFSSLPPEVKHEILTDM KEFTKRRRTLFEAMPEESDDFSQYQLKGLLKKNYLNQHIEH VQKEMNQQHSGHIRRQYEDEGGFLKEVESRRVVSEDTSHYI LIKGIQAKTVAEVDSESLPSSSKMHGMSFDVKSSPCEKLKTE KEPDATPPSPRTLLAMQAALLGSSSEEELESENRRQARGRN APAAVDEGSISPRTLSAIKRALDDDEDVKVCAGDDVQTGGP GAEEMRINSSTENSDEGLKVRDGKGIPFTATLASSSVNSAEE	SEQ ID NO: 203

	HVASTNEGREPTDSVPKEQMSLVHVGTEAFPISDESMIKDR KDRLPLESAVVRHSDAPGLPNGRELTPASPTCTNSVSKNETH AEVLEQQNELCPYESKFDSSLLSSDDETKCKPNSASEVIGPV SLQETSSIVSVPSEAVDNVENVVSFNAKEHENFLETIQEQQT TESAGQDLISIPKAVEPMEIDSEESESDGSFIEVQSVISDEELQ AEFPETSKPPSEQGEEELVGTREGEAPAESESLLRDNSERDD VDGEPQEAEKDAEDSLHEWQDINLEELETLESNLLAQQNSL KAQKQQQERIAATVTGQMFLESQELLRLFGIPYIQAPMEAE AQCAILDLTDQTSGTITDDSDIWLFGARHVYRNFFNKNKFV EYYQYVDFHNQLGLDRNKLINLAYLLGSDYTEGIPTVGCVT AMEILNEFPGHGLEPLLKFSEWWHEAQKNPKIRPNPHDTKV KKKLRTLQLTPGFPNPAVAEAYLKPVVDDSKGSFLWGKPD LDKIREFCQRYFGWNRTKTDESLFPVLKQLDAQQTQLRIDSF FRLAQQEKEDAKRIKSQRLNRAVTCMLRKEKEAAASEIEAV SVAMEKEFELLDKAKRKTQKRGITNTLEESSSLKRKRLSDS KRKNTCGGFLGETCLSESSDGSSSEDAESSSLMNVQRRTAA KEPKTSASDSQNSVKEAPVKNGGATTSSSSDSDDDGGKEK MVLVTARSVFGKKRRKLRRARGRKRKT	

[0620] Согласно различным вариантам осуществления слитые белки мультилоскутного праймирующего редактора, охватываемые настоящим изобретением, могут включать любой вариант эндонуклеазы лоскутов вышеуказанных последовательностей, имеющий аминокислотную последовательность, которая на по меньшей мере приблизительно 70% идентична, по меньшей мере приблизительно 80% идентична, по меньшей мере приблизительно 90% идентична, по меньшей мере приблизительно 95% идентична, по меньшей мере приблизительно 96% идентична, по меньшей мере приблизительно 97% идентична, по меньшей мере приблизительно 98% идентична, по меньшей мере приблизительно 99% идентична, по меньшей мере приблизительно 99.5% идентична или по меньшей мере приблизительно 99.9% идентична любой из вышеуказанных последовательностей.

[0621] Другие эндонуклеазы, которые можно использовать в способах согласно настоящему изобретению для облегчения удаления 5'-лоскута одноцепочечной ДНК, включают без ограничения (1) эндонуклеазу trex 2, (2) эндонуклеазу exo1 (например, Keijzers et al., *Biosci Rep.* 2015, 35(3): e00206)

Trex 2

[0622] 3’ экзонуклеаза 2 с тремя праймированными репарациями (TREX2) - человек

[0623] № доступа NM_080701

MSEAPRAETFVFLDLEATGLPSVEPEIAELSLFAVHRSSLENPEHDESGALVLPRV LDKLTLCMCPERPFTAKASEITGLSSEGLARCRKAGFDGAVVRTLQAFLSRQAGPICLVA HNGFDYDFPLLCAELRRLGARLPRDTVCLDTLPALRGLDRAHSHGTRARGRQGYSLGS LFHRYFRAEPSAAHSAEGDVHTLLLIFLHRAAELLAWADEQARGWAHIEPMYLPPDDPS LEA (SEQ ID NO: 3865).

[0624] 3’ 'экзонуклеаза 2 с тремя праймированными репарациями (TREX2) - мышь

[0625] № доступа NM_011907

[0626] MSEPPRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSLV LPRVLDKLTLCMCPERPFTAKASEITGLSSESLMHCGKAGFNGAVVRTLQGFLSRQEGPI CLVAHNGFDYDFPLLCTELQRLGAHLPQDTVCLDTLPALRGLDRAHSHGTRAQGRKSY SLASLFHRYFQAEPSAAHSAEGDVHTLLLIFLHRAPELLAWADEQARSWAHIEPMYVPP DGPSLEA (SEQ ID NO: 3866).

[0627] 3’ 'экзонуклеаза 2 с тремя праймированными репарациями (TREX2) - крыса

[0628] № доступа NM_001107580

[0629] MSEPLRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSL VLPRVLDKLTLCMCPERPFTAKASEITGLSSEGLMNCRKAAFNDAVVRTLQGFLSRQEG PICLVAHNGFDYDFPLLCTELQRLGAHLPRDTVCLDTLPALRGLDRVHSHGTRAQGRKS YSLASLFHRYFQAEPSAAHSAEGDVNTLLLIFLHRAPELLAWADEQARSWAHIEPMYVP PDGPSLEA (SEQ ID NO: 3867).

ExoI

[0630] Экзонуклеаза 1 человека (EXO1) участвует во многих различных метаболических процессах ДНК, включая репарацию ошибочно спаренных оснований ДНК (MMR), микроопосредованное соединение концов, гомологичную рекомбинацию (HR) и репликацию. EXO1 человека принадлежит к семейству эукариотических нуклеаз Rad2/XPG, которое также включает FEN1 и GEN1. Семейство Rad2/XPG консервативно в нуклеазном домене для всех видов от фага до человека. Продукт гена EXO1 проявляет активность как 5'-экзонуклеазы, так и 5'-лоскута. Кроме того, EXO1 содержит внутреннюю активность 5'-РНКазы H. Человеческий EXO1 обладает высокой аффинностью к процессингу двухцепочечной ДНК (дцДНК), разрывам, гэпам, псевдо Y-структурам и может разрешать соединения Холлидея, с использованием их наследственной активности в отношении лоскутов. EXO1 человека участвует в MMR и содержит консервативные связывающие домены, непосредственно взаимодействующие с MLH1 и MSH2. Нуклеолитическую активность EXO1 положительно стимулируют PCNA, MutSα (комплекс MSH2/MSH6), 14-3-3, MRN и комплекс 9-1-1.

[0631] экзонуклеаза 1 (EXO1) № доступа NM_003686 (экзонуклеаза 1 Homo sapiens

(EXO1), вариант транскрипта 3) – изоформа A

MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDR
YVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVS
EARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQAIITEDSD
LLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILSGCDYLSSLRG
IGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLYQLVFDPIKR
KLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTAMPAHSRS
HSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVISTKGLNLPR
KSSIVKRPRSAELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSV
STPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDSTDCVSNKVSIQPLDETAVTDK
ENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDHIPDKATVFTDEESYSFESSKF
TRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENNMSDVSQL
KSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNIKLLDSQS
DQTSKLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKPASIQKR
KHHNAENKPGLQIKLNELWKNFGFKKF (SEQ ID NO: 3868).

[0632] экзонуклеаза 1 (EXO1) № доступа NM_006027 (экзонуклеаза 1 из Homo sapiens (EXO1), вариант транскрипта 3) – изоформа B

[0633] MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKG
EPTDRYVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLR
EGKVSEARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQAI
ITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILSGCDY
LSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLYQL
VFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTA
MPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVIST
KGLNLPRKSSIVKRPRSAELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNG
PTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDSTDCVSNKVSIQPLD
ETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDHIPDKATVFTDEES
YSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENN
MSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCN
IKLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKK
PASIQKRKHHNAENKPGLQIKLNELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEE
DIFNKPECGRVQRAIFQ (SEQ ID NO: 3869).

[0634] экзонуклеаза 1 (EXO1) № доступа NM_001319224 (экзонуклеаза 1 из Homo sapiens (EXO1), вариант транскрипта 4) – изоформа C

[0635] MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKG

EPTDRYVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLR
EGKVSEARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAGIVQAI
ITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRYMCILSGCDY
LSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLYQL
VFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTA
MPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVIST
KGLNLPRKSSIVKRPRSELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGP
TNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPGTRSRFFCSSDSTDCVSNKVSIQPLDE
TAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDHIPDKATVFTDEESY
SFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENNM
SDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNI
KLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKK
PASIQKRKHHNAENKPGLQIKLNELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEE
DIFNKPECGRVQRAIFQ (SEQ ID NO: 3870).

D. Интеины и расщепленные интеины

[0636] Следует понимать, что согласно некоторым вариантам осуществления (например, доставка мультилоскутного праймирующего редактора *in vivo* с использованием частиц AAV) может быть предпочтительным для расщепления полипептида (например, деаминазы или napDNAbp) или слитого белка (например, мультилоскутного праймирующего редактора) на N-концевую половину и C-концевую половину, доставляя их по отдельности, а затем позволяя их совместную локализацию с образованием полного белка (или слитого белка, в зависимости от случая) внутри клетки. Каждая из отдельных половин белка или слитого белка может содержать расщепленный интеин для облегчения реформирования полного белка или слитого белка по механизму транс-сплайсинга белка.

[0637] Транс-сплайсинг белков, катализируемый расщепленными интеинами, представляет собой полностью ферментативный способ лигирования белков. Расщепленный 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 расщепленного интеина могут нековалентно связываться с образованием активного интеина и катализировать реакцию сплайсинга, по существу, таким же образом, как это делает непрерывный интеин. Расщепленные интеины встречаются в природе, а также созданы в лабораториях. В контексте настоящего изобретения термин «расщепленный интеин» относится к любому интеину, в котором существует один или несколько разрывов пептидной связи между N-концевой и C-концевой аминокислотными последовательностями, так что N-концевая и C-концевая

последовательности становятся отдельными молекулами, которые могут нековалентно повторно ассоциироваться или восстанавливаться в интеин, который является функциональным для реакций транс-сплайсинга. Любой каталитически активный интеин или его фрагмент можно использовать для получения расщепленного интеина для применения в способах согласно настоящему изобретению. Например, согласно одному аспекту расщепленный интеин может быть получен из эукариотического интеина. Согласно другому аспекту расщепленный интеин может быть получен из бактериального интеина. Согласно другому аспекту расщепленный интеин может быть получен из интеина архей. Предпочтительно полученный таким образом расщепленный интеин будет иметь только те аминокислотные последовательности, которые необходимы для катализа реакций транс-сплайсинга.

[0638] В контексте настоящего изобретения «N-концевой расщепленный интеин (In)» относится к любой последовательности интеина, которая содержит N-концевую аминокислотную последовательность, функциональную для реакций транс-сплайсинга. Таким образом, In также содержит последовательность, которая подвергается сплайсингу, когда происходит транс-сплайсинг. In может содержать последовательность, которая является модификацией N-концевой части последовательности интеина, встречающейся в природе. Например, In может содержать дополнительные аминокислотные остатки и/или мутированные остатки при условии, что включение таких дополнительных и/или мутантных остатков не делает In нефункциональным при транс-сплайсинге. Предпочтительно включение дополнительных и/или мутантных остатков улучшает или усиливает активность транс-сплайсинга In.

[0639] В контексте настоящего изобретения «C-концевой расщепленный интеин (Ic)» относится к любой последовательности интеина, которая содержит C-концевую аминокислотную последовательность, функциональную для реакций транс-сплайсинга. Согласно одному аспекту Ic содержит от 4 до 7 смежных аминокислотных остатков, по меньшей мере 4 аминокислоты из которых происходят из последней β-цепи интеина, из которого он был получен. Таким образом, Ic также содержит последовательность, которая подвергается сплайсингу, когда происходит транс-сплайсинг. Ic может содержать последовательность, которая является модификацией C-концевой части встречающейся в природе последовательности интеина. Например, Ic может содержать дополнительные аминокислотные остатки и/или мутированные остатки при условии, что включение таких дополнительных и/или мутантных остатков не делает In нефункциональным при транс-сплайсинге. Предпочтительно включение дополнительных и/или мутантных остатков улучшает или усиливает транс-сплайсинговую активность Ic.

[0640] Согласно некоторым вариантам осуществления настоящего изобретения пептид, связанный с Ic или In, может содержать дополнительный химический фрагмент, включая без ограничения флуоресцентные группы, биотин, полиэтиленгликоль (PEG), аналоги аминокислот, неприродные аминокислоты, фосфатные группы, гликозильные группы, радиоизотопные метки и фармацевтические молекулы. Согласно другим вариантам осуществления пептид, связанный с Ic, может содержать одну или несколько химически реакционноспособных групп, включая без ограничения кетон, альдегид, остатки Cys и остатки Lys. N-интеин и C-интеин расщепленного интеина могут нековалентно связываться с образованием активного интеина и катализировать реакцию сплайсинга, когда присутствует «интеин-сплайсинговый полипептид (ISP)». В контексте настоящего изобретения «интеин-сплайсинговый полипептид (ISP)» относится к части аминокислотной последовательности расщепленного интеина, которая сохраняется, когда Ic, In или оба удаляются из расщепленного интеина. Согласно определенным вариантам осуществления In содержит ISP. Согласно другому варианту осуществления Ic содержит ISP. Согласно другому варианту осуществления ISP представляет собой отдельный пептид, который не связан ковалентно ни с In, ни с Ic.

[0641] Расщепленные интеины могут быть созданы из непрерывных интеинов путем создания одного или нескольких сайтов расщепления в неструктурированной петле или промежуточной аминокислотной последовательности между -12 консервативными бета-цепями, обнаруживаемыми в структуре мини-интеинов. Может существовать некоторая гибкость в расположении сайта расщепления в областях между бета-цепями при условии, что создание расщепления не нарушит структуру интеина, в частности структурированными бета-цепями, в достаточной степени, чтобы сплайсинговая активность белка не была потеряна.

[0642] При транс-сплайсинге белка один белок-предшественник состоит из части N-экстеина, за которой следует N-интеин, другой белок-предшественник состоит из C-интеина, за которым следует часть C-экстеина, и реакция транс-сплайсинга (катализируемая N- и C-интеинами вместе) вырезает две последовательности интеина и связывает две последовательности экстеина пептидной связью. Транс-сплайсинг белков, являясь ферментативной реакцией, может работать с очень низкими (например, микромолярными) концентрациями белков и может осуществляться в физиологических условиях.

[0643] Иллюстративные последовательности являются следующими:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
2-4 ИНТЕИН:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 8)
3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYTNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 9)
30R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 10)
30R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 11)

30R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 12)
37R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC ((SEQ ID NO: 13)
37R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 14)
37R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 15)

[0644] Хотя интеины чаще всего обнаруживаются как непрерывные домены, некоторые из них существуют в естественно расщепленной форме. В этом случае два

фрагмента экспрессируются как отдельные полипептиды и должны ассоциироваться до того, как произойдет сплайсинг, так называемый, транс-сплайсинг белка.

[0645] Иллюстративным расщепленным интеином является интеин *Ssp* DnaE, который содержит две субъединицы, а именно DnaE-N и DnaE-C. Две разные субъединицы кодируются отдельными генами, а именно *dnaE-n* и *dnaE-c*, которые кодируют субъединицы DnaE-N и DnaE-C соответственно. DnaE представляет собой природный расщепленный интеин PCC6803 *Synechocytis sp.* и способен управлять транс-сплайсингом двух отдельных белков, каждый содержащий слияние либо с DnaE-N, либо с DnaE-C.

[0646] Дополнительные встречающиеся в природе или сконструированные последовательности расщепленного интеина известны или могут быть получены из последовательностей цельного интеина, описанных в настоящем документе или доступных в данной области техники. Примеры последовательностей расщепленных интеинов 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, каждая из которых включена в настоящий документ посредством ссылки. Дополнительные последовательности расщепленного интеина можно найти, например, в WO 2013/045632, WO 2014/055782, WO 2016/069774 и EP 2877490, содержание каждой из которых включено в настоящий документ посредством ссылки.

[0647] Кроме того, транс-сплайсинг белка 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 дает возможность экспрессировать белок в виде двух неактивных фрагментов, которые затем подвергают лигированию с образованием функционального продукта, например, как показано на фиг. 66 и 67 в отношении образования полного слитого белка PE из двух отдельно экспрессируемых половин.

E. Домен взаимодействия РНК-белок

[0648] Согласно различным вариантам осуществления два отдельных белковых домена (например, домен Cas9 и домен полимеразы) могут быть совместно локализованы друг с другом с образованием функционального комплекса (подобного функции слитого белка, содержащего два отдельных белковых домена) с использованием «система

рекрутирования РНК-белок», например, «способ мечения MS2». Такие системы обычно помечают один белковый домен «доменом взаимодействия РНК-белок» (он же «домен рекрутирования РНК-белок»), а другой — «РНК-связывающим белком», который специфически распознает и связывается с доменом взаимодействия РНК-белок, например, специфической структурой шпильки. Эти типы систем можно использовать для совместной локализации доменов мультилоскутного праймирующего редактора, а также для привлечения дополнительных функций к мультилоскутному праймирующему редактору, например, домену UGI. В одном примере способ мечения MS2 основан на естественном взаимодействии белка оболочки бактериофага MS2 («MCP» или «MS2cp») со структурой «петля на стебле» или «шпилька», присутствующей в геноме фага, т.е. «MS2 шпилька." В случае шпильки MS2 она распознается и связывается белком оболочки бактериофага MS2 (MCP). Таким образом, в одном примерном сценарии слияние дезаминаза-MS2 может рекрутировать слияние Cas9-MCP.

[0649] Обзор других модульных доменов взаимодействия РНК-белок описан в данной области техники, 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, каждая из которых включена в настоящий документ посредством ссылки во всей своей полноте. Другие системы включают шпильку PP7, которая специфически рекрутирует белок PCP, и шпильку «com», которая специфически рекрутирует белок Com. См. Zalatan et al.

[0650] Нуклеотидная последовательность шпильки MS2 (или эквивалентно называемой «аптамером MS2») представляет: GCCAACATGAGGATCACCCATGTCTGCAGGGCC (SEQ ID NO: 763).

[0651] Аминокислотная последовательность MCP или MS2cp представляет собой: GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDG NPIPSAIAANSGIY (SEQ ID NO: 764).

F. UGI домен

[0652] Согласно другим вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут содержать один или

несколько доменов ингибиторов урацилгликозилазы. Термин «ингибитор урацилгликозилазы (UGI)» или «домен UGI» в контексте настоящего изобретения относится к белку, который способен ингибировать фермент репарации путем вырезания оснований урацил-ДНК-гликозилазы. Согласно некоторым вариантам осуществления домен UGI содержит UGI дикого типа или UGI, как указано в SEQ ID NO: 3873. Согласно некоторым вариантам осуществления белки UGI, представленные в настоящем документе, включают фрагменты UGI и белки, гомологичные UGI или фрагменту UGI. Например, согласно некоторым вариантам осуществления домен UGI содержит фрагмент аминокислотной последовательности, представленной в SEQ ID NO: 3873. Согласно некоторым вариантам осуществления фрагмент UGI содержит 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. Согласно некоторым вариантам осуществления UGI содержит аминокислотную последовательность, гомологичную аминокислотной последовательности, представленной в SEQ ID NO: 3873, или аминокислотную последовательность, гомологичную фрагменту аминокислотной последовательности, представленной в SEQ ID NO: 3873. Согласно некоторым вариантам осуществления белки, содержащие UGI или фрагменты UGI, или гомологи UGI или фрагменты UGI, называются «вариантами UGI». Вариант UGI гомологичен UGI или его фрагменту. Например, вариант UGI на по меньшей мере 70% идентичен, по меньшей мере 75% идентичен, по меньшей мере 80% идентичен, по меньшей мере 85% идентичен, по меньшей мере 90% идентичен, по меньшей мере 95% идентичен, по меньшей мере 96% идентичен, по меньшей мере 97% идентичен, по меньшей мере 98% идентичен, по меньшей мере 99% идентичен, по меньшей мере 99.5% идентичен или по меньшей мере 99.9% идентичен UGI дикого типа или UGI, как представлено в SEQ ID NO: 3873. Согласно некоторым вариантам осуществления вариант UGI содержит фрагмент UGI, такой как фрагмент, который на по меньшей мере 70% идентичен, по меньшей мере 80% идентичен, по меньшей мере 90% идентичен, по меньшей мере 95% идентичен, по меньшей мере 96% идентичен, по меньшей мере 97% идентичен, по меньшей мере 98% идентичен, по меньшей мере 99% идентичен, по меньшей мере 99.5% идентичен или по меньшей мере 99.9% соответствующему фрагменту UGI дикого типа или UGI, как представлено в SEQ ID NO: 3873. Согласно некоторым вариантам осуществления UGI содержит следующую аминокислотную последовательность:

[0653] Ингибитор урацил-ДНК-гликозилазы:

[0654] >sp|P14739|UNGI_BPPB2
MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSD
APEYKPWALVIQDSNGENKIKML (SEQ ID NO: 3873).

[0655] Мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут содержать более одного домена UGI, который может быть разделен одним или несколькими линкерами, как описано в настоящем документе.

G. Дополнительные элементы PE

[0656] Согласно определенным вариантам осуществления мультилоскутные праймирующие редакторы, описанные в настоящем документе, могут содержать ингибитор репарации оснований. Термин «ингибитор репарации оснований» или «IBR» относится к белку, который способен ингибировать активность фермента репарации нуклеиновой кислоты, например, фермента эксцизионной репарации оснований. В некоторых вариантах реализации IBR представляет собой ингибитор репарации путем вырезания оснований OGG. Согласно некоторым вариантам осуществления IBR представляет собой ингибитор репарации путем вырезания оснований («iBER»). Примеры ингибиторов репарации путем вырезания оснований включают ингибиторы APE1, Endo III, Endo IV, Endo V, Endo VIII, Fpg, hOGG1, hNEIL1, T7 EndoI, T4PDG, UDG, hSMUG1 и hAAG. Согласно некоторым вариантам осуществления IBR представляет собой ингибитор Endo V или hAAG. Согласно некоторым вариантам осуществления IBR представляет собой iBER, который может представлять собой каталитически неактивную гликозилазу или каталитически неактивную диоксигеназу, или низкомолекулярный или пептидный ингибитор оксидазы, или их варианты. Согласно некоторым вариантам осуществления IBR представляет собой iBER, который может представлять собой ингибитор TDG, ингибитор MBD4 или ингибитор фермента AlkBH. Согласно некоторым вариантам осуществления IBR представляет собой iBER, который содержит каталитически неактивный TDG или каталитически неактивный MBD4. Примером каталитически неактивного TDG является мутант N140A SEQ ID NO: 3872 (человеческий TDG).

[0657] Некоторые иллюстративные гликозилазы представлены ниже. Каталитически инактивированные варианты любого из этих доменов гликозилазы представляют собой iBER, которые могут быть слиты с доменом napDNAbp или полимеразы мультилоскутных праймирующих редакторов, представленных в настоящем документе.

[0658] OGG (человека)

[0659] MPARALLPRRMGHRTLASTPALWASIPCPRSELRLDLVLPSGQSFRWREQ SPAHWSGVLADQVWTLTQTEEQLHCTVYRGDKSQASRPTPDELEAVRKYFQLDVTLA QLYHHWGSVDSHFQEVAQKFQGVRLLRQDPIECLFSFICSSNNNIARITGMVERLCQAFG PRLIQLDDVTYHGFPSLQALAGPEVEAHLRKLGLGYRARYVSASARAILEEQGGLAWLQ QLRESSYEEAHKALCILPGVGTKVADCICLMALDKPQAVPVDVHMWHIAQRDYSWHPT TSQAKGPSPQTNKELGNFFRSLWGPYAGWAQAVLFSADLRQSRHAQEPPAKRRKGSKG PEG (SEQ ID NO: 3937)

[0660] MPG (человека)

[0661] MVTPALQMKKPKQFCRRMGQKKQRPARAGQPHSSSDAAQAPAEQPHSS SDAAQAPCPRERCLGPPTTPGPYRSIYFSSPKGHLTRLGLEFFDQPAVPLARAFLGQVLV RRLPNGTELRGRIVETEAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFC MNISSQGDGACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALA INKSFDQRDLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSP WVSVVDRVAEQDTQA (SEQ ID NO: 3938)

[0662] MBD4 (человека)

[0663] MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGED EEQMMIKRSSECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDV YFISPQGLKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQ NQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRK VRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDR TVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEK YEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQ IERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRT SGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKFSDEYLTKQ WKYPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENHEKLSLS (SEQ ID NO: 3871)

[0664] TDG (человека)

[0665] MEAENAGSYSLQQAQAFYTFPFQQLMAEAPNMAVVNEQQMPEEVPAPA PAQEPVQEAPKGRKRKPRTTEPKQPVEPKKPVESKKSGKSAKSKEKQEKITDTFKVKRK VDRFNGVSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSG LSEVQLNHMDDHTLPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRI AVFNGKCIYEIFSKEVFGVKVKNLEFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKV HYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLAQEDAKKMAVKEEKYDPGYEAAYG GAYGENPCSSEPCGFSSNGLIESVELRGESAFSGIPNGQWMTQSFTDQIPSFSNHCGTQEQ EEESHA (SEQ ID NO: 3872)

[0666] Согласно некоторым вариантам осуществления описанные в настоящем документе слитые белки могут содержать один или несколько гетерологичных доменов белка (например, приблизительно или более приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, или более доменов в дополнение к другим компонентам праймирующего редактора). Слитый белок может содержать 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.

[0667] Примеры доменов белка, которые могут быть слиты с мультилоскутным праймирующим 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 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). Мультилоскутный праймирующий 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 праймирующего редактора, описаны в публикации патента США 2011/0059502, опубликованной 10 марта 2011 и включенной в настоящий документ посредством ссылки во всей своей полноте.

[0668] Согласно аспекту настоящего 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. Согласно определенным вариантам осуществления настоящего изобретения генный продукт представляет собой люциферазу. Согласно дополнительным вариантам осуществления настоящего изобретения экспрессия генного продукта является пониженной.

[0669] Подходящие предусмотренные в настоящем документе 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-8-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. Согласно некоторым вариантам осуществления слитый белок содержит одну или несколько His меток.

[0670] Согласно некоторым вариантам осуществления настоящего изобретения активность системы мультилоскутного праймированного редактирования может временно регулироваться посредством изменения времени нахождения, количества, и/или активности экспрессируемых компонентов системы PE. Например, как описано в настоящем документе, PE может быть слит с доменом белка, который способен модифицировать внутриклеточное время полужизни PE. Согласно определенным вариантам осуществления с использованием двух или более векторов (например, векторной системы, в которой описанные в настоящем документе компоненты кодируются на двух или более отдельных векторах), активность системы PE может временно регулироваться путем контролирования времени, за которое доставляются векторы. Например, согласно некоторым вариантам осуществления вектор, кодирующий нуклеазную систему, может доставлять PE перед вектором, кодирующим матрицу. Согласно другим вариантам осуществления вектор, кодирующий PEгРНК, может доставлять направляющую перед вектором, кодирующим систему PE. Согласно некоторым вариантам осуществления векторы, кодирующие систему PE и PEгРНК, доставляются одновременно. Согласно определенным вариантам осуществления одновременно доставляемый векторы временно

доставляют, например, РЕ, РЕгРНК, и/или компоненты гидовой РНК второй цепи. Согласно другим вариантам осуществления РНК (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. Согласно некоторым вариантам осуществления время полужизни РНК может быть повышено. Согласно некоторым вариантам осуществления время полужизни РНК может быть снижено. Согласно некоторым вариантам осуществления элемент может быть способен повышать стабильность РНК. Согласно некоторым вариантам осуществления элемент может быть способен понижать стабильность РНК. Согласно некоторым вариантам осуществления элемент может располагаться в 3' UTR РНК. Согласно некоторым вариантам осуществления элемент может включать сигнал полиаденилирования (PA). Согласно некоторым вариантам осуществления элемент может включать кэп, например, конец иРНК или РЕгРНК против хода транскрипции. Согласно некоторым вариантам осуществления РНК может не включать PA, так что она подвергается более быстрому расщеплению в клетке после транскрипции. Согласно некоторым вариантам 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. Согласно некоторым вариантам осуществления дестабилизирующий элемент может промотировать распад РНК, влиять на стабильность РНК, или активировать трансляцию. Согласно некоторым вариантам осуществления ARE может иметь от 50 до 150 нуклеотидов в длину. Согласно некоторым вариантам осуществления ARE может включать по меньшей мере одну копию последовательности AUUUA. Согласно некоторым вариантам осуществления по меньшей мере один ARE может быть добавлен к 3' UTR РНК. Согласно некоторым вариантам осуществления элемент может представлять собой вирус гепатита сурков (WHP).

[0671] Посттрансляционный регуляторный элемент (WPRE), который создает третичную структуру для стимулирования экспрессии из транскрипта. Согласно другим вариантам 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). Согласно некоторым вариантам осуществления WPRE или его эквивалент может быть добавлен к 3' UTR РНК. Согласно некоторым вариантам осуществления элемент может быть выбран из других мотивов

последовательности РНК с высоким содержанием быстро или медленно разлагающихся транскриптов.

[0672] Согласно некоторым вариантам осуществления вектор, кодирующий PE или PEгРНК, может подвергаться саморазрушению посредством расщепления последовательности- мишени, присутствующей на векторе, посредством системы PE. Расщепление может устранять непрерывную транскрипцию PE или PEгРНК 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] PEгРНК

[0673] Мультилоскутное система праймированного редактирования, описанные в настоящем документе, охватывают применение любых подходящих PEгРНК. Механизм нацеленной-праймированной обратной транскрипции (TPRT) можно использовать или адаптировать для проведения точного и универсального редактирования генома на основе CRISPR/Cas посредством использования специально сконфигурированной гидовой РНК, содержащей матричную последовательность обратной транскрипции (RT), которая кодирует желаемое изменение нуклеотида. В настоящей заявке эта специально сконфигурированная гидовая РНК называется «удлиненной гидовой РНК» или «PEгРНК», поскольку матричная последовательность RT может быть предоставлена как расширение стандартной или традиционной молекулы гидовой РНК. В настоящей заявке рассматривается любая подходящая конфигурация или расположение удлиненных годовых РНК для использования при двухлоскутном и четырехлоскутном праймированном редактировании.

[0674] Общие конструкции PEгРНК, при двухлоскутном и четырехлоскутном праймированном редактировании, показаны на **фиг. 92**. PEгРНК, используемые для при двухлоскутном и четырехлоскутном праймированном редактировании, имеют дизайн, аналогичный тем, которые используются для классического праймированном редактировании, однако для матричной области RT не обязательно кодировать какую-либо гомологию с локусом-мишенью. Вместо этого две PEгРНК могут согласно различным вариантам осуществления содержать RT-матрицы, которые кодируют синтез 3'-лоскутов, 3'-концы которых являются обратными комплементарными последовательностями друг друга. Эта комплементарность между 3'-лоскутами способствует их отжигу и замене

эндогенной последовательности ДНК предполагаемой новой последовательностью ДНК. Это требует, чтобы 5'-области RT-матриц в двух PEгРНК представляли собой последовательности, обратно комплементарные друг другу, и эта степень комплементарности может варьироваться (**фин. 92**).

Структура PEгРНК

[0675] На фиг. 3A показан один вариант осуществления удлиненной гидовой РНК, используемой в системе мультилоскутного праймированного редактирования, описанной в настоящем документе, при этом традиционная гидовая РНК (зеленая часть) включает последовательность протоспейсера длиной ~20 нуклеотидов и область ядра гРНК, которая связывается с napDNAbp. Согласно этому варианту осуществления гидовая РНК включает удлиненный сегмент РНК на 5'-конце, т.е. 5'- удлинение. Согласно этому варианту осуществления 5'-удлинение включает последовательность матрицы обратной транскрипции, участок связывания праймера обратной транскрипции и необязательную линкерную последовательность из 5-20 нуклеотидов. Как показано на фиг. 1A-1B, участок связывания праймера RT гибридизуется со свободным 3'-концом, который образуется после образования разрыва в нецелевой цепи R-петли, тем самым праймируя обратную транскриптазу для полимеризации ДНК в 5'-3' направлении.

[0676] На фиг. 3B показан другой вариант удлиненной гидовой РНК, используемой в системе праймированного редактирования, описанной в настоящем документе, где традиционная гидовая РНК (зеленая часть) включает последовательность протоспейсера длиной ~20 нуклеотидов и ядро гРНК, которое связывается с napDNAbp. Согласно этому варианту осуществления гидовая РНК включает удлиненный сегмент РНК на 3'-конце, т.е. 3'-удлинение. Согласно этому варианту осуществления 3'-удлинение включает последовательность матрицы обратной транскрипции и участок связывания праймера обратной транскрипции. Как показано на фиг. 1C-1D, участок связывания праймера RT гибридизуется со свободным 3'-концом, который образуется после образования разрыва в нецелевой цепи R-петли, тем самым праймируя обратную транскриптазу для полимеризации ДНК в 5'-3' направление.

[0677] На фиг. 3C показан другой вариант осуществления удлиненной гидовой РНК, используемой в описанной в настоящем документе системе праймированного редактирования, где традиционная гидовая РНК (зеленая часть) включает последовательность протоспейсера ~20 нуклеотидов и ядро гРНК, которое связывается с napDNAbp. В этом варианте осуществления гидовая РНК включает удлиненный сегмент РНК в межмолекулярном положении внутри ядра гРНК, т.е. внутримолекулярное удлинение. Согласно этому варианту осуществления внутримолекулярное удлинение

включает последовательность матрицы обратной транскрипции и участок связывания праймера обратной транскрипции. Участок связывания праймера RT гибридизуется со свободным 3'-концом, который образуется после образования разрыва в нецелевой нити R-петли, таким образом праймируя обратную транскриптазу для полимеризации ДНК в направлении 5'-3'.

[0678] Согласно одному варианту осуществления участок удлинения межмолекулярной РНК находится не в протоспейсерной последовательности гидовой РНК. Согласно другому варианту осуществления положение внутримолекулярного удлинения РНК находится в ядре гРНК. Согласно другому варианту осуществления положение внутримолекулярного удлинения РНК является любым относительно гидовой молекулы РНК, кроме как в пределах последовательности протоспейсера, или в положении, которое нарушает последовательность протоспейсера.

[0679] Согласно одному варианту осуществления внутримолекулярное удлинение РНК встроено по ходу транскрипции от 3′-конца протоспейсерной последовательности. Согласно другому варианту осуществления внутримолекулярное удлинение РНК встроено в пределах по меньшей мере 1 нуклеотида, по меньшей мере 2 нуклеотидов, по меньшей мере 3 нуклеотидов, по меньшей мере 4 нуклеотидов, 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 ходу транскрипции от 3′-конца протоспейсерной последовательности.

[0680] Согласно другим вариантам осуществления внутримолекулярное удлинение РНК встроено в гРНК, что относится к части гидовой РНК, соответствующей или содержащей tracrРНК, которая связывается и/или взаимодействует с белком Cas9 или его эквивалентом (т.е. с другим napDNAbp). Предпочтительно вставка внутримолекулярного удлинения РНК не нарушает или минимально нарушает взаимодействие между частью tracrRNA и napDNAbp.

[0681] Длина удлинения РНК (которое включает по меньшей мере матрицу RT и участок связывания праймера, *например*, см. фиг. 92) может быть любой подходящей длины. Согласно различным вариантам осуществления удлинение РНК составляет 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 200 нуклеотидов, по меньшей мере 300 нуклеотидов, по меньшей мере 400 нуклеотидов или по меньшей мере 500 нуклеотидов в длину.

[0682] Последовательность матрицы RT также может иметь любую подходящую длину. Например, последовательность матрицы RT может составлять по меньшей мере 3 нуклеотидов, по меньшей мере 4 нуклеотидов, 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 в длину.

[0683] Согласно другим вариантам осуществления последовательность участка связывания праймера обратной транскрипции составляет по меньшей мере 3 нуклеотидов, по меньшей мере 4 нуклеотидов, 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 в длину.

[0684] Согласно другим вариантам осуществления необязательный линкер или спейсерная последовательность составляет по меньшей мере 3 нуклеотидов, по меньшей мере 4 нуклеотидов, 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 в длину.

[0685] Последовательность RT матрицы согласно определенным вариантам осуществления кодирует молекулу одноцепочечной ДНК, которая является гомологичной не являющейся мишенью цепи (и, таким образом, комплементарной соответствующему участку целевой цепи), 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.

[0686] Как изображено на Фиг. 1G, синтезированный продукт в виде одноцепочечной ДНК последовательности матрицы RT является гомологичным не являющейся мишенью цепи и содержит одно или несколько нуклеотидных изменений. Продукт в виде одноцепочечной ДНК последовательности матрицы RT гибридизируется равновесным образом с комплементарной последовательностью целевой цепи, тем самым вытесняя гомологичную последовательность эндогенной целевой цепи. Вытесненная эндогенная цепь может быть обозначена согласно некоторым вариантам осуществления как

5'-эндогенный лоскут ДНК (например, см. фиг. 1E). Данный 5'-эндогенный лоскут ДНК может быть удален с помощью эндонуклеазы 5'-лоскутов (например, FEN1), при этом продукт в виде одноцепочечной ДНК, являющийся теперь гибридизированным с эндогенной целевой цепью, 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.

[0687] Согласно различным вариантам осуществления нуклеотидная последовательность последовательности матрицы RT соответствует нуклеотидной последовательность нецелевой цепи, которая становится вытесненной в виде 5'-лоскутов и которая перекрывает подвергаемый редактированию участок.

[0688] Согласно различным вариантам осуществления удлиненных гидовых РНК, последовательность матрицы обратной транскрипции может кодировать одноцепочечный лоскут ДНК, который комплементарен последовательности эндогенной ДНК, являющейся смежной относительно участка разрыва, при этом одноцепочечный лоскут ДНК включает требуемое нуклеотидное изменение. Одноцепочечный лоскут ДНК может вытеснять эндогенную одноцепочечную ДНК в участке разрыва. Вытесненная эндогенная одноцепочечная ДНК в участке разрыва может иметь 5 -конец и образовывать эндогенный лоскут, который может быть подвергнут вырезанию клеткой. Согласно различным вариантам осуществления удаление 5-концевого эндогенного лоскута может обеспечивать стимулирование образования продукта, поскольку удаление 5-концевого эндогенного лоскута стимулирует гибридизирование одноцепочечного 3'-лоскута ДНК с соответствующей комплементарной цепью ДНК и встраивание или ассимилирование требуемого нуклеотидного изменения, которое несет одноцепочечный 3'-лоскут ДНК, в ДНК-мишени.

[0689] Согласно различным вариантам осуществления удлиненных гидовых РНК клеточная репарация лоскута одноцепочечной ДНК приводит к внесению требуемого нуклеотидного изменения, тем самым обеспечивая образование желаемого продукта.

[0690] Согласно другим вариантам осуществления желаемое нуклеотидное изменение вносится в окне редактирования, которое соответствует приблизительно от -10 до +10 от участка разрыва, или между приблизительно от -20 до +20 от участка разрыва, или между приблизительно от -30 до +30 от участка разрыва, или между приблизительно от -40 до + 40 от участка разрыва, или между приблизительно от -50 до +50 от участка разрыва, или между приблизительно от -60 до +60 от участка разрыва, или между приблизительно от -70 до +70 от участка разрыва, или между приблизительно от -80 до +80

от участка разрыва, или между приблизительно от -90 до +90 от участка разрыва, или между приблизительно от -100 до +100 от участка разрыва, или между приблизительно от -200 до +200 от участка разрыва. Согласно другим вариантам осуществления желаемое нуклеотидное изменение вносится в окне редактирования, которое составляет 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.

[0691] Согласно другим вариантам осуществления желаемое 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.

[0692] Согласно различным аспектам удлиненные гидовые РНК представляют собой модифицированные варианты гидовой РНК. Гидовые РНК могут являться встречающимися в природе, экспрессируемыми из кодирующей нуклеиновой кислоты или

химически синтезированными. В данной области техники хорошо известны способы получения или иным образом синтезирования гидовых РНК и определения подходящей последовательности гидовой РНК, включая протоспейсерную последовательность, которая взаимодействует и гибридизируется с целевой цепью представляющего интерес целевого участка гена.

[0693] Согласно различным вариантам осуществления конкретные аспекты дизайна последовательности гидовой РНК будут зависеть от нуклеотидной последовательности представляющего интерес целевого участка генома (т.е. желаемого сайта, подлежащего редактированию) и типа napDNAbp (например, белка Cas9), присутствующего в системах мультилоскутного праймированного редактирования, описанных в настоящем документе, среди других факторов, таких как расположение последовательности PAM, процентное содержание G/C в целевой последовательности, степень микрогомологии областей, вторичные структуры и т.д.

[0694] В общем, гидовая последовательность представляет собой любую полинуклеотидную последовательность, имеющую достаточную комплементарность с целевой полинуклеотидной последовательностью для гибридизации с целевой последовательностью и прямого специфического для последовательности связывания napDNAbp (например, Cas9, гомолога Cas9 или варианта Cas9) с целевой последовательностью. Согласно некоторым вариантам осуществления степень комплементарности между гидовой последовательностью и соответствующей целевой последовательность. при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет приблизительно или более приблизительно 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99% или более. Оптимальное выравнивание можно определить с использованием любого подходящего алгоритма выравнивания последовательностей, не ограничивающий пример которого включает алгоритм Смита-Уотермана, алгоритм Нидлмана-Вунша, алгоритмы, основанные на преобразовании Берроуза-Уилера (например, алгоритм выравнивания Берроуза-Уилера), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (доступный на soap.genomics.org.cn) и Maq (доступный на maq.sourceforge.net). Согласно некоторым вариантам осуществления гидовая последовательность составляет приблизительно или более 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 в длину.

[0695] Согласно некоторым вариантам осуществления гидовая последовательность составляет менее приблизительно 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 или менее нуклеотидов в длину. Способность гидовой последовательности направлять специфическое к

последовательности связывание мультилоскутного праймирующего редактора с целевой последовательностью можно оценить с помощью любого подходящего анализа. Например, компоненты мультилоскутного праймирующего редактора, включая тестируемую гидовую последовательность, могут быть введены в клетку-хозяин, имеющую соответствующую целевую последовательность, например, путем трансфекции векторами, кодирующими компоненты мультилоскутного праймирующего редактора, раскрытого в настоящем документе, с последующей оценкой предпочтительного расщепления в целевой последовательности, например, с помощью анализа Surveyor, как описано в настоящем документе. Подобным образом, расщепление целевой полинуклеотидной последовательности можно оценить в пробирке, предоставив целевую последовательность, компоненты мультилоскутного праймирующего редактора, включая тестируемую гидовую последовательность и контрольную гидовую последовательность, отличную от тестируемой гидовой последовательности, и сравнения связывание или скорость расщепления целевой последовательности между реакциями тестовой и контрольной последовательностями. Возможны и другие анализы, и это будет понятно специалистам в данной области техники.

[0696] Гидовая последовательность может быть выбрана для нацеливания на любую целевую последовательность. Согласно некоторым вариантам осуществления целевая последовательность представляет собой последовательность в геноме клетки. Иллюстративные целевые последовательности включают те, которые являются уникальными в целевом геноме. Например, для Cas9 *S. pyogenes* уникальная целевая последовательность в геноме может включать целевой участок Cas9 в форме MMMMMMMMNNNNNNNNNNNNXGG (SEQ ID NO: 204), где NNNNNNNNNNNNXGG (SEQ ID NO: 205) (N представляет собой A, G, T или C, и X может быть любым). Уникальная целевая последовательность в геноме может включать целевой участок Cas9 S. pyogenes в форме MMMMMMMMMNNNNNNNNNNNXGG (SEQ ID NO: 206), где NNNNNNNNNNNXGG (SEQ ID NO: 207) (N представляет собой A, G, T или C, и X может быть любым). Для CRISPR1Cas9 S. thermophilus уникальная целевая последовательность в геноме может включать целевой участок Cas9 в форме MMMMMMMMNNNNNNNNNNNNXXAGAAW (SEQ ID NO: 208), где NNNNNNNNNNNNXXAGAAW (SEQ ID NO: 209) (N представляет собой A, G, T или C, и X может быть любым, и W представляет собой A или T). Уникальная целевая последовательность в геноме может включать целевой участок CRISPR 1 Cas9 S. thermophilus в форме MMMMMMMMMNNNNNNNNNNNXXAGAAW (SEQ ID NO: 210), где NNNNNNNNNNNXXAGAAW (SEQ ID NO: 211) (N представляет собой A, G, T или C,

и X может быть любым, и W представляет собой A или T). Для Cas9 *S. pyogenes* уникальная целевая последовательность в геноме может включать целевой участок Cas9 в форме MMMMMMMMNNNNNNNNNNNNXGGXG (SEQ ID NO: 212), где NNNNNNNNNNNNXGGXG (SEQ ID NO: 213) (N представляет собой A, G, T или C, и X может быть любым). Уникальная целевая последовательность в геноме может включать целевой участок Cas9 S. pyogenes в форме MMMMMMMMMNNNNNNNNNNNXGGXG (SEQ ID NO: 214), где NNNNNNNNNNNXGGXG (SEQ ID NO: 215) (N представляет собой A, G, T или C, и X может быть любым). В каждой из этих последовательностей «M» может быть A, G, T или C, и его не нужно рассматривать при идентификации последовательности как уникальной.

[0697] В контексте настоящего изобретения в последовательности РЕгРНК или гидовой РНК, если не указано иное, следует понимать, что буква «T» или «тимин» указывает на основание в последовательности ДНК, которая кодирует последовательность РЕгРНК или гидовой РНК, и предназначена для обозначения основания урацил (U) в РЕгРНК или гидовой РНК, или любого химически модифицированного основания урацил, известного в данной области, такого как 5-метоксиурацил.

[0698] Согласно некоторым вариантам осуществления гидовая последовательность выбирается для уменьшения степени вторичной структуры внутри гидовой последовательности. Вторичная структура может быть определена любым подходящим алгоритмом сворачивания полинуклеотидов. Некоторые программы основаны на вычислении минимальной свободной энергии Гиббса. Примером одного из таких алгоритмов является mFold, описанный Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Другим примером алгоритма фолдинга является он-лайн сервис, разработанный при Institute for Theoretical Chemistry at the University of Vienna, с использованием алгоритма предсказания структуры центроида (см., например, A. R. Gruber et al., 2008, Cell 106(1): 23-24, and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62). Дополнительные алгоритмы можно найти в заявке США № 61/836,080, общая ссылка BI-2013/004A), которая включена в настоящий документ посредством ссылки.

[0699] В общем, парная последовательность tracr включает любую последовательность, которая имеет достаточную комплементарность с последовательностью tracr для обеспечения одного или нескольких из: (1) вырезания гидовой последовательности, фланкированной парными последовательностями tracr, в клетке, содержащей соответствующую последовательность tracr, и (2) образование комплекса с целевой последовательностью, где комплекс содержит сопряженную последовательность tracr, гибридизованную с последовательностью tracr. В общем, степень

комплементарности относится к оптимальному выравниванию парной последовательности tracr и последовательности tracr по длине более короткой из двух последовательностей. Оптимальное выравнивание может быть определено с помощью любого подходящего алгоритма выравнивания и может дополнительно учитывать вторичные структуры, такие как самокомплементарность либо в последовательности tracr, либо в парной последовательности tracr. Согласно некоторым вариантам осуществления степень комплементарности между последовательностью tracr и парной последовательностью tracr по длине более короткой из двух при оптимальном выравнивании составляет приблизительно или более приблизительно 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99% или выше. Согласно некоторым вариантам осуществления последовательность tracr составляет приблизительно или более 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 длину. Согласно некоторым вариантам осуществления последовательность tracr и парная последовательность tracr содержатся в одном транскрипте, так что гибридизация между ними дает транскрипт, имеющий вторичную структуру, такую как шпилька. Предпочтительные образующие петлю последовательности для использования в структурах шпилек имеют длину четыре нуклеотида и наиболее предпочтительно имеют последовательность GAAA. Однако можно использовать более длинные или более короткие последовательности, образующие петель, а также альтернативные последовательности. Последовательности предпочтительно включают триплет нуклеотидов (например, AAA) и дополнительный нуклеотид (например, C или G). Примеры последовательностей, образующих петлю, включают CAAA и AAAG. Согласно варианту осуществления настоящего изобретения транскрипт или транскрибируемая полинуклеотидная последовательность имеет по меньшей мере две или более шпильки. Согласно предпочтительным вариантам транскрипт имеет две, три, четыре или пять шпилек. Согласно другому варианту осуществления настоящего изобретения транскрипт имеет не более пяти шпилек. Согласно некоторым вариантам осуществления отдельный транскрипт дополнительно включает последовательность терминации транскрипции, предпочтительно это полиТ-последовательность, например, шесть Т-нуклеотидов. Дополнительные неограничивающие примеры одиночных полинуклеотидов, содержащих гидовую последовательность, парную последовательность tracr и последовательность tracr, представлены ниже (перечислены с 5' по 3'), где «N» представляет собой основание гидовой последовательности, первый блок строчных букв представляет собой парную последовательность tracr, а второй блок строчных букв представляет собой последовательность tracr, а конечная поли-Т-последовательность представляет собой терминатор транскрипции:

(1)NNNNNNNNGTTTTTGTACTCTCAAGATTTAGAAATAAATCTTGCAGAAGCTACAAA GATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTCGTTAT TTAATTTTTT (SEQ ID NO: 216),

[0700] (2)NNNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAAGC TACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTT TCGTTATTTAATTTTTT (SEQ ID NO: 217),

[0701] (3)NNNNNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAA GCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGT TTTTT (SEQ ID NO: 218),

[0702] (4)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTA AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 219),

[0703] (5)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTA AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGTTTTTTT (SEQ ID NO: 220), И

[0704] (6) NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT AGTCCGTTATCATTTTTTTT (SEQ ID NO: 221).

[0705] Согласно некоторым вариантам осуществления последовательности (1)-(3) используются в сочетании с Cas9 из CRISPR1 S. thermophilus. Согласно некоторым вариантам осуществления последовательности (4)-(6) используются в сочетании с Cas9 из S. pyogenes. Согласно некоторым вариантам осуществления последовательность tracr представляет собой транскрипт, отдельный от транскрипта, содержащего парную последовательность tracr.

[0706] Специалисту в данной области техники очевидно, что для нацеливания любого из слитых белков, содержащих домен Cas9 и одноцепочечный ДНК-связывающий белок, как раскрыто в настоящем документе, на целевой участок, например, участок, содержащий точку мутацию, которую необходимо отредактировать, обычно необходимо совместно экспрессировать слитый белок вместе с гидовой РНК, например, сгРНК. Как более подробно объяснено в другом месте настоящего документа, гидовая РНК обычно содержит каркас tracrRNA, обеспечивающий связывание Cas9, и гидовую последовательность, которая придает специфичность к последовательности слитому белку Cas9:редактирующий нуклеиновую кислоту фермент/домен.

[0707] Согласно некоторым вариантам осуществления гидовая РНК содержит структуру 5′-[гидовая последовательность]-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAAGGCUAGUCCGUUAUCAACUUG

AAAAAGUGGCACCGAGUCGGUGCUUUUU-3′ (SEQ ID NO: 222), где гидовая последовательность содержит последовательность, комплементарную целевой последовательности. Длина гидовой последовательности обычно составляет 20 нуклеотидов. Последовательности подходящих годовых РНК для нацеливания слитых белков Cas9:редактирующий нуклеиновую кислоту фермент/домен на конкретные целевые участки генома будут очевидны специалистам в данной области техники на основании настоящего раскрытия. Такие подходящие годовые последовательности РНК обычно содержат годовые последовательности, которые комплементарны нуклеиновой последовательности в пределах 50 нуклеотидов против хода транскрипции или по ходу транскрипции от целевого редактируемого нуклеотида. В настоящем документе представлены некоторые иллюстративные годовые РНК-последовательности, подходящие для нацеливания любого из предусмотренных слитых белков на конкретные целевые последовательности. Дополнительные годовые последовательности хорошо известны в данной области техники и могут использоваться с мультилоскутными праймирующим редактором, описанным в настоящем документе.

[0708] Согласно другим вариантам осуществления РЕгРНК включают приведенные на фиг. 3D.

[0709] Согласно другим вариантам осуществления, РЕгРНК могут включать приведенные на фиг. 3E.

[0710] На фиг. 3D представлена структура варианта осуществления РЕгРНК, охватываемого настоящим изобретением, которая может быть сконструирована в соответствии с методологией, определенной в примере 2. РЕгРНК содержит три основных составных элемента, расположенных в направлении от 5' к 3', а именно: спейсер, ядро гРНК и плечо удлинения на 3′ конце. Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5' к 3', а именно: сайт связывания праймера (A), матрица редактирования (B) и плечо гомологии (C). Кроме того, РЕгРНК может содержать необязательную область модификатора 3'-конца (e1) и необязательную область модификатора 5'-конца (e2). Кроме того, РЕгРНК может содержать сигнал терминации транскрипции на 3'-конце РЕгРНК (не показано). Эти структурные элементы дополнительно определены в данном документе. Изображение структуры РЕгРНК не является ограничивающим и охватывает варианты расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми другими показанными областями, не ограничиваясь тем, что они расположены на 3'- и 5'-концах.

[0711] На фиг. 3E представлена структура другого варианта осуществления

охватываемой настоящим изобретением РЕгРНК, которая может быть сконструирована в соответствии с методологией, определенной в примере 2. РЕгРНК содержит три основных составных элемента, расположенных в направлении от 5' к 3', а именно: спейсер, ядро гРНК и плечо удлинения на 3′ конце. Плечо удлинения может быть дополнительно разделено на следующие структурные элементы в направлении от 5' к 3', а именно: сайт связывания праймера (A), матрица редактирования (B) и плечо гомологии (C). Кроме того, РЕгРНК может содержать необязательную область модификатора 3'-конца (e1) и необязательную область модификатора 5'-конца (e2). Кроме того, РЕгРНК может содержать сигнал терминации транскрипции на 3'-конце РЕгРНК (не показан). Эти структурные элементы дополнительно определены в данном документе. Изображение структуры РЕгРНК не является ограничивающим и охватывает варианты расположения элементов. Например, необязательные модификаторы последовательности (e1) и (e2) могут быть расположены внутри или между любыми другими показанными областями, не ограничиваясь тем, что они расположены на 3'- и 5'-концах.

Улучшения РЕгРНК

[0712] РЕгРНК также могут включать дополнительные усовершенствования конструкции, которые могут модифицировать свойства и/или характеристики РЕгРНК, тем самым повышая эффективность мультилоскутного праймированного редактирования. Согласно различным вариантам осуществления эти усовершенствования могут относиться к одной или нескольким из множества различных категорий, включая без ограничения: (1) конструкции, обеспечивающие эффективную экспрессию функциональных РЕгРНК из промоторов, не являющихся полимеразой III (pol III), что позволит экспрессию более длинных РЕгРНК без обременительных требований к последовательности, (2) усовершенствования основного Cas9-связывающего каркаса РЕгРНК, которые могут повысить эффективность, (3) модификации РЕгРНК для улучшения процессивности RT, позволяющие вставлять более длинные последовательности в целевые геномные локусы, и (4) добавление мотивов РНК к 5'- или 3'-концам РЕгРНК, которые улучшают стабильность РЕгРНК, усиливают процессивность RT, предотвращают неправильную укладку РЕгРНК или привлекают дополнительные факторы, важные для редактирования генома

[0713] Согласно одному варианту осуществления РЕгРНК может быть сконструирована с промоторами polIII для улучшения экспрессии более длинной РЕгРНК с более длинными плечами удлинения. сгРНК обычно экспрессируются с промотором U6 snRNA. Этот промотор рекрутирует pol III для экспрессии ассоциированной РНК и полезен для экспрессии коротких РНК, которые сохраняются в ядре. Однако pol III не является высокопроцессивным и не может экспрессировать РНК длиной более нескольких сотен

нуклеотидов на уровнях, необходимых для эффективного редактирования генома. Кроме того, pol III может останавливаться или заканчиваться на участках U, что потенциально ограничивает разнообразие последовательностей, которые могут быть вставлены с использованием РЕгРНК. Другие промоторы, которые привлекают полимеразу II (например, pCMV) или полимеразу I (например, промотор U1 snRNA), исследовали на их способность экспрессировать более длинные сгРНК. Однако эти промоторы обычно частично транскрибируются, что может привести к дополнительной последовательности 5' спейсера в экспрессируемой РЕгРНК, что, как было показано, приводит к заметному снижению активности Cas9:сгРНК сайт-зависимым образом. Кроме того, в то время как РЕгРНК, транскрибируемые pol III, могут просто терминироваться серией из 6-7 U, РЕгРНК, транскрибируемые из pol II или pol I, требуют другого сигнала терминации. Часто такие сигналы также приводят к полиаденилированию, что приводит к нежелательному транспорту РЕгРНК из ядра. Точно так же РНК, экспрессируемые с промоторов pol II, таких как pCMV, обычно имеют 5'-кэпирование, что также приводит к их экспорту в ядро.

[0714] Ранее, Rinn и соавторы провели скрининг различных платформ экспрессии для продукции длинных некодирующих РНК-(lncRNA)меченных сгРНК[183]. Эти платформы включают РНК, экспрессируемые из pCMV и заканчивающиеся элементом ENE из ncRNA MALAT1 человека[184] , элементом PAN ENE из KSHV[185] или 3'-боксом из U1 snRNA[186] . Примечательно, что MALAT1 ncRNA и PAN ENE образуют тройные спирали, защищающие полиА-хвост[184, 187]. Эти конструкции могут также повышать стабильность РНК. Предполагается, что эти системы экспрессии также будут способствовать экспрессии более длинных РЕгРНК.

[0715] Кроме того, был разработан ряд методов для расщепления части промотора pol II, которая будет транскрибироваться как часть РЕгРНК, с добавлением либо саморасщепляющегося рибозима, такого как рибозимы hammerhead[188], pistol[189], hatchet[189], шпилька [190], VS[191], twister[192] или twister sister[192], или другие саморасщепляющиеся элементы для процессинга транскрибируемой направляющей, или шпилька, которая распознается Csy4[193] и также приводит к процессингу направляющей. Кроме того, предполагается, что включение множественных мотивов ENE может приводить к улучшенной экспрессии и стабильности РЕгРНК, как ранее продемонстрировано для KSHV PAN РНК и элемента[185]. Также ожидается, что циркуляризация РЕгРНК в форме кольцевой интронной РНК (ciRNA) может также приводить к усиленной экспрессии и стабильности РНК, а также к локализации в ядре[194].

[0716] Согласно различным вариантам осуществления РЕгРНК может включать различные указанные выше элементы, что иллюстрируется следующей

последовательностью.

[0717] Неограничивающий пример 1 - Платформа экспрессии PEгРНК, состоящая из pCMV, шпильки Csy4, PEгРНК и MALAT1 ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGAC GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTG ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT GGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGA CCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTAGGGTCATGAAGGTTTTTCTT TTCCTGAGAAAACAACACGTATTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAA AAAAAAAAAAGCAAAAGATGCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATC CCTGCGGCGTCTTTGCTTTGACT (SEQ ID NO: 223)

[0718] Неограничивающий пример 2 - Платформа для экспрессии PEгРНК, состоящая из pCMV, Csy4, PEгРНК и PAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGAC GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTG ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT GGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGA CCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTGTTTTGGCTGGGTTTTTCCTTG TTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAA CATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATG CAACTTACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 224)

[0719] Неограничивающий пример 3 - Платформа для экспрессии PEгРНК, состоящая из pCMV, Csy4, PEгРНК и 3xPAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG
CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGAC
GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG
GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA
CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTG
ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT
CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
GGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG
TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGA
CCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTGTTTTGGCTGGGTTTTTCCTTG
TTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAA
CATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATG
CAACTTACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAACACA
CTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTT
TTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAA
ATTAAACTAATACCATAACGTAATGCAACTTACAACATAAATAAAGGTCAATGTTTAATCCATA
AAAAAAAAAAAAAAAAAATCTCTCTGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACC
TCCAGTGACCAGACGGCAAGGTTTTTATCCCAGTGTATATTGGAAAAACATGTTATACTTTTGA
CAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACTTACAACATAA
ATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAA (SEQ ID NO: 225)

[0720] Неограничивающий пример 4 - Платформа для экспрессии РЕгРНК, состоящая из pCMV, Csy4, РЕгРНК и 3'-бокса

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG
CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGAC
GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG
GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGA
CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTG
ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACT
CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT
GGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAGACTGAGCACGTGAG
TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGA
CCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTGTTTCAAAAGTAGACTGTACGC
TAAGGGTCATATCTTTTTTTGTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 226)

[0721] Неограничивающий пример 5 - Платформа экспрессии РЕгРНК, состоящая из

pU1, Csy4, РЕгРНК и 3'-бокса

CTAAGGACCAGCTTCTTTGGGAGAGAACAGACGCAGGGGCGGGAGGGAAAAAGGGA GAGGCAGACGTCACTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCGGTTGAGTGGCAGAAA GGCAGACGGGGACTGGGCAAGGCACTGTCGGTGACATCACGGACAGGGCGACTTCTATGTAG ATGAGGCAGCGCAGAGGCTGCTGCTTCGCCACTTGCTGCTTCACCACGAAGGAGTTCCCGTG CCCTGGGAGCGGGTTCAGGACCGCTGATCGGAAGTGAGAATCCCAGCTGTGTGTCAGGGCTG GAAAGGGCTCGGGAGTGCGCGGGGCAAGTGACCGTGTGTGTAAAGAGTGAGGCGTATGAGG CTGTGTCGGGGCAGAGGCCCAAGATCTCAGTTCACTGCCGTATAGGCAGGGCCCAGACTGAG CACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTCAGCAAGTTCAGA GAAATCTGAACTTGCTGGATTTTTGGAGCAGGGAGATGGAATAGGAGCTTGCTCCGTCCACTC CACGCATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGCACCCACTTTCTGGAGTTTCAAA AGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTGGTTTGTGTCTTGGTTGGCGTCTTAA A (SEQ ID NO: 227).

[0722] РЕгРНК может быть улучшена путем внесения улучшений в последовательности каркаса или ядра. Это можно сделать известным образом.

[0723] Ядро, Cas9-связывающий каркас РЕгРНК, вероятно, могут быть улучшены для повышения активности РЕ. Несколько таких подходов уже были продемонстрированы. Например, первый парный элемент каркаса (P1) содержит элемент сопряжения GTTTT-AAAAC (SEQ ID NO: 3939). Было показано, что такие запуски Ts приводят к паузе pol III и преждевременной терминации транскрипта РНК. Было показано, что рациональная мутация одной из пар TA в пару GC в этой части P1 усиливает активность сгРНК, указывая на то, что этот подход также возможен для РЕгРНК[195] . Кроме того, увеличение длины P1 также, как было показано, усиливает сворачивание сгРНК и ведет к повышению активности[195], указывая на то, что это еще один путь для улучшения активности РЕгРНК. Примеры улучшений ядра могут включать:

[0724] РЕгРНК, содержащая удлинение 6 нуклеотидов в P1

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGCTCATGAAAATGAGCTAGCA AGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTC TGCCATCAAAGCGTGCTCAGTCTGTTTTTTT (SEQ ID NO: 228)

[0725] РЕгРНК, содержащая мутацию TA в GC в P1

GGCCCAGACTGAGCACGTGAGTTTGAGAGCTAGAAATAGCAAGTTTAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAG CGTGCTCAGTCTGTTTTTTT (SEQ ID NO: 229)

[0726] Согласно различным другим вариантам осуществления РЕгРНК может быть улучшена путем внесения модификаций в область матрицы редактирования. По мере

увеличения размера вставки, матрицей для которой является РЕгРНК, она с большей вероятностью будет расщепляться эндонуклеазами, подвергаться спонтанному гидролизу или складываться во вторичные структуры, которые не могут быть обратно транскрибированы с помощью RT, или которые нарушают сворачивание каркаса РЕгРНК и последующее связывание Cas9-RT. Соответственно, вполне вероятно, что модификация матрицы РЕгРНК может быть необходима для воздействия на большие вставки, такие как вставки целых генов. Некоторые стратегии для этого включают включение модифицированных нуклеотидов в синтетические или полусинтетические РЕгРНК, которые делают РНК более устойчивыми к деградации или гидролизу или с меньшей вероятностью принимают ингибиторные вторичные структуры[196]. Такие модификации могут включать 8-аза-7-деазагуанозин, который может уменьшить вторичную структуру РНК в G-богатых последовательностях, закрытые нуклеиновые кислоты (LNA), которые уменьшают деградацию и улучшают некоторые виды вторичной структуры РНК, Модификации 2'-О-метил, 2'-фтор или 2'-О-метоксиэтокси повышающие стабильность РНК. Такие модификации также могут быть включены в другие части РЕгРНК для повышения стабильности и активности. В качестве альтернативы или дополнительно матрица РЕгРНК может быть сконструирована таким образом, чтобы она одновременно кодировала желаемый белковый продукт, а также с большей вероятностью принимала простые вторичные структуры, которые могут быть развернуты с помощью RT. Такие простые структуры будут действовать как термодинамические, снижая вероятность появления более сложных структур, препятствующих обратной транскрипции. Наконец, можно также разделить матрицу на две отдельные РЕгРНК. В такой конструкции РЕ будет использоваться для инициации транскрипции, а также рекрутирования отдельной матричной РНК в целевой сайт с помощью РНК-связывающего белка, слитого с Cas9, или элемента распознавания РНК на самой РЕгРНК, такого как аптамер MS2. RT может либо напрямую связываться с этой отдельной матричной РНК, либо инициировать обратную транскрипцию на исходной РЕгРНК перед переключением на вторую матрицу. Такой подход может сделать возможными длинные вставки, предотвращая неправильную укладку РЕгРНК при добавлении длинной матрицы, а также не требуя диссоциации Cas9 от генома для возникновения длинных вставок, которые, возможно, могут ингибировать длинные вставки на основе РЕ.

[0727] Согласно другим вариантам осуществления РЕгРНК может быть улучшена путем введения дополнительных мотивов РНК на 5'- и 3'-концах РЕгРНК или даже в положениях между ними (например, в ядре гРНК или в спейсере). Несколько таких мотивов, таких как PAN ENE из KSHV и ENE из MALAT1, обсуждались выше как

возможные средства терминации экспрессии более длинных РЕгРНК с промоторов, отличных от pol III. Эти элементы образуют тройные спирали РНК, которые охватывают хвост полиА, в результате чего они остаются внутри ядра[184,187]. Однако, образуя сложные структуры на 3'-конце РЕгРНК, которые перекрывают концевой нуклеотид, эти структуры, вероятно, также могут помочь предотвратить опосредованную экзонуклеазой деградацию РЕгРНК.

[0728] Другие структурные элементы, вставленные на 3'-конце, также могут повышать стабильность РНК, хотя и не допуская терминации с промоторов, не являющихся pol III. Такие мотивы могут включать шпильки или квадруплексы РНК, закрывающие 3'-конец[197], или саморасщепляющиеся рибозимы, такие как HDV, которые приводят к образованию 2'-3'-циклического фосфата на 3'-конце, а также потенциально делают РЕгРНК менее вероятно подверженной расщеплению экзонуклеазами[198]. Индуцирование циклизации РЕгРНК посредством неполного сплайсинга с образованием ciRNA может также повышать стабильность РЕгРНК и приводить к сохранению РЕгРНК внутри ядра[194].

[0729] Дополнительные мотивы РНК могут также улучшать процессивность RT или повышать активность РЕгРНК посредством усиления связывания RT с дуплексом ДНК-РНК. Добавление нативной последовательности, связанной с RT, в его родственный ретровирусный геном может усиливать активность RT[199]. Это может включать нативный участок связывания праймеров (PBS), полипуриновый тракт (PPT) или петли kissing, участвующие в димеризации ретровирусного генома и инициации транскрипции[199].

[0730] Добавление мотивов димеризации, таких как петли kissing или пара рецепторов тетрапетля/тетрапетля GNRA[200], на 5'- и 3'-концах РЕгРНК может также привести к эффективной циркуляризации РЕгРНК, улучшая стабильность. Кроме того, предполагается, что добавление этих мотивов может сделать возможным физическое разделение спейсера и праймера РЕгРНК, предотвращая окклюзию спейсера, которая будет препятствовать активности РЕ. Короткие 5'-удлинения или 3'- удлинения РЕгРНК, которые образуют небольшую шпильку в области спейсера или вдоль участка связывания праймера, также могут выгодно конкурировать с отжигом интракомплементарных областей по длине РЕгРНК, как например, взаимодействие между спейсером и участком связывания праймера, которое может возникнуть. Наконец, петли kissing также могут быть использованы для рекрутирования других матричных РНК в геномный сайт и обеспечения переключения активности RT с одной РНК на другую. В качестве иллюстративных вариантов различных вторичных структур РЕгРНК, изображенная на фиг. 3D и фиг. 3E, показывает ряд вторичных структур РНК, которые можно встроить в любую область РЕгРНК, в том числе в концевые части плеча удлинения (т.е. e1 и e2), как показано.

[0731] Примеры улучшения включают без ограничения:

[0732] Слияние PEгРНК-HDV

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG GCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGC GTGCTCAGTCTGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACA TGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 230)

[0733] Петля kissing PEгРНК-MMLV

GGTGGGAGACGTCCCACCGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTC GGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCTT TTTTT (SEQ ID NO: 231)

[0734] Петля kissing рибозима PEгРНК-VS

GAGCAGCATGGCGTCGCTGCTCACGGCCCAGACTGAGCACGTGAGTTTTAGA GCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGAC CGAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTCCATCAGTTGACACC CTGAGGTTTTTTT (SEQ ID NO: 232)

[0735] Рецептор тетрапетля/тетрапетля PEгРНК-GNRA

GCAGACCTAAGTGGUGACATATGGTCTGGGCCCAGACTGAGCACGTGAGTTTT AGAGCTAUACGTAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTUACGAAGTGG GACCGAGTCGGTCCTCTGCCATCAAAGCTTCGACCGTGCTCAGTCTGCATGCGATTAG AAATAATCGCATGTTTTTTT (SEQ ID NO: 233)

[0736] Вторичное слияние РНК-HDV с переключением матрицы PEгРНК

TCTGCCATCAAAGCTGCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCGG CCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGC GAATGGGACTTTTTTT (SEQ ID NO: 234)

[0737] Каркасы PEгРНК могут быть дополнительно улучшены с помощью направленной эволюции, аналогично тому, как были улучшены SpCas9 и праймирующие редакторы (PE). Направленная эволюция может усилить распознавание PEгРНК с помощью Cas9 или эволюционировавших вариантов Cas9. Кроме того, вполне вероятно, что разные последовательности каркаса PEгРНК будут оптимальными для разных геномных локусов, либо повышая активность PE в рассматриваемом участке, либо снижая нецелевые активности, либо и то, и другое. Наконец, эволюция каркасов PEгРНК, к которым были добавлены другие мотивы РНК, с большой вероятностью улучшит активность слитых PEгРНК по сравнению с неэволюционировавшими слитыми РНК. Например, эволюция аллостерических рибозимов, состоящих из c-di-GMP-I аптамеров и рибозимов hammerhead,

привела к значительному улучшению активности[202], указывая на то, что эволюция также улучшит активность слияний hammerhead-PEгРНК. Кроме того, хотя Cas9 в настоящее время обычно не переносит 5'-удлинение сгРНК, направленная эволюция, вероятно, вызовет мутации, смягчающие эту непереносимость, позволяя использовать дополнительные мотивы РНК.

[0738] Настоящее изобретение охватывает любые такие способы дальнейшего повышения эффективности раскрытых в настоящем документе систем мультилоскутного праймированного редактирования.

[0739] Согласно различным вариантам осуществления может быть предпочтительно ограничить появление непрерывной последовательности T из плеча удлинения, поскольку последовательные серии T могут ограничивать способность PEгРНК к транскрипции. Например, необходимо избегать цепочек, содержащих по меньшей мере три последовательных T, по меньшей мере четыре последовательных T, по меньшей мере пять последовательных T, по меньшей мере шесть последовательных T, по меньшей мере семь последовательных T, по меньшей мере восемь последовательных T, по меньшей мере девять последовательных T, по меньшей мере десять последовательных T, по меньшей мере одиннадцать последовательных T, по меньшей мере двенадцать последовательных T, по меньшей мере тринадцать последовательных T, по меньшей мере четырнадцать последовательных T или по меньшей мере пятнадцать последовательных T, при конструировании PEгРНК, или они должны быть по меньшей мере удалены из конечной сконструированной последовательности. Согласно одному варианту осуществления можно избежать включения нежелательных цепочек последовательных T в плече удлинения PEгРНК, избегая целевых участков, которые богаты последовательными парами нуклеотидных оснований A:T.

Конструкции расщепленной PEгРНК для транс-праймированного редактирования

[0740] Настоящее изобретение также охватывает *транс*-праймированное редактирование, которое относится к модифицированному варианту праймированного редактирования, которое осуществляется путем разделения PEгРНК на две отдельные молекулы: гидовую РНК и молекулу tPERT. Молекула tPERT запрограммирована на совместную локализацию с комплексом праймированного редактирования в целевом участке ДНК, перенося участок связывания праймера и матрицу для синтеза ДНК в транс относительно праймирующего редактора. Например, см. фиг. 3G для варианта осуществления транс-праймирующего редактора (tPE), на котором показана двухкомпонентная система, содержащая (1) комплекс рекрутирующего белка (RP)-

PE:гРНК и (2) tPERT, который включает участок связывания праймера и матрицу для синтеза ДНК., соединенную с доменом рекрутирования РНК-белок (например, петля на стебле или шпилька), при этом рекрутирующий белковый компонент комплекса RP-PE:гРНК рекрутирует tPERT в редактируемый целевой участок, тем самым связывая PBS и матрицей для синтеза ДНК с праймирующим редактором *в транс*. Другими словами, tPERT сконструирован таким образом, чтобы содержать (полностью или частично) плечо удлинения РЕгРНК, которое включает участок связывания праймера и матрицу для синтеза ДНК. Одним из преимуществ этого подхода является отделение плеча удлинения РЕгРНК от гидовой РНК, тем самым сводя к минимуму взаимодействия отжига, которые имеют тенденцию возникать между PBS плеча удлинения и спейсерной последовательностью гидовой РНК.

[0741] Ключевой особенностью *транс*-праймированного редактирования является способность *транс*-праймирующего редактора рекрутировать tPERT в участок редактирования ДНК, тем самым эффективно совместно локализуя все функции РЕгРНК в участке праймированного редактирования. Рекрутирование может быть достигнуто путем установки домена рекрутирования РНК-белок, такого как аптамер MS2, в tPERT и слияния соответствующего рекрутирующего белка с праймирующим редактором (например, через линкер с napDNAbp или через линкер с полимеразой), который способен специфически связываться с доменом рекрутирования РНК-белок, тем самым рекрутируя молекулу tPERT в комплекс праймированного редактирования. Как показано в процессе, описанном на фиг. 3H, комплекс RP-PE:гРНК связывается с целевой последовательностью ДНК и разрывает ее. Затем рекрутирующий белок (RP) рекрутирует tPERT для совместной локализации с комплексом праймированного редактора, связанным с целевым участком ДНК, тем самым позволяя участку связывания праймера, расположенному на tPERT, связываться с последовательностью праймера на цепи с разрывом, и впоследствии позволяя полимеразе (например, RT) синтезировать одну цепь ДНК против матрицы для синтеза ДНК, расположенной на tPERT, вплоть до 5'-конца tPERT.

[0742] Хотя tPERT показан на фиг. 3G и фиг. 3H, как содержащий PBS и матрицу для синтеза ДНК на 5'-конце домена рекрутирования РНК-белка, tPERT в других конфигурациях может быть разработан с PBS и матрицей для синтеза ДНК, расположенной на 3'-конце домена рекрутирования РНК-белка. Однако tPERT с 5'-удлинением имеет то преимущество, что синтез одноцепочечной ДНК естественным образом заканчивается на 5'-конце tPERT и, таким образом, не возникает риска использования какой-либо части домена рекрутирования РНК-белок в качестве матрицы в ходе стадии синтеза ДНК праймированного редактирования.

Способ конструирования РЕгРНК

[0743] Настоящее изобретение также относится к способам конструирования РЕгРНК.

[0744] Согласно одному аспекту конструирования подход конструирование может учитывать конкретное применение, для которого используется праймированное редактирование. Например, как проиллюстрировано и раскрыто в настоящем документе, праймированное редактирование может быть использовано без ограничений, для (a) внесения корректирующих мутацию изменений в нуклеотидную последовательность, (b) установки белковых и РНК-меток, (c) установки иммуноэпитопов на представляющих интерес белках, (d) установки индуцируемых доменов димеризации в белках, (e) установки или удаления последовательностей, чтобы изменить эту активность биомолекулы, (f) установки сайтов-мишеней рекомбиназы, чтобы направить определенные генетические изменения, и (g) мутагенеза целевой последовательности с использованием RT сниженной точности. В дополнение к этим способам, которые, как правило, вставляют, изменяют или удаляют нуклеотидные последовательности в представляющих интерес целевых участках, праймирующие редакторы также можно использовать для конструирования высокопрограммируемых библиотек, а также для проведения записи клеточных данных и исследований по отслеживанию клонов. В этих различных применениях, как описано в настоящем документе, конкретные аспекты конструкции относятся к получению РЕгРНК, которая особенно полезна для любого данного из этих применений.

[0745] При конструировании РЕгРНК для любого конкретного применения или использования праймированного редактирования может быть принят во внимание ряд факторов, которые включают без ограничения:

(a) целевую последовательность, т.е. нуклеотидную последовательность, в которой праймирующий редактор желает установить одну или несколько модификаций нуклеотидных оснований,

(b) расположение участка разреза в целевой последовательности, т.е. конкретное положение нуклеотидного основания, в котором праймирующий редактор будет индуцировать одноцепочечный разрыв с созданием 3'-концевой последовательности праймера RT на одной стороне разрыва и 5'-концового эндогенного лоскута на другой стороне разрыва (который в конечном итоге удаляется с помощью FEN1 или ее эквивалента и заменяется 3'-лоскутом оцДНК. Участок разреза аналогичен «положению редактирования», поскольку он создает 3'-концевую последовательность праймера RT, которая удлиняется посредством RT в ходе РНК-зависимой полимеризации ДНК с созданием 3'-лоскута одноцепочечной ДНК, содержащего желаемое редактирование,

которое затем заменяет 5'-лоскут эндогенной ДНК в целевой последовательности.

(c) доступные последовательности PAM (включая канонические участки PAM SpCas9, а также неканонические участки PAM, распознаваемые вариантами и эквивалентами Cas9 с удлиненными или отличающимися специфичностями PAM),

(d) расстояние между доступными последовательностями PAM и расположение участка разреза в целевой последовательности,

(e) конкретный Cas9, вариант Cas9 или эквивалент Cas9 используемого праймирующего редактора,

(f) последовательность и длина участка связывания праймера,

(g) последовательность и длина матрицы редактирования,

(h) последовательность и длина плеча гомологии,

(i) спейсерная последовательность и длина и

(j) последовательность ядра.

[0746] Эти аспекты раскрыты в настоящем документе выше.

[0747] Согласно одному варианту осуществления настоящее изобретение относится к подходу к конструированию подходящей PEгРНК и необязательно руководству по конструированию вводящей разрыв сгРНК для разрыва второго участка. Согласно этому варианту осуществления представлен пошаговый набор инструкций по конструированию PEгРНК и вносящей разрыв сгРНК для праймированного редактирования, при котором учитывается один или несколько из приведенных выше аспектов. Стадии относятся к примерам, показанным на фиг. 70A-70I.

1. **Определение целевой последовательности и редактирование**. Получите последовательность целевого участка ДНК (~ 200 п.о.), сосредоточенную вокруг местоположения желаемого редактирования (точечная мутация, вставка, удаление или их комбинация). См. фиг. 70A.

2. **Положение целевых PAM**. Определите PAM в непосредственной близости от желаемого места редактирования. PAM можно идентифицировать на любой цепи ДНК, проксимально к желаемому месту редактирования. В то время как PAM, близкие к положению редактирования, являются предпочтительными (т. е. сайт разрыва находится менее чем в пределах 30 нуклеотидов от положения редактирования или менее чем в пределах 29 нуклеотидов, 28 нуклеотидов, 27 нуклеотидов, 26 нуклеотидов, 25 нуклеотидов, 24 нуклеотидов, 23 нуклеотидов, 22 нуклеотидов, 21 нуклеотидов, 20 нуклеотидов, 19 нуклеотидов, 18 нуклеотидов, 17 нуклеотидов, 16 нуклеотидов, 15 нуклеотидов, 14 нуклеотидов, 13 нуклеотидов, 12 нуклеотидов, 11 нуклеотидов, 10 нуклеотидов, 9 нуклеотидов, 8 нуклеотидов, 7 нуклеотидов, 6 нуклеотидов, 5 нуклеотидов,

4 нуклеотидов, 3 нуклеотидов или 2 нуклеотидов от положения редактирования до участка разрыва), можно установить редактирования с помощью протоспейсеров и PAM, которые помещают разрыв на ≥ 30 нуклеотидов от положения редактирования. См. фиг. 70B.

3. **Положение участков разрыва**. Для каждого рассматриваемого PAM определите соответствующий участок разрыва и на какой цепи он находится. Для никазы Sp Cas9 H840A расщепление происходит в цепи, содержащей PAM, между 3-м и 4-м основаниями 5' по отношению к NGG PAM. Все отредактированные нуклеотиды должны располагаться на 3'-конце участка разрыва, поэтому соответствующие PAM должны размещать участок разрыва на 5'-конце целевого редактирования на цепи, содержащей PAM. В примере, показанном ниже, есть два возможных PAM. Для простоты оставшиеся стадии демонстрируют конструкцию PEгРНК с использованием только PAM 1. См. фиг. 70C.

4. **Конструирование спейсерной последовательности**. Протоспейсер Sp Cas9 соответствует 20 нуклеотидам 5' по отношению к NGG PAM на цепи, содержащей PAM. Для эффективной инициации транскрипции Pol III требуется, чтобы G был первым транскрибируемым нуклеотидом. Если первым нуклеотидом протоспейсера является G, последовательность спейсера для PEгРНК представляет собой просто последовательность протоспейсера. Если первый нуклеотид протоспейсера не представляет собой G, последовательность спейсера PEгРНК представляет собой G, за которой следует последовательность протоспейсера. См. фиг. 70D.

5. **Конструирование участка связывания праймера (PBS)**. Используя исходную последовательность аллеля, определите праймер ДНК на цепи, содержащей PAM. 3'-конец ДНК-праймера представляет собой нуклеотид непосредственно перед участком разрыва (т.е. 4-е основание 5' по отношению к NGG PAM для Sp Cas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера PEгРНК (PBS), содержащий от 12 до 13 нуклеотидов, комплементарных ДНК-праймеру, можно использовать для последовательностей, которые содержат ~ 40-60% GC. Для последовательностей с низким содержанием GC следует тестировать более длинные (от 14 до 15 нуклеотидов) PBS. Для последовательностей с более высоким содержанием GC следует тестировать более короткие (от 8 до 11 нуклеотидов) PBS. Оптимальные последовательности PBS должны быть определены эмпирически, независимо от содержания GC. Чтобы сконструировать последовательность PBS длины *p*, возьмите обратный комплемент первых *p* нуклеотидов 5' от участка разрыва в цепи, содержащей PAM, используя последовательность *исходного* аллеля. См. фиг. 70E.

6. **Конструирование матрицы RT (или матрицы для синтеза ДНК)**. Матрица

RT (или матрица для синтеза ДНК, где полимераза не является обратной транскриптазой) кодирует сконструированное редактирование и гомологию с последовательностью, соседней с редактированием. Согласно одному варианту осуществления эти области соответствуют матрице для синтеза ДНК на фиг. 3D и фиг. 3E, где матрица для синтеза ДНК содержит «матрицу редактирования» и «плечо гомологии». Оптимальная длина матрицы RT зависит от целевого участка. Для коротких редактирований (положения от +1 до +6) рекомендуется протестировать короткую (от 9 до 12 нуклеотидов), среднюю (от 13 до 16 нуклеотидов) и длинную (от 17 до 20 нуклеотидов) матрицу RT. Для редактирований дальнего диапазона (положения +7 и далее) рекомендуется использовать матрицы RT, которые простираются как минимум на 5 нуклеотидов (предпочтительно на 10 или более нуклеотидов) от положения редактирования, чтобы обеспечить достаточную гомологию 3' лоскута ДНК. Для редактирований дальнего диапазона следует скринировать несколько матриц RT, чтобы определить функциональные конструкции. Для более крупных вставок и делеций (≥5 нуклеотидов) рекомендуется включение большей 3'-гомологии (~ 20 нуклеотидов или более) в матрицу RT. Эффективность редактирования обычно снижается, когда матрица RT кодирует синтез G в качестве последнего нуклеотида в продукте обратной транскрипции ДНК (соответствует C в матрице RT PEгРНК). Поскольку многие матрицы RT поддерживают эффективное праймированное редактирование, при разработке матриц RT рекомендуется избегать использования G в качестве конечного синтезированного нуклеотида. Чтобы сконструировать последовательность матрицы RT длины *r*, используйте *желаемую* последовательность аллеля и возьмите обратный комплемент первых нуклеотидов *r* 3’ участка разрыва в цепи, которая первоначально содержала PAM. Обратите внимание, что по сравнению с редактированиями SNP, редактирования вставкой или делецией с использованием матриц RT той же длины не будет содержать идентичной гомологии. См. фиг. 70F.

7. **Сборка полной последовательности PEгРНК**. Соедините компоненты PEгРНК в следующем порядке (от 5 до 3): спейсер, каркас, матрица RT и PBS. См. фиг. 70G.

8. **Конструирование вносящих разрыв сгРНК для PE3**. Идентифицируйте PAM в не редактируемой цепи до и после редактирования. Оптимальные положения разрыва сильно зависят от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии от 40 до 90 нуклеотидов в 5' положении от PEгРНК-индуцированного разрыва, приводят к более высокому выходу редактирования и меньшему количеству вставок-делеций. Вносящая разрыв сгРНК имеет спейсерную последовательность, которая соответствует 20-нуклеотидному протоспейсеру в *исходном* аллеле, с добавлением 5'-G, если протоспейсер не начинается с G. См. фиг. 70H.

9. **Конструирование вносящих разрыв сгРНК PE3b**. Если PAM существует в комплементарной цепи и соответствующий ему протоспейсер перекрывается с последовательностью, предназначенной для редактирования, это редактирование может быть кандидатом для системы PE3b. В системе PE3b спейсерная последовательность вносящей разрыв сгРНК соответствует последовательности желаемого редактируемого аллеля, но не исходного аллеля. Система PE3b работает эффективно, когда отредактированный нуклеотид (нуклеотиды) попадает в область затравки (~ 10 нуклеотидов, примыкающую к PAM) протоспейсера вносящей затравку сгРНК. Это предотвращает разрыв комплементарной цепи до установки отредактированной цепи, предотвращая конкуренцию между PEгРНК и сгРНК за связывание ДНК-мишени. PE3b также позволяет избежать образования одновременных разрывов на обеих цепях, что значительно снижает образование дефектов при сохранении высокой эффективности редактирования. сгРНК PE3b должны иметь спейсерную последовательность, которая соответствует 20-нуклеотидному протоспейсеру в *желаемом* аллеле, с добавлением 5'G, если это необходимо. См. фиг. 70I.

[0748] Приведенный выше постадийный способ конструирования подходящей PEгРНК и сгРНК с разрывом второго сайта никоим образом не является ограничивающим. В описании рассматриваются варианты вышеописанного постадийного способа, которые могут быть получены специалистом в данной области техники.

[7] Применения с использованием мультилоскутного праймированного редактирования

[0749] Двухлоскутное и четырехлоскутное праймированное редактирование (*m.e.*, мультилоскутное праймированное редактирование) имеет множество потенциальных применений, таких как установка пептидных меток, РНК-меток, иммуноэпитопов, доменов димеризации и сайтов-мишеней рекомбиназы. Одним из таких применений двухлоскутного праймированного редактирования является установка последовательностей рекомбиназы или интегразы в указанные пользователем места в геноме. На **фиг. 93** проиллюстрирована установку сайтов *attB* (38 п.о.) и *attP* (50 п.о.) рекомбиназы Bxb1 в целевом участке генома человека (*HEK293T сайт 3* или *HEK3*) с одновременной делецией 90 п.о. промежуточной последовательности между двумя участками разрыва. Различная степень комплементарности между 3′-лоскутами обеспечивает успешное редактирование, хотя более длинные последовательности комплементарности дают более благоприятное соотношение желаемых редактирований и инделей. Другие применения двухлоскутного праймированного редактирования включают эндогенную маркировку генов пептидными

или белковыми последовательностями или замену экзонов новыми последовательностями ДНК, которые потенциально могут заменить несколько вариантов, для которых мутация попадает в последовательность экзона.

[0750] Двухлоскутное праймированное редактирование можно использовать для введения одного или двух сайтов рекомбиназы в целевые положения генома человека. Если вставляются одиночные сайты рекомбиназные, их можно использовать в качестве основных сайтов для опосредованной рекомбиназой реакции между геномным сайтом рекомбиназны и вторым сайтом рекомбиназны в экзогенно поставляемой ДНК, такой как плазмида. Это позволяет целенаправленно интегрировать ДНК. Если два сайта рекомбиназы встроены в соседние области ДНК, в зависимости от ориентации сайтов рекомбиназы, их можно использовать для опосредованного рекомбиназой вырезания или инверсии промежуточной последовательности, или для опосредованного рекомбиназой обмена кассет с экзогенной ДНК для интеграции ДНК. Интеграция совместимых сайтов рекомбиназы на разных хромосомах делает возможной целевую и направленную хромосомную транслокацию. Двухлоскутное праймированное редактирование можно использовать для эффективного введения сайтов рекомбиназы в ряд локусов в геноме человека (**фиг. 94**). Таким образом, сочетание интеграции сайта рекомбиназы с помощью двухлоскутного праймированного редактирования с ферментами рекомбиназы ДНК представляет собой мощный подход для достижения многих типов редактирования SV и целевой интеграции ДНК.

[0751] Мультилоскутные праймирующие редакторы (например, вариант осуществления на фиг. 90), описанные в настоящем документе, можно использовать для точной вставки новой последовательности ДНК, точной делеции последовательности эндогенной геномной ДНК или замены последовательности эндогенной геномной ДНК на новую последовательность ДНК.

[0752] Например, как проиллюстрировано и обсуждается в настоящем документе, двойное праймированное редактирование можно использовать для (a) внесения корректирующих мутацию изменений в нуклеотидную последовательность, (b) введения белковых и РНК-меток, (в) введения иммуноэпитопов в представляющие интерес белки, (d) установки индуцируемых доменов димеризации в белках, (e) установки сайтов-мишеней рекомбиназы, чтобы направлять специфические генетические изменения. В дополнение к этим способам, которые, как правило, вставляют, изменяют или удаляют нуклеотидные последовательности в целевых участках, двойные праймирующие редакторы также могут использоваться для создания высокопрограммируемых библиотек, а также для проведения записи клеточных данных и исследований по отслеживанию клонов.

[0753] Эти конкретные примеры использования двойного праймированного редактирования никоим образом не предназначены для ограничения. Настоящее изобретение охватывает любое применение двойного праймированного редактирования, которое включает, в общем, некоторую форму установки, делеции и/или замены одного или нескольких нуклеотидных оснований в целевом участке в нуклеотидной последовательности, например, геномной ДНК.

[0754] Для любого из приведенных в качестве примера применений двойного праймированного редактирования можно использовать любой раскрытый в настоящем документе праймирующий редактор, включая PE1, PE2, PE3 и PE3b или PE-короткий.

А. Праймированное редактирование по сравнению с мультилоскутным праймированным редактированием

Классическое праймированное редактирование

[0755] Согласно различным вариантам осуществления праймированное редактирование (или «праймированное редактирование») осуществляется путем контактирования молекулы ДНК-мишени (для которой желательно внести изменение в нуклеотидную последовательность) с программируемым нуклеиновой кислотой ДНК-связывающим белком (napDNAbp), образующим комплекс с удлиненной гидовой РНК. Со ссылкой на фиг. 1G, удлиненная гидовая РНК содержит удлинение на 3’- или 5’-конце гидовой РНК или во внутримолекулярном участке гидовой РНК и кодирует желаемое изменение нуклеотида (например, изменение одного нуклеотида, вставку или делецию). На стадии (a) комплекс napDNAbp/удлиненная гРНК связывается с молекулой ДНК, и удлиненная гРНК направляет napDNAbp для связывания с локусом-мишенью. На стадии (b) в одну из цепей ДНК целевого локуса вводят разрыв (например, с помощью нуклеазы или химического агента), тем самым создавая доступный 3'-конец в одной из цепей целевого локуса. Согласно определенным вариантам осуществления разрыв создается в цепи ДНК, которая соответствует цепи R-петли, т.е. в цепи, которая не гибридизуется с последовательностью гидовой РНК, т.е. в «нецелевой цепи». Разрыв, однако, может быть введен в любую из цепей. То есть разрыв может быть введен в «целевую цепь» R-петли (т.е. в цепь, гибридизованную с протоспейсерной последовательностью удлиненной гРНК) или в «нецелевую цепь» (т.е. в цепь, образующую одноцепочечную часть R-петли, которая комплементарна целевой цепи). На стадии (c) 3’-конец цепи ДНК (образованный разрывом) взаимодействует с удлиненной частью гидовой РНК, чтобы праймировать обратную транскрипцию (т.е. «нацеленная-праймированная RT»). Согласно некоторым вариантам осуществления 3’-концевая цепь ДНК гибридизуется со специфической праймирующей

последовательностью RT на удлиненной части гидовой РНК, т.е. «праймирующей последовательностью обратной транскриптазы». На стадии (d) вводится обратная транскриптаза (в виде слитого белка с napDNAbp или *в транс*), которая синтезирует одну цепь ДНК от 3'-конца праймированного сайта к 5'-концу удлиненной гидовой РНК. Это формирует лоскут одноцепочечной ДНК, содержащий желаемое изменение нуклеотида (например, замену одного основания, вставку или делецию, или их комбинацию), и который в остальном гомологичен эндогенной ДНК в участке разрыва или рядом с ним. На стадии (e) высвобождаются napDNAbp и гидовая РНК. Стадии (f) и (g) относятся к разрешению лоскута одноцепочечной ДНК таким образом, чтобы желаемое изменение нуклеотида было включено в целевой локус. Этот способ может быть направлен на образование желаемого продукта путем удаления соответствующего 5'-лоскута эндогенной ДНК (например, с помощью FEN1 или аналогичного фермента, который предоставляется *в транс*, в виде слияния с праймирующим редактором или эндогенно), который образует после того, как 3′ -лоскут одноцепочечной ДНК проникает в эндогенную последовательность ДНК и гибридизуется с ней. Без ограничения теорией, процессы репарации и репликации эндогенной ДНК клетки разрешают ДНК с ошибочным спариванием, с включением изменения (изменений) нуклеотидов с образованием желаемого измененного продукта. Способ также может быть направлен на образованием продукта с «разрывом второй цепи», как показано на фиг. 1G, или «временным разрывом второй цепи», как показано на фиг. 1I и раскрывается в настоящем документе.

[0756] Способ праймированного редактирования может вносить по меньшей мере одно или несколько из следующих генетических изменений: трансверсии, трансзиции, делеции и вставки. Кроме того, праймированное редактирование может быть реализовано для конкретных применений. Например, как проиллюстрировано и раскрыто в настоящем документе, праймированное редактирование можно использовать для (a) внесения корректирующих мутацию изменений в нуклеотидную последовательность, (b) установки белковых и РНК-меток, (c) введения иммуноэпитопов в представляющие интерес белки, (d) установки индуцируемых доменов димеризации в белках, (e) установки или удаления последовательностей, чтобы изменить эту активность биомолекулы, (f) установки сайтов-мишеней рекомбиназы, чтобы направить определенные генетические изменения, и (g) мутагенеза целевой последовательности с использованием RT сниженной точности. В дополнение к этим способам, которые в общем вставляют, изменяют или удаляют нуклеотидные последовательности в целевых участках, праймирующие редакторы также можно использовать для создания высокопрограммируемых библиотек, а также для проведения записи клеточных данных и исследований по отслеживанию клонов. Авторы

настоящего изобретения также предусмотрели дополнительные конструктивные особенности РЕгРНК, направленные на повышение эффективности праймированного редактирования. Кроме того, авторы настоящего изобретения разработали способы успешной доставки праймирующих редакторов с использованием векторных систем доставки, которые включают расщепление napDNAbp с использованием интеиновых доменов.

[0757] Термин «система праймированного редактирования» или «праймирующий редактор (PE)» относятся к композициям, участвующим в способе редактирования генома с использованием нацеленной-праймированной обратной транскрипции (TPRT), как описано в настоящем документе, включающим без ограничений napDNAbp, обратные транскриптазы, слитые белки (например, содержащие napDNAbp и обратные транскриптазы), удлиненные гидовые РНК и комплексы, содержащие слитые белки и удлиненные гидовые РНК, а также вспомогательные элементы, такие как компоненты разрыва второй цепи и эндогенные эндонуклеазы для удаления 5'-лоскута ДНК (например, FEN1) для помощи в направлении способа праймированного редактирования к образованию отредактированного продукта.

[0758] Согласно другому варианту осуществления схема на фиг. 3F показывает взаимодействие типичной РЕгРНК с целевым участком двухцепочечной ДНК и сопутствующее образование 3'-лоскута одноцепочечной ДНК, содержащего представляющее интерес генетическое изменение. Двухцепочечная ДНК показана с верхней цепью в направлении от 3’ к 5’, а нижней цепью в направлении от 5’ к 3’. Верхняя цепь содержит «протоспейсер» и последовательность PAM и называется «целевой цепью». Комплементарная нижняя цепь называется «нецелевой цепью». Хотя это не показано, изображенная РЕгРНК будет находиться в комплексе с Cas9 или его эквивалентом. Как показано на схеме, спейсер РЕгРНК отжигается с комплементарной областью на целевой цепи, которая называется протоспейсером, который расположен сразу по ходу транскрипции от последовательности PAM и имеет длину приблизительно 20 нуклеотидов. Это взаимодействие формируется как гибрид ДНК/РНК между спейсерной РНК и протоспейсерной ДНК и индуцирует образование R-петли в области, противоположной протоспейсеру. Как указано в другом месте настоящего документа, белок Cas9 (не показан) затем индуцирует разрыв в нецелевой цепи, как показано. Затем это приводит к образованию области 3'-лоскута оцДНК, которая, в соответствии с *z*, взаимодействует с 3'-концом РЕгРНК в участке связывания праймера. 3'-конец лоскута оцДНК (т.е. последовательность праймера обратной транскриптазы) отжигается с участком связывания праймера (A) на РЕгРНК, тем самым праймируя обратную транскриптазу. Затем обратная

транскриптаза (например, представленная *в транс-* или *цис* как слитый белок, присоединенный к конструкции Cas9) затем полимеризует одну цепь ДНК, кодируемую матрицей редактирования (B) и плечом гомологии (C). Полимеризация продолжается по направлению к 5'-концу плеча удлинения. Полимеризованная цепь оцДНК образует 3'-концевой лоскут оцДНК, который, как описано в другом месте (например, как показано на фиг. 1G), вторгается в эндогенную ДНК, вытесняя соответствующую эндогенную цепь (которая удаляется как 5'-лоскут эндогенной ДНК), и устанавливает желаемое редактирование нуклеотидов (изменение одной пары нуклеотидов, делеция, вставка (включая целые гены)) посредством естественных циклов репарации/репликации ДНК.

[0759] Это применение праймированного редактирования может быть дополнительно описано в Примере 1.

Двойное праймированное редактирование

[0760] В настоящем описании раскрыта система двойного праймированного редактирования (или система двухлоскутного праймированного редактирования), которая решает проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в нередактируемую комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. В системе двухлоскутного праймированного редактирования две РЕгРНК используются для нацеливания на противоположные цепи геномного сайта и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 91**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная цепь. Поскольку обе цепи дуплекса синтезируются как отредактированная ДНК, система двухлоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК нужно только вырезать парные 5'-лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные цепи ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое содержит минимальную геномную гомологию. Нуклеазоактивные варианты праймирующих редакторов, которые разрезают обе цепи ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК. Соответственно, согласно

определенным вариантам осуществления система двухлоскутного праймированного редактирования включает пару вновь синтезированных цепей ДНК (например, 3'-лоскуты), которые обладают гомологией с последовательностью эндогенной ДНК в редактируемом целевом участке. Согласно некоторым вариантам осуществления система двухлоскутного праймированного редактирования включает пару вновь синтезированных цепей ДНК, где по меньшей мере одна из вновь синтезированных цепей ДНК не обладает гомологией с последовательностью эндогенной ДНК в редактируемом целевом участке. Согласно некоторым вариантам осуществления одна из вновь синтезированных цепей ДНК имеет гомологию с эндогенной последовательностью ДНК в целевом участке, а не с другой вновь синтезированной цепью ДНК. Согласно некоторым вариантам осуществления каждая из вновь синтезированных цепей ДНК гомологична эндогенной последовательности ДНК в целевом участке, а не другой вновь синтезированной цепи ДНК. Например, вновь синтезированный 3'-лоскут, кодируемый одной из двухлоскутных РЕгРНК, может содержать область, комплементарную последовательности протоспейсера другой двухлоскутной РЕгРНК. Соответственно, согласно некоторым вариантам осуществления пара двухлоскутных РЕгРНК, каждая из которых комплементарна спейсерной последовательности другой РЕгРНК, может привести к делеции последовательности эндогенной ДНК, расположенной между последовательностями протоспейсеров пары двухлоскутных РЕгРНК. Согласно некоторым вариантам осуществления система двухлоскутного праймированного редактирования включает пару вновь синтезированных цепей ДНК, где ни одна из вновь синтезированных цепей ДНК не имеет гомологии с последовательностью эндогенной ДНК в целевом участке, подлежащем редактированию. Наоборот, каждая из двух вновь синтезированных нитей ДНК (например, 3'-лоскуты) содержит область комплементарности друг другу и может образовывать дуплекс посредством комплементарности. Желаемая отредактированная часть по сравнению с редактируемым целевым участком эндогенной последовательности ДНК в дуплексе может быть затем включена в целевой участок. Подобно классическому праймированному редактированию, двойное праймированное редактирование представляет собой универсальный и точный способ редактирования генома, при котором новая генетическая информация напрямую записывается в указанный участок ДНК с использованием программируемого нуклеиновой кислотой ДНК-связывающего белка («napDNAbp»), работающего совместно с полимеразой (т.е. в форме слитого белка или иным образом представленной *в транс* по отношению к napDNAbp), где система праймированного редактирования запрограммирована с помощью гидовой РНК праймированного редактирования (РЕ) («РЕгРНК»), которая одновременно определяет целевой участок и

является матрицей для синтеза желаемого редактирование в форме замещающей цепи ДНК посредством удлинения (либо ДНК, либо РНК), сконструированного на гидовой РНК (например, на 5'- или 3'-конце или на внутренней части гидовой РНК). Цепь замещения, содержащая желаемое редактирование (например, замену одного нуклеотидного основания), имеет ту же последовательность, что и эндогенная цепь целевого участка, подлежащего редактированию (за исключением того, что она включает желаемое редактирование). С помощью механизмов репарации и/или репликации ДНК эндогенная цепь целевого участка заменяется вновь синтезированной замещающей цепью, содержащей желаемое редактирование. В некоторых случаях праймированное редактирование можно рассматривать как технологию редактирования генома «поиск и замена», поскольку праймирующие редакторы, как описано в настоящем документе, не только ищут и находят желаемый целевой участок для редактирования, но в то же время кодируют замещающую цепь, содержащую желаемое редактирование, которая устанавливается вместо соответствующей цепи эндогенной ДНК целевого участка.

B. Применение двойного праймированного редактирования для пептидного мечения

[0761] Согласно другому аспекту настоящее изобретение относится к способу применениях, описанных в настоящем документе двойных праймирующих редакторов для генетической прививки одной или нескольких пептидных меток в белок с использованием праймированного редактирования (как вариант осуществления, показанный на фиг. 90). Более конкретно, настоящее изобретение относится к способу генетической установки одной или нескольких пептидных меток в белок, предусматривающему: приведение в контакт целевой нуклеотидной последовательности, кодирующей белок, с праймирующим редактором, сконструированным для вставки в нее второй нуклеотидной последовательности, кодирующей одну или несколько пептидных меток, с получением в результате рекомбинантной нуклеотидной последовательности, которая кодирует слитый белок, содержащий белок, слитый с белковой меткой.

[0762] Согласно другим вариантам осуществления настоящее изобретение относится к способу получения слитого белка, содержащего представляющий интерес пептид и одну или более пептидных меток, причем способ предусматривает: приведение в контакт целевой нуклеотидной последовательности, кодирующей белок, с праймирующим редактором, сконструированным для вставки в нее второй нуклеотидной последовательности, кодирующей одну или несколько пептидных меток, с получением рекомбинантной нуклеотидной последовательности, которая кодирует слитый белок,

содержащий белок, слитый с белковой меткой.

[0763] Согласно различным вариантам осуществления целевая нуклеотидная последовательность представляет собой специфический 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 т.д.). Представляющий интерес ген может также кодировать молекулу РНК, включая без ограничения информационную РНК (иРНК), транспортную РНК (тРНК), рибосомную РНК (рРНК), малую ядерную РНК (мяРНК), антисмысловую РНК, гидовую РНК, микроРНК. (миРНК), малую интерферирующую РНК (siРНК) и внеклеточную РНК (cfRNA).

[0764] Пептидная метка может представлять собой любую пептидную метку или ее вариант, который придает белку одну или несколько функций для таких целей, как разделение, очистка, визуализация, солюбилизация или обнаружение. Метки пептидов могут включать «метки аффинности» (для облегчения очистки белков), «метки солюбилизации» (для помощи в правильном фолдинге белков), «метки хроматографии» (для изменения хроматографических свойств белков), «метки эпитопов» (для связывания с высокоаффинными антителами) и «флуоресцентные метки» (для облегчения визуализации белков в клетке или in vitro). Примеры пептидных меток включают без ограничения следующие метки:

НАЗВАНИЕ	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ**	**SEQ ID NO:**
AVITAG™	GLNDIFEAQKIEWHE	SEQ ID NO: 245
С-МЕТКА	EPEA	SEQ ID NO: 246
КАЛЬМОДУЛИН-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: 250
HA-МЕТКА	YPYDVPDYA	SEQ ID NO: 251
HIS-МЕТКА	H (HIS_1) HH (HIS_2) HHH (HIS_3) HHHH (HIS_4) HHHHH (HIS_5) HHHHHH (HIS_6) HHHHHHH (HIS_7)	SEQ ID NO: 252 SEQ ID NO: 253 SEQ ID NO: 254 SEQ ID NO: 255 SEQ ID NO: 256 SEQ ID NO: 257 SEQ ID NO: 258

	HHHHHHHH (HIS_8) HHHHHHHHH (HIS_9) HHHHHHHHHH (HIS_{10}) HHHHHHHHHH…H…(HIS_N, ГДЕ N = 1-25)	SEQ ID NO: 259 SEQ ID NO: 260 SEQ ID NO: 261 SEQ ID NO: 262
MYC-МЕТКА	EQKLISEEDL	SEQ ID NO: 263
NE-МЕТКА	TKENPRSNQEESYDDNES	SEQ ID NO: 264
RHO1D4-МЕТКА	TETSQVAPA	SEQ ID NO: 265
S-МЕТКА	KETAAAKFERQHMDS	SEQ ID NO: 266
SBP-МЕТКА	MDEKTTGWRGGHVVEGLAGELEQLR ARLEHHPQGQREP	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: 274
VSV-МЕТКА	YTDIEMNRLGK	SEQ ID NO: 275
XPRESS МЕТКА	DLYDDDDK	SEQ ID NO: 276

[0765] Пептидными метками также могут быть следующие аффинные метки (для разделения и/или очистки белков) (как описано в таблице 9.9.1 в Kimple et al., «Overview of Affinity Tags for Protein Purification,» *Curr Protoc Protein Sci*, 2013, 73:Unit-9.9, которая включена в настоящий документ посредством ссылки).

НАЗВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	
ЭПИТОП AU1	DTYRYI	SEQ ID NO: 278
ЭПИТОП AU5	TDFYLK	SEQ ID NO: 279
ЭПИТОП БАКТЕРИОФАГА T7 (T7-TAG)	MASMTGGQQMG	SEQ ID NO: 280
МЕТКА ВИРУСА ИНФЕКЦИОННОЙ КАТАРАЛЬНОЙ ЛИХОРАДКИ ОВЕЦ (B-МЕТКА)	QYPALT	SEQ ID NO: 281
ЭПИТОП E2	SSTSSDFRDR	SEQ ID NO: 282
ГИСТИДИНОВАЯ МЕТКА АФФИННОСТИ (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-МЕТКА	KETAAAKFERQHMDS	SEQ ID NO: 289
VSV-G	YTDIEMNRLGK	SEQ ID NO: 290

[0766] Согласно конкретным вариантам осуществления пептидные метки могут включать метку His^6, FLAG-метку, V5-метку, GCN4-метку, HA-метку, Myc-метку, FIAsH/ReAsH-метку, субстрат сортазы, pi-кольцо.

[0767] Согласно различным вариантам осуществления пептидные метки можно использовать для применений, которые включают флуоресцентное мечение белков, иммунопреципитацию, иммуноблоттинг, иммуногистохимию, рекрутирование белков, индуцибельные белковые дегроны и полногеномный скрининг.

[0768] Согласно различным другим вариантам осуществления пептидная метка может включать последовательность интеина для введения функции самосплайсинга белка. В контексте настоящего изобретения термин «интеин» относится к полипептидным доменам самопроцессинга, обнаруживаемым в организмах из всех доменов жизни. Интеин (*промежуточный белок*) проводит уникальное событие самопроцессинга, известное как сплайсинг белка, при котором он вырезается из более крупного полипептида-предшественника посредством расщепления двух пептидных связей и в процессе лигирует фланкирующе последовательности экстеина (внешний белок) посредством образования новой пептидной связи. Эта перестройка происходит посттрансляционно (или, возможно, котрансляционно), поскольку интеиновые гены обнаруживаются встроенными в рамку внутри других генов, кодирующих белок. Кроме того, интеин-опосредованный сплайсинг белков является спонтанным, для этого не требуется никаких внешних факторов или источников энергии, только свертывание интеинового домена. Этот процесс также известен как *цис*-сплайсинг белков, в отличие от естественного процесса *транс*-сплайсинга белков с «расщепленными интеинами». Интеины представляют собой белковый эквивалент интронов самосплайсинга РНК (см. Perler et al., Nucleic Acids Res. 22:1125-1127 (1994)), которые катализируют свое собственное вырезание из белка-предшественника с сопутствующим слиянием фланкирующих последовательностей белка, известных как экстеины (обзор 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)).

[0769] Механизм процесса сплайсинга белка изучен очень подробно (Chong, et al., J. Biol. Chem. 1996, 271, 22159-22168, Xu, M-Q & Perler, F. B. EMBO Journal, 1996, 15, 5146-5153), и консервативные аминокислоты были обнаружены в точках сплайсинга интеина и экстеина (Xu, et al., EMBO Journal, 1994, 13 5517-522).

[0770] Интеины также могут существовать в виде двух фрагментов, кодируемых двумя отдельно транскрибируемыми и транслируемыми генами. Эти так называемые расщепленные интеины самоассоциируются и катализируют активность сплайсинга белков в транс. Расщепленные интеины были идентифицированы у различных 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.), and Zettler J. et al, FEBS Letters. 553:909-914 (2009)), но до сих пор не были обнаружены у эукариот. Недавно биоинформатический анализ метагеномных данных об окружающей среде выявил 26 различных локусов с новым расположением генома. В каждом локусе консервативная область, кодирующая фермент, прерывается расщепленным интеином с автономным геном эндонуклеазы, вставленным между частями, кодирующими субдомены интеина. Среди них полностью собраны пять локусов: ДНК-хеликазы (gp41-1, gp41-8), инозин-5'-монофосфатдегидрогеназы (IMPDH-1) и каталитических субъединиц рибонуклеотидредуктазы (NrdA-2 и NrdJ-1). Эта разорванная генная организация, по-видимому, присутствует в основном у фагов (Dassa et al, Nucleic Acids Research. 57:2560-2573 (2009)).

[0771] Согласно определенным вариантам осуществления праймирующие редакторы, описанные в настоящем документе, можно использовать для встраивания меток расщепленных интеинов в два разных белка, вызывая их внутриклеточное лигирование при совместной экспрессии с образованием слитого белка. При транс-сплайсинге белка один белок-предшественник состоит из части N-экстеина, за которой следует N-интеин, другой белок-предшественник состоит из С-интеина, за которым следует часть С-экстеина, и реакция транс-сплайсинга (катализируемая N- и С-интеины вместе) вырезает две последовательности интеина и связывает две последовательности экстеина пептидной связью. Транс-сплайсинг белков, являясь ферментативной реакцией, может работать с очень низкими (например, микромолярными) концентрациями белков и может осуществляться в физиологических условиях.

[0772] Расщепленный интеин Npu DnaE был охарактеризован как имеющий самую высокую скорость реакции транс-сплайсинга белка. Кроме того, реакция сплайсинга белка

Npu DnaE считается надежной и высокоэффективной в отношении различных последовательностей экстеина, температура от 6 до 37°C и в присутствии до 6 M мочевины (Zettler J. et al, FEBS Letters. 553:909-914 (2009), Iwai I. et al, FEBS Letters 550: 1853-1858 (2006)). Как и ожидалось, когда была введена мутация Cysl Ala в N-домен этих интеинов, начальный сдвиг N в S-ацил и, следовательно, сплайсинг белка блокировался. К сожалению, реакция C-концевого расщепления также была практически полностью ингибирована. Зависимость циклизации аспарагина на C-концевом соединении сплайсинга от ацильного сдвига на N-концевой разрывающейся пептидной связи, по-видимому, является уникальным свойством, общим для естественно расщепленных аллелей интеина DnaE (Zettler J. et al. FEBS Letters. 555:909-914 (2009)).

[0773] Транс-сплайсинг белков, катализируемый расщепленными интеинами, представляет собой полностью ферментативный способ лигирования белков. Расщепленный 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 расщепленного интеина могут нековалентно связываться с образованием активного интеина и катализировать реакцию сплайсинга, по существу, таким же образом, как это делает непрерывный интеин. Расщепленные интеины были обнаружены в природе, а также созданы в лабораториях. В контексте настоящего изобретения термин «расщепленный интеин» относится к любому интеину, в котором существует один или несколько разрывов пептидной связи между N-концевой и C-концевой аминокислотными последовательностями, так что N-концевая и C-концевая последовательности становятся отдельными молекулами, которые могут нековалентно повторно ассоциироваться или восстанавливаться в интеин, который является функциональным для реакций транс-сплайсинга. Любой каталитически активный интеин или его фрагмент можно использовать для получения расщепленного интеина для применения в способах согласно настоящему изобретению. Например, согласно одному аспекту расщепленный интеин может быть получен из эукариотического интеина. Согласно другому аспекту расщепленный интеин может быть получен из бактериального интеина. Согласно другому аспекту расщепленный интеин может быть получен из интеина архей. Предпочтительно полученный таким образом расщепленный интеин будет иметь только те аминокислотные последовательности, которые необходимы для катализа реакций транс-сплайсинга.

[0774] Расщепленные интеины могут быть созданы из непрерывных интеинов путем создания одного или нескольких участков расщепления в неструктурированной петле или промежуточной аминокислотной последовательности между -12 консервативными бета-

цепями, обнаруженными в структуре мини-интеинов. Может существовать некоторая гибкость в расположении сайта расщепления в областях между бета-цепями при условии, что создание расщепления не нарушит структуру интеина, в частности структурированных бета-цепей, в достаточной степени, чтобы активность сплайсинга белка была потеряна.

[0775] Праймирующие редакторы, описанные в настоящем документе, могут включать пептидные метки (включая интеины) на С-конце представляющего интерес белка. Согласно другим вариантам осуществления пептидные метки (включая интеины) могут быть включены на N-конце представляющего интерес белка. Пептидные метки также могут быть включены внутрь представляющего интерес белка. Полученные слитые белки, созданные с помощью описанных в настоящем документе праймирующих редакторов, могут иметь следующие структуры:

[0776] [представляющий интерес белок] – [пептидная метка],

[0777] [пептидная метка] – [представляющего интерес белка], или

[0778] [представляющего интерес белка– N-концевая область] – [пептидная метка] – [представляющего интерес белка – С-концевая область].

[0779] Принципы конструирования гидовой РНК для использования при пептидном мечении могут быть применены для пептидного мечения. Например, согласно одному варианту осуществления структура PEgRNA для пептидного мечения может иметь следующую структуру: 5'-[спейсерная последовательность]-[ядро или каркас гРНК]-[плечо удлинения]-3', где плечо удлинения содержит в 5' - 3' направлении плечо гомологии, матрицу редактирования (содержащую последовательность, кодирующую пептидную метку) и участок связывания праймера. Эта конфигурация изображена на фиг. 3D и на фиг. 24.

[0780] Согласно другому варианту осуществления структура РЕгРНК для пептидного мечения может иметь следующую структуру: 5′-[плечо удлинения]-[спейсерная последовательность]-[ядро или каркас гРНК]-3′, где плечо удлинения содержит в 5' - 3' направлении плечо гомологии, матрицу редактирования (содержащую последовательность, кодирующую пептидную метку) и участок связывания праймера. Эта конфигурация изображена на фиг. 3E.

[0781] Варианты осуществления пептидного мечения с применением праймированного редактирования представлены на фиг. 25 и 26 и описаны в Примере 4.

<u>С. Применение двойного праймированного редактирования для мечения РНК</u>

[0782] Двойное праймированное редактирование также можно использовать для манипулирования, изменения и иной модификации последовательностей ДНК,

кодирующих функции РНК, посредством мечения РНК, и таким образом обеспечивает средства для непрямой модификации структуры и функции РНК. Например, двойное РЕ можно использовать для вставки мотивов, которые функционируют на уровне РНК (далее мотивы РНК), для мечения или иного манипулирования некодирующими РНК или иРНК. Эти мотивы могут служить для увеличения экспрессии генов, снижения экспрессии генов, изменения сплайсинга, изменения посттранскрипционной модификации, воздействия на субклеточное расположение РНК, обеспечения выделения или определения внутри- или внеклеточного расположения РНК (с использованием, например, флуоресцентных РНК-аптамеров, такие как Spinach, Spinach2, Baby Spinach или Broccoli), рекрутинга эндогенных или экзогенных белковых или РНК-связывающих, введения сгРНК или индукции процессинга РНК посредством либо саморасщепления, либо РНКаз (см. фиг. 28B и пример 6 для получения дополнительной информации).

[0783] Следующие метки или мотивы РНК могут быть вставлены в представляющий интерес ген с использованием двойного праймированного редактирования (например, с помощью методологии двойного РЕ на фиг. 90) с соответствующей РЕгРНК (сконструированной с использованием приведенных в настоящем документе руководств) для воздействия на различные свойства РНК, включая транспорт РНК, уровень экспрессии, сплайсинг и обнаружение.

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
ВИРУС ПОЛИОМЫ ОБЕЗЬЯНЫ 40 (SV40) ТИПА 1	AACTTGTTTATTGCAGCTTAT AATGGTTACAAATAAAGCAA TAGCATCACAAATTTCACAA ATAAAGCATTTTTTTCACTGC ATTCTAGTTGTGGTTTGTCCA AACTCATCAATGTATCTTA (SEQ ID NO: 331)	ТЕРМИНАЦИЯ ТРАНСКРИПЦИИ МЕЧЕНОГО ГЕНА, ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ, ПОВЫШЕННАЯ СТАБИЛЬНОС	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***TAAGATACA TTGATGAGTTTGGA CAAACCACAACTAG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
		ТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	**AATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTT**CTATG*GCCCTGACTGGAA* (SEQ ID NO: 332)
ВИРУС ПОЛИОМЫ ОБЕЗЬЯНЫ 40 (SV40) ТИПА 2	CCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCTC (SEQ ID NO: 333)	ТЕРМИНАЦИЯ ТРАНСКРИПЦИИ МЕЧЕНОГО ГЕНА, ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ, ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GAGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTGGGCCATGGCTATG*GCCCTGACTGGAA* (SEQ ID NO: 334)
ВИРУС ПОЛИОМЫ	TGATCATAATCAAGCCATATCACATCTGTAGAGGTTTACTT	ТЕРМИНАЦИЯ ТРАНСКРИПЦ	ATCCTTCCAGTCAGGGCCATGTTTGAGAGC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
ОБЕЗЬЯНЫ 40 (SV40) ТИПА 3	GCTTTAAAAAACCTCCACAC CTCCCCCTGAACCTGAAACA TAAAATGAATGCAATTGTTG TTGTTAACTTGTTTATTGCAG CTTATAATGGTTACAAATAAA GCAATAGCATCACAAATTTC ACAAATAAAGCATTTTTTTC ACTGCATTCTAGTTGTGGTTT GTCCAAACTCATCAATGTAT CTTATCATGTCTGGATCTGC (SEQ ID NO: 335)	ИИ МЕЧЕНОГО ГЕНА, ТРАНСПОРТ ИРНК В ЦИТОЗОЛЬ, ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***GCAGATCC AGACATGATAAGATA CATTGATGAGTTTG GACAAACCACAACT AGAATGCAGTGAAA AAAATGCTTTATTTG TGAAATTTGTGATG CTATTGCTTTATTTG TAACCATTATAAGCT GCAATAAACAAGTT AACAACAACAATTG CATTCATTTTATGTT TCAGGTTCAGGGGG AGGTGTGGAGGTTT TTTAAAGCAAGTAA ACCTCTACAGATGT GATATGGCTTGATTA TGATCA**CTATG*GCCC TGACTGGAA* (SEQ ID NO: 336)
ГОРМОН РОСТА ЧЕЛОВЕКА (HGH)	ACGGGTGGCATCCCTGTGAC CCCTCCCCAGTGCCTCTCCT GGCCCTGGAAGTTGCCACTC CAGTGCCCACCAGCCTTGTC CTAATAAAATTAAGTTGCATC	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ, ПОВЫШЕННАЯ	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	ATTTTGTCTGACTAGGTGTC CTTCTATAATATTATGGGGTG GAGGGGGGTGGTATGGAGC AAGGGGCAAGTTGGGAAGA CAACCTGTAGGGCCTGCGGG GTCTATTGGGAACCAAGCTG GAGTGCAGTGGCACAATCTT GGCTCACTGCAATCTCCGCC TCCTGGGTTCAAGCGATTCT CCTGCCTCAGCCTCCCGAGT TGTTGGGATTCCAGGCATGC ATGACCAGGCTCAGCTAATT TTTGTTTTTTTGGTAGAGAC GGGGTTTCACCATATTGGCC AGGCTGGTCTCCAACTCCTA ATCTCAGGTGATCTACCCAC CTTGGCCTCCCAAATTGCTG GGATTACAGGCGTGAACCAC TGCTCCCTTCCCTGTCCTT (SEQ ID NO: 337)	СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***AAGGACAG GGAAGGGAGCAGT GGTTCACGCCTGTA ATCCCAGCAATTTG GGAGGCCAAGGTG GGTAGATCACCTGA GATTAGGAGTTGGA GACCAGCCTGGCCA ATATGGTGAAACCC CGTCTCTACCAAAA AAACAAAAATTAGC TGAGCCTGGTCATG CATGCCTGGAATCC CAACAACTCGGGAG GCTGAGGCAGGAGA ATCGCTTGAACCCA GGAGGCGGAGATTG CAGTGAGCCAAGAT TGTGCCACTGCACT CCAGCTTGGTTCCC AATAGACCCCGCAG GCCCTACAGGTTGT CTTCCCAACTTGCC CCTTGCTCCATACC ACCCCCCTCCACCC CATAATATTATAGAA GGACACCTAGTCAG ACAAAATGATGCAA

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			CTTAATTTTATTAGG ACAAGGCTGGTGGG CACTGGAGTGGCAA CTTCCAGGGCCAGG AGAGGCACTGGGGA GGGGTCACAGGGAT GCCACCCGTCTATG*G CCCTGACTGGAA* (SEQ ID NO: 338)
ГОРМОН РОСТА КРУПНОГО РОГАТОГО СКОТА (BGH)	CGACTGTGCCTTCTAGTTGC CAGCCATCTGTTGTTTGCCC CTCCCCCGTGCCTTCCTTGA CCCTGGAAGGTGCCACTCCC ACTGTCCTTTCCTAATAAAAT GAGGAAATTGCATCGCATTG TCTGAGTAGGTGTCATTCTAT TCTGGGGGGTGGGGTGGGG CAGGACAGCAAGGGGGAGG ATTGGGAAGACAATAGCAGG CATGCTGGGGATGCGGTGGG CTCTATGG (SEQ ID NO: 339)	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ, ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***CCATAGAGC CCACCGCATCCCCA GCATGCCTGCTATT GTCTTCCCAATCCT CCCCCTTGCTGTCC TGCCCCACCCCACC CCCCAGAATAGAAT GACACCTACTCAGA CAATGCGATGCAAT TTCCTCATTTTATTA GGAAAGGACAGTGG GAGTGGCACCTTCC AGGGTCAAGGAAGG CACGGGGGAGGGG

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			CAAACAACAGATGG CTGGCAACTAGAAG GCACAGTCGCTATG *GCCCTGACTGGAA* (SEQ ID NO: 340)
БЕТА-ГЛОБИН КРОЛИКА (RBGLOB)	TTCACTCCTCAGGTGCAGGC TGCCTATCAGAAGGTGGTGG CTGGTGTGGCCAATGCCCTG GCTCACAAATACCACTGAGA TCTTTTTCCCTCTGCCAAAA ATTATGGGGACATCATGAAG CCCCTTGAGCATCTGACTTC TGGCTAATAAAGGAAATTTA TTTTCATTGCAATAGTGTGTT GGAATTTTTTGTGTCTCTCAC TCGGAAGGACATATGGGAGG GCAAATCATTTAAAACATCA GAATGAGTATTTGGTTTAGA GTTTGGCAACATATGCCCATA TGCTGGCTGCCATGAACAAA GGTTGGCTATAAAGAGGTCA TCAGTATATGAAACAGCCCC CTGCTGTCCATTCCTTATTCC ATAGAAAAGCCTTGACTTGA GGTTAGATTTTTTTTATATTTT GTTTTGTGTTATTTTTTTCTT TAACATCCCTAAAATTTTCCT TACATGTTTTACTAGCCAGAT TTTTCCTCCTCTCCTGACTAC TCCCAGTCATAGCTGTCCCT	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ, ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ КОДИРУЕМОГО БЕЛКА	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***GATCTCCAT AAGAGAAGAGGGAC AGCTATGACTGGGA GTAGTCAGGAGAGG AGGAAAAATCTGGC TAGTAAAACATGTAA GGAAAATTTTAGGG ATGTTAAAGAAAAA AATAACACAAAACA AAATATAAAAAAAAT CTAACCTCAAGTCA AGGCTTTTCTATGG AATAAGGAATGGAC AGCAGGGGGCTGTT TCATATACTGATGAC CTCTTTATAGCCAAC CTTTGTTCATGGCA GCCAGCATATGGGC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	CTTCTCTTATGGAGATC (SEQ ID NO: 341)		**ATATGTTGCCAAACT CTAAACCAAATACTC ATTCTGATGTTTTAA ATGATTTGCCCTCC CATATGTCCTTCCGA GTGAGAGACACAAA AAATTCCAACACAC TATTGCAATGAAAAT AAATTTCCTTTATTA GCCAGAAGTCAGAT GCTCAAGGGGCTTC ATGATGTCCCCATAA TTTTTGGCAGAGGG AAAAAGATCTCAGT GGTATTTGTGAGCC AGGGCATTGGCCAC ACCAGCCACCACCT TCTGATAGGCAGCC TGCACCTGAGGAGT GAA**CTATG*GCCCTGA CTGGAA* (SEQ ID NO: 342)
ТИМИДИНКИНАЗА (TK)	GGGGGAGGCTAACTGAAAC ACGGAAGGAGACAATACCG GAAGGAACCCGCGCTATGAC GGCAATAAAAAGACAGAATA AAACGCACGGGTGTTGGGT CGTTTGTTCATAAACGCGGG GTTCGGTCCCAGGGCTGGCA CTCTGTCGATACCCCACCGA	ТРАНСПОРТ РНК В ЦИТОПЛАЗМУ, ПОВЫШЕННАЯ СТАБИЛЬНОСТЬ РНК И ЭКСПРЕССИЯ	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***CTATGGCAG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	GACCCCATTGGGGCCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAGGCCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCTGCCATAG (SEQ ID NO: 343)	КОДИРУЕМОГО БЕЛКА	**GGCCTGCCGCCCCGACGTTGGCTGCGAGCCCTGGGCCTTCACCCGAACTTGGGGGGTGGGGTGGGGAAAAGGAAGAAACGCGGGCGTATTGGCCCCAATGGGGTCTCGGTGGGGTATCGACAGAGTGCCAGCCCTGGGACCGAACCCCGCGTTTATGAACAAACGACCCAACACCCGTGCGTTTTATTCTGTCTTTTTATTGCCGTCATAGCGCGGGTTCCTTCCGGTATTGTCTCCTTCCGTGTTTCAGTTAGCCTCCCCCCTATG*GCCCTGACTGGAA* (SEQ ID NO: 344)
MALAT1 ENE-MASCRNA	TAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAAGATGCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCGTCTTTGCTTTGACT (SEQ ID NO:	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И СТАБИЛИЗАЦИИ ТРАНСКРИПТА	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***AGTCAAAGCAAAGACGCCGCAG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	345)		**GGATTTGAACCCCG TCCTGGAAACCAGG AGTGCCAACCACCA GCATCTTTTGCTTTT TTTTTTTTTAAGCTA GAAAAAGGCCAAAA AGCAAAACCTGAGA AAACAATACGTGTT GTTTTCTCAGGAAA AGAAAAACCTTCAT GACCCTA**CTATG*GCC CTGACTGGAA* (SEQ ID NO: 346)
KSHV PAN ENE	TGTTTTGGCTGGGTTTTTCCT TGTTCGCACCGGACACCTCC AGTGACCAGACGGCAAGGT TTTTATCCCAGTGTATATTGG AAAAACATGTTATACTTTTG ACAATTTAACGTGCCTAGAG CTCAAATTAAACTAATACCAT AACGTAATGCAACTTACAAC ATAAATAAAGGTCAATGTTT AATCCATAAAAAAAAAAAA AAAAAAA (SEQ ID NO: 347)	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И СТАБИЛИЗАЦИИ ТРАНСКРИПТА	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***TTTTTTTTT TTTTTTTTTATGGA TTAAACATTGACCTT TATTTATGTTGTAAG TTGCATTACGTTATG GTATTAGTTTAATTT GAGCTCTAGGCACG TTAAATTGTCAAAA GTATAACATGTTTTT CCAATATACACTGG**

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			GATAAAAACCTTGC CGTCTGGTCACTGG AGGTGTCCGGTGCG AACAAGGAAAAACC CAGCCAAAACACTAT G*GCCCTGACTGGAA* (SEQ ID NO: 348)
ТРИ ПОСЛЕДОВАТЕЛЬНЫХ KSHV PAN ENES С КОРОТКИМИ НЕКОНСЕРВАТИВНЫМИ РНК ЛИНКЕРАМИ	TGTTTTGGCTGGGTTTTTCCT TGTTCGCACCGGACACCTCC AGTGACCAGACGGCAAGGT TTTTATCCCAGTGTATATTGG AAAAACATGTTATACTTTTG ACAATTTAACGTGCCTAGAG CTCAAATTAAACTAATACCAT AACGTAATGCAACTTACAAC ATAAATAAAGGTCAATGTTT AATCCATAAAAAAAAAAAA AAAAAAAACACACTGTTTTG GCTGGGTTTTTCCTTGTTCG CACCGGACACCTCCAGTGAC CAGACGGCAAGGTTTTTATC CCAGTGTATATTGGAAAAAC ATGTTATACTTTTGACAATTT AACGTGCCTAGAGCTCAAAT TAAACTAATACCATAACGTA ATGCAACTTACAACATAAAT AAAGGTCAATGTTTAATCCA TAAAAAAAAAAAAAAAAAA ATCTCTCTGTTTTGGCTGGGT TTTTCCTTGTTCGCACCGGA	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ И СТАБИЛИЗАЦИИ ТРАНСКРИПТА. ПО ПРОГНОЗУ БОЛЬШЕ ОДНОГО PAN ENE	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***TTTTTTTTT TTTTTTTTTATGGA TTAAACATTGACCTT TATTTATGTTGTAAG TTGCATTACGTTATG GTATTAGTTTAATTT GAGCTCTAGGCACG TTAAATTGTCAAAA GTATAACATGTTTTT CCAATATACACTGG GATAAAAACCTTGC CGTCTGGTCACTGG AGGTGTCCGGTGCG AACAAGGAAAAACC CAGCCAAAACAGAG AGATTTTTTTTTTT

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	CACCTCCAGTGACCAGACG GCAAGGTTTTTATCCCAGTG TATATTGGAAAAACATGTTAT ACTTTTGACAATTTAACGTG CCTAGAGCTCAAATTAAACT AATACCATAACGTAATGCAA CTTACAACATAAATAAAGGT CAATGTTTAATCCATAAAAA AAAAAAAAAAAAAA (SEQ ID NO: 349)		**TTTTTTTATGGATTA AACATTGACCTTTAT TTATGTTGTAAGTTG CATTACGTTATGGTA TTAGTTTAATTTGAG CTCTAGGCACGTTA AATTGTCAAAAGTAT AACATGTTTTTCCA ATATACACTGGGATA AAAACCTTGCCGTC TGGTCACTGGAGGT GTCCGGTGCGAACA AGGAAAAACCCAGC CAAAACAGTGTGTT TTTTTTTTTTTTTT TTTATGGATTAAACA TTGACCTTTATTTAT GTTGTAAGTTGCAT TACGTTATGGTATTA GTTTAATTTGAGCTC TAGGCACGTTAAAT TGTCAAAAGTATAA CATGTTTTTCCAATA TACACTGGGATAAA AACCTTGCCGTCTG GTCACTGGAGGTGT CCGGTGCGAACAAG GAAAAACCCAGCCA AAACA**CTATG*GCCCT GACTGGAA*

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			(SEQ ID NO: 350)
SMBOX/U1 SNRNA БОКС	CAGCAAGTTCAGAGAAATCT GAACTTGCTGGATTTTTGGA GCAGGGAGATGGAATAGGA GCTTGCTCCGTCCACTCCAC GCATCGACCTGGTATTGCAG TACCTCCAGGAACGGTGCAC CCACTTTCTGGAGTTTCAAA AGTAGACTGTACGCTAAGGG TCATATCTTTTTTTGTTTGGT TTGTGTCTTGGTTGGCGTCT TAAA (SEQ ID NO: 351)	RESULTS IN ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ ТРАНСКРИПТА	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***TTTAAGACG CCAACCAAGACACA AACCAAACAAAAAA AGATATGACCCTTA GCGTACAGTCTACT TTTGAAACTCCAGA AAGTGGGTGCACCG TTCCTGGAGGTACT GCAATACCAGGTCG ATGCGTGGAGTGGA CGGAGCAAGCTCCT ATTCCATCTCCCTGC TCCAAAAATCCAGC AAGTTCAGATTTCT CTGAACTTGCTGC**T ATG*GCCCTGACTGGAA* (SEQ ID NO: 352)
U1 SNRNA 3′ БОКС	GTTTCAAAAGTAGACTGTAC GCTAAGGGTCATATCTTTTTT TGTTTGGTTTGTGTCTTGGTT GGCGTCTTAAA (SEQ ID NO: 353)	ПРИВОДИТ К УДЕРЖАНИЮ РНК В ЯДРЕ, ТЕРМИНАЦИИ ТРАНСКРИПТА	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
		.	AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC***TTTAAGACG CCAACCAAGACACA AACCAAACAAAAAA AGATATGACCCTTA GCGTACAGTCTACT TTTGAAAC**CTATG*GC CCTGACTGGAA* (SEQ ID NO: 354)
TRNA-ЛИЗИН	GCCCGGCTAGCTCAGTCGGT AGAGCATGAGACTCTTAATC TCAGGGTCGTGGGTTCGAGC CCCACGTTGGGCG (SEQ ID NO: 355)	СООБЩАЕТСЯ, ЧТО ОН ОБЕСПЕЧИВАЕТ ТРАНСПОРТ РНК В МИТОХОНДРИИ	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC CGTTATCAACTTGAA AAAGTGGGACCGAG TCGGTCC*ACCTGAACC GTATATC* **CGCCCAACGTGGGG CTCGAACCCACGAC CCTGAGATTAAGAG TCTCATGCTCTACC GACTGAGCTAGCCG GGC**CTATG*GCCCTGA CTGGAA* (SEQ ID NO: 356)
АПТАМЕР BROCCOLI	GAGACGGTCGGGTCCAGATA TTCGTATCTGTCGAGTAGAG TGTGGGCTC (SEQ ID NO: 357)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAGG GCCATGTTTGAGAGC TAGAAATAGCAAGTT TAAATAAGGCTAGTC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			CGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GAGCCCACACTCTACTCGACAGATACGAATATCTGGACCCGACCGTCTCC**CTATG*GCCCTGACTGGAA*(SEQ ID NO: 358)
АПТАМЕР SPINACH	GACGCAACTGAATGAAATGGTGAAGGACGGGTCCAGGTGTGGCTGCTTCGGCAGTGCAGCTTGTTGAGTAGAGTGTGAGCTCCGTAACTAGTCGCGTC (SEQ ID NO: 359)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GACGCGACTAGTTACGGAGCTCACACTCTACTCAACAAGCTGCACTGCCGAAGCAGCCACACCTGGACCCGTCCTTCACCATTTCATTCAGTTGCGTC**CTATG*GCCCTGACTGGAA*(SEQ ID NO: 360)
АПТАМЕР SPINACH2	GATGTAACTGAATGAAATGGTGAAGGACGGGTCCAGTAGGCTGCTTCGGCAGCCTACTTGTTGAGTAGAGTGTGAGCTC	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
	CGTAACTAGTTACATC (SEQ ID NO: 361)		CGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GATGTAACTAGTTACGGAGCTCACACTCTACTCAACAAGTAGGCTGCCGAAGCAGCCTACTGGACCCGTCCTTCACCATTTCATTCAGTTACAT**CCTATG*GCCCTGACTGGAA* (SEQ ID NO: 362)
АПТАМЕР MANGO	GGCACGTACGAAGGGACGGTGCGGAGAGGAGAGTACGTGC (SEQ ID NO: 363)	ВИЗУАЛИЗАЦИЯ (ФЛУОРЕСЦЕНЦИЯ)	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GCACGTACTCTCCTCTCCGCACCGTCCCTTCGTACG**TGCC*CTATG*GCCCTGACTGGAA* (SEQ ID NO: 364)
РИБОЗИМ	GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC (SEQ ID NO: 365)	ПРОЦЕССИНГ 3′-КОНЦЕВОЙ РНК	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTC

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
			CGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***GTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGCC**CTATG*GCCCTGACTGGAA* (SEQ ID NO: 366)
N^6-МЕТИЛАДЕНОЗИНОВЫЙ МАРКЕР (M^6A)	GGACTCTAGGACTGGACTTTGGACT (SEQ ID NO: 367)	МИШЕНЬ ДЛЯ МЕТИЛИРОВАНИЯ (МЕТИЛИРОВАНИЕ ПОДЧЕРКНУТО). МЕТИЛИРОВАНИЕ M6A МОЖЕТ ПРИВОДИТЬ К ПОВЫШЕННОЙ СТАБИЛЬНОСТИ И ЭКСПРЕССИИ РНК, НО ЕЩЕ НЕ ПОЛНОСТЬЮ	ATCCTTCCAGTCAGGGCCATGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC*ACCTGAACCGTATATC***AGTCCAAAGTCCAGTCCTAGAG****TCC**CTATG*GCCCTGACTGGAA* (SEQ ID NO: 368)

МОТИВ РНК	ПОСЛЕДОВАТЕЛЬНОСТЬ МОТИВА РНК	ФУНКЦИЯ/ ЭФФЕКТ	ИЛЛЮСТРАТИВНАЯ РЕГРНК ДЛЯ ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ ВСТАВКИ МОТИВА РНК В ИЛЛЮСТРАТИВНЫЙ ГЕН HEXA *
		ИЗУЧЕНО	

* каждая РЕгРНК показана в направлении от 5' к 3' и имеет следующие структурные элементы на фиг. 3F в соответствии с типом шрифта следующим образом: 5′ – спейсерная последовательность (обычный шрифт) – ядро гРНК (подчеркнутая последовательность) – плечо гомологии (выделено курсивом) – матрица RT (выделено жирным шрифтом) – участок связывания праймера (выделено курсивом) – 3’.

[0784] РЕгРНК из приведенной выше таблицы сконструированы для сайт-специфической вставки примеров вышеуказанных мотивов в ген *HEXA* (дефектный при болезни Тея-Сакса) (например, GenBank № KR710351.1 (SEQ ID NO: 369), однако, только в целях иллюстрации. Использование праймированного редактирования при мечении РНК не ограничивается геном *HEXA* и действительно может быть любым. иРНК *HEXA* имеет следующую нуклеотидную последовательность:

GTTCGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAA
GTTGGCATGACAAGTTCCAGGCTTTGGTTTTCGCTGCTGCTGGCGGCAGCGTTCGCAGGACG
GGCGACGGCCCTCTGGCCCTGGCCTCAGAACTTCCAAACCTCCGACCAGCGCTACGTCCTTT
ACCCGAACAACTTTCAATTCCAGTACGATGTCAGCTCGGCCGCGCAGCCCGGCTGCTCAGTC
CTCGACGAGGCCTTCCAGCGCTATCGTGACCTGCTTTTCGGTTCCGGGTCTTGGCCCCGTCCT
TACCTCACAGGGAAACGGCATACACTGGAGAAGAATGTGTTGGTTGTCTCTGTAGTCACACCT
GGATGTAACCAGCTTCCTACTTTGGAGTCAGTGGAGAATTATACCCTGACCATAAATGATGACC
AGTGTTTACTCCTCTCTGAGACTGTCTGGGGAGCTCTCCGAGGTCTGGAGACTTTTAGCCAGC
TTGTTTGGAAATCTGCTGAGGGCACATTCTTTATCAACAAGACTGAGATTGAGGACTTTCCCC
GCTTTCCTCACCGGGGCTTGCTGTTGGATACATCTCGCCATTACCTGCCACTCTCTAGCATCCT
GGACACTCTGGATGTCATGGCGTACAATAAATTGAACGTGTTCCACTGGCATCTGGTAGATGA
TCCTTCCTTCCCATATGAGAGCTTCACTTTTCCAGAGCTCATGAGAAAGGGGTCCTACAACCC
TGTCACCCACATCTACACAGCACAGGATGTGAAGGAGGTCATTGAATACGCACGGCTCCGGG
GTATCCGTGTGCTTGCAGAGTTTGACACTCCTGGCCACACTTTGTCCTGGGGACCAGGTATCC
CTGGATTACTGACTCCTTGCTACCCTGGGTCTGAGCCCTCTGGCACCTTTGGACCAGTGAATC
CCAGTCTCAATAATACCTATGAGTTCATGAGCACATTCTTCTTAGAAGTCAGCTCTGTCTTCCC
AGATTTTTATCTTCATCTTGGAGGAGATGAGGTTGATTTCACCTGCTGGAAGTCCAACCCAGA

GATCCAGGACTTTATGAGGAAGAAAGGCTTCGGTGAGGACTTCAAGCAGCTGGAGTCCTTCT
ACATCCAGACGCTGCTGGACATCGTCTCTTCTTATGGCAAGGGCTATGTGGTGTGGCAGGAGG
TGTTTGATAATAAAGTAAAGATTCAGCCAGACACAATCATACAGGTGTGGCGAGAGGATATTC
CAGTGAACTATATGAAGGAGCTGGAACTGGTCACCAAGGCCGGCTTCCGGGCCCTTCTCTCT
GCCCCCTGGTACCTGAACCGTATATCCTATGGCCCTGACTGGAAGGATTTCTACGTAGTGGAAC
CCCTGGCATTTGAAGGTACCCCTGAGCAGAAGGCTCTGGTGATTGGTGGAGAGGCTTGTATGT
GGGGAGAATATGTGGACAACACAAACCTGGTCCCCAGGCTCTGGCCCAGAGCAGGGGCTGTT
GCCGAAAGGCTGTGGAGCAACAAGTTGACATCTGACCTGACATTTGCCTATGAACGTTTGTC
ACACTTCCGCTGTGAGTTGCTGAGGCGAGGTGTCCAGGCCCAACCCCTCAATGTAGGCTTCT
GTGAGCAGGAG
TTTGAACAGACCTGCCCAACTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAA
TTTGTTGCAACGAAC (SEQ ID NO: 369).

[0785] Соответствующий белок HEXA имеет следующую аминокислотную последовательность:

MTSSRLWFSLLLAAAFAGRATALWPWPQNFQTSDQRYVLYPNNFQFQYDVSSAAQPGCS
VLDEAFQRYRDLLFGSGSWPRPYLTGKRHTLEKNVLVVSVVTPGCNQLPTLESVENYTLTINDDQ
CLLLSETVWGALRGLETFSQLVWKSAEGTFFINKTEIEDFPRFPHRGLLLDTSRHYLPLSSILDTLD
VMAYNKLNVFHWHLVDDPSFPYESFTFPELMRKGSYNPVTHIYTAQDVKEVIEYARLRGIRVLAE
FDTPGHTLSWGPGIPGLLTPCYPGSEPSGTFGPVNPSLNNTYEFMSTFFLEVSSVFPDFYLHLGGD
EVDFTCWKSNPEIQDFMRKKGFGEDFKQLESFYIQTLLDIVSSYGKGYVVWQEVFDNKVKIQPD
TIIQVWREDIPVNYMKELELVTKAGFRALLSAPWYLNRISYGPDWKDFYVVEPLAFEGTPEQKAL
VIGGEACMWGEYVDNTNLVPRLWPRAGAVAERLWSNKLTSDLTFAYERLSHFRCELLRRGVQAQ
PLNVGFCEQEFEQT (SEQ ID NO: 370).

[0786] Примечательно, что полученные мотивы РНК будут включены в транслируемую область гена *HEXA*, нарушая функцию гена, кодирующего белок. Вставленные мотивы полиаденилирования приводят к преждевременной терминации транскрипта. Этот сайт просто иллюстрирует потенциальные РЕгРНК, которые могут привести к вставке перечисленных мотивов РНК из приведенной выше таблицы в геномный сайт, который транскрибируется и, таким образом, будет продуцировать продукт РНК.

[0787] РЕгРНК для использования с РЕ для мечения РНК могут быть экспрессированы с промотора U6 (в этом случае один гуанозин будет добавлен на 5'-конце РЕгРНК для направляющих, которые включают протоспейсеры, которые не начинаются с G, и 6-7 тимина будет добавлен на 3'-конце) или промотора pol II, такого как pCMV (в этом случае может оказаться необходимым удалить внутренне транскрибируемую последовательность этого промотора с 5'-конца РНК с помощью саморасщепляющегося элемента или мотива Csy4, и к 3'-концу РНК необходимо добавить мотив терминации, который не приводит к экспорту РНК из ядра, например, мотив 3'-бокса, перечисленный

выше. Необходимо отметить, что этот мотив не будет вставлен в геном в результате РЕ, так как он будет на 3' стороне отжига). Основной каркас РЕгРНК подчеркнут, области гомологии и отжига выделены курсивом, а вставленная последовательность выделена жирным шрифтом. Необходимо отметить, что вставленная последовательность является обратным комплементом приведенных выше примеров, как описано ниже, и поэтому эти РЕгРНК должны быть нацелены на кодирующую цепь.

[0788] Также необходимо отметить, что саморасщепляющиеся рибозимы, отличные от HDV, согласно некоторым вариантам осуществления, должны быть связаны с данным целевым участком, то есть, в то время как HDV расщепляет кодируемый транскрипт непосредственно на 5' своей последовательности, сайты разрезания для всех других саморасщепляющихся рибозимов находятся внутри самого рибозима. Следовательно, первые и последние приблизительно 5-10 нуклеотидов (а в некоторых случаях потенциально более 10) фактически будут частью кодируемой последовательности. Например, чтобы расщепить последовательность 5′NNNNNTCATCCTGATAAACTGCAAA3′ (SEQ ID NO: 371) после 5 N, где N представляет собой любой нуклеотид, с использованием саморасщепляющегося рибозима hammerhead может быть вставлена следующая последовательность, где подчеркнутые последовательности образуют несовершенный элемент спаривания РНК.

[0789] 5′NNNNN<u>CAGTTTGTACGGA</u>TGACTGATGAGTCCCAAATAGGACGAAAC GCGCTTCGGTGCGTCTCATCCTGATAAACTGCAAA-3′ (SEQ ID NO: 372).

[0790] Существует значительная гибкость с точки зрения длины и природы этого элемента спаривания, и это было бы верно для любого из саморасщепляющихся рибозимов, отличных от HDV, перечисленных в исходной заявке. Чтобы установить рибозим hammerhead для расщепления иРНК hexA с использованием РЕгРНК с тем же протоспейсером, что и в перечисленных выше конструкциях, можно использовать следующую последовательность РЕгРНК (метки такие же, как указано выше):

[0791] 5′ACCTGAACCGTATATCGACGCACCGAAGCGCGTTTCGTCCTATTTGGG ACTCATCAGGATATACGGTTCAGGTGATATACGGTTCAGGTGACGCACCGAAGCGCGT TTCGTCCTATTTGGGACTCATCAGACCTGAACCGTATATCATCCTTCCAGTCAGGGCCA T<u>GTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA AGTGGGACCGAGTCGGTCC</u>ACCTGAACCGTATATC**GACGCACCGAAGCGCGTTTCG TCCTATTTGGGACTCATCAGGATATACGGT**TCAGGTCTATGGCCCTGACTGGAA-3′ (где основной каркас РЕгРНК подчеркнут, области гомологии и отжига выделены курсивом, а встроенная последовательность выделена жирным шрифтом) (SEQ ID NO: 373).

[0792] Конструирование других РЕгРНК для вставки мотивов РНК может следовать общему принципу, описанному в настоящем документе. Однако отмечается, что многие мотивы РНК потенциально высоко структурированы, что может затруднить их обратную транскрипцию и вставку в геном. Хотя для некоторых последовательностей РНК, таких как простые шпильки, структурирована как сама последовательность РНК, так и ее комплемент. Однако это вряд ли верно для последовательностей, отмеченных выше. Следовательно, при вставке этих мотивов РЕгРНК, скорее всего, будет кодировать обратный комплемент этих последовательностей, что приведет к вставке последовательности ДНК, фактически кодирующей мотив, в геном. Точно так же включение саморасщепляющегося рибозима в область матрицы РЕгРНК приведет к процессингу и неэффективной активности, в то время как включение его обратного комплемента не приведет. Таким образом, эти РЕгРНК, вероятно, должны нацеливаться на кодирующую цепь, тогда как РЕгРНК, кодирующие другие типы вставок (такие как терапевтическая коррекция), теоретически способны нацеливаться на любую цепь.

[0793] Также необходимо отметить, что для многих вставленных мотивов полученная РЕгРНК может быть неспособна транскрибироваться с промотора U6, что требует использования других промоторов, таких как pCMV. Точно так же более длинные РЕгРНК также могут быть менее стабильными. Более короткие мотивы, такие как маркеры m^6A, не будут приводить к такой проблеме.

D. Применение двойного праймированного редактирования для вставки иммуноэпитопов

[0794] Двойное праймированное редактирование (например, такой как вариант осуществления, показанный на фиг. 90) также можно использовать в качестве средства для вставки известных эпитопов иммуногенности в эндогенную или чужеродную геномную ДНК, что приводит к модификации соответствующих белков для терапевтических или биотехнологических применений (см. фиг. 31 и 32). До изобретения классического праймированного редактирования или двойного праймированного редактирования, такие вставки могли быть достигнуты только неэффективно и с высокими скоростями образования делеций из DSB. Праймированное редактирование решает проблему образования большого количества дефектов из-за редактирований вставки, в то же время предлагая более высокую эффективность, чем HDR. Эта более низкая скорость образования вставки-делеции представляет собой главное преимущество праймированного редактирования по сравнению с HDR как способа направленных вставок ДНК, особенно в описанном применении вставки эпитопов иммуногенности. Праймирующий редактор

представляет собой эффективный подход для достижения таких целенаправленных вставок в клетки млекопитающих.

[0795] Ключевой концепцией настоящего изобретения является использование двойных праймирующих редакторов для вставки нуклеотидной последовательности, содержащей ранее описанные эпитопы иммуногенности, в эндогенную или чужеродную геномную ДНК для подавления и/или разрушения их белковых продуктов и/или типов экспрессирующих клеток. Нуклеотидные последовательности для вставки иммуногенного эпитопа должны быть нацелены на гены таким образом, чтобы продуцировать слитые белки белка, кодируемого целевым геном, и транслировать соответствующий белок вставленного иммуногенного эпитопа. Иммунная система пациента должна быть предварительно обучена распознавать эти эпитопы в результате стандартной предварительной иммунизации путем плановой вакцинации, например, против столбняка, дифтерии или кори. В результате иммуногенной природы слитых эпитопов можно ожидать, что иммунная система пациента распознает и отключит праймированно отредактированный белок (а не только вставленный эпитоп) и, возможно, клетки, из которых он экспрессируется.

[0796] Технологии точного нацеливания на геном с использованием системы CRISPR/Cas недавно были исследованы в широком диапазоне применений, включая вставку сконструированных последовательностей ДНК в целевые геномные локусы. Ранее для этого применения использовали репарацию, направляемую гомологией (HDR), требующую донорной матрицы одноцепочечной ДНК и инициации репарации посредством двухцепочечного разрыва ДНК (DSB). Эта стратегия предлагает широчайший диапазон возможных изменений в клетках и является единственным доступным способом для встраивания больших последовательностей ДНК в клетки млекопитающих. Тем не менее, HDR затруднена из-за нежелательных клеточных побочных эффектов, связанных с ее инициированием DSB, таких как высокий уровень образования инделей (делеций-вставок), транслокации ДНК, большие делеции и активация P53. В дополнение к этим недостаткам HDR ограничена низкой эффективностью во многих типах клеток (Т-клетки являются заметным исключением из этого наблюдения). Недавние усилия по преодолению этих недостатков включают слияние мутантов Rad51 человека с никазой Cas9 D10A (RDN), в результате чего получается система HDR без DSB, которая отличается улучшенным соотношением продукт HDR: индел и более низким нецелевым редактированием, но все еще связана с зависимостью от клеточного типа и лишь умеренной эффективностью редактирования HDR.

[0797] Недавно разработанные слияния Cas9 с обратными транскриптазами («праймирующие редакторы») в сочетании с РЕгРНК представляют собой новую

технологию редактирования генома, которая предлагает ряд преимуществ по сравнению с существующими способами редактирования генома, включая возможность установки любой замены одного нуклеотида, а также вставки или удаления любого короткого участка нуклеотидов (по крайней мере, до нескольких десятков оснований) сайт-специфическим образом. Примечательно, что редактирования PE достигаются, как правило, с низким уровнем непреднамеренных вставок-делеций. Таким образом, PE обеспечивает применения редактирования на основе нацеленной вставки, которые ранее были невозможны или непрактичны.

[0798] Этот конкретный аспект описывает способ использования праймированного редактирования в качестве средства для вставки эпитопов известной иммуногенности в эндогенную или чужеродную геномную ДНК, что приводит к модификации соответствующих белков для терапевтических или биотехнологических применений (см. фиг. 31 и 32). До изобретения праймированного редактирования такие вставки могли быть достигнуты только неэффективно и с высокой скоростью образования вставок-делеций из DSB. Праймированное редактирование решает проблему образования большого количества вставок-делеций из-за редактирования вставки, в то же время предлагая более высокую эффективность, чем HDR. Эта более низкая скорость образования вставок-делеций представляет собой главное преимущество праймированного редактирования по сравнению с HDR как способа направленных вставок ДНК, особенно в описанном применении вставки эпитопов иммуногенности. Длина эпитопов находится в диапазоне от нескольких оснований до сотен оснований. Праймирующий редактор является наиболее эффективной и точной методикой для достижения таких нацеленных вставок в клетки млекопитающих.

[0799] Ключевой концепцией этого аспекта является использование праймирующих редакторов для вставки нуклеотидной последовательности, содержащей ранее описанные эпитопы иммуногенности, в эндогенную или чужеродную геномную ДНК для подавления и/или разрушения их белковых продуктов и/или типов экспрессирующих клеток. Нуклеотидные последовательности для вставки иммуногенного эпитопа должны быть нацелены на гены таким образом, чтобы продуцировать слитые белки белка, кодируемого целевым геном, и транслировать соответствующий белок вставленного эпитопа иммуногенности. Иммунная система пациента должна быть предварительно обучена распознавать эти эпитопы в результате стандартной предварительной иммунизации путем плановой вакцинации, например, против столбняка, дифтерии или кори. В результате иммуногенной природы слитых эпитопов можно ожидать, что иммунная система пациента распознает и отключит праймированно отредактированный белок (а не только вставленный

эпитоп) и, возможно, клетки, из которых он экспрессируется.

[0800] Слияния с генами-мишенями могут быть сконструированы по мере необходимости, чтобы гарантировать, что трансляция белка вставленного эпитопа открыта для распознавания иммунной системой. Это может включать целевые вставки нуклеотидов, приводящие к трансляции белков, с достижением С-концевых слияний эпитопов иммуногенности с генами-мишенями, N-концевые слияния эпитопов иммуногенности с генами-мишенями или вставки нуклеотидов в гены таким образом, что эпитопы иммуногенности кодируются в поверхностно-экспонированных участках белковой структуры.

[0801] Белковые линкеры, кодируемые в виде нуклеотидов, вставленных между последовательностью гена-мишени и вставленной нуклеотидной последовательностью эпитопа иммуногенности, могут быть сконструированы как часть настоящего изобретения для облегчения распознавания иммунной системой, клеточного переноса, функции белка или фолдинга белка гена-мишени. Эти вставленные кодируемые нуклеотидом белковые линкеры могут включать (но без ограничения) линкер XTEN различной длины и последовательности или линкеры глицин-серин различной длины и последовательности. Эти сконструированные линкеры ранее использовали для успешного слияния белков. Примеры линкеров могут включать любые из описанных в настоящем документе, 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 и 30, и где X представляет собой любую аминокислоту. Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность (GGS)n (SEQ ID NO: 176), где n представляет собой 1, 3 или 7. Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGSETPGTSESATPES (SEQ ID NO: 171). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 172). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGSGGSGGS (SEQ ID NO: 173). Согласно некоторым вариантам осуществления линкер содержит аминокислотную последовательность SGGS (SEQ ID NO: 174).

[0802] Отличительные признаки этого аспекта включают способность использовать ранее приобретенные иммунные ответы на конкретные аминокислотные последовательности в качестве средства для индукции иммунного ответа против неиммуногенных белков. Другой отличительной чертой является способность вставлять

нуклеотидные последовательности этих иммуногенных эпитопов нацеленным образом, который не вызывает высоких уровней нежелательных делеций-ставок в качестве побочного продукта редактирования и является эффективным при вставке. Это специфическое применение PE позволяет комбинировать способы доставки, специфичные для типа клеток (такие как серотипы AAV), для вставки эпитопов в типы клеток, которые представляют интерес для запуска иммунного ответа.

[0803] Праймированное редактирование как средство вставки иммуногенных эпитопов в патогенные гены можно использовать для программирования иммунной системы пациента на борьбу с широким спектром заболеваний (без ограничения раком, как в иммуноонкологических стратегиях). Непосредственное применение этой технологии может быть в качестве средства для лечения рака, поскольку она могла бы подорвать механизм ускользания опухоли от иммунного ответа, вызывая иммунный ответ на соответствующий онкоген, такой как HER2, или факторы роста, такие как EGFR. Такой подход может оказаться подобным конструированию Т-клеток, но одним из новых достижений этого подхода является то, что его можно использовать во многих типах клеток и для лечения заболеваний, помимо рака, без необходимости генерировать и вводить пациентам сконструированные Т-клетки.

[0804] Применение PE для вставки эпитопа иммуногенности, против которого большинство людей уже вакцинировано (столбняк, коклюш, дифтерия, корь, свинка, краснуха и т.д.), в чужеродный или эндогенный ген, вызывающий заболевание, чтобы иммунная система пациента научилась отключать этот белок.

[0805] Заболевания, которые могут иметь потенциальную терапевтическую пользу от вышеупомянутой стратегии, включают заболевания, вызванные агрегацией токсичных белков, как например, при фатальной семейной бессоннице. Другие заболевания, которые могут принести пользу, включают заболевания, вызванные патогенной сверхэкспрессией нетоксичного эндогенного белка, а также заболевания, вызванные чужеродными патогенами.

[0806] Первичные терапевтические показания включают упомянутые выше, такие как терапия рака, прионное и другие нейродегенеративные заболевания, инфекционные заболевания и профилактическая медицина. Вторичные терапевтические показания могут включать профилактическую помощь пациентам с запущенными генетическими заболеваниями. Ожидается, что текущие стандартные лекарственные средства могут использоваться в сочетании с праймированным редактированием для некоторых заболеваний, таких как особенно агрессивные виды рака, или в случаях, когда лекарственные средства помогают облегчить симптомы болезни до полного излечения

болезни. Ниже приведены примеры иммуноэпитопов, которые могут быть вставлены в гены раскрытыми в настоящем документе праймирующими редакторами:

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
1	СТОЛБНЯЧНЫЙ ТОКСИН	QYIKANSKFIGITEL (SEQ ID NO: 396)	CATGATATAAAAGCAAATTCTAAATTTATA GGTATAACTGAACTA (SEQ ID NO: 397)
2	МУТАНТ ДИФТЕРИЙНОГО ТОКСИНА CRM197	GADDVVDSSKSFV MENFSSYHGTKPG YVDSIQKGIQKPKS GTQGNYDDDWKEF YSTDNKYDAAGYS VDNENPLSGKAGG VVKVTYPGLTKVL ALKVDNAETIKKEL GLSLTEPLMEQVGT EEFIKRFGDGASRV VLSLPFAEGSSSVEY INNWEQAKALSVE LEINFETRGKRGQD AMYEYMAQACAG NRVRRSVGSSLSCI NLDWDVIRDKTKT KIESLKEHGPIKNK MSESPNKTVSEEKA KQYLEEFHQTALEH PELSELKTVTGTNP VFAGANYAAWAVN VAQVIDSETADNLE KTTAALSILPGIGSV MGIADGAVHHNTE EIVAQSIALSSLMVA QAIPLVGELVDIGFA AYNFVESIINLFQVV	GGCGCCGACGACGTGGTGGACAGCAGCA AGAGCTTCGTGATGGAGAACTTCAGCAGC TACCACGGCACCAAGCCCGGCTACGTGGA CAGCATCCAGAAGGGCATCCAGAAGCCCA AGAGCGGCACCCAGGGCAACTACGACGA CGACTGGAAGGAGTTCTACAGCACCGAC AACAAGTACGACGCCGCCGGCTACAGCGT GGACAACGAGAACCCCCTGAGCGGCAAG GCCGGCGGCGTGGTGAAGGTGACCTACCC CGGCCTGACCAAGGTGCTGGCCCTGAAG GTGGACAACGCCGAGACCATCAAGAAGG AGCTGGGCCTGAGCCTGACCGAGCCCCTG ATGGAGCAGGTGGGCACCGAGGAGTTCAT CAAGAGGTTCGGCGACGGCGCCAGCAGG GTGGTGCTGAGCCTGCCCTTCGCCGAGGG CAGCAGCAGCGTGGAGTACATCAACAACT GGGAGCAGGCCAAGGCCCTGAGCGTGGA GCTGGAGATCAACTTCGAGACCAGGGGC AAGAGGGGCCAGGACGCCATGTACGAGT ACATGGCCCAGGCCTGCGCCGGCAACAG GGTGAGGAGGAGCGTGGGCAGCAGCCTG AGCTGCATCAACCTGGACTGGGACGTGAT CAGGGACAAGACCAAGACCAAGATCGAG AGCCTGAAGGAGCACGGCCCCATCAAGA ACAAGATGAGCGAGAGCCCCAACAAGAC CGTGAGCGAGGAGAAGGCCAAGCAGTAC CTGGAGGAGTTCCACCAGACCGCCCTGGA

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
		HNSYNRPAYSPGHK TQPFLHDGYAVSW NTVEDSIIRTGFQGE SGHDIKITAENTPLP IAGVLLPTIPGKLDV NKSKTHISVNGRKI RMRCRAIDGDVTF CRPKSPVYVGNGV HANLHVAFHRSSSE KIHSNEISSDSIGVL GYQKTVDHTKVNS KLSLFFEIKS (SEQ ID NO: 398)	GCACCCCGAGCTGAGCGAGCTGAAGACC GTGACCGGCACCAACCCCGTGTTCGCCGG CGCCAACTACGCCGCCTGGGCCGTGAACG TGGCCCAGGTGATCGACAGCGAGACCGCC GACAACCTGGAGAAGACCACCGCCGCCC TGAGCATCCTGCCCGGCATCGGCAGCGTG ATGGGCATCGCCGACGGCGCCGTGCACCA CAACACCGAGGAGATCGTGGCCCAGAGC ATCGCCCTGAGCAGCCTGATGGTGGCCCA GGCCATCCCCCTGGTGGGCGAGCTGGTGG ACATCGGCTTCGCCGCCTACAACTTCGTG GAGAGCATCATCAACCTGTTCCAGGTGGT GCACAACAGCTACAACAGGCCCGCCTACA GCCCCGGCCACAAGACCCAGCCCTTCCTG CACGACGGCTACGCCGTGAGCTGGAACA CCGTGGAGGACAGCATCATCAGGACCGGC TTCCAGGGCGAGAGCGGCCACGACATCA AGATCACCGCCGAGAACACCCCCCTGCCC ATCGCCGGCGTGCTGCTGCCCACCATCCC CGGCAAGCTGGACGTGAACAAGAGCAAG ACCCACATCAGCGTGAACGGCAGGAAGAT CAGGATGAGGTGCAGGGCCATCGACGGC GACGTGACCTTCTGCAGGCCCAAGAGCCC CGTGTACGTGGGCAACGGCGTGCACGCCA ACCTGCACGTGGCCTTCCACAGGAGCAGC AGCGAGAAGATCCACAGCAACGAGATCA GCAGCGACAGCATCGGCGTGCTGGGCTAC CAGAAGACCGTGGACCACACCAAGGTGA ACAGCAAGCTGAGCCTGTTCTTCGAGATC AAGAGC (SEQ ID NO: 399)
3	ИММУНОЭПИТОП	GTYRLIPNARANLTA (SEQ ID NO: 400)	GGCACCTACAGGCTGATCCCCAACGCCAG GGCCAACCTGACCGCC (SEQ ID NO: 401)

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
	ПАРОТИТА 1		
4	ИММУНОЭПИТОП ПАРОТИТА 2 ИММУНОЭПИТОП ПАРОТИТА 1	PSKFFTISDSATFAPGPVSNA (SEQ ID NO: 402) PSKLFIMLDNATFAPGPVVNA (SEQ ID NO: 404)	CCGAGCAAATTTTTTACCATTAGCGATAGCGCGACCTTTGCGCCGGGCCCGGTGAGCAACGCG (SEQ ID NO: 403) CCGAGCAAACTGTTTATTATGCTGGATAACGCGACCTTTGCGCCGGGCCCGGTGGTGAACGCG (SEQ ID NO: 405) ВЫБРАННЫЕ ПРИМЕРЫ ИЗ МНОГООБРАЗИЯ ГЕМАГЛЮТИНИН–НЕЙРАМИНИДАЗЫ (HN) СРЕДИ ВСПЫШЧНЫХ ШТАММОВ (ТАБЛИЦА 1), ОТКЛОНЕНИЕ МЕЖДУ ВАКЦИННЫМ ШТАММОМ JL5 И ВСПЫШЧНЫМИ ШТАММАМИ (ТАБЛИЦА 2)
5	ВИРУС КРАСНУХИ (RV)	TPPPYQVSCGGESDRASARVIDPAAQS (SEQ ID NO: 406)	ACCCCCCCCCCCTACCAGGTGAGCTGCGGCGGCGAGAGCGACAGGGCCAGCGCCAGGGTGATCGACCCCGCCGCCCAGAGC (SEQ ID NO: 407)
6	ГЕМАГГЛЮТИНИН	PEYAYKIVKNKKMEDGFLQGMVDGWYGHHSNEQGSGLMENERTLDKANPNNDLCSWSDHEASSNNTNQEDLLQRESRRKKRIGTSTLNQRGNCNTKCQTEEARLKREEVSLVKSDQCSNGSLQCRANNSTEQVD	CCCGAGTACGCCTACAAGATCGTGAAGAACAAGAAGATGGAGGACGGCTTCCTGCAGGGCATGGTGGACGGCTGGTACGGCCACCACAGCAACGAGCAGGGCAGCGGCCTGATGGAGAACGAGAGGACCCTGGACAAGGCCAACCCCAACAACGACCTGTGCAGCTGGAGCGACCACGAGGCCAGCAGCAACAACACCAACCAGGAGGACCTGCTGCAGAGGGAGAGCAGGAGGAAGAAGAGGATCGGCACCAGCACCCTGAACCAGAGGGGCAACTGCAAC

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
		(SEQ ID NO: 408)	ACCAAGTGCCAGACCGAGGAGGCCAGGC TGAAGAGGGAGGAGGTGAGCCTGGTGAA GAGCGACCAGTGCAGCAACGGCAGCCTG CAGTGCAGGGCCAACAACAGCACCGAGC AGGTGGAC (SEQ ID NO: 409)
7	НЕЙРАМИНИДАЗА	TKSTNSRSGGISGP DNEAPVGEAPSPYG DNPRPNDGNNIRIG SKGYNGIITDTIEES CSCYPDAKVVKSV ELDSTIWTSGSSPN QKIITIGWDPNGWT GTPMSPNGAYGTD GPSNGQANQHQAE SISAGNSSLCPIRDN WHGSNRSWSWPD GAE (SEQ ID NO: 410)	ACCAAGAGCACCAACAGCAGGAGCGGCG GCATCAGCGGCCCCGACAACGAGGCCCCC GTGGGCGAGGCCCCCAGCCCCTACGGCG ACAACCCCAGGCCCAACGACGGCAACAA CATCAGGATCGGCAGCAAGGGCTACAACG GCATCATCACCGACACCATCGAGGAGAGC TGCAGCTGCTACCCCGACGCCAAGGTGGT GAAGAGCGTGGAGCTGGACAGCACCATC TGGACCAGCGGCAGCAGCCCCAACCAGA AGATCATCACCATCGGCTGGGACCCCAAC GGCTGGACCGGCACCCCCATGAGCCCCAA CGGCGCCTACGGCACCGACGGCCCCAGC AACGGCCAGGCCAACCAGCACCAGGCCG AGAGCATCAGCGCCGGCAACAGCAGCCT GTGCCCCATCAGGGACAACTGGCACGGCA GCAACAGGAGCTGGAGCTGGCCCGACGG CGCCGAG (SEQ ID NO: 411)
8	TAP1 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА ГЕМАГГЛЮТИНИНЕ ВИРУСА	EKIVLLLAMMEKIV LLLAKCQTPMGAIK AVDGVTNKCPYLG SPSF (SEQ ID NO: 412)	GAGAAGATCGTGCTGCTGCTGGCCATGAT GGAGAAGATCGTGCTGCTGCTGGCCAAGT GCCAGACCCCCATGGGCGCCATCAAGGCC GTGGACGGCGTGACCAACAAGTGCCCCTA CCTGGGCAGCCCCAGCTTC (SEQ ID NO: 413)

ВАКЦИНА	**ЗАБОЛЕВАНИЕ**	**АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА**	**ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)**
	H5N1		
9	TAP2 (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА НЕЙРАМИНИДАЗЕ ВИРУСА H5N1	IRPCFWVELNPNQKIITIRPCFWVELICYPDAGEIT (SEQ ID NO: 414)	ATCAGGCCCTGCTTCTGGGTGGAGCTGAACCCCAACCAGAAGATCATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGATCTGCTACCCCGACGCCGGCGAGATCACC (SEQ ID NO: 415)
10	ЭПИТОПЫ ГЕМАГГЛЮТИНИНА 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 ГЕМАГГЛЮТИНИНА К HLA КЛАССА II	MVSLVKSDQIGTSTLNQR (SEQ ID NO: 420)	ATGGTGAGCCTGGTGAAGAGCGACCAGATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 421)
13	ЭПИТОПЫ НЕЙРАМИН	YNGIITDTI (SEQ ID NO: 422)	TACAACGGCATCATCACCGACACCATC (SEQ ID NO: 423)

ВАКЦИНА	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	ИЛЛЮСТРАТИВНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕИНОВОЙ КИСЛОТЫ (8)
	ИДАЗЫ К HLA КЛАССА II		
14	ГЕМАГГЛЮТИНИНОВЫЙ ЭПИТОП H5N1, СВЯЗАННЫЙ С HLA КЛАССА I И КЛАССА II	MEKIVLLLAEKIVLLLAMMVSLVKSDQCPYLGSPSFIGTSTLNQRKCQTPMGAIKAVDGVTNK (SEQ ID NO: 424)	ATGGAGAAGATCGTGCTGCTGCTGGCCGAGAAGATCGTGCTGCTGCTGGCCATGATGGTGAGCCTGGTGAAGAGCGACCAGTGCCCCTACCTGGGCAGCCCCAGCTTCATCGGCACCAGCACCCTGAACCAGAGG (SEQ ID NO: 425)
15	ЭПИТОП НЕЙРАМИНИДАЗЫ, СВЯЗАННЫЙ С H5N1 КЛАССА I И КЛАССА II HLA	NPNQKIITIYNGIITDTICYPDAGEITIRPCFWVELRPCFWVELI (SEQ ID NO: 426)	AACCCCAACCAGAAGATCATCACCATCTACAACGGCATCATCACCGACACCATCTGCTACCCCGACGCCGGCGAGATCACCATCAGGCCCTGCTTCTGGGTGGAGCTGAGGCCCTGCTTCTGGGTGGAGCTGATC (SEQ ID NO: 427)

[0807] Также могут быть установлены дополнительные иммуноэпитопы, известные в данной области техники. Любой из иммуноэпитопов, доступных в базе данных и ресурсе для анализа иммунных эпитопов (iedb.org/epitopedetails_v3.php) (содержание которых включено в настоящий документ посредством ссылки), может быть установлен праймирующими редакторами, раскрытыми в настоящем документе.

[0808] Согласно некоторым вариантам осуществления иммуноэпитопы, которые могут быть установлены праймирующими редакторами, раскрытыми в настоящем документе, могут включать любой из следующих эпитопов:

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
1	*СТОЛБНЯЧНЫЙ ТОКСИН*	*QYIKANSKFIGITEL*	*396*	*NA*
2	*МУТАНТ ДИФТЕРИЙНОГО ТОКСИНА CRM197*	*GADDVVDSSKSFVMENFSSYHGT KPGYVDSIQKGIQKPKSGTQGNY DDDWKEFYSTDNKYDAAGYSVD NENPLSGKAGGVVKVTYPGLTKV LALKVDNAETIKKELGLSLTEPL MEQVGTEEFIKRFGDGASRVVLS LPFAEGSSSVEYINNWEQAKALSV ELEINFETRGKRGQDAMYEYMA QACAGNRVRRSVGSSLSCINLDW DVIRDKTKTKIESLKEHGPIKNK MSESPNKTVSEEKAKQYLEEFHQ TALEHPELSELKTVTGTNPVFAGA NYAAWAVNVAQVIDSETADNLEK TTAALSILPGIGSVMGIADGAVHH NTEEIVAQSIALSSLMVAQAIPLVG ELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYA VSWNTVEDSIIRTGFQGESGHDI KITAENTPLPIAGVLLPTIPGKLD VNKSKTHISVNGRKIRMRCRAID GDVTFCRPKSPVYVGNGVHANL HVAFHRSSSEKIHSNEISSDSIGVL GYQKTVDHTKVNSKLSLFFEIKS*	*428*	*NA*
3	*ПАРОТИТ*	*GTYRLIPNARANLTA*	*400*	*NA*
4	*ПАРОТИТ*	*PSKFFTISDSATFAPGPVSNA, PSKLFIMLDNATFAPGPVVNA*	*402, 404*	*NA*
5	ВИРУС КРАСНУХИ *(RV)*	*TPPPYQVSCGGESDRASARVIDPA AQS*	*406*	*NA*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
6	*ГЕМАГГЛЮТИНИН*	*PEYAYKIVKNKKMEDGFLQGMV DGWYGHHSNEQGSGLMENERTL DKANPNNDLCSWSDHEASSNNT NQEDLLQRESRRKKRIGTSTLNQ RGNCNTKCQTEEARLKREEVSLV KSDQCSNGSLQCRANNSTEQVD*	*408*	*NA*
7	*НЕЙРАМИНИДАЗА*	*TKSTNSRSGGISGPDNEAPVGEA PSPYGDNPRPNDGNNIRIGSKGY NGIITDTIEESCSCYPDAKVVKSV ELDSTIWTSGSSPNQKIITIGWDP NGWTGTPMSPNGAYGTDGPSNG QANQHQAESISAGNSSLCPIRDN WHGSNRSWSWPDGAE*	*410*	*NA*
8	*ТАР (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА ГЕМАГГЛЮТИНИНЕ ВИРУСА H5N1*	*EKIVLLLAMMEKIVLLLAKCQTP MGAIKAVDGVTNKCPYLGSPSF*	*412*	*NA*
9	*ТАР (ПРЕЗЕНТАЦИЯ ТРАНСПОРТНОГО АНТИГЕНА) НА НЕЙРАМИНИДАЗЕ ВИРУСА H5N1*	*IRPCFWVELNPNQKIITIRPCFWV ELICYPDAGEIT*	*414*	*NA*
10	*ЭПИТОПЫ ГЕМАГГЛЮТИНИНА К HLA КЛАССА I*	*MEKIVLLLAEKIVLLLAMCPYLGS PSFKCQTPMGAIKAVDGVTNK*	*416*	*NA*
11	*ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА I*	*NPNQKIITICYPDAGEITIRPCFW VELRPCFWVELI*	*418*	*NA*
12	*ЭПИТОПЫ ГЕМАГГЛЮТИНИНА*	*MVSLVKSDQIGTSTLNQR*	*420*	*NA*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	К HLA КЛАССА II			
13	ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА II	YNGIITDTI	422	NA
14	ЭПИТОПЫ ГЕМАГГЛЮТИНИНА К HLA КЛАССА I И КЛАССА II	MEKIVLLLAEKIVLLLAMMVSLVK SDQCPYLGSPSFIGTSTLNQRKC QTPMGAIKAVDGVTNK	424	NA
15	ЭПИТОПЫ НЕЙРАМИНИДАЗЫ К HLA КЛАССА I И КЛАССА II, H5N1-СВЯЗАННЫЙ	NPNQKIITIYNGIITDTICYPDAGE ITIRPCFWVELRPCFWVELI	426	NA
16	ДИФТЕРИЙНАЯ ПАЛОЧКА	AACAGNRVRRSVGSSLKC	899	SRC280292
17	ШТАММ EDMONSTON ВИРУСА КОРИ	AADHCPVVEVNGVTI	900	P69353.1
18	ШТАММ EDMONSTON ВИРУСА КОРИ	AAHLPTGTPLDID	901	P04851.1
19	ПАЛОЧКА КОКЛЮША	AALAVWAGLAVQ	902	Q00879.1
20	ШТАММ EDMONSTON ВИРУСА КОРИ	AALGVATAAQITAGI	903	P69353.1
21	ВИРУС КРАСНУХИ STRAIN THERIEN	AALLNTPPPYQVSCGGESDRATA R	904	P07566.1
22	ВИРУС КРАСНУХИ	AAQSFTGVVYGTHTT	905	BAA28178.1
23	ВИРУС КРАСНУХИ	ACEVEPAFGHSDAAC	906	BAA28178.1
24	ВИРУС КРАСНУХИ	ACTFWAVNAYSSGGY	907	BAA28178.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
25	ВИРУС КРАСНУХИ	ADDPLLR	908	BAA19893.1
26	ВИРУС КРАСНУХИ	ADDPLLRT	909	CAJ88851.1
27	ШТАММ EDMONSTON ВИРУСА КОРИ	AEMICDIDTYIVEAG	910	P04851.1
28	ШТАММ EDMONSTON ВИРУСА КОРИ	AEMICDIDTYIVEAGLASFI	911	P04851.1
29	ШТАММ EDMONSTON ВИРУСА КОРИ	AEPLLSC	912	P04851.1
30	ПАЛОЧКА КОКЛЮША	AFVSTSSSRRYTEVY	913	CAD44970.1
31	ПАЛОЧКА КОКЛЮША	AGFIYRETFCITTIYKTGQPAADH YYSKVTA	914	P04979.1
32	ВИРУС КРАСНУХИ	AGLLACCAKCLYYLR	915	BAA28178.1
33	ВИРУС КРАСНУХИ	AHTTSDPWHPPG	916	BAA19893.1
34	ШТАММ EDMONSTON ВИРУСА КОРИ	AIAKLEDAKELLESS	917	P69353.1
35	ШТАММ EDMONSTON ВИРУСА КОРИ	AIDNLRASLETTNQA	918	P69353.1
36	ПАЛОЧКА КОКЛЮША	AKGVEFR	919	ACI16088.1
37	ШТАММ EDMONSTON ВИРУСА КОРИ	AKWAVPTTRTDDKLR	920	P08362.1
38	ШТАММ EDMONSTON ВИРУСА КОРИ	ALAEVLKKPV	921	ABO69699.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
39	ШТАММ EDMONSTON ВИРУСА КОРИ	ALGVINTLEWIPRFK	922	P08362.1
40	ШТАММ EDMONSTON ВИРУСА КОРИ	ALHQSMLNSQAIDNL	923	P69353.1
41	ШТАММ EDMONSTON ВИРУСА КОРИ	ALIGILSLFV	924	ABI54110.1
42	ВИРУС КРАСНУХИ	ALLNTPPPYQVSCGGESDRA	925	CAJ88851.1
43	ВИРУС КРАСНУХИ STRAIN M33	ALLNTPPPYQVSCGGESDRASARV	926	CAJ88851.1
44	ВИРУС КРАСНУХИ	ALVEGLAPGGGNCHL	927	BAA28178.1
45	ПАЛОЧКА КОКЛЮША	AMAAWSERAGEA	928	P04977.1
46	ШТАММ EDMONSTON ВИРУСА КОРИ	ANCASILCKCYTTGT	929	P69353.1
47	ПАЛОЧКА КОКЛЮША	ANPNPYTSRRSV	930	P04977.1
48	ВИРУС КРАСНУХИ	APGPGEVW	931	CAJ88851.1
49	ВИРУС КРАСНУХИ	APLPPHTTERIETRSARHPWRIR	932	ABD64214.1
50	ВИРУС КРАСНУХИ VACCINE STRAIN RA27/3	APPMPPQPPRAHGQHYGHHHHQLPFLG	933	CAA33016.1
51	ВИРУС КРАСНУХИ STRAIN THERIEN	APPTLPQPPCAHGQHYGHHHHQLPFLG	934	P07566.1
52	ВИРУС КРАСНУХИ	APPTLPQPPRAHGQHYGHHHHQLPFLG	935	ABD64214.1
53	ПАЛОЧКА КОКЛЮША	APQPGPQPPQPPQPQPEAPAPQ	936	P14283.3

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
54	ВИРУС КРАСНУХИ	AQLASYFNPGGSYYK	937	BAA28178.1
55	ШТАММ EDMONSTON ВИРУСА КОРИ	ARAAHLPTGTPLD	938	P04851.1
56	ШТАММ EDMONSTON ВИРУСА КОРИ	ARLVSEIAMHTTEDK	939	P04851.1
57	ШТАММ EDMONSTON ВИРУСА КОРИ	ARLVSEIAMHTTEDKISRAV	940	P04851.1
58	ШТАММ EDMONSTON ВИРУСА КОРИ-В	ASDVETAEGGEIHELLR	941	P03422.1
59	МОРБИЛЛИВИРУС КОРИ	ASDVETAEGGEIHELLRLQ	942	ABO69699.1
60	ШТАММ EDMONSTON ВИРУСА КОРИ-В	ASDVETAEGGEIHELLRLQSR	943	P03422.1
61	МОРБИЛЛИВИРУС КОРИ	ASDVETAEGGEIHKLLRLQ	944	AAA63285.1
62	ВИРУС КРАСНУХИ STRAIN M33	ASDVLPGHWLQG	945	NP_740663.1
63	ШТАММ EDMONSTON ВИРУСА КОРИ	ASELGITAEDARLVS	946	P04851.1
64	ШТАММ EDMONSTON ВИРУСА КОРИ	ASELGITAEDARLVSEIAMH	947	P04851.1
65	ШТАММ EDMONSTON ВИРУСА КОРИ	ASESSQDPQDSRR	948	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
66	ШТАММ EDMONSTON ВИРУСА КОРИ	ASILCKCYTTGTIIN	949	P69353.1
67	ВИРУС КРАСНУХИ	ASPVCQRHSPDCSRL	950	BAA28178.1
68	ПАЛОЧКА КОКЛЮША	ASQARWTGATRA	951	BAF35031.1
69	ВИРУС КРАСНУХИ STRAIN THERIEN	ASYFNPGGSYYKQYHPTACEVEP AFGHS	952	P07566.1
70	ШТАММ EDMONSTON ВИРУСА КОРИ	ASYKVMTRSSHQSLV	953	P69353.1
71	ПАЛОЧКА КОКЛЮША	ASYVKKPKEDVD	954	ACI16088.1
72	ШТАММ EDMONSTON ВИРУСА КОРИ	ATAAQITAGIALHQS	955	P69353.1
73	ВИРУС КРАСНУХИ	ATPERPRL	956	CAJ88851.1
74	ШТАММ EDMONSTON ВИРУСА КОРИ	AVCLGGLIGIPALIC	957	P69353.1
75	ШТАММ EDMONSTON ВИРУСА КОРИ	AVGPRQAQVSF	958	P04851.1
76	ВИРУС КРАСНУХИ	AVNAYSSGGYAQLAS	959	BAA28178.1
77	ВИРУС КРАСНУХИ	AVSETRQTWAEWAAA	960	BAA28178.1
78	ШТАММ EDMONSTON ВИРУСА КОРИ	AVTAPDTAADSELRR	961	P04851.1
79	ШТАММ EDMONSTON ВИРУСА КОРИ	AVTAPDTAADSELRRWIKYT	962	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
80	ПАЛОЧКА КОКЛЮША	AYGGIIKDAPPGAGFIYRETFC	963	P04979.1
81	ВИРУС КРАСНУХИ	CALPLAGLLACCAKC	964	BAA28178.1
82	ВИРУС КРАСНУХИ	CARIWNGTQRACTFW	965	BAA28178.1
83	ШТАММ EDMONSTON ВИРУСА КОРИ	CARTLVSGSFGNRFI	966	P69353.1
84	ПАЛОЧКА КОКЛЮША	CASPYEGRYRDMYDALRBRLLY	967	SRC280066
85	ПАЛОЧКА КОКЛЮША	CAVFVRSGQPVIGA	968	AAA83981.1
86	ВИРУС КРАСНУХИ	CCAKCLYYLRGAIAPR	969	BAA28178.1
87	ШТАММ EDMONSTON ВИРУСА КОРИ	CCRGRCNKKGEQVGM	970	P69353.1
88	ВИРУС КРАСНУХИ	CEIPTDVSCEGLGAW	971	BAA28178.1
89	ПАЛОЧКА КОКЛЮША	CFGKDLKRPGSSPMEV	972	P0A3R5.1
90	МОРБИЛЛИВИРУС КОРИ	CFQQACKGKIQALCE	973	P06830.1
91	МОРБИЛЛИВИРУС КОРИ	CFQQACKGKIQALCENPEWAPL KDNRIPS	974	AAR89413.1
92	ВИРУС КРАСНУХИ	CGGESDRASARVIDP	975	BAA28178.1
93	ПАЛОЧКА КОКЛЮША	CITTIYKTGQPAADHYYSKVTA	976	P04979.1
94	МОРБИЛЛИВИРУС КОРИ	CKGKIQALCENPEWA	977	AAR89413.1
95	ШТАММ EDMONSTON ВИРУСА КОРИ	CKPWQESRKNKAQ	978	P04851.1
96	ШТАММ	CNKKGEQVGMSRPGL	979	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
97	ВИРУС КРАСНУХИ	CNVTTEHPFCNTPHG	980	BAA28178.1
98	ПАЛОЧКА КОКЛЮША	CQVGSSNSAF	981	P04977.1
99	МОРБИЛЛИВИРУС КОРИ	CSGPTTIRGQFS	982	P08362.1
100	ПАЛОЧКА КОКЛЮША	CTSPYDGKYWSMYSRL	983	AAA83981.1
101	ШТАММ EDMONSTON ВИРУСА КОРИ	CVLADSESGGHITHS	984	P08362.1
102	МОРБИЛЛИВИРУС КОРИ	CYTTGTIINQDPDKILTYIAADHC	985	AAF02706.1
103	ВИРУС КРАСНУХИ	DADDPLLR	986	CAJ88851.1
104	ШТАММ EDMONSTON ВИРУСА КОРИ	DARAAHLPTGTPLDI	987	P04851.1
105	ШТАММ EDMONSTON ВИРУСА КОРИ	DARAAHLPTGTPLDIDTASE	988	P04851.1
106	ШТАММ EDMONSTON ВИРУСА КОРИ	DCHAPTYLPAEVDGD	989	P08362.1
107	ВИРУС КРАСНУХИ	DCSRLVGATPERPRL	990	BAA28178.1
108	ШТАММ EDMONSTON ВИРУСА КОРИ	DDKLRMETCFQQACK	991	P08362.1
109	ВИРУС КРАСНУХИ	DDPLLRTA	992	CAJ88851.1
110	ВИРУС КРАСНУХИ	DDPLLRTAPGPGEVW	993	BAA28178.1
111	ПАЛОЧКА	DDPPATVYRYD	994	P04977.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
112	ПАЛОЧКА КОКЛЮША	DDPPATVYRYDSRPPED	995	CAD44970.1
113	ПАЛОЧКА КОКЛЮША	DDPPATVYRYDSRPPEDV	996	ACI04548.1
114	МОРБИЛЛИВИРУС КОРИ	DEVGLRTPQRFTDLV	997	P06830.1
115	ШТАММ EDMONSTON ВИРУСА КОРИ	DHCPVVEVNGVTIQV	998	P69353.1
116	ШТАММ EDMONSTON ВИРУСА КОРИ	DIDTASESSQDPQ	999	P04851.1
117	ШТАММ EDMONSTON ВИРУСА КОРИ	DINKVLEKLGYSGGD	1000	P69353.1
118	ПАЛОЧКА КОКЛЮША	DLIAYKQ	1001	ACI16088.1
119	ШТАММ EDMONSTON ВИРУСА КОРИ	DLIGQKLGLKLLRYY	1002	P69353.1
120	ВИРУС КРАСНУХИ	DLQKALEAQSRALRAELAA	1003	P07566.1
121	МОРБИЛЛИВИРУС КОРИ	DLQYVLATYDTSRVE	1004	P06830.1
122	ШТАММ EDMONSTON ВИРУСА КОРИ	DLSLRRFMV	1005	P04851.1
123	ШТАММ EDMONSTON ВИРУСА КОРИ	DLSNCMVALGELKLA	1006	P08362.1
124	ВИРУС КРАСНУХИ	DLVEYIMNYTGNQQSRWGLGSP	1007	CAJ88851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
		NC		
125	МОРБИЛЛИВИРУС КОРИ	DLVKFISDKIKFLNP	1008	AAR89413.1
126	МОРБИЛЛИВИРУС КОРИ	DLVKFISTKIKFLNP	1009	SRC280117
127	ШТАММ EDMONSTON ВИРУСА КОРИ	DLYKSNHNNV	1010	P08362.1
128	ПАЛОЧКА КОКЛЮША	DNVLDHLTGR	1011	ACI04548.1
129	ПАЛОЧКА КОКЛЮША	DNVLDHLTGRSC	1012	P04977.1
130	ПАЛОЧКА КОКЛЮША	DNVLDHLTGRSCQ	1013	P04977.1
131	МОРБИЛЛИВИРУС КОРИ	DPDKILTYIAA	1014	AAF02706.1
132	ВИРУС КРАСНУХИ STRAIN THERIEN	DPGDLVEYIMNYTGNQQSR	1015	P07566.1
133	ВИРУС КРАСНУХИ	DPLLRTAP	1016	CAJ88851.1
134	ВИРУС КРАСНУХИ	DPLLRTAPGPGEVWVTPVIGSQ	1017	CAJ88851.1
135	ШТАММ EDMONSTON ВИРУСА КОРИ	DPQDSRRSAEPLL	1018	P04851.1
136	ШТАММ EDMONSTON ВИРУСА КОРИ	DPVIDRLYLSSHRGV	1019	P08362.1
137	ШТАММ EDMONSTON ВИРУСА КОРИ	DQILRSMKGLSSTSI	1020	P69353.1
138	МОРБИЛЛИВИРУС КОРИ	DQYCADVAAEELMNA	1021	P06830.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
139	ШТАММ EDMONSTON ВИРУСА КОРИ	DSESGGHITH	1022	P08362.1
140	ШТАММ EDMONSTON ВИРУСА КОРИ	DTASESSQDPQDS	1023	P04851.1
141	ВИРУС КРАСНУХИ	DTVMSVFALASYVQH	1024	BAA28178.1
142	ПАЛОЧКА КОКЛЮША	DVFQNGFTAWGNND	1025	P04977.1
143	ВИРУС КРАСНУХИ	DVGAVPPGKFVTAAL	1026	BAA28178.1
144	ДИФТЕРИЙНАЯ ПАЛОЧКА	DVNKSKTHISVNGRKI	1027	CAE11230.1
145	ВИРУС КРАСНУХИ	DVSCEGLGAWVPAAP	1028	BAA28178.1
146	ВИРУС КРАСНУХИ STRAIN THERIEN	DWASPVCQRHSPDCSRLVGATPE RPRL	1029	P07566.1
147	ШТАММ EDMONSTON ВИРУСА КОРИ	EARESYRETGPSR	1030	P04851.1
148	ПАЛОЧКА КОКЛЮША	EAVEAERAGRGTG	1031	ACI04548.1
149	ШТАММ EDMONSTON ВИРУСА КОРИ	EDAKELLESSDQILR	1032	P69353.1
150	ШТАММ EDMONSTON ВИРУСА КОРИ	EDRRVKQSRGEAR	1033	P04851.1
151	ШТАММ EDMONSTON ВИРУСА КОРИ	EDSITIPYQGSGKGV	1034	P08362.1
152	ВИРУС КРАСНУХИ	EEAFTYLCTAPGCAT	1035	BAA28178.1
153	ШТАММ	EGFNMILGTILAQIW	1036	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
154	ШТАММ EDMONSTON ВИРУСА КОРИ	EGFNMILGTILAQIWVLLAK	1037	P04851.1
155	ВИРУС КРАСНУХИ	EHPFCNTPHGQLEVQ	1038	BAA28178.1
156	ШТАММ EDMONSTON ВИРУСА КОРИ	EISDIEVQDPEGFNM	1039	P04851.1
157	ШТАММ EDMONSTON ВИРУСА КОРИ	EISDIEVQDPEGFNMILGTI	1040	P04851.1
158	ШТАММ EDMONSTON ВИРУСА КОРИ	EKPNLSSKRSE	1041	P08362.1
159	ШТАММ EDMONSTON ВИРУСА КОРИ	ELKLAALCHGEDSIT	1042	P08362.1
160	МОРБИЛЛИВИРУС КОРИ	ELMNALVNSTLLETR	1043	P06830.1
161	ШТАММ EDMONSTON ВИРУСА КОРИ	ELPRL	1044	P04851.1
162	МОРБИЛЛИВИРУС КОРИ	ENPEWAPLKDNRIPSYGVLSVDL	1045	AAR89413.1
163	ШТАММ EDMONSTON ВИРУСА КОРИ	EPIRDALNAMTQNIR	1046	P69353.1
164	ШТАММ EDMONSTON ВИРУСА КОРИ	EQVGMSRPGLKPDLT	1047	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
165	ВИРУС КРАСНУХИ	ERPRLRLV	1048	CAJ88851.1
166	ВИРУС КРАСНУХИ	ERPRLRLVDADDPLL	1049	BAA28178.1
167	MEASLES VIRUS CAM/RB	ESPGQLIQRITDDPDVS	1050	P04851.1
168	ШТАММ EDMONSTON ВИРУСА КОРИ	ESRGIKARITHVDTE	1051	P69353.1
169	ШТАММ EDMONSTON ВИРУСА КОРИ	ESSCTFMPEGTVCSQ	1052	P69353.1
170	ШТАММ EDMONSTON ВИРУСА КОРИ	ESSQDPQDSRRSA	1053	P04851.1
171	ШТАММ EDMONSTON ВИРУСА КОРИ	ETRTTNQFLAVSKGN	1054	P08362.1
172	ШТАММ EDMONSTON ВИРУСА КОРИ	EVDGDVKLSSNLVIL	1055	P08362.1
173	ШТАММ EDMONSTON ВИРУСА КОРИ-В	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	ПАЛОЧКА КОКЛЮША	FEYVDTYGDNAG	1061	P04977.1
179	МОРБИЛЛИВИРУС КОРИ	FGPLITHGSGMDLYK	1062	P06830.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
180	ШТАММ EDMONSTON ВИРУСА КОРИ	FIFDALAEV	1063	ABK40531.1
181	ШТАММ EDMONSTON ВИРУСА КОРИ	FISDKIKFL	1064	P08362.1
182	ШТАММ EDMONSTON ВИРУСА КОРИ	FKRNKDKPPITSGSG	1065	P04851.1
183	ШТАММ EDMONSTON ВИРУСА КОРИ	FKRNKDKPPITSGSGGAIRG	1066	P04851.1
184	ВИРУС КРАСНУХИ	FKTVRPVALPRTLAP	1067	BAA28178.1
185	ШТАММ EDMONSTON ВИРУСА КОРИ	FLMDRHIIV	1068	ABK40531.1
186	ШТАММ EDMONSTON ВИРУСА КОРИ	FMAVLLTLQTPTGQI	1069	P69353.1
187	ШТАММ EDMONSTON ВИРУСА КОРИ	FMPEGTVCSQNALYP	1070	P69353.1
188	МОРБИЛЛИВИРУС КОРИ	FMYMSLLGV	1071	AAN09804.1
189	ШТАММ EDMONSTON ВИРУСА КОРИ	FNVPIKEAGEDCHAP	1072	P08362.1
190	МОРБИЛЛИВИРУС КОРИ	FRDLTWCINPPERIK	1073	AAC35876.2
191	ШТАММ EDMONSTON	FSHDDPISSDQSRFG	1074	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
192	ШТАММ EDMONSTON ВИРУСА КОРИ	FSHDDPISSDQSRFGWFENK	1075	P04851.1
193	МОРБИЛЛИВИРУС КОРИ	FTDLVKFISDKIKFL	1076	P06830.1
194	МОРБИЛЛИВИРУС КОРИ	FTWDQKLWCRHFCVL	1077	P06830.1
195	ПАЛОЧКА КОКЛЮША	FVRDGQSVIGACASPYEGRYRDL YDALRRLLY	1078	SRC280066
196	ПАЛОЧКА КОКЛЮША	FVRSGQPVIGACTSPYDGKYWSIL YSRLRKMLY	1079	SRC280066
197	МОРБИЛЛИВИРУС КОРИ	FYKDNPHPKGSRIVI	1080	P06830.1
198	ПАЛОЧКА КОКЛЮША	GAASSYFEYVDTYG	1081	ACI04548.1
199	ПАЛОЧКА КОКЛЮША	GAFDLKTTFCIMTTRNTGQPA	1082	AAA83981.1
200	ШТАММ EDMONSTON ВИРУСА КОРИ	GALIGILSLFVESPG	1083	P04851.1
201	ШТАММ EDMONSTON ВИРУСА КОРИ	GALIGILSLFVESPGQLIQR	1084	P04851.1
202	ВИРУС КРАСНУХИ	GATPERPR	1085	CAJ88851.1
203	ПАЛОЧКА КОКЛЮША	GAYGRCPNGTRALTVAELRGNAE L	1086	P04979.1
204	ВИРУС КРАСНУХИ	GCFAPWDLEATGACI	1087	BAA28178.1
205	МОРБИЛЛИВИРУС КОРИ	GDINKVLEKLGYS	1088	BAB60865.1
206	МОРБИЛЛИВИРУС	GDINKVLEKLGYSGGDLLG	1089	AAL29688.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОРИ			
207	ВИРУС КРАСНУХИ	GDLRAVHHRPVPA	1090	CAA28880.1
208	ВИРУС КРАСНУХИ	GDLVEYIMNYTGNQQ	1091	BAA28178.1
209	ШТАММ EDMONSTON ВИРУСА КОРИ	GDSSITTRSRLLDRL	1092	P04851.1
210	ШТАММ EDMONSTON ВИРУСА КОРИ	GDSSITTRSRLLDRLVRLIG	1093	P04851.1
211	МОРБИЛЛИВИРУС КОРИ	GEDCHAPTYLPAEVD	1094	P06830.1
212	ШТАММ EDMONSTON ВИРУСА КОРИ	GELSTLESLMNLYQQ	1095	P04851.1
213	ШТАММ EDMONSTON ВИРУСА КОРИ	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	ПАЛОЧКА КОКЛЮША	GEYGGVIKDGTPGGA	1101	AAA83981.1
219	ВИРУС КРАСНУХИ	GFLSGVGPMRLRHGADT	1102	SRC265968
220	ШТАММ EDMONSTON ВИРУСА КОРИ-B	GFRASDVETAEGGEIHELLRLQ	1103	P03422.1
221	ПАЛОЧКА КОКЛЮША	GGAVPGGAVPGGAVPGGFGPG GFGP	1104	P14283.3
222	ПАЛОЧКА	GGAVPGGAVPGGFGPGGFGPG	1105	CAA09475.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША	GFGP		
223	ПАЛОЧКА КОКЛЮША	GGAVPGGAVPGGFGPGGFGPG GFGPGGFGP	1106	CAA09474.1
224	МОРБИЛЛИВИРУС КОРИ	GGHITHSGMVGMGVS	1107	P06830.1
225	МОРБИЛЛИВИРУС КОРИ	GILESRGIKARITHVDTESY	1108	P26032.1
226	ПАЛОЧКА КОКЛЮША	GITGETTTTEYSNARYV	1109	CAD44970.1
227	ШТАММ EDMONSTON ВИРУСА КОРИ	GKEDRRVKQSRGE	1110	P04851.1
228	ПАЛОЧКА КОКЛЮША	GKVTNGS	1111	ACI16088.1
229	ВИРУС КРАСНУХИ	GLGAWVPAAPCARIW	1112	BAA28178.1
230	ШТАММ EDMONSTON ВИРУСА КОРИ	GLIGIPALICCCRGR	1113	P69353.1
231	ВИРУС КРАСНУХИ STRAIN 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	ПАЛОЧКА КОКЛЮША	GNAELQTYLRQITPGWSIYGLYD GTYLG	1119	P04979.1
237	ВИРУС КРАСНУХИ	GNCHLTVNGEDVGAV	1120	BAA28178.1
238	ПАЛОЧКА	GNNDNVLDHLTGR	1121	P04977.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
239	ПАЛОЧКА КОКЛЮША	GNNDNVLDHLTGRSC	1122	P04977.1
240	ШТАММ EDMONSTON ВИРУСА КОРИ	GNRFILSQGNLIANC	1123	P69353.1
241	ВИРУС КРАСНУХИ	GNRGRGQRRDWSRAPPPPEERQ ETRSQTPAPKPS	1124	P07566.1
242	ВИРУС КРАСНУХИ	GPGEVWVT	1125	CAJ88851.1
243	ВИРУС КРАСНУХИ STRAIN THERIEN	GPMRLRHGADTRCGRLI	1126	P07566.1
244	ПАЛОЧКА КОКЛЮША	GPNHTKV	1127	ACI16083.1
245	MEASLES VIRUS STRAIN HALLE	GPRQAQVSF	1128	P10050.1
246	ШТАММ EDMONSTON ВИРУСА КОРИ	GPRQAQVSFLQGDQS	1129	P04851.1
247	ШТАММ EDMONSTON ВИРУСА КОРИ	GPRQAQVSFLQGDQSENELP	1130	P04851.1
248	МОРБИЛЛИВИРУС КОРИ	GRGYNVSSIVTMTSQ	1131	P06830.1
249	ПАЛОЧКА КОКЛЮША	GRTPFII	1132	ACI16083.1
250	ВИРУС КРАСНУХИ	GSPNCHGPDWASPVC	1133	BAA28178.1
251	ВИРУС КРАСНУХИ	GSQARKCGLHIRAGP	1134	BAA28178.1
252	ПАЛОЧКА КОКЛЮША	GSSNSAFVSTSSSRR	1135	P04977.1
253	ШТАММ EDMONSTON	GSTKSCARTLVSGSF	1136	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
254	ВИРУС КРАСНУХИ	GSYYKQYHPTACEVE	1137	BAA28178.1
255	ВИРУС КРАСНУХИ	GTHTTAVSETRQTWA	1138	BAA28178.1
256	ШТАММ EDMONSTON ВИРУСА КОРИ	GTIINQDPDKILTYI	1139	P69353.1
257	ПАЛОЧКА КОКЛЮША	GTLVRMAPVIG	1140	ADA85124.1
258	ШТАММ EDMONSTON ВИРУСА КОРИ	GTPLDIDTASESS	1141	P04851.1
259	ПАЛОЧКА КОКЛЮША	GTYLGQAYGGIIKDAPPGAGFIYR ETFC	1142	P04979.1
260	ПАЛОЧКА КОКЛЮША	GVATKGLGVHAKSSDWG	1143	P15318.2
261	ДИФТЕРИЙНАЯ ПАЛОЧКА	GVLLPTIPGKLDVNKSKTHI	1144	AAV70486.1
262	МОРБИЛЛИВИРУС КОРИ	GVLSVDLSLTVELKI	1145	P06830.1
263	МОРБИЛЛИВИРУС КОРИ	GVPIELQVECFTWDQ	1146	P06830.1
264	ШТАММ EDMONSTON ВИРУСА КОРИ	GVSCTVTREDGTNRR	1147	P08362.1
265	ШТАММ EDMONSTON ВИРУСА КОРИ	GVSYNIGSQEWYTTV	1148	P69353.1
266	ШТАММ EDMONSTON ВИРУСА КОРИ	GYNVSSIVTMTSQGM	1149	P08362.1
267	МОРБИЛЛИВИРУС	HFCVLADSESGGHIT	1150	P06830.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОРИ			
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	ВИРУС КРАСНУХИ STRAIN M33	HGPDWASPVCQRHSPDCSRLVG ATPERPRLRLV	1155	CAJ88851.1
273	ШТАММ EDMONSTON ВИРУСА КОРИ	HITHSGMEGMGVSCT	1156	P08362.1
274	МОРБИЛЛИВИРУС КОРИ	HKSLSTNLDVTNSIE	1157	P06830.1
275	ШТАММ EDMONSTON ВИРУСА КОРИ	HLMIDRPYV	1158	P08362.1
276	ШТАММ EDMONSTON ВИРУСА КОРИ	HLPTGTPLDIDTA	1159	P04851.1
277	ШТАММ EDMONSTON ВИРУСА КОРИ-В	HLPTGTPLDIDTATESSQDPQDS R	1160	Q77M43.1
278	МОРБИЛЛИВИРУС КОРИ	HMTNYLEQPVSNDLS	1161	P06830.1
279	ШТАММ EDMONSTON ВИРУСА КОРИ	HQSLVIKLMPNITLL	1162	P69353.1
280	МОРБИЛЛИВИРУС КОРИ	HRAAIYTAEIHKSLS	1163	P06830.1
281	ПАЛОЧКА КОКЛЮША	HRMQEAVEAERAGRGTGH	1164	P04977.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
282	ШТАММ EDMONSTON ВИРУСА КОРИ	HVDTESYFIVLSIAY	1165	P69353.1
283	ШТАММ EDMONSTON ВИРУСА КОРИ	HWGNLSKIGVVGIGS	1166	P69353.1
284	ВИРУС КРАСНУХИ	HWWQLTLGATCALPL	1167	BAA28178.1
285	ВИРУС КРАСНУХИ	HYRNASDVLPGHWLQGGWGCYNL	1168	NP_740663.1
286	ШТАММ EDMONSTON ВИРУСА КОРИ	IDLGPPISLERLDVG	1169	P69353.1
287	ШТАММ EDMONSTON ВИРУСА КОРИ	IEAIRQAGQEMILAV	1170	P69353.1
288	ВИРУС КРАСНУХИ STRAIN M33	IETRSARHP	1171	CAA28880.1
289	ШТАММ EDMONSTON ВИРУСА КОРИ	IGSQEWYTTVPKYVA	1172	P69353.1
290	ШТАММ EDMONSTON ВИРУСА КОРИ	IKGVIVHRLEGVSYN	1173	P69353.1
291	ШТАММ EDMONSTON ВИРУСА КОРИ	IKHIIIVPIPGDSSI	1174	P04851.1
292	ШТАММ EDMONSTON ВИРУСА КОРИ	IKHIIIVPIPGDSSITTRSR	1175	P04851.1
293	ПАЛОЧКА КОКЛЮША	IKLKDCP	1176	ACI16083.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
294	ШТАММ EDMONSTON ВИРУСА КОРИ	IKLMPNITLLNNCTR	1177	P69353.1
295	МОРБИЛЛИВИРУС КОРИ	ILLERLDVGT	1178	AAF85664.1
296	МОРБИЛЛИВИРУС КОРИ	ILPGQDLQYV	1179	P08362.1
297	ШТАММ EDMONSTON ВИРУСА КОРИ	ILTYIAADHCPVVEV	1180	P69353.1
298	МОРБИЛЛИВИРУС КОРИ	INQDPDKILTY	1181	AAL29688.1
299	ШТАММ EDMONSTON ВИРУСА КОРИ	IPRFKVSPYLFNVPI	1182	P08362.1
300	ШТАММ EDMONSTON ВИРУСА КОРИ	IQALSYALGGDINKV	1183	P69353.1
301	ВИРУС КРАСНУХИ	IRAGPYGHATVEMPE	1184	BAA28178.1
302	ПАЛОЧКА КОКЛЮША	IRMGTDK	1185	ACI16088.1
303	ШТАММ EDMONSTON ВИРУСА КОРИ	ISNFDESSCTFMPEG	1186	P69353.1
304	ШТАММ EDMONSTON ВИРУСА КОРИ	ITAGIALHQSMLNSQ	1187	P69353.1
305	ШТАММ EDMONSTON ВИРУСА КОРИ	ITDDPDVSIRLLEVV	1188	P04851.1
306	ШТАММ	ITDDPDVSIRLLEVVQSDQS	1189	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
307	ПАЛОЧКА КОКЛЮША	ITTYV	1190	ACI16083.1
308	ШТАММ EDMONSTON ВИРУСА КОРИ	IVEAGLASFILTIKF	1191	P04851.1
309	ШТАММ EDMONSTON ВИРУСА КОРИ	IVEAGLASFILTIKFGIETM	1192	P04851.1
310	ШТАММ EDMONSTON ВИРУСА КОРИ	IVEAGLASFILTIKFGIETMYPALG	1193	P04851.1
311	ВИРУС КРАСНУХИ	KALEAQSRALRAELAA	1194	P07566.1
312	ШТАММ EDMONSTON ВИРУСА КОРИ	KARITHVDTESYFIV	1195	P69353.1
313	ВИРУС КРАСНУХИ	KCGLHIRAGPYGHAT	1196	BAA28178.1
314	ШТАММ EDMONSTON ВИРУСА КОРИ	KCYTTGTIINQDPDK	1197	P69353.1
315	ШТАММ EDMONSTON ВИРУСА КОРИ	KDNPHPKGSR	1198	P08362.1
316	ШТАММ EDMONSTON ВИРУСА КОРИ	KDNRIPSYGVLSVDL	1199	P08362.1
317	МОРБИЛЛИВИРУС КОРИ	KFLNPDREYDFRDLT	1200	AAC35876.2
318	ВИРУС КРАСНУХИ	KFVTAALLN	1201	BAA28178.1
319	ШТАММ	KGNCSGPTTIR	1202	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
320	МОРБИЛЛИВИРУС КОРИ	KIKFLNPDREYDFRD	1203	P06830.1
321	ВИРУС КРАСНУХИ	KIVDGGCFAPWDLEA	1204	BAA28178.1
322	ШТАММ EDMONSTON ВИРУСА КОРИ	KLGLKLLRYYTEILS	1205	P69353.1
323	ШТАММ EDMONSTON ВИРУСА КОРИ	KLGVWKSPTDMQSWV	1206	P08362.1
324	ПАЛОЧКА КОКЛЮША	KLKECPQ	1207	ACI16088.1
325	МОРБИЛЛИВИРУС КОРИ	KLLRYYTEI	1208	P26031.1
326	ШТАММ EDMONSTON ВИРУСА КОРИ	KLMPFSGDFV	1209	ABK40531.1
327	МОРБИЛЛИВИРУС КОРИ	KLMPNITLL	1210	P26031.1
328	МОРБИЛЛИВИРУС КОРИ	KLRMETCFQQACKGKIQALCEN PEWA	1211	AAR89413.1
329	МОРБИЛЛИВИРУС КОРИ	KLWCRHFCV	1212	P08362.1
330	ШТАММ EDMONSTON ВИРУСА КОРИ	KLWCRHFCVL	1213	P08362.1
331	МОРБИЛЛИВИРУС КОРИ	KLWESPQEI	1214	BAB60863.1
332	ШТАММ EDMONSTON	KMSSAVGFV	1215	ABO69699.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
333	ШТАММ EDMONSTON ВИРУСА КОРИ-В	KMSSAVGFVPDTGPASR	1216	P03422.1
334	ПАЛОЧКА КОКЛЮША	KMVYATN	1217	ACI16083.1
335	ШТАММ EDMONSTON ВИРУСА КОРИ	KPDLTGTSKSYVRSL	1218	P69353.1
336	МОРБИЛЛИВИРУС КОРИ	KPNLSSKRSELSQLS	1219	P08362.1
337	ШТАММ EDMONSTON ВИРУСА КОРИ	KPNLSSKRSELSQLSMYRVF	1220	P08362.1
338	ШТАММ EDMONSTON ВИРУСА КОРИ	KQSRGEARESYRETG	1221	P04851.1
339	ШТАММ EDMONSTON ВИРУСА КОРИ	KQSRGEARESYRETGPSRAS	1222	P04851.1
340	ШТАММ EDMONSTON ВИРУСА КОРИ	KRFAGVVLAGAALGV	1223	P69353.1
341	ШТАММ EDMONSTON ВИРУСА КОРИ	KRTPGNKPRIAEMIC	1224	P04851.1
342	ШТАММ EDMONSTON ВИРУСА КОРИ	KRTPGNKPRIAEMICDIDTY	1225	P04851.1
343	МОРБИЛЛИВИРУС КОРИ	KSNHNNVYWLTIPPMKNLALGVI NTL	1226	AAR89413.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
344	МОРБИЛЛИВИРУС КОРИ	KVSPYLFNV	1227	P08362.1
345	ПАЛОЧКА КОКЛЮША	KVVQLPKISKNALKANG	1228	ACI16083.1
346	ПАЛОЧКА КОКЛЮША	KVVQLPKISKNALRNDG	1229	ACI16087.1
347	ПАЛОЧКА КОКЛЮША	LAHRRIPPENIR	1230	P04977.1
348	ПАЛОЧКА КОКЛЮША	LALALWAGFALS	1231	P11092.1
349	ВИРУС КРАСНУХИ	LAPGGGNCHLTVNGE	1232	BAA28178.1
350	ШТАММ EDMONSTON ВИРУСА КОРИ	LAQIWVLLAKAVTAP	1233	P04851.1
351	ШТАММ EDMONSTON ВИРУСА КОРИ	LAQIWVLLAKAVTAPDTAAD	1234	P04851.1
352	ВИРУС КРАСНУХИ	LASYFNPGGSYYKQYHPTACEVE PAFGHS	1235	BAA19893.1
353	МОРБИЛЛИВИРУС КОРИ	LAVSKGNCSGPTTIR	1236	P06830.1
354	ШТАММ EDMONSTON ВИРУСА КОРИ	LCENPEWAPLKDNRI	1237	P08362.1
355	МОРБИЛЛИВИРУС КОРИ	LDRLVRLIG	1238	ABI54110.1
356	ДИФТЕРИЙНАЯ ПАЛОЧКА	LEEEGVTPL	1239	P33120.2
357	ПАЛОЧКА КОКЛЮША	LEHRMQEAVEAERAGRGTGHFI	1240	CAD44970.1
358	ШТАММ	LEKLGYSGGDLLGIL	1241	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
359	ШТАММ EDMONSTON ВИРУСА КОРИ	LEQPVSNDLS	1242	P08362.1
360	ШТАММ EDMONSTON ВИРУСА КОРИ	LERKWLDVVRNIIAE	1243	P04851.1
361	ШТАММ EDMONSTON ВИРУСА КОРИ	LERKWLDVVRNIIAEDLSLR	1244	P04851.1
362	ШТАММ EDMONSTON ВИРУСА КОРИ	LFGPSLRDPISAEIS	1245	P69353.1
363	МОРБИЛЛИВИРУС КОРИ	LGELKLAALCHGEDS	1246	P06830.1
364	ШТАММ EDMONSTON ВИРУСА КОРИ	LGGKEDRRVKQSR	1247	P04851.1
365	ВИРУС КРАСНУХИ	LGHDGHHGGTLRVGQHHRNASDVL	1248	ABD64214.1
366	ВИРУС КРАСНУХИ STRAIN THERIEN	LGSPNCHGPDWASPVCQRHS	1249	P07566.1
367	ВИРУС КРАСНУХИ STRAIN THERIEN	LGSPNCHGPDWASPVCQRHSPDCSRLV	1250	P07566.1
368	ВИРУС КРАСНУХИ	LHDPDTEAPTEACVTSWL	1251	ABD64214.1
369	ШТАММ EDMONSTON ВИРУСА КОРИ	LIANCASILCKCYTT	1252	P69353.1
370	МОРБИЛЛИВИРУС КОРИ	LIGLLAIAGIRLHRAAIYTAEIHK	1253	AAR89413.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
371	ШТАММ EDMONSTON ВИРУСА КОРИ	LIPSMNQLSCDLIGQ	1254	P69353.1
372	ШТАММ EDMONSTON ВИРУСА КОРИ	LKIKIASGFGPLITH	1255	P08362.1
373	МОРБИЛЛИВИРУС КОРИ	LKIKIASGFGPLITHGSGMDLYK	1256	AAR89413.1
374	ПАЛОЧКА КОКЛЮША	LKLYFEP	1257	ACI16088.1
375	ШТАММ EDMONSTON ВИРУСА КОРИ	LLAVLFVMFL	1258	P08362.1
376	ШТАММ EDMONSTON ВИРУСА КОРИ	LLDRLVRLIGNPDVS	1259	P04851.1
377	ШТАММ EDMONSTON ВИРУСА КОРИ	LLDRLVRLIGNPDVSGPKLT	1260	P04851.1
378	ШТАММ EDMONSTON ВИРУСА КОРИ	LLESSDQILRSMKGL	1261	P69353.1
379	МОРБИЛЛИВИРУС КОРИ	LLETRTTNQFLAVSK	1262	P06830.1
380	ШТАММ EDMONSTON ВИРУСА КОРИ	LLEVVQSDQSQSGLT	1263	P04851.1
381	ШТАММ EDMONSTON ВИРУСА КОРИ	LLEVVQSDQSQSGLTFASR	1264	P04851.1
382	ШТАММ	LLEVVQSDQSQSGLTFASRG	1265	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
383	МОРБИЛЛИВИРУС КОРИ	LLGILESRGIKARIT	1266	AAL29688.1
384	ВИРУС КРАСНУХИ	LLRTAPGP	1267	CAJ88851.1
385	ШТАММ EDMONSTON ВИРУСА КОРИ	LLRYYTEILSLFGPS	1268	P69353.1
386	ВИРУС КРАСНУХИ STRAIN THERIEN	LLVPWVLIFMVCRRACRRRG	1269	P07566.1
387	ШТАММ EDMONSTON ВИРУСА КОРИ	LLWRSRCKIV	1270	ABK40528.1
388	МОРБИЛЛИВИРУС КОРИ	LLWSYAMGV	1271	P04851.1
389	ШТАММ EDMONSTON ВИРУСА КОРИ	LLWSYAMGVGVELEN	1272	P04851.1
390	ШТАММ EDMONSTON ВИРУСА КОРИ	LLWSYAMGVGVELENSMGGL	1273	P04851.1
391	МОРБИЛЛИВИРУС КОРИ	LMIDRPYVL	1274	P08362.1
392	ШТАММ EDMONSTON ВИРУСА КОРИ	LNAMTQNIRPVQSVA	1275	P69353.1
393	ВИРУС КРАСНУХИ	LNTPPPYQVSCGGES	1276	BAA28178.1
394	ШТАММ EDMONSTON ВИРУСА КОРИ	LRDPISAEISIQALS	1277	P69353.1
395	ПАЛОЧКА	LRGSGDLQEYLRHVTR	1278	AAA83981.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
396	ВИРУС КРАСНУХИ	LRLVDADD	1279	CAJ88851.1
397	ВИРУС КРАСНУХИ	LRLVDADDPLLR	1280	BAA19893.1
398	ВИРУС КРАСНУХИ	LRLVDADDPLLRTAPGPGEVWVTPVIGSQAR	1281	BAA19893.1
399	ПАЛОЧКА КОКЛЮША	LRRLLYMIYMSGLAVRVHVSKEEQYYDY	1282	P04979.1
400	ВИРУС КРАСНУХИ	LRTAPGPG	1283	CAJ88851.1
401	ВИРУС КРАСНУХИ	LRVGQHYRNASDVLPGHWLQ	1284	NP_740663.1
402	ПАЛОЧКА КОКЛЮША	LRYLA	1285	ACI16088.1
403	ШТАММ EDMONSTON ВИРУСА КОРИ	LSCKPWQESRKNK	1286	P04851.1
404	МОРБИЛЛИВИРУС КОРИ	LSEIKGVIVHRLEGV	1287	AAL29688.1
405	ШТАММ EDMONSTON ВИРУСА КОРИ	LSIAYPTLSEIKGVI	1288	P69353.1
406	ШТАММ EDMONSTON ВИРУСА КОРИ	LSLLDLYLGRGYNVS	1289	P08362.1
407	ШТАММ EDMONSTON ВИРУСА КОРИ	LSQGNLIANCASILC	1290	P69353.1
408	МОРБИЛЛИВИРУС КОРИ	LSSHRGVIADNQAKW	1291	P06830.1
409	ШТАММ EDMONSTON ВИРУСА КОРИ	LSVDLSLTVELKIKI	1292	P08362.1
410	ПАЛОЧКА	LTGISICNPGSSLC	1293	AAA83981.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
411	ШТАММ EDMONSTON ВИРУСА КОРИ	LTIKFGIETMYPALG	1294	P04851.1
412	ШТАММ EDMONSTON ВИРУСА КОРИ	LTIKFGIETMYPALGLHEFA	1295	P04851.1
413	ШТАММ EDMONSTON ВИРУСА КОРИ	LTLQTPTGQIHWGNL	1296	P69353.1
414	ВИРУС КРАСНУХИ	LVDADDPL	1297	CAJ88851.1
415	ВИРУС КРАСНУХИ	LVDADDPLLR	1298	BAA19893.1
416	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ-В	LVSGSFGNRFILSQGNLI	1303	P26031.1
421	ШТАММ EDMONSTON ВИРУСА КОРИ	LYKSNHNNVYWLTIP	1304	P08362.1
422	ШТАММ EDMONSTON ВИРУСА КОРИ	LYPMSPLLQECLRGSTKSCARTLVS	1305	P69353.1
423	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEA	1306	P07566.1
424	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSR	1307	ABD64200.1
425	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSRAL	1308	P07566.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	STRAIN THERIEN	RAELAA		
426	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSRAL RAGLAA	1309	ABD64200.1
427	ВИРУС КРАСНУХИ VACCINE STRAIN RA27/3	MASTTPITMEDLQKALETQSRVL RAGLTA	1310	CAA33016.1
428	ШТАММ EDMONSTON ВИРУСА КОРИ	MATLLRSLALFKRNK	1311	P04851.1
429	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ	MGLKVNVSAIFMAVL	1316	P69353.1
434	МОРБИЛЛИВИРУС КОРИ	MIDRPYVLLAVLFVM	1317	P06830.1
435	ШТАММ EDMONSTON ВИРУСА КОРИ	MILAVQGVQDYINNE	1318	P69353.1
436	ШТАММ EDMONSTON ВИРУСА КОРИ	MLNSQAIDNLRASLE	1319	P69353.1
437	ШТАММ EDMONSTON ВИРУСА КОРИ	MNALVNSTLLETRTT	1320	P08362.1
438	ВИРУС КРАСНУХИ	MNYTGNQQSRWGLGSPNCHGP	1321	BAA19893.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
		DWASPVCQRHS		
439	ШТАММ EDMONSTON ВИРУСА КОРИ	MQSWVPLSTDDPVID	1322	P08362.1
440	МОРБИЛЛИВИРУС КОРИ	MSLSLLDLYLGRGYN	1323	P06830.1
441	ШТАММ EDMONSTON ВИРУСА КОРИ	MSPLLQECLRGSTKS	1324	P69353.1
442	МОРБИЛЛИВИРУС КОРИ	MSPQRDRINAFYKDN	1325	P06830.1
443	МОРБИЛЛИВИРУС КОРИ	MYRVFEVSVIRNPGL	1326	P06830.1
444	ШТАММ EDMONSTON ВИРУСА КОРИ	NALYPMSPLLQECLR	1327	P69353.1
445	ВИРУС КРАСНУХИ STRAIN THERIEN	NCHGPDWASPVCQRHSPDCSRL VGAT	1328	P07566.1
446	ШТАММ EDMONSTON ВИРУСА КОРИ	NFGRSYFDPAYFRLG	1329	P04851.1
447	ШТАММ EDMONSTON ВИРУСА КОРИ	NFGRSYFDPAYFRLGQEMVR	1330	P04851.1
448	ВИРУС КРАСНУХИ	NGTQRACTFWAVNAY	1331	BAA28178.1
449	ШТАММ EDMONSTON ВИРУСА КОРИ	NGVTIQVGSRRYPDA	1332	P69353.1
450	ШТАММ EDMONSTON ВИРУСА КОРИ	NIIAEDLSLRRFMVA	1333	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
451	ШТАММ EDMONSTON ВИРУСА КОРИ	NIIAEDLSLRRFMVALILDI	1334	P04851.1
452	ШТАММ EDMONSTON ВИРУСА КОРИ	NITLLNNCTRVEIAE	1335	P69353.1
453	МОРБИЛЛИВИРУС КОРИ	NLALGVINTLEWIPR	1336	P06830.1
454	ДИФТЕРИЙНАЯ ПАЛОЧКА	NLFQVVHWSYNRPAYSPG	1337	SRC280292
455	ШТАММ EDMONSTON ВИРУСА КОРИ	NLVILPGQDLQYVLA	1338	P08362.1
456	ШТАММ EDMONSTON ВИРУСА КОРИ	NLYQQMGKPAPYMVN	1339	P04851.1
457	ШТАММ EDMONSTON ВИРУСА КОРИ	NLYQQMGKPAPYMVNLENSI	1340	P04851.1
458	ШТАММ EDMONSTON ВИРУСА КОРИ	NNCTRVEIAEYRRLL	1341	P69353.1
459	ШТАММ EDMONSTON ВИРУСА КОРИ	NPDREYDFRD	1342	P08362.1
460	MEASLES VIRUS STRAIN HALLE	NPDVSGPKL	1343	P10050.1
461	ШТАММ EDMONSTON ВИРУСА КОРИ	NPDVSGPKLTGALIG	1344	P04851.1
462	ШТАММ	NPDVSGPKLTGALIGILSLF	1345	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
463	МОРБИЛЛИВИРУС КОРИ	NPPERIKLDYDQYCA	1346	P06830.1
464	МОРБИЛЛИВИРУС КОРИ	NQAKWAVPTTRTDDK	1347	P06830.1
465	МОРБИЛЛИВИРУС КОРИ	NQDPDKILTYIAADH	1348	AAF02706.1
466	ШТАММ EDMONSTON ВИРУСА КОРИ	NQLSCDLIGQKLGLK	1349	P69353.1
467	ВИРУС КРАСНУХИ	NQQSRWGLGSPNCHGPDWASPV CQRHS	1350	ABD64214.1
468	РУБУЛАВИРУС ПАРОТИТА	NSTLGVKSAREF	1351	ABP48111.1
469	ВИРУС КРАСНУХИ	NTPHGQLEVQVPPDP	1352	BAA28178.1
470	ШТАММ EDMONSTON ВИРУСА КОРИ	NVSAIFMAVLLTLQT	1353	P69353.1
471	МОРБИЛЛИВИРУС КОРИ	PAEVDGDVKLSSNLV	1354	P06830.1
472	ВИРУС КРАСНУХИ	PAFGHSDAACWGFPT	1355	BAA28178.1
473	ШТАММ EDMONSTON ВИРУСА КОРИ	PALICCCRGRCNKKG	1356	P69353.1
474	МОРБИЛЛИВИРУС КОРИ	PDKILTYIAADHC	1357	AAF02706.1
475	ШТАММ EDMONSTON ВИРУСА КОРИ	PERIKLDYDQYCADV	1358	P08362.1
476	ВИРУС КРАСНУХИ	PERPRLRL	1359	CAJ88851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
477	ВИРУС КРАСНУХИ	PGCATQAPVPVRLAG	1360	BAA28178.1
478	ВИРУС КРАСНУХИ VACCINE STRAIN RA27/3	PGCATQAPVPVRLAGVRFESKIV DGGCFA	1361	CAJ88851.1
479	ВИРУС КРАСНУХИ	PGEVWVTP	1362	CAJ88851.1
480	ВИРУС КРАСНУХИ	PGEVWVTPVIGSQAR	1363	BAA28178.1
481	ДИФТЕРИЙНАЯ ПАЛОЧКА	PGKLDVNKSKTHISVN	1364	CAE11230.1
482	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ-В	PISLERLDVG	1369	P26031.1
487	ШТАММ EDMONSTON ВИРУСА КОРИ	PISLERLDVGTNLGN	1370	P69353.1
488	ПАЛОЧКА КОКЛЮША	PKALFTQQGGAYGRC	1371	P04979.1
489	ШТАММ EDMONSTON ВИРУСА КОРИ	PKYVATQGYLISNFD	1372	P69353.1
490	ШТАММ EDMONSTON ВИРУСА КОРИ	PLDIDTASESSQD	1373	P04851.1
491	ВИРУС КРАСНУХИ	PLGLKFKTVRPVALP	1374	BAA28178.1
492	ШТАММ	PLITHGSGMDLYKSN	1375	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
493	МОРБИЛЛИВИРУС КОРИ	PLKDNRIPSYGVLSV	1376	P06830.1
494	ВИРУС КРАСНУХИ	PLLRTAPG	1377	CAJ88851.1
495	ШТАММ EDMONSTON ВИРУСА КОРИ	PLLSCKPWQESRK	1378	P04851.1
496	ПАЛОЧКА КОКЛЮША	PPATVYRYDSRPPE	1379	P04977.1
497	МОРБИЛЛИВИРУС КОРИ	PPISLERLDVGT	1380	AAL29688.1
498	ВИРУС КРАСНУХИ	PPPPEERQETRSQTPAPKPS	1381	P07566.1
499	ПАЛОЧКА КОКЛЮША	PQEQITQHGSPYGRC	1382	AAA83981.1
500	ПАЛОЧКА КОКЛЮША	PQEQITQHGSPYGRCANK	1383	AAA83981.1
501	ПАЛОЧКА КОКЛЮША	PQPGPQPPQPPQPQPEAPAPQP PAG	1384	P14283.3
502	ШТАММ EDMONSTON ВИРУСА КОРИ	PRLGGKEDRRVKQ	1385	P04851.1
503	ВИРУС КРАСНУХИ	PRLRLVDA	1386	CAJ88851.1
504	ВИРУС КРАСНУХИ	PRNVRVTGCYQCGTP	1387	BAA28178.1
505	ШТАММ EDMONSTON ВИРУСА КОРИ	PSRASDARAAHLP	1388	P04851.1
506	ШТАММ EDMONSTON ВИРУСА КОРИ	PTGQIHWGNLSKIGV	1389	P69353.1
507	ШТАММ	PTGTPLDIDTASE	1390	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
508	ШТАММ EDMONSTON ВИРУСА КОРИ	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	ДИФТЕРИЙНАЯ ПАЛОЧКА	PVFAGANYAAWAVNVAQVI	1396	AAV70486.1
514	ВИРУС КРАСНУХИ	PVIGSQAR	1397	CAJ88851.1
515	МОРБИЛЛИВИРУС КОРИ	PVVEVNGVTIQVGSR	1398	AAL29688.1
516	ВИРУС КРАСНУХИ STRAIN THERIEN	PWELVVLTARPEDGWTCRGV	1399	P07566.1
517	ШТАММ EDMONSTON ВИРУСА КОРИ	PWQESRKNKAQTR	1400	P04851.1
518	ШТАММ EDMONSTON ВИРУСА КОРИ	PYMVNLENSIQNKFS	1401	P04851.1
519	ШТАММ EDMONSTON ВИРУСА КОРИ	PYMVNLENSIQNKFSAGSYP	1402	P04851.1
520	ШТАММ EDMONSTON ВИРУСА КОРИ	PYQGSGKGVS	1403	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
521	ВИРУС КРАСНУХИ	PYQVSCGGESDRASA	1404	BAA28178.1
522	МОРБИЛЛИВИРУС КОРИ	QACKGKIQALCEN	1405	P08362.1
523	ШТАММ EDMONSTON ВИРУСА КОРИ	QAGQEMILAVQGVQD	1406	P69353.1
524	МОРБИЛЛИВИРУС КОРИ	QALCENPECVPLKDN	1407	P06830.1
525	ПАЛОЧКА КОКЛЮША	QALGALK	1408	ACI16088.1
526	ВИРУС КРАСНУХИ	QAPVPVRLAGVRFES	1409	BAA28178.1
527	ВИРУС КРАСНУХИ	QCGTPALVEGLAPGG	1410	BAA28178.1
528	ШТАММ EDMONSTON ВИРУСА КОРИ	QDPDKILTYIAADHC	1411	P69353.1
529	ШТАММ EDMONSTON ВИРУСА КОРИ	QECLRGSTKSCARTL	1412	P69353.1
530	ПАЛОЧКА КОКЛЮША	QEQITQHGSPYGRC	1413	AAA83981.1
531	ШТАММ EDMONSTON ВИРУСА КОРИ	QESRKNKAQTRTP	1414	P04851.1
532	ШТАММ EDMONSTON ВИРУСА КОРИ	QGDQSENELPRLGGK	1415	P04851.1
533	ШТАММ EDMONSTON ВИРУСА КОРИ	QGDQSENELPRLGGKEDRRV	1416	P04851.1
534	ДИФТЕРИЙНАЯ ПАЛОЧКА	QGESGHDIKITAENTPLPIA	1417	AAV70486.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
535	МОРБИЛЛИВИРУС КОРИ	QGSGKGVSFQLVKLG	1418	P06830.1
536	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ	QLSMYRVFEV	1423	P08362.1
541	ПАЛОЧКА КОКЛЮША	QLSNIT	1424	ACI16083.1
542	ШТАММ EDMONSTON ВИРУСА КОРИ	QNIRPVQSVASSRRH	1425	P69353.1
543	ШТАММ EDMONSTON ВИРУСА КОРИ	QNKFSAGSYPLLWSY	1426	P04851.1
544	ШТАММ EDMONSTON ВИРУСА КОРИ	QNKFSAGSYPLLWSYAMGVG	1427	P04851.1
545	ШТАММ EDMONSTON ВИРУСА КОРИ	QNKFSAGSYPLLWSYAMGVGVELEN	1428	P04851.1
546	ШТАММ EDMONSTON ВИРУСА КОРИ	QQACKGKIQALCENP	1429	P08362.1
547	ШТАММ	QQRRVVGEFRLERKW	1430	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
548	ШТАММ EDMONSTON ВИРУСА КОРИ	QQRRVVGEFRLERKWLDVVR	1431	P04851.1
549	ПАЛОЧКА КОКЛЮША	QQTRANPNPYTSRRSVAS	1432	P04977.1
550	ВИРУС КРАСНУХИ	QRHSPDCSRLVGATP	1433	BAA28178.1
551	ШТАММ EDMONSTON ВИРУСА КОРИ	QSGLTFASRGTNMED	1434	P04851.1
552	ШТАММ EDMONSTON ВИРУСА КОРИ	QSGLTFASRGTNMEDEADQY	1435	P04851.1
553	ДИФТЕРИЙНАЯ ПАЛОЧКА	QSIALSSLMVAQAIPLVGEL	1436	AAV70486.1
554	ШТАММ EDMONSTON ВИРУСА КОРИ	QSRFGWFENKEISDI	1437	P04851.1
555	ШТАММ EDMONSTON ВИРУСА КОРИ	QSRFGWFENKEISDIEVQDP	1438	P04851.1
556	ШТАММ EDMONSTON ВИРУСА КОРИ	QSRGEAR	1439	P04851.1
557	ШТАММ EDMONSTON ВИРУСА КОРИ	QSRGEARESYRETGPSRA	1440	P04851.1
558	ВИРУС КРАСНУХИ STRAIN THERIEN	QTGRGGSAPRPELGPPTN	1441	P07566.1
559	ВИРУС КРАСНУХИ	QTPAPKPSRAPPQQPQPPRMQT	1442	P07566.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	STRAIN THERIEN	GRG		
560	ШТАММ EDMONSTON ВИРУСА КОРИ	QVGSRRYPDAVYLHR	1443	P69353.1
561	ШТАММ EDMONSTON ВИРУСА КОРИ	QVSFLQGDQSENE	1444	P04851.1
562	ВИРУС КРАСНУХИ	QYHPTACEVEPAFGH	1445	BAA28178.1
563	ШТАММ EDMONSTON ВИРУСА КОРИ	QYVLATYDTSRVEHA	1446	P08362.1
564	ПАЛОЧКА КОКЛЮША	RANPNPYTSRRSV	1447	ACI04548.1
565	ШТАММ EDMONSTON ВИРУСА КОРИ	RASDARAAHLPTG	1448	P04851.1
566	ШТАММ EDMONSTON ВИРУСА КОРИ	RASLETTNQAIEAIR	1449	P69353.1
567	ВИРУС КРАСНУХИ STRAIN THERIEN	RCGRLICGLSTTAQYPPTRF	1450	P07566.1
568	ПАЛОЧКА КОКЛЮША	RDGQSVIGACASPYEGRYR	1451	P04979.1
569	ШТАММ EDMONSTON ВИРУСА КОРИ	RESYRETGPSRAS	1452	P04851.1
570	ШТАММ EDMONSTON ВИРУСА КОРИ	RETGPSRASDARA	1453	P04851.1
571	РУБУЛАВИРУС ПАРОТИТА	RFAKYQQQGRLEAR	1454	P21186.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
572	ВИРУС КРАСНУХИ STRAIN THERIEN	RFGAPQAFLAGLLLATVAVGTARA	1455	P07566.1
573	ШТАММ EDMONSTON ВИРУСА КОРИ	RFMVALILDIKRTPG	1456	P04851.1
574	ШТАММ EDMONSTON ВИРУСА КОРИ	RFMVALILDIKRTPGNKPRI	1457	P04851.1
575	ШТАММ EDMONSTON ВИРУСА КОРИ	RGEARESYRETGP	1458	P04851.1
576	ВИРУС КРАСНУХИ	RGTTPPAYG	1459	CAA28880.1
577	ВИРУС КРАСНУХИ STRAIN M33	RIETRSARH	1460	ABD64214.1
578	ПАЛОЧКА КОКЛЮША	RILAGALATYQ	1461	P04977.1
579	ПАЛОЧКА КОКЛЮША	RIPPENIRRVT	1462	ACI04548.1
580	ПАЛОЧКА КОКЛЮША	RISNLND	1463	ACI16083.1
581	ПАЛОЧКА КОКЛЮША	RLANLNG	1464	ACI16088.1
582	ШТАММ EDMONSTON ВИРУСА КОРИ	RLDVGTNLGNAIAKL	1465	P69353.1
583	ШТАММ EDMONSTON ВИРУСА КОРИ	RLERKWLDV	1466	P04851.1
584	ШТАММ EDMONSTON ВИРУСА КОРИ	RLGGKEDRRVKQSRG	1467	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
585	ШТАММ EDMONSTON ВИРУСА КОРИ	RLGGKEDRRVKQSRGEARES	1468	P04851.1
586	ШТАММ EDMONSTON ВИРУСА КОРИ	RLLDRLVRL	1469	ABI54110.1
587	ВИРУС КРАСНУХИ	RLRLVDAD	1470	CAJ88851.1
588	ВИРУС КРАСНУХИ	RLRLVDADDPLLR	1471	BAA19893.1
589	ВИРУС КРАСНУХИ	RLRLVDADDPLLRTAPGPGEVWV TPVIGSQA	1472	BAA19893.1
590	ВИРУС КРАСНУХИ	RLRLVQDADDPLLRIAPGPGEVW VTPVIGSQA	1473	SRC265968
591	ШТАММ EDMONSTON ВИРУСА КОРИ	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	RMQTGRGGSAPRPELGPPTNPF QAAVA	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	ПАЛОЧКА КОКЛЮША	RQAESSEAMAAWSERAGEA	1482	P04977.1
600	ВИРУС КРАСНУХИ	RQTWAEWAAAHWWQL	1483	BAA28178.1
601	ШТАММ EDMONSTON ВИРУСА КОРИ	RRVKQSRGEARES	1484	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
602	МОРБИЛЛИВИРУС КОРИ	RRYPDAVYL	1485	ACA09725.1
603	ШТАММ EDMONSTON ВИРУСА КОРИ	RRYPDAVYLHRIDLG	1486	P69353.1
604	ШТАММ EDMONSTON ВИРУСА КОРИ	RSAGKVSSTLASELG	1487	P04851.1
605	ШТАММ EDMONSTON ВИРУСА КОРИ	RSAGKVSSTLASELGITAED	1488	P04851.1
606	ШТАММ EDMONSTON ВИРУСА КОРИ	RSELSQLS	1489	P08362.1
607	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ	RTVLEPIRDALNAMT	1496	P69353.1
614	ШТАММ EDMONSTON ВИРУСА КОРИ	RVEHAVVYYVYSPSR	1497	P08362.1
615	ШТАММ	RVFEVGVIRNPGLGA	1498	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
616	ПАЛОЧКА КОКЛЮША	RVHVSKEEQYYDYEDATFE	1499	P04978.2
617	ВИРУС КРАСНУХИ	RVIDPAAQ	1500	BAA28178.1
618	ВИРУС КРАСНУХИ	RVKFHTETRTVWQLS	1501	BAA28178.1
619	ПАЛОЧКА КОКЛЮША	RVYHNGITGET	1502	ACI04548.1
620	ПАЛОЧКА КОКЛЮША	RYDSRPPEDVF	1503	ACI04548.1
621	ШТАММ EDMONSTON ВИРУСА КОРИ	RYPDAVYLHRIDLGP	1504	P69353.1
622	ШТАММ EDMONSTON ВИРУСА КОРИ	SAEISIQALSYALGG	1505	P69353.1
623	ШТАММ EDMONSTON ВИРУСА КОРИ	SAEPLLSCKPWQESR	1506	P04851.1
624	ШТАММ EDMONSTON ВИРУСА КОРИ	SAEPLLSCKPWQESRKNKAQ	1507	P04851.1
625	МОРБИЛЛИВИРУС КОРИ	SAGKVSSTLASELG	1508	P04851.1
626	ШТАММ EDMONSTON ВИРУСА КОРИ	SAGKVSSTLASELGITAEDARLVS	1509	P04851.1
627	ШТАММ EDMONSTON ВИРУСА КОРИ	SCTVTREDGT	1510	P08362.1
628	ВИРУС КРАСНУХИ	SDAACWGFPTDTVMS	1511	BAA28178.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
629	ШТАММ EDMONSTON ВИРУСА КОРИ	SDARAAHLPTGTP	1512	P04851.1
630	ВИРУС КРАСНУХИ STRAIN THERIEN	SDWHQGTHVCHTKHMDFWCVEHD	1513	P07566.1
631	ШТАММ EDMONSTON ВИРУСА КОРИ	SELRRWIKYTQQRRV	1514	P04851.1
632	ШТАММ EDMONSTON ВИРУСА КОРИ	SELRRWIKYTQQRRVVGEFR	1515	P04851.1
633	ШТАММ EDMONSTON ВИРУСА КОРИ	SELSQL	1516	P08362.1
634	ШТАММ EDMONSTON ВИРУСА КОРИ	SELSQLS	1517	P08362.1
635	ПАЛОЧКА КОКЛЮША	SEYLAHRRIPPENIRRVTRV	1518	CAD44970.1
636	ШТАММ EDMONSTON ВИРУСА КОРИ	SFLQGDQSENELP	1519	P04851.1
637	ШТАММ EDMONSTON ВИРУСА КОРИ	SGKGVSFQLVKLGVW	1520	P08362.1
638	ШТАММ EDMONSTON ВИРУСА КОРИ	SHRGVIADNQAKWAV	1521	P08362.1
639	ШТАММ EDMONSTON ВИРУСА КОРИ	SIEHQVKDVLTPLFK	1522	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
640	ШТАММ EDMONSTON ВИРУСА КОРИ	SKIGVVGIGSASYKV	1523	P69353.1
641	МОРБИЛЛИВИРУС КОРИ	SKRSELSQLSMYRVF	1524	P06830.1
642	МОРБИЛЛИВИРУС КОРИ	SLFVESPGQLIQRITDDPDVS	1525	ABI54110.1
643	ШТАММ EDMONSTON ВИРУСА КОРИ	SLSTNLDVTNSIEHQ	1526	P08362.1
644	ШТАММ EDMONSTON ВИРУСА КОРИ	SLSTNLDVTNSIEHQVKDVLTPLFK	1527	P08362.1
645	ШТАММ EDMONSTON ВИРУСА КОРИ	SLWGSGLLML	1528	BAE98296.1
646	ШТАММ EDMONSTON ВИРУСА КОРИ	SMKGLSSTSIVYILI	1529	P69353.1
647	МОРБИЛЛИВИРУС КОРИ	SMYRVFEVGV	1530	P08362.1
648	МОРБИЛЛИВИРУС КОРИ	SNDLSNCMVALGELK	1531	P06830.1
649	MEASLES VIRUS STRAIN HALLE	SPGQLIQR	1532	P10050.1
650	ШТАММ EDMONSTON ВИРУСА КОРИ	SQDPQDSRRSAEP	1533	P04851.1
651	МОРБИЛЛИВИРУС КОРИ	SRIVINREHLMIDRP	1534	P06830.1
652	ШТАММ	SRKNKAQTRTPLQ	1535	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
653	ВИРУС КРАСНУХИ	SRLVGATP	1536	CAJ88851.1
654	ВИРУС КРАСНУХИ	SRLVGATPERPRLRLVDADDPLLR	1537	CAJ88851.1
655	ШТАММ EDMONSTON ВИРУСА КОРИ	SRPGLKPDLTGTSKS	1538	P69353.1
656	ПАЛОЧКА КОКЛЮША	SRRSVASIVGTLVRM	1539	CAD44970.1
657	ВИРУС КРАСНУХИ	SRWGLGSPNCHGPDW	1540	BAA28178.1
658	ПАЛОЧКА КОКЛЮША	SSATK	1541	ACI16088.1
659	ВИРУС КРАСНУХИ	SSGGYAQLASYFNPG	1542	BAA28178.1
660	ПАЛОЧКА КОКЛЮША	SSLGNGV	1543	ACI16083.1
661	МОРБИЛЛИВИРУС КОРИ	SSNLVILPGQDLQYV	1544	P06830.1
662	ШТАММ EDMONSTON ВИРУСА КОРИ	SSQDPQDSRRSAEPL	1545	P04851.1
663	ШТАММ EDMONSTON ВИРУСА КОРИ	SSQDPQDSRRSAEPLLSCKP	1546	P04851.1
664	ШТАММ EDMONSTON ВИРУСА КОРИ	SSRRHKRFAGVVLAG	1547	P69353.1
665	ШТАММ EDMONSTON ВИРУСА КОРИ	SSTSIVYILIAVCLG	1548	P69353.1
666	ПАЛОЧКА	STPGIVI	1549	AAA83981.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
667	ПАЛОЧКА КОКЛЮША	STPGIVIPPQEQITQHGSPYGRC	1550	AAA83981.1
668	ПАЛОЧКА КОКЛЮША	STSSSRRYTEVY	1551	P04977.1
669	ШТАММ EDMONSTON ВИРУСА КОРИ	SYFIVLSIAYPTLSE	1552	P69353.1
670	ШТАММ EDMONSTON ВИРУСА КОРИ	SYRETGPSRASDA	1553	P04851.1
671	ПАЛОЧКА КОКЛЮША	SYVK	1554	ACI16083.1
672	ВИРУС КРАСНУХИ	SYVQHPHKTVRVKFH	1555	BAA28178.1
673	ВИРУС КРАСНУХИ	TAPGPGEV	1556	CAJ88851.1
674	ПАЛОЧКА КОКЛЮША	TATRLLSSTNSRLC	1557	AAA83981.1
675	МОРБИЛЛИВИРУС КОРИ	TDDPVIDRLYLSSHR	1558	P06830.1
676	ШТАММ EDMONSTON ВИРУСА КОРИ	TEILSLFGPSLRDPI	1559	P69353.1
677	ВИРУС КРАСНУХИ	TETRTVWQLSVAGVS	1560	BAA28178.1
678	ПАЛОЧКА КОКЛЮША	TEVYLEHRMQEAVE	1561	P04977.1
679	ШТАММ EDMONSTON ВИРУСА КОРИ	TFMPEGTVCSQNALY	1562	P69353.1
680	ВИРУС КРАСНУХИ	TGACICEIPTDVSCE	1563	BAA28178.1
681	ПАЛОЧКА КОКЛЮША	TGDLRAY	1564	ACI16083.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
682	МОРБИЛЛИВИРУС КОРИ	TGMQSWVPLSTDDPV	1565	P06830.1
683	ВИРУС КРАСНУХИ	TGNQQSRWGLGSPNC	1566	BAA28178.1
684	ШТАММ EDMONSTON ВИРУСА КОРИ	TGPSRASDARAAH	1567	P04851.1
685	МОРБИЛЛИВИРУС КОРИ	TGTIINQDPDKILTY	1568	AAF02706.1
686	ВИРУС КРАСНУХИ	TGVVYGTHTTAVSET	1569	BAA28178.1
687	ШТАММ EDMONSTON ВИРУСА КОРИ	TIRGQFSNMSLSLLD	1570	P08362.1
688	ВИРУС КРАСНУХИ	TLGATCALPLAGLLA	1571	BAA28178.1
689	МОРБИЛЛИВИРУС КОРИ	TLLNNCTRV	1572	P26031.1
690	ШТАММ EDMONSTON ВИРУСА КОРИ-В	TLNVPPPPDPGR	1573	P03422.1
691	ШТАММ EDMONSTON ВИРУСА КОРИ-В	TLNVPPPPDPGRASTSGTPIKK	1574	P03422.1
692	МОРБИЛЛИВИРУС КОРИ	TMTSQGMYGGTYPVE	1575	P06830.1
693	ШТАММ EDMONSTON ВИРУСА КОРИ	TNLGNAIAKLEDAKE	1576	P69353.1
694	ШТАММ EDMONSTON ВИРУСА КОРИ	TNMEDEADQYFSHDD	1577	P04851.1
695	ШТАММ EDMONSTON	TNMEDEADQYFSHDDPISSD	1578	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
696	ВИРУС КРАСНУХИ STRAIN M33	TNPFQAAVARGLRPP	1579	CAA28880.1
697	МОРБИЛЛИВИРУС КОРИ	TNSIEHQVKDVLTPL	1580	P06830.1
698	ШТАММ EDMONSTON ВИРУСА КОРИ	TNYLEQPVSNDLSNC	1581	P08362.1
699	МОРБИЛЛИВИРУС КОРИ	TNYLEQPVSNDLSNCMVALGELK LAAL	1582	AAR89413.1
700	ВИРУС КРАСНУХИ	TPERPRLR	1583	CAJ88851.1
701	ВИРУС КРАСНУХИ STRAIN THERIEN	TPERPRLRLVDADDPLLRTA	1584	P07566.1
702	MEASLES VIRUS STRAIN HALLE	TPGNKPRIA	1585	P10050.1
703	ШТАММ EDMONSTON ВИРУСА КОРИ	TPLDIDTASESSQDP	1586	P04851.1
704	ШТАММ EDMONSTON ВИРУСА КОРИ	TPLDIDTASESSQDPQDSRR	1587	P04851.1
705	ШТАММ EDMONSTON ВИРУСА КОРИ	TPLFKIIGDEVGLRT	1588	P08362.1
706	ШТАММ EDMONSTON ВИРУСА КОРИ	TPLQCTM	1589	P04851.1
707	ВИРУС КРАСНУХИ	TPVIGSQA	1590	CAJ88851.1
708	ВИРУС КРАСНУХИ	TPVIGSQARK	1591	BAA19893.1
709	ШТАММ EDMONSTON	TQGYLISNFDESSCT	1592	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
710	ПАЛОЧКА КОКЛЮША	TRANPNPYTSRRSVASIVGTLVRM	1593	P04977.1
711	ВИРУС КРАСНУХИ	TRFGCAMRWGLPP	1594	NP_740663.1
712	ПАЛОЧКА КОКЛЮША	TRNTGQPATDHYYSNV	1595	AAA83981.1
713	ШТАММ EDMONSTON ВИРУСА КОРИ	TRTPLQCTMTEIF	1596	P04851.1
714	ВИРУС КРАСНУХИ	TRWHRLLRMPVR	1597	ABD64216.1
715	ШТАММ EDMONSTON ВИРУСА КОРИ	TSGSGGAIRGIKHII	1598	P04851.1
716	ШТАММ EDMONSTON ВИРУСА КОРИ	TSGSGGAIRGIKHIIIVPIP	1599	P04851.1
717	ШТАММ EDMONSTON ВИРУСА КОРИ	TSQGMYGGTYLVEKP	1600	P08362.1
718	МОРБИЛЛИВИРУС КОРИ	TSRVEHAVVYYVYSP	1601	P06830.1
719	ПАЛОЧКА КОКЛЮША	TSSSRRYTEVYL	1602	ACI04548.1
720	ПАЛОЧКА КОКЛЮША	TSYVG	1603	ACI16088.1
721	ШТАММ EDMONSTON ВИРУСА КОРИ	TTEDKISRAVGPRQA	1604	P04851.1
722	ШТАММ EDMONSTON ВИРУСА КОРИ	TTEDKISRAVGPRQAQVSFL	1605	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
723	ВИРУС КРАСНУХИ STRAIN M33	TTERIETRSARHP	1606	ABD64214.1
724	ШТАММ EDMONSTON ВИРУСА КОРИ	TTNQAIEAIRQAGQE	1607	P69353.1
725	ВИРУС КРАСНУХИ	TTSDPWHPPGPLGLK	1608	BAA28178.1
726	ШТАММ EDMONSTON ВИРУСА КОРИ	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	ПАЛОЧКА КОКЛЮША	VAPGIVIPPKALFTQQGGAYGRC	1613	P04979.1
731	ВИРУС КРАСНУХИ	VCHTKHMDFWCVEHDRPPPATP TPL	1614	NP_740663.1
732	ВИРУС КРАСНУХИ	VCQRHSPDCSRLVGATPER	1615	BAA19893.1
733	ВИРУС КРАСНУХИ	VDADDPLL	1616	CAJ88851.1
734	ВИРУС КРАСНУХИ	VDADDPLLRTAPGPGEVWVT	1617	BAA19893.1
735	ДИФТЕРИЙНАЯ ПАЛОЧКА	VDIGFAAYNFVESIINLFQV	1618	AAV70486.1
736	ШТАММ EDMONSTON ВИРУСА КОРИ	VEIAEYRRLLRTVLE	1619	P69353.1
737	ШТАММ EDMONSTON ВИРУСА КОРИ	VELENSMGGLNFGRS	1620	P04851.1
738	ШТАММ EDMONSTON ВИРУСА КОРИ	VELENSMGGLNFGRSYFDPA	1621	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
739	МОРБИЛЛИВИРУС КОРИ	VELKIKIASGFGPLI	1622	P06830.1
740	ВИРУС КРАСНУХИ	VEMDEWIHAHTTSD	1623	SRC265968
741	ВИРУС КРАСНУХИ	VEMPEWIHAHTTSDP	1624	BAA28178.1
742	ДИФТЕРИЙНАЯ ПАЛОЧКА	VERRLVKVL	1625	P33120.2
743	МОРБИЛЛИВИРУС КОРИ	VESPGQLI	1626	ABI54110.1
744	ШТАММ EDMONSTON ВИРУСА КОРИ	VESPGQLIQRITDDP	1627	P04851.1
745	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ	VGIGSASYKVMTRSS	1633	P69353.1
751	ШТАММ EDMONSTON ВИРУСА КОРИ	VGLRTPQRFTDLVKF	1634	P08362.1
752	ДИФТЕРИЙНАЯ ПАЛОЧКА	VHHNTEEIVAQSIALSSLMV	1635	AAV70486.1
753	ШТАММ EDMONSTON ВИРУСА КОРИ	VHRLEGVSYNIGSQE	1636	P69353.1
754	ВИРУС КРАСНУХИ	VIGSQARK	1637	CAJ88851.1
755	ПАЛОЧКА	VITGSI	1638	ACI16088.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
756	ПАЛОЧКА КОКЛЮША	VITGTI	1639	ACI16083.1
757	ШТАММ EDMONSTON ВИРУСА КОРИ	VKQSRGEA	1640	P04851.1
758	ШТАММ EDMONSTON ВИРУСА КОРИ	VLFVMFLSLI	1641	P08362.1
759	МОРБИЛЛИВИРУС КОРИ	VLFVMFLSLIGLLAI	1642	P06830.1
760	МОРБИЛЛИВИРУС КОРИ	VLTPLFKIIGDEVGL	1643	P06830.1
761	ШТАММ EDMONSTON ВИРУСА КОРИ	VMTRSSHQSLVIKLM	1644	P69353.1
762	ВИРУС КРАСНУХИ	VPAAPCARIWNGTQR	1645	BAA28178.1
763	ВИРУС КРАСНУХИ	VPPDPGDLVEYIMNY	1646	BAA28178.1
764	ШТАММ EDMONSTON ВИРУСА КОРИ	VQSVASSRRHKRFAG	1647	P69353.1
765	ПАЛОЧКА КОКЛЮША	VQTGGTSRTVTMRYLAS	1648	ACI16083.1
766	ПАЛОЧКА КОКЛЮША	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 EDMONSTON ВИРУСА КОРИ	VSGSFGNRFILSQGN	1653	P69353.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
771	ПАЛОЧКА КОКЛЮША	VSKEEQYYDYEDAT	1654	AAA83981.1
772	ШТАММ EDMONSTON ВИРУСА КОРИ	VSKGNCSGPTTIRGQ	1655	P08362.1
773	ВИРУС КРАСНУХИ	VTAALLNTPPPYQVS	1656	BAA28178.1
774	ВИРУС КРАСНУХИ	VTGCYQCGTPALVEG	1657	BAA28178.1
775	ВИРУС КРАСНУХИ	VTPVIGSQ	1658	CAJ88851.1
776	ВИРУС КРАСНУХИ STRAIN THERIEN	VTTEHPFCNTPHGQLEVQVPPD	1659	P07566.1
777	ШТАММ EDMONSTON ВИРУСА КОРИ	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 EDMONSTON ВИРУСА КОРИ	VYILIAVCLGGLIGI	1664	P69353.1
782	ШТАММ EDMONSTON ВИРУСА КОРИ	VYLHRIDLGPPISLE	1665	P69353.1
783	ПАЛОЧКА КОКЛЮША	VYRYDSRP	1666	P04977.1
784	ПАЛОЧКА КОКЛЮША	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

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
788	ВИРУС КРАСНУХИ	WGFPTDTVMSVFALA	1671	BAA28178.1
789	ВИРУС КРАСНУХИ	WHPPGPLGLKFKTVR	1672	BAA28178.1
790	ВИРУС КРАСНУХИ	WIHAHTTSDPWHPPG	1673	BAA28178.1
791	ШТАММ EDMONSTON ВИРУСА КОРИ	WLTIPPMKNLALGVI	1674	P08362.1
792	ШТАММ EDMONSTON ВИРУСА КОРИ	WQESRKNKAQTRTPLQCTMT	1675	P04851.1
793	ВИРУС КРАСНУХИ	WVCIFMVCRRACR	1676	SRC265968
794	ВИРУС КРАСНУХИ	WVTPVIGS	1677	CAJ88851.1
795	ШТАММ EDMONSTON ВИРУСА КОРИ	WYTTVPKYVATQGYL	1678	P69353.1
796	ШТАММ EDMONSTON ВИРУСА КОРИ	YALGGDINKVLEKLG	1679	P69353.1
797	ШТАММ EDMONSTON ВИРУСА КОРИ	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	ПАЛОЧКА КОКЛЮША	YFEPGPT	1684	ACI16083.1
802	ВИРУС КРАСНУХИ	YFNPGGSYYKQYHPT	1685	BAA28178.1
803	ШТАММ EDMONSTON	YFRLGQEMVRRSAGK	1686	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
804	ШТАММ EDMONSTON ВИРУСА КОРИ	YFRLGQEMVRRSAGKVSSTL	1687	P04851.1
805	МОРБИЛЛИВИРУС КОРИ	YFYPFRLPIKGVPIE	1688	P06830.1
806	ПАЛОЧКА КОКЛЮША	YGDNAGRILAGALAT	1689	P04977.1
807	ВИРУС КРАСНУХИ	YGHATVEMPEWIHAH	1690	BAA28178.1
808	ВИРУС КРАСНУХИ	YIMNYTGNQQSRWGL	1691	BAA28178.1
809	ШТАММ EDMONSTON ВИРУСА КОРИ	YINNELIPSMNQLSC	1692	P69353.1
810	ВИРУС КРАСНУХИ	YLCTAPGCATQAPVP	1693	BAA28178.1
811	МОРБИЛЛИВИРУС КОРИ	YLFTVPIKEAGEDCH	1694	P06830.1
812	ШТАММ EDMONSTON ВИРУСА КОРИ	YLHDPEFNL	1695	ABK40531.1
813	ШТАММ EDMONSTON ВИРУСА КОРИ	YLNMSRLFV	1696	ABK40531.1
814	MEASLES VIRUS STRAIN HALLE	YPALGLHEF	1697	P10050.1
815	МОРБИЛЛИВИРУС КОРИ	YPALGLHEFA	1698	ABI54110.1
816	ШТАММ EDMONSTON ВИРУСА КОРИ	YPALGLHEFAGELST	1699	P04851.1
817	ШТАММ EDMONSTON	YPALGLHEFAGELSTLESLM	1700	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ			
818	ШТАММ EDMONSTON ВИРУСА КОРИ	YPFRLPIKGVPIELQ	1701	P08362.1
819	MEASLES VIRUS STRAIN HALLE	YPLLWSYAM	1702	P10050.1
820	РУБУЛАВИРУС ПАРОТИТА	YQQQGRL	1703	P21186.1
821	ПАЛОЧКА КОКЛЮША	YQSEYLAHRR	1704	P04977.1
822	ШТАММ EDMONSTON ВИРУСА КОРИ	YRETGPSRASDARAA	1705	P04851.1
823	ШТАММ EDMONSTON ВИРУСА КОРИ	YRETGPSRASDARAAHLPTG	1706	P04851.1
824	ВИРУС КРАСНУХИ	YRNASDVLPGHWLQGGWGCYNL SDW	1707	NP_740663.1
825	ШТАММ EDMONSTON ВИРУСА КОРИ	YRRLLRTVLEPIRDA	1708	P69353.1
826	ПАЛОЧКА КОКЛЮША	YRYDSRPP	1709	P04977.1
827	ШТАММ EDMONSTON ВИРУСА КОРИ	YSGGDLLGILESRGI	1710	P69353.1
828	ПАЛОЧКА КОКЛЮША	YSKVTATBLLASTNSRLCAVFVRD G	1711	SRC280066
829	ШТАММ EDMONSTON ВИРУСА КОРИ	YSPSRSFSYFYPFRL	1712	P08362.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
830	ВИРУС КРАСНУХИ STRAIN THERIEN	YTGNQQSRWGLGSPNCHGPDWASPV	1713	P07566.1
831	МОРБИЛЛИВИРУС КОРИ	YVLLAVLFV	1714	P08362.1
832	МОРБИЛЛИВИРУС КОРИ	YVYSPGRSFSYFYPF	1715	P06830.1
833	ПАЛОЧКА КОКЛЮША	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	ПАЛОЧКА КОКЛЮША	ATYQSEYLAHRRIPP	1721	ACI04548.1
839	ПАЛОЧКА КОКЛЮША	CMARQAESSEAMAAWSERAGEAMVLVYYESIAYSF	1722	ACI04548.1
840	ПАЛОЧКА КОКЛЮША	CQVGSSNSAFVSTSSSRRYTEVYL	1723	ACI04548.1
841	ПАЛОЧКА КОКЛЮША	DDPPATVYRYDSRPP	1724	ACI04548.1
842	ПАЛОЧКА КОКЛЮША	GALATYQSEYLAHRRIPP	1725	ACI04548.1
843	ПАЛОЧКА КОКЛЮША	MAAWSERAGEAMVLVYYESIAYSF	1726	ACI04548.1
844	ПАЛОЧКА КОКЛЮША	MVLVYYESIAYSF	1727	ACI04548.1
845	ПАЛОЧКА	PATVYRYDSRPPEDV	1728	ACI04548.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
846	ПАЛОЧКА КОКЛЮША	YDSRPPEDV	1729	ACI04548.1
847	ПАЛОЧКА КОКЛЮША	EPGITTNYDT	1730	ACI16087.1
848	ПАЛОЧКА КОКЛЮША	GDLRAYKMVYATNPQTQLSN	1731	ACI16083.1
849	ПАЛОЧКА КОКЛЮША	KNGDVEASAITTYVGFSVVYP	1732	ACI16083.1
850	ПАЛОЧКА КОКЛЮША	KVTNGSKSYTLRYLASYVK	1733	ACI16088.1
851	ПАЛОЧКА КОКЛЮША	QALGALKLYFEPGITTNYDTGDLI AYKQTYNASGN	1734	ACI16088.1
852	ПАЛОЧКА КОКЛЮША	YATNPQTQLS	1735	ACI16083.1
853	ДИФТЕРИЙНАЯ ПАЛОЧКА	DNENPLSGKAGGVVKVTYPGLT KV	1736	AAV70486.1
854	ДИФТЕРИЙНАЯ ПАЛОЧКА	ENFSSYHGTKPGYVDSI	1737	AAV70486.1
855	ДИФТЕРИЙНАЯ ПАЛОЧКА	KVDNAETIKKELGLSLTEP	1738	AAV70486.1
856	ВИРУС КРАСНУХИ STRAIN M33	MEDLQKALEAQSRALRAGLAA	1739	CAA28880.1
857	ДИФТЕРИЙНАЯ ПАЛОЧКА	QKGIQKPKSGTQGNYDDDWKG FY	1740	AAV70486.1
858	ВИРУС КРАСНУХИ STRAIN M33	RTGAWQRKDWSRAPPPPEERQE SRSQTPAPKPSR	1741	CAA28880.1
859	ВИРУС КРАСНУХИ	AAGASQSRRPRPPRHARAQHLPE MTPAVT	1742	SRC265968
860	ВИРУС КРАСНУХИ	CVTSWLWSEGEGAVFYRVDLHFI NLGTP	1743	CAA28880.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
861	ВИРУС КРАСНУХИ	FRVGGTRWHRLLRMPVRGLDGD TAPLP	1744	CAA28880.1
862	ДИФТЕРИЙНАЯ ПАЛОЧКА	GRKIRMRCRAIDGDVTFCRPKSP VYVGN	1745	1007216A
863	ВИРУС КРАСНУХИ	GTPPLDEDGRWDPALMYNPCGP EPPAHV	1746	CAA28880.1
864	ДИФТЕРИЙНАЯ ПАЛОЧКА	GVHANLHVAFHRSSSEKIHSNEIS SDSIGVLGYQKTVDHTKVNSKLS LFFEIKS	1747	AAV70486.1
865	ВИРУС КРАСНУХИ	MASTTPITMEDLQKALEAQSRAL RAGLAAG	1748	ABD64200.1
866	ВИРУС КРАСНУХИ	PELGPPTNPFQAAVARGLRPPLH DPDTEAPTEAC	1749	CAA28880.1
867	ВИРУС КРАСНУХИ	PLPPHTTERIETRSARHPWRIRFG AP	1750	CAA28880.1
868	ВИРУС КРАСНУХИ	SRAPPPPEERQESRSQTPAPKPSR APP	1751	CAA28880.1
869	ВИРУС КРАСНУХИ	SRAPPQQPQPPRMQTGRGGSAP RPELGP	1752	CAA28880.1
870	ВИРУС КРАСНУХИ	TPAVTPEGPAPPRTGAWQRKDWS RAPP	1753	CAA28880.1
871	ВИРУС КРАСНУХИ	VRAYNQPAGDVRGVWGKGERTY AEQDFRV	1754	CAA28880.1
872	ВИРУС КРАСНУХИ	AFGHSDAACWGFPTDTVMSV	1755	CAA28880.1
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

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
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	ПАЛОЧКА КОКЛЮША	AAASSPDAHVPF	1778	AAA22983.1
896	ПАЛОЧКА КОКЛЮША	AASSPDA	1779	AAA22983.1
897	ПАЛОЧКА КОКЛЮША	AKLGAAASSPDA	1780	AAA22983.1
898	ПАЛОЧКА КОКЛЮША	AMKPYEVTPTRM	1781	AAA22983.1
899	ПАЛОЧКА КОКЛЮША	AMTHLSPALADVPYVLVKTNMVVTS	1782	AAA22983.1
900	ПАЛОЧКА КОКЛЮША	ASSPDAHVPFCF	1783	AAA22983.1
901	ПАЛОЧКА КОКЛЮША	ASSPDAHVPFCFGKDLKRPGSSPME	1784	AAA22983.1
902	ПАЛОЧКА КОКЛЮША	CFGKDLKRPGSS	1785	AAA22983.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
903	ПАЛОЧКА КОКЛЮША	CFGKDLKRPGSSPMEVMLRAVFMQQ	1786	AAA22983.1
904	ПАЛОЧКА КОКЛЮША	CGIAAKLGAAAS	1787	AAA22983.1
905	ПАЛОЧКА КОКЛЮША	CGIAAKLGAAASSPDAHVPFCFGKD	1788	AAA22983.1
906	ПАЛОЧКА КОКЛЮША	DAHVPFCFGKDL	1789	AAA22983.1
907	ПАЛОЧКА КОКЛЮША	DLKRPGSSPMEV	1790	AAA22983.1
908	ПАЛОЧКА КОКЛЮША	DVPYVLVKTNMV	1791	AAA22983.1
909	ПАЛОЧКА КОКЛЮША	DVPYVLVKTNMVVTSVAMKPYEVTPT	1792	AAA22983.1
910	ПАЛОЧКА КОКЛЮША	EVMLRAVFMQQR	1793	AAA22983.1
911	ПАЛОЧКА КОКЛЮША	FEGKPALELIRM	1794	AAA22983.1
912	ПАЛОЧКА КОКЛЮША	FLGPKQLTFEGK	1795	AAA22983.1
913	ПАЛОЧКА КОКЛЮША	FLGPKQLTFEGKPALELIRMVECSG	1796	AAA22983.1
914	ПАЛОЧКА КОКЛЮША	FMQQRPLRMFLGPKQLT	1797	AAA22983.1
915	ПАЛОЧКА КОКЛЮША	GKDLKRPGSSPM	1798	AAA22983.1
916	ПАЛОЧКА КОКЛЮША	GKDLKRPGSSPME	1799	AAA22983.1
917	ПАЛОЧКА КОКЛЮША	GKPALELIRMVE	1800	AAA22983.1
918	ПАЛОЧКА	GPKQLTFEGKPA	1801	AAA22983.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
919	*ПАЛОЧКА КОКЛЮША*	*HVPFCFGKDLKR*	*1802*	*AAA22983.1*
920	*ПАЛОЧКА КОКЛЮША*	*IAAKLGAAASSP*	*1803*	*AAA22983.1*
921	*ПАЛОЧКА КОКЛЮША*	*KPYEVTPTRMLV*	*1804*	*AAA22983.1*
922	*ПАЛОЧКА КОКЛЮША*	*KQLTFEGKPALE*	*1805*	*AAA22983.1*
923	*ПАЛОЧКА КОКЛЮША*	*KRPGSSPMEVML*	*1806*	*AAA22983.1*
924	*ПАЛОЧКА КОКЛЮША*	*LELIRMVECSGK*	*1807*	*AAA22983.1*
925	*ПАЛОЧКА КОКЛЮША*	*LGAAASSPDAHV*	*1808*	*AAA22983.1*
926	*ПАЛОЧКА КОКЛЮША*	*LIRMVECSGKQD*	*1809*	*AAA22983.1*
927	*ПАЛОЧКА КОКЛЮША*	*LVCGIAAKLGAA*	*1810*	*AAA22983.1*
928	*ПАЛОЧКА КОКЛЮША*	*MKPYEVTPTRM*	*1811*	*AAA22983.1*
929	*ПАЛОЧКА КОКЛЮША*	*MLRAVFMQQRPL*	*1812*	*AAA22983.1*
930	*ПАЛОЧКА КОКЛЮША*	*MQQRPLRM*	*1813*	*AAA22983.1*
931	*ПАЛОЧКА КОКЛЮША*	*MQQRPLRMFLGP*	*1814*	*AAA22983.1*
932	*ПАЛОЧКА КОКЛЮША*	*MVVTSVAMKPYE*	*1815*	*AAA22983.1*
933	*ПАЛОЧКА КОКЛЮША*	*MVVTSVAMKPYEVTPTRMLVCGIAA*	*1816*	*AAA22983.1*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
934	*ПАЛОЧКА КОКЛЮША*	*PALELIRMVECS*	*1817*	*AAA22983.1*
935	*ПАЛОЧКА КОКЛЮША*	*PALELIRMVECSGK*	*1818*	*AAA22983.1*
936	*ПАЛОЧКА КОКЛЮША*	*PFCFGKDLKRPG*	*1819*	*AAA22983.1*
937	*ПАЛОЧКА КОКЛЮША*	*PGSSPMEVMLRA*	*1820*	*AAA22983.1*
938	*ПАЛОЧКА КОКЛЮША*	*PGSSPMEVMLRAVF*	*1821*	*AAA22983.1*
939	*ПАЛОЧКА КОКЛЮША*	*PKQLTFEGK*	*1822*	*AAA22983.1*
940	*ПАЛОЧКА КОКЛЮША*	*PLRMFLGPKQLT*	*1823*	*AAA22983.1*
941	*ПАЛОЧКА КОКЛЮША*	*PTRMLVCGIAAK*	*1824*	*AAA22983.1*
942	*ПАЛОЧКА КОКЛЮША*	*PYVLVKTNMVVT*	*1825*	*AAA22983.1*
943	*ПАЛОЧКА КОКЛЮША*	*QLTFEGKPALELIRMVECSGKQDCP*	*1826*	*AAA22983.1*
944	*ПАЛОЧКА КОКЛЮША*	*QRPLRMFLGPKQ*	*1827*	*AAA22983.1*
945	*ПАЛОЧКА КОКЛЮША*	*RAVFMQQRPLRM*	*1828*	*AAA22983.1*
946	*ПАЛОЧКА КОКЛЮША*	*RMFLGPKQLTFE*	*1829*	*AAA22983.1*
947	*ПАЛОЧКА КОКЛЮША*	*RMLVCGIAAKLG*	*1830*	*AAA22983.1*
948	*ПАЛОЧКА КОКЛЮША*	*RMVECSGKQDCP*	*1831*	*AAA22983.1*
949	*ПАЛОЧКА*	*SPDAHVPFCFGK*	*1832*	*AAA22983.1*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
950	*ПАЛОЧКА КОКЛЮША*	*SSPMEVMLRAVF*	*1833*	*AAA22983.1*
951	*ПАЛОЧКА КОКЛЮША*	*SSPMEVMLRAVFMQQRPLRMFLGPK*	*1834*	*AAA22983.1*
952	*ПАЛОЧКА КОКЛЮША*	*SVAMKPYEVTPT*	*1835*	*AAA22983.1*
953	*ПАЛОЧКА КОКЛЮША*	*VECSGKQDCP*	*1836*	*AAA22983.1*
954	*ПАЛОЧКА КОКЛЮША*	*VFMQQRPLRMFL*	*1837*	*AAA22983.1*
955	*ПАЛОЧКА КОКЛЮША*	*VFMQQRPLRMFLGPKQLTFEGKPAL*	*1838*	*AAA22983.1*
956	*ПАЛОЧКА КОКЛЮША*	*VKTNMVVTSVAM*	*1839*	*AAA22983.1*
957	*ПАЛОЧКА КОКЛЮША*	*VLVKTNMVVTSV*	*1840*	*AAA22983.1*
958	*ПАЛОЧКА КОКЛЮША*	*VTPTRMLVCGIA*	*1841*	*AAA22983.1*
959	*ПАЛОЧКА КОКЛЮША*	*VTSVAMKPYEVT*	*1842*	*AAA22983.1*
960	*ПАЛОЧКА КОКЛЮША*	*YEVTPTRMLVCG*	*1843*	*AAA22983.1*
961	*ПАЛОЧКА КОКЛЮША*	*YEVTPTRMLVCGIAAKLGAAASSPD*	*1844*	*AAA22983.1*
962	*ПАЛОЧКА КОКЛЮША*	*CASPYEGRYRDMYDALR*	*1845*	*P04979.1*
963	*ПАЛОЧКА КОКЛЮША*	*CAVFVRDGQSV*	*1846*	*P04979.1*
964	*ПАЛОЧКА КОКЛЮША*	*CITTIYKTG*	*1847*	*P04979.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
965	ПАЛОЧКА КОКЛЮША	CPNGTRALTV	1848	P04979.1
966	ПАЛОЧКА КОКЛЮША	DALRRLLYMIYMSG	1849	P04979.1
967	ВИРУС КРАСНУХИ STRAIN THERIEN	GNRGRGQRRDWSRAPPPPEERQETRS	1850	P07566.1
968	ПАЛОЧКА КОКЛЮША	GQPAADHYYSKVT	1851	P04979.1
969	ВИРУС КРАСНУХИ	GSPNCHGPDWASPVCQRHS	1852	ABD64214.1
970	ШТАММ EDMONSTON ВИРУСА КОРИ	HKSLSTNLDVTNSIEHQ	1853	P08362.1
971	ПАЛОЧКА КОКЛЮША	LFTQQGGAYGRC	1854	P04979.1
972	ШТАММ EDMONSTON ВИРУСА КОРИ	LIGLLAIAGIRLHRAAIYTAEI	1855	P08362.1
973	МОРБИЛЛИВИРУС КОРИ	PDTAADSELRRWIKY	1856	ABI54110.1
974	ВИРУС КРАСНУХИ STRAIN THERIEN	PNCHGPDWASPVCQRHS	1857	P07566.1
975	ВИРУС КРАСНУХИ	QTPAPKPSRAPPQQPQPPRMQTGR	1858	ABD64216.1
976	ВИРУС КРАСНУХИ VACCINE STRAIN RA27/3	RAGLTAGASQSRRPRPPR	1859	CAA33016.1
977	ВИРУС КРАСНУХИ VACCINE STRAIN RA27/3	RFGAPQAFLAGLLLAAVAVGTARA	1860	ABD64214.1
978	ПАЛОЧКА КОКЛЮША	RGNAELQTYLRQITPG	1861	P04979.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
979	ПАЛОЧКА КОКЛЮША	RVHVSKEEQYYDYED	1862	P04979.1
980	ПАЛОЧКА КОКЛЮША	SIYGLYDGTYL	1863	P04979.1
981	ПАЛОЧКА КОКЛЮША	SKVTATRLLASTNS	1864	P04979.1
982	ПАЛОЧКА КОКЛЮША	TQQGGAYGRCPNGTRA	1865	P04979.1
983	ПАЛОЧКА КОКЛЮША	VAPGIVIPPKAL	1866	P04979.1
984	ПАЛОЧКА КОКЛЮША	DSRPPEDVFQNGFTAWG	1867	ACI04548.1
985	ПАЛОЧКА КОКЛЮША	EHRMQEAVEAERAGR	1868	ACI04548.1
986	ШТАММ EDMONSTON ВИРУСА КОРИ	ETCFQQACKGKIQALCENPEWA	1869	P08362.1
987	ПАЛОЧКА КОКЛЮША	EYVDTYGDNAGRILAGALATYQ	1870	ACI04548.1
988	ПАЛОЧКА КОКЛЮША	HRRIPPENIRRVTR	1871	ACI04548.1
989	ПАЛОЧКА КОКЛЮША	MARQAESSE	1872	ACI04548.1
990	ПАЛОЧКА КОКЛЮША	MQEAVEAERAGR	1873	ACI04548.1
991	ПАЛОЧКА КОКЛЮША	SQQTRANPNPYTSRR	1874	ACI04548.1
992	ПАЛОЧКА КОКЛЮША	TRANPNPYTSRRSVASIVGTLVHG	1875	SRC280066
993	ПАЛОЧКА КОКЛЮША	TVYRYDSRPPED	1876	ACI04548.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
994	МОРБИЛЛИВИРУС КОРИ	NDRNLLD	1877	P10050.1
995	МОРБИЛЛИВИРУС КОРИ	NMEDEADQYFSHDDPISSDQSRFGWFENK	1878	P04851.1
996	МОРБИЛЛИВИРУС КОРИ	SRASDARAAHLPTGTPLDID	1879	P04851.1
997	ПАЛОЧКА КОКЛЮША	EDVFQNGFTAW	1880	ACI04548.1
998	ДИФТЕРИЙНАЯ ПАЛОЧКА	AEGSSSVEYINNWEQAK	1881	AAV70486.1
999	ДИФТЕРИЙНАЯ ПАЛОЧКА	GPIKNKMSESPNKT	1882	AAV70486.1
1000	MEASLES VIRUS STRAIN HALLE	GPKLTGALIGILSLFVESPGQLIQRITDDPDV	1883	P10050.1
1001	ДИФТЕРИЙНАЯ ПАЛОЧКА	GYQKTVDHTKVNSK	1884	AAV70486.1
1002	ДИФТЕРИЙНАЯ ПАЛОЧКА	KTVDH	1885	AAV70486.1
1003	ДИФТЕРИЙНАЯ ПАЛОЧКА	SESPNK	1886	AAV70486.1
1004	ДИФТЕРИЙНАЯ ПАЛОЧКА	AEGSSSVEYINNWEQAKALS	1887	AAV70486.1
1005	ДИФТЕРИЙНАЯ ПАЛОЧКА	AQAIPLVGELVDIGFAAYNF	1888	AAV70486.1
1006	ДИФТЕРИЙНАЯ ПАЛОЧКА	ASRVVLSLPFAEGSSSVEYI	1889	AAV70486.1
1007	ДИФТЕРИЙНАЯ ПАЛОЧКА	CINLDWDVIRDKTKTKIESL	1890	AAV70486.1
1008	ДИФТЕРИЙНАЯ ПАЛОЧКА	CRAIDGDVTFCRPKSPVYVG	1891	AAV70486.1
1009	ДИФТЕРИЙНАЯ	CRPKSPVYVGNGVHANLHVA	1892	AAV70486.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	ПАЛОЧКА			
1010	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*DAAGYSVDNENPLSGKAGGV*	*1893*	*AAV70486.1*
1011	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*DKTKTKIESLKEHGPIKNKM*	*1894*	*AAV70486.1*
1012	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*EEFIKRFGDGASRVVLSLPF*	*1895*	*AAV70486.1*
1013	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*EKAKQYLEEFHQTALEHPEL*	*1896*	*AAV70486.1*
1014	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*FHRSSSEKIHSNEISSDSIG*	*1897*	*AAV70486.1*
1015	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*GADDVVDSSKSFVMENFSSY*	*1898*	*CAE11230.1*
1016	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*GKRGQDAMYEYMAQACAGNR*	*1899*	*AAV70486.1*
1017	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*GSVMGIADGAVHHNTEEIVA*	*1900*	*AAV70486.1*
1018	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*GTQGNYDDDWKGFYSTDNKY*	*1901*	*AAV70486.1*
1019	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*HDGYAVSWNTVEDSIIRTGF*	*1902*	*AAV70486.1*
1020	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*HGTKPGYVDSIQKGIQKPKS*	*1903*	*AAV70486.1*
1021	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*HQTALEHPELSELKTVTGTN*	*1904*	*AAV70486.1*
1022	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*IQKGIQKPKSGTQGNYDDDW*	*1905*	*AAV70486.1*
1023	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*KEHGPIKNKMSESPNKTVSE*	*1906*	*AAV70486.1*
1024	*ДИФТЕРИЙНАЯ ПАЛОЧКА*	*KGFYSTDNKYDAAGYSVDNE*	*1907*	*AAV70486.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1025	ДИФТЕРИЙНАЯ ПАЛОЧКА	LDVNKSKTHISVNGRKIRMR	1908	AAV70486.1
1026	ВИРУС КРАСНУХИ	MASTIPITMEDLQKALEA	1909	SRC265968
1027	ДИФТЕРИЙНАЯ ПАЛОЧКА	NGVHANLHVAFHRSSSEKIH	1910	AAV70486.1
1028	ДИФТЕРИЙНАЯ ПАЛОЧКА	NNWEQAKALSVELEINFETR	1911	AAV70486.1
1029	ДИФТЕРИЙНАЯ ПАЛОЧКА	NPLSGKAGGVVKVTYPGLTK	1912	AAV70486.1
1030	ДИФТЕРИЙНАЯ ПАЛОЧКА	PVFAGANYAAWAVNVAQVID	1913	AAV70486.1
1031	ДИФТЕРИЙНАЯ ПАЛОЧКА	SELKTVTGTNPVFAGANYAA	1914	AAV70486.1
1032	ДИФТЕРИЙНАЯ ПАЛОЧКА	SESPNKTVSEEKAKQYLEEF	1915	AAV70486.1
1033	ДИФТЕРИЙНАЯ ПАЛОЧКА	SETADNLEKTTAALSILPGI	1916	AAV70486.1
1034	ДИФТЕРИЙНАЯ ПАЛОЧКА	SFVMENFSSYHGTKPGYVDS	1917	AAV70486.1
1035	ДИФТЕРИЙНАЯ ПАЛОЧКА	SNEISSDSIGVLGYQKTVDH	1918	AAV70486.1
1036	ДИФТЕРИЙНАЯ ПАЛОЧКА	SPGHKTQPFLHDGYAVSWNT	1919	AAV70486.1
1037	ДИФТЕРИЙНАЯ ПАЛОЧКА	SVNGRKIRMRCRAIDGDVTF	1920	AAV70486.1
1038	ДИФТЕРИЙНАЯ ПАЛОЧКА	TAALSILPGIGSVMGIADGA	1921	AAV70486.1
1039	ДИФТЕРИЙНАЯ ПАЛОЧКА	TAENTPLPIAGVLLPTIPGK	1922	AAV70486.1
1040	ДИФТЕРИЙНАЯ ПАЛОЧКА	TEPLMEQVGTEEFIKRFGDG	1923	AAV70486.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1041	ДИФТЕРИЙНАЯ ПАЛОЧКА	TIKKELGLSLTEPLMEQVGT	1924	AAV70486.1
1042	ДИФТЕРИЙНАЯ ПАЛОЧКА	TKVNSKLSLFFEIKS	1925	AAV70486.1
1043	ДИФТЕРИЙНАЯ ПАЛОЧКА	VEDSIIRTGFQGESGHDIKI	1926	AAV70486.1
1044	ДИФТЕРИЙНАЯ ПАЛОЧКА	VELEINFETRGKRGQDAMYE	1927	AAV70486.1
1045	ДИФТЕРИЙНАЯ ПАЛОЧКА	VESIINLFQVVHNSYNRPAY	1928	AAV70486.1
1046	ДИФТЕРИЙНАЯ ПАЛОЧКА	VHNSYNRPAYSPGHKTQPFL	1929	AAV70486.1
1047	ДИФТЕРИЙНАЯ ПАЛОЧКА	VKVTYPGLTKVLALKVDNAE	1930	AAV70486.1
1048	ДИФТЕРИЙНАЯ ПАЛОЧКА	VLALKVDNAETIKKELGLSL	1931	AAV70486.1
1049	ДИФТЕРИЙНАЯ ПАЛОЧКА	VLGYQKTVDHTKVNSKLSLF	1932	AAV70486.1
1050	ДИФТЕРИЙНАЯ ПАЛОЧКА	VRRSVGSSLSCINLDWDVIR	1933	AAV70486.1
1051	ДИФТЕРИЙНАЯ ПАЛОЧКА	WAVNVAQVIDSETADNLEKT	1934	AAV70486.1
1052	ДИФТЕРИЙНАЯ ПАЛОЧКА	YMAQACAGNRVRRSVGSSLS	1935	AAV70486.1
1053	ПАЛОЧКА КОКЛЮША	AKAPPAPKPAPQPGP	1936	ABO77783.1
1054	ПАЛОЧКА КОКЛЮША	APKPAPQPGP	1937	ABO77783.1
1055	ПАЛОЧКА КОКЛЮША	APKPAPQPGPQPPQP	1938	ABO77783.1
1056	ШТАММ	AQTRTPLQCTMTEIF	1939	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
1057	ШТАММ EDMONSTON ВИРУСА КОРИ	ASRGTNMEDEADQYFSHDD	1940	P04851.1
1058	ПАЛОЧКА КОКЛЮША	ATIRR	1941	ABO77783.1
1059	ПАЛОЧКА КОКЛЮША	DNRAG	1942	ABO77783.1
1060	ПАЛОЧКА КОКЛЮША	EAPAPQPPAGRELSA	1943	ABO77783.1
1061	ШТАММ EDMONSTON ВИРУСА КОРИ	EMVRRSAGKVSSTLASELGI	1944	P04851.1
1062	ПАЛОЧКА КОКЛЮША	GASEL	1945	ABO77783.1
1063	ПАЛОЧКА КОКЛЮША	GDALAGGAVP	1946	AAA22980.1
1064	ПАЛОЧКА КОКЛЮША	GDAPAGGAVP	1947	ABO77783.1
1065	ПАЛОЧКА КОКЛЮША	GDTWDDD	1948	ABO77783.1
1066	ПАЛОЧКА КОКЛЮША	GERQH	1949	ABO77783.1
1067	ПАЛОЧКА КОКЛЮША	GGAVP	1950	ABO77783.1
1068	ПАЛОЧКА КОКЛЮША	GGFGP	1951	P14283.3
1069	ПАЛОЧКА КОКЛЮША	GGFGPGGFGP	1952	BAF35031.1
1070	ПАЛОЧКА	GGFGPVLDGW	1953	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1071	ШТАММ EDMONSTON ВИРУСА КОРИ	GGKEDRRVKQSRGEARESYR	1954	P04851.1
1072	ПАЛОЧКА КОКЛЮША	GILLEN	1955	ABO77783.1
1073	ПАЛОЧКА КОКЛЮША	GIRRFL	1956	ABO77783.1
1074	ШТАММ EDMONSTON ВИРУСА КОРИ	HDDPISSDQSRFGWFENKEI	1957	P04851.1
1075	ШТАММ EDMONSTON ВИРУСА КОРИ	HEFAGELSTLESLMNLY	1958	P04851.1
1076	ПАЛОЧКА КОКЛЮША	HLGGLAGY	1959	ABO77783.1
1077	ШТАММ EDMONSTON ВИРУСА КОРИ	HTTEDKISRAVGPRQAQVSFL	1960	P04851.1
1078	ШТАММ EDMONSTON ВИРУСА КОРИ	ICDIDTYIVEAGLASFILTI	1961	P04851.1
1079	ШТАММ EDMONSTON ВИРУСА КОРИ	IIIVPIPGDSSITTRSRLLD	1962	P04851.1
1080	ШТАММ EDMONSTON ВИРУСА КОРИ	ILDIKRTPGNKPRIAEMICD	1963	P04851.1
1081	ПАЛОЧКА КОКЛЮША	KALLYR	1964	ABO77783.1
1082	ШТАММ	KPPITSGSGGAIRGIKHIII	1965	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
1083	ПАЛОЧКА КОКЛЮША	LAGSGL	1966	ABO77783.1
1084	ШТАММ EDMONSTON ВИРУСА КОРИ	LGITAEDARLVSEIAMHTTE	1967	P04851.1
1085	ШТАММ EDMONSTON ВИРУСА КОРИ	LGTILAQIWVLLAKAVTA	1968	P04851.1
1086	ШТАММ EDMONSTON ВИРУСА КОРИ	LPTGTPLDIDTASESSQD	1969	P04851.1
1087	ШТАММ EDMONSTON ВИРУСА КОРИ	MATLLRSLALFKRNKDK	1970	P04851.1
1088	ПАЛОЧКА КОКЛЮША	PAPQPP	1971	ABO77783.1
1089	ШТАММ EDMONSTON ВИРУСА КОРИ	PDTAADSELRRWIKYTQQRR	1972	P04851.1
1090	ШТАММ EDMONSTON ВИРУСА КОРИ	PKLTGALIGILSLFVESPGQ	1973	P04851.1
1091	ПАЛОЧКА КОКЛЮША	PQP	1974	ABO77783.1
1092	ПАЛОЧКА КОКЛЮША	PQPGP	1975	ABO77783.1
1093	ПАЛОЧКА КОКЛЮША	PQPGPQPPQPPQPQP	1976	ABO77783.1
1094	ШТАММ	QDPQDSRRSAEPLLSCKPWQ	1977	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
1095	ШТАММ EDMONSTON ВИРУСА КОРИ	QRRVVGEFRLERKWLDVVR	1978	P04851.1
1096	ПАЛОЧКА КОКЛЮША	RELSA	1979	ABO77783.1
1097	ПАЛОЧКА КОКЛЮША	RFAPQ	1980	ABO77783.1
1098	ШТАММ EDMONSTON ВИРУСА КОРИ	SIQNKFSAGSYPLLWSYAMG	1981	P04851.1
1099	ПАЛОЧКА КОКЛЮША	SITLQAGAH	1982	ABO77783.1
1100	ПАЛОЧКА КОКЛЮША	SLQPED	1983	ABO77783.1
1101	ПАЛОЧКА КОКЛЮША	SNALSKRL	1984	ABO77783.1
1102	ШТАММ EDMONSTON ВИРУСА КОРИ	SPGQLIQRITDDPDVSIRLL	1985	P04851.1
1103	ПАЛОЧКА КОКЛЮША	TELPSIPG	1986	ABO77783.1
1104	ПАЛОЧКА КОКЛЮША	TFTLANK	1987	ABO77783.1
1105	ПАЛОЧКА КОКЛЮША	TWDDD	1988	ABO77783.1
1106	ШТАММ EDMONSTON ВИРУСА КОРИ	VSFLQGDQSENELPRLGGKE	1989	P04851.1
1107	ШТАММ	WQESRKNKAQTRTPLQC	1990	P04851.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ			
1108	ШТАММ EDMONSTON ВИРУСА КОРИ	YAMGVGVELENSMGGLNFGR	1991	P04851.1
1109	ШТАММ EDMONSTON ВИРУСА КОРИ	YQQMGKPAPYMVNLENSI	1992	P04851.1
1110	MEASLES VIRUS STRAIN HALLE	MTRSSHQSLVIKLMP	1993	P69355.1
1111	MEASLES VIRUS STRAIN HALLE	PIRDALNAMTQNIRP	1994	P69355.1
1112	ВИРУС КРАСНУХИ STRAIN M33	ALLNTPPPYQVSCGGESDRASAGH	1995	CAA28880.1
1113	ВИРУС КРАСНУХИ STRAIN THERIEN	GLGSPNCHGPDWASPVCQRHS	1996	P07566.1
1114	ВИРУС КРАСНУХИ STRAIN THERIEN	GLGSPNCHGPDWASPVCQRHSPDCSRLV	1997	P07566.1
1115	ВИРУС КРАСНУХИ STRAIN THERIEN	NYTGNQQSRWGLGSPNCHGPDWASPV	1998	P07566.1
1116	ВИРУС КРАСНУХИ STRAIN THERIEN	TLPQPPCAHGQHYGHHHHQL	1999	P07566.1
1117	ВИРУС КРАСНУХИ STRAIN 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
1123	ВИРУС КРАСНУХИ	LVGATPERPRLRLVDADDPLLRTAP	2006	BAA28178.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1124	ВИРУС КРАСНУХИ	TAPGPGEVWVTPVI	2007	BAA28178.1
1125	МОРБИЛЛИВИРУС КОРИ	FIVLSIAYPTLSEIK	2008	AAL29688.1
1126	ШТАММ EDMONSTON ВИРУСА КОРИ-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

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1140	МОРБИЛЛИВИРУС КОРИ	VRVIDPSLGDRKDE	2023	SRC280148
1141	ПАЛОЧКА КОКЛЮША	DLSDGDLLV	2024	AAC31207.1
1142	ПАЛОЧКА КОКЛЮША	EAERAGRGTG	2025	ACI04548.1
1143	ПАЛОЧКА КОКЛЮША	YRYDSRPPEDV	2026	ACI04548.1
1144	ПАЛОЧКА КОКЛЮША	YVDTYGDNAG	2027	ACI04548.1
1145	МОРБИЛЛИВИРУС КОРИ	SFSYFYPFR	2028	CAB43772.1
1146	РУБУЛАВИРУС ПАРОТИТА	DIFIVSPR	2029	ADF49557.1
1147	МОРБИЛЛИВИРУС КОРИ	QDSRRSADALLRLQAMAGISEEQ GSDTDTPIVYNDRN	2030	CAA59302.1
1148	МОРБИЛЛИВИРУС КОРИ	SAEALLRLQA	2031	BAH22350.1
1149	МОРБИЛЛИВИРУС КОРИ	RIVINREHL	2032	BAB39835.1
1150	MEASLES VIRUS GENOTYPE A	IPRFK	2033	BAB39848.1
1151	ПАЛОЧКА КОКЛЮША	AAALSPMEI	2034	P15318.2
1152	ПАЛОЧКА КОКЛЮША	AAASVVGAPV	2035	P15318.2
1153	ПАЛОЧКА КОКЛЮША	AALGRQDSI	2036	P15318.2
1154	ПАЛОЧКА КОКЛЮША	AAQRLVHAIA	2037	P15318.2
1155	ПАЛОЧКА	AAVEAAEL	2038	P15318.2

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1156	ПАЛОЧКА КОКЛЮША	AGANVLNGL	2039	P15318.2
1157	ПАЛОЧКА КОКЛЮША	AGYANAAD	2040	P15318.2
1158	ПАЛОЧКА КОКЛЮША	AGYEQFEFRV	2041	P15318.2
1159	ПАЛОЧКА КОКЛЮША	AITGNADNL	2042	P15318.2
1160	ПАЛОЧКА КОКЛЮША	AKEKNATLM	2043	P15318.2
1161	ПАЛОЧКА КОКЛЮША	AKGVFLSL	2044	P15318.2
1162	ПАЛОЧКА КОКЛЮША	APHEYGFGI	2045	P15318.2
1163	ПАЛОЧКА КОКЛЮША	ARQGNDLEI	2046	P15318.2
1164	ПАЛОЧКА КОКЛЮША	ASVVGAPV	2047	P15318.2
1165	ПАЛОЧКА КОКЛЮША	ATLMFRLV	2048	P15318.2
1166	ПАЛОЧКА КОКЛЮША	AVAAAQRL	2049	P15318.2
1167	ПАЛОЧКА КОКЛЮША	DAGANVLNGL	2050	P15318.2
1168	ПАЛОЧКА КОКЛЮША	DALLAQLYR	2051	P15318.2
1169	ПАЛОЧКА КОКЛЮША	DANGVLKHSI	2052	P15318.2
1170	ПАЛОЧКА КОКЛЮША	DGDMNIGVI	2053	P15318.2

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
1171	*ПАЛОЧКА КОКЛЮША*	*DHVKNIENL*	*2054*	*P15318.2*
1172	*ПАЛОЧКА КОКЛЮША*	*DIDMFAIM*	*2055*	*P15318.2*
1173	*ПАЛОЧКА КОКЛЮША*	*DMFAIMPHL*	*2056*	*P15318.2*
1174	*ПАЛОЧКА КОКЛЮША*	*DNVRNVENV*	*2057*	*P15318.2*
1175	*ПАЛОЧКА КОКЛЮША*	*DNVRNVENVI*	*2058*	*P15318.2*
1176	*ПАЛОЧКА КОКЛЮША*	*DTVDYSAM*	*2059*	*P15318.2*
1177	*ПАЛОЧКА КОКЛЮША*	*DTVDYSAMI*	*2060*	*P15318.2*
1178	*ПАЛОЧКА КОКЛЮША*	*DYLRQAGL*	*2061*	*P15318.2*
1179	*ПАЛОЧКА КОКЛЮША*	*DYYDNVRNV*	*2062*	*P15318.2*
1180	*ПАЛОЧКА КОКЛЮША*	*EFTTFVEI*	*2063*	*P15318.2*
1181	*ПАЛОЧКА КОКЛЮША*	*EFTTFVEIV*	*2064*	*P15318.2*
1182	*ПАЛОЧКА КОКЛЮША*	*EGYVFYEN*	*2065*	*P15318.2*
1183	*ПАЛОЧКА КОКЛЮША*	*ENVQYRHV*	*2066*	*P15318.2*
1184	*ПАЛОЧКА КОКЛЮША*	*EQLANSDGL*	*2067*	*P15318.2*
1185	*ПАЛОЧКА КОКЛЮША*	*FGVGYGHDTI*	*2068*	*P15318.2*
1186	*ПАЛОЧКА*	*FSPDVLETVP*	*2069*	*P15318.2*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1187	*ПАЛОЧКА КОКЛЮША*	*FSVDHVKNI*	*2070*	*P15318.2*
1188	*ПАЛОЧКА КОКЛЮША*	*GDDTYLFGV*	*2071*	*P15318.2*
1189	*ПАЛОЧКА КОКЛЮША*	*GDDVFLQDL*	*2072*	*P15318.2*
1190	*ПАЛОЧКА КОКЛЮША*	*GEDGNDIFL*	*2073*	*P15318.2*
1191	*ПАЛОЧКА КОКЛЮША*	*GERFNVRKQL*	*2074*	*P15318.2*
1192	*ПАЛОЧКА КОКЛЮША*	*GGAGNDTLV*	*2075*	*P15318.2*
1193	*ПАЛОЧКА КОКЛЮША*	*GGDDFEAV*	*2076*	*P15318.2*
1194	*ПАЛОЧКА КОКЛЮША*	*GKSEFTTFV*	*2077*	*P15318.2*
1195	*ПАЛОЧКА КОКЛЮША*	*GKSLFDDGL*	*2078*	*P15318.2*
1196	*ПАЛОЧКА КОКЛЮША*	*GNADNLKSV*	*2079*	*P15318.2*
1197	*ПАЛОЧКА КОКЛЮША*	*GQLVEVDTL*	*2080*	*P15318.2*
1198	*ПАЛОЧКА КОКЛЮША*	*GRSKFSPDV*	*2081*	*P15318.2*
1199	*ПАЛОЧКА КОКЛЮША*	*GSSAYDTV*	*2082*	*P15318.2*
1200	*ПАЛОЧКА КОКЛЮША*	*GTVEKWPAL*	*2083*	*P15318.2*
1201	*ПАЛОЧКА КОКЛЮША*	*GVDYYDNV*	*2084*	*P15318.2*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
1202	*ПАЛОЧКА КОКЛЮША*	*GYEQFEFRV*	*2085*	*P15318.2*
1203	*ПАЛОЧКА КОКЛЮША*	*HAVGAQDVV*	*2086*	*P15318.2*
1204	*ПАЛОЧКА КОКЛЮША*	*IAAGRIGLGI*	*2087*	*P15318.2*
1205	*ПАЛОЧКА КОКЛЮША*	*IGDAQANTL*	*2088*	*P15318.2*
1206	*ПАЛОЧКА КОКЛЮША*	*IGLGILADL*	*2089*	*P15318.2*
1207	*ПАЛОЧКА КОКЛЮША*	*IGNAAGIPL*	*2090*	*P15318.2*
1208	*ПАЛОЧКА КОКЛЮША*	*IGTSMKDVL*	*2091*	*P15318.2*
1209	*ПАЛОЧКА КОКЛЮША*	*IGVTTDFEL*	*2092*	*P15318.2*
1210	*ПАЛОЧКА КОКЛЮША*	*IPLTADIDM*	*2093*	*P15318.2*
1211	*ПАЛОЧКА КОКЛЮША*	*ISKSALEL*	*2094*	*P15318.2*
1212	*ПАЛОЧКА КОКЛЮША*	*ITGNADNL*	*2095*	*P15318.2*
1213	*ПАЛОЧКА КОКЛЮША*	*KIFVVSAT*	*2096*	*P15318.2*
1214	*ПАЛОЧКА КОКЛЮША*	*KQLNNANVYR*	*2097*	*P15318.2*
1215	*ПАЛОЧКА КОКЛЮША*	*KVIGNAAGI*	*2098*	*P15318.2*
1216	*ПАЛОЧКА КОКЛЮША*	*LAKVVSQL*	*2099*	*P15318.2*
1217	*ПАЛОЧКА*	*LANDYARKI*	*2100*	*P15318.2*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1218	*ПАЛОЧКА КОКЛЮША*	*LDYLRQAGL*	*2101*	*P15318.2*
1219	*ПАЛОЧКА КОКЛЮША*	*LGKGFASL*	*2102*	*P15318.2*
1220	*ПАЛОЧКА КОКЛЮША*	*LGKGFASLM*	*2103*	*P15318.2*
1221	*ПАЛОЧКА КОКЛЮША*	*LGVDYYDN*	*2104*	*P15318.2*
1222	*ПАЛОЧКА КОКЛЮША*	*LGVDYYDNV*	*2105*	*P15318.2*
1223	*ПАЛОЧКА КОКЛЮША*	*LKHSIKLDVI*	*2106*	*P15318.2*
1224	*ПАЛОЧКА КОКЛЮША*	*LQAGYIPV*	*2107*	*P15318.2*
1225	*ПАЛОЧКА КОКЛЮША*	*LQLTGGTVE*	*2108*	*P15318.2*
1226	*ПАЛОЧКА КОКЛЮША*	*LSAAVFGL*	*2109*	*P15318.2*
1227	*ПАЛОЧКА КОКЛЮША*	*LSLGKGFASL*	*2110*	*P15318.2*
1228	*ПАЛОЧКА КОКЛЮША*	*LSPMEIYGL*	*2111*	*P15318.2*
1229	*ПАЛОЧКА КОКЛЮША*	*NAHDNFLAGG*	*2112*	*P15318.2*
1230	*ПАЛОЧКА КОКЛЮША*	*NANVYREGV*	*2113*	*P15318.2*
1231	*ПАЛОЧКА КОКЛЮША*	*NDTLYGGL*	*2114*	*P15318.2*
1232	*ПАЛОЧКА КОКЛЮША*	*NGLAGNDVL*	*2115*	*P15318.2*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
1233	*ПАЛОЧКА КОКЛЮША*	*NNANVYREGV*	*2116*	*P15318.2*
1234	*ПАЛОЧКА КОКЛЮША*	*NTVSYAAL*	*2117*	*P15318.2*
1235	*ПАЛОЧКА КОКЛЮША*	*NVLRNIENAV*	*2118*	*P15318.2*
1236	*ПАЛОЧКА КОКЛЮША*	*PALTFITPL*	*2119*	*P15318.2*
1237	*ПАЛОЧКА КОКЛЮША*	*PETSNVLRNI*	*2120*	*P15318.2*
1238	*ПАЛОЧКА КОКЛЮША*	*PMEIYGLV*	*2121*	*P15318.2*
1239	*ПАЛОЧКА КОКЛЮША*	*PQAYFEKNL*	*2122*	*P15318.2*
1240	*ПАЛОЧКА КОКЛЮША*	*PVNPNLSKL*	*2123*	*P15318.2*
1241	*ПАЛОЧКА КОКЛЮША*	*QAGWNASSV*	*2124*	*P15318.2*
1242	*ПАЛОЧКА КОКЛЮША*	*QAGWNASSVI*	*2125*	*P15318.2*
1243	*ПАЛОЧКА КОКЛЮША*	*QDAANAGNL*	*2126*	*P15318.2*
1244	*ПАЛОЧКА КОКЛЮША*	*QDAANAGNLL*	*2127*	*P15318.2*
1245	*ПАЛОЧКА КОКЛЮША*	*QDSGYDSL*	*2128*	*P15318.2*
1246	*ПАЛОЧКА КОКЛЮША*	*QQSHYADQL*	*2129*	*P15318.2*
1247	*ПАЛОЧКА КОКЛЮША*	*RALQGAQAV*	*2130*	*P15318.2*
1248	*ПАЛОЧКА*	*RGGLGLDTL*	*2131*	*P15318.2*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1249	ПАЛОЧКА КОКЛЮША	RKQLNNANV	2132	P15318.2
1250	ПАЛОЧКА КОКЛЮША	RQDSGYDSL	2133	P15318.2
1251	ПАЛОЧКА КОКЛЮША	RQFRYDGDM	2134	P15318.2
1252	ПАЛОЧКА КОКЛЮША	RSKFSPDVL	2135	P15318.2
1253	ПАЛОЧКА КОКЛЮША	SAGAAAGAL	2136	P15318.2
1254	ПАЛОЧКА КОКЛЮША	SAHWGQRAL	2137	P15318.2
1255	ПАЛОЧКА КОКЛЮША	SAMIHPGRI	2138	P15318.2
1256	ПАЛОЧКА КОКЛЮША	SAMIHPGRIV	2139	P15318.2
1257	ПАЛОЧКА КОКЛЮША	SAYDTVSGI	2140	P15318.2
1258	ПАЛОЧКА КОКЛЮША	SAYGYEGD	2141	P15318.2
1259	ПАЛОЧКА КОКЛЮША	SGGAGDDVL	2142	P15318.2
1260	ПАЛОЧКА КОКЛЮША	SGLQVAGA	2143	P15318.2
1261	ПАЛОЧКА КОКЛЮША	SGYDSLDGV	2144	P15318.2
1262	ПАЛОЧКА КОКЛЮША	SLLTGALNGI	2145	P15318.2
1263	ПАЛОЧКА КОКЛЮША	SPMEIYGL	2146	P15318.2

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
1264	*ПАЛОЧКА КОКЛЮША*	*SQMLTRGQL*	*2147*	*P15318.2*
1265	*ПАЛОЧКА КОКЛЮША*	*SSAYDTVSGI*	*2148*	*P15318.2*
1266	*ПАЛОЧКА КОКЛЮША*	*SSLAHGHTA*	*2149*	*P15318.2*
1267	*ПАЛОЧКА КОКЛЮША*	*SSVTSGDSV*	*2150*	*P15318.2*
1268	*ПАЛОЧКА КОКЛЮША*	*SVIGVQTTEI*	*2151*	*P15318.2*
1269	*ПАЛОЧКА КОКЛЮША*	*TNTVSYAAL*	*2152*	*P15318.2*
1270	*ПАЛОЧКА КОКЛЮША*	*TSLIAEGV*	*2153*	*P15318.2*
1271	*ПАЛОЧКА КОКЛЮША*	*TSLLTGAL*	*2154*	*P15318.2*
1272	*ПАЛОЧКА КОКЛЮША*	*TVPASPGL*	*2155*	*P15318.2*
1273	*ПАЛОЧКА КОКЛЮША*	*VAKEKNATL*	*2156*	*P15318.2*
1274	*ПАЛОЧКА КОКЛЮША*	*VAPHEYGFGI*	*2157*	*P15318.2*
1275	*ПАЛОЧКА КОКЛЮША*	*VAVVTSLL*	*2158*	*P15318.2*
1276	*ПАЛОЧКА КОКЛЮША*	*VFYENRAYG*	*2159*	*P15318.2*
1277	*ПАЛОЧКА КОКЛЮША*	*VFYENRAYGV*	*2160*	*P15318.2*
1278	*ПАЛОЧКА КОКЛЮША*	*VITDFELEV*	*2161*	*P15318.2*
1279	*ПАЛОЧКА*	*VNPHSTSL*	*2162*	*P15318.2*

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1280	*ПАЛОЧКА КОКЛЮША*	*VNPHSTSLI*	*2163*	*P15318.2*
1281	*ПАЛОЧКА КОКЛЮША*	*VNPNLSKL*	*2164*	*P15318.2*
1282	*ПАЛОЧКА КОКЛЮША*	*VNPNLSKLF*	*2165*	*P15318.2*
1283	*ПАЛОЧКА КОКЛЮША*	*VQQPIIEKL*	*2166*	*P15318.2*
1284	*ПАЛОЧКА КОКЛЮША*	*VQYRHVEL*	*2167*	*P15318.2*
1285	*ПАЛОЧКА КОКЛЮША*	*VSIAAAASV*	*2168*	*P15318.2*
1286	*ПАЛОЧКА КОКЛЮША*	*VSIAAAASVV*	*2169*	*P15318.2*
1287	*ПАЛОЧКА КОКЛЮША*	*VTSLLTGAL*	*2170*	*P15318.2*
1288	*ПАЛОЧКА КОКЛЮША*	*VVLANASRI*	*2171*	*P15318.2*
1289	*ПАЛОЧКА КОКЛЮША*	*WPALNLFSV*	*2172*	*P15318.2*
1290	*ПАЛОЧКА КОКЛЮША*	*WVRKASAL*	*2173*	*P15318.2*
1291	*ПАЛОЧКА КОКЛЮША*	*YAVQYRRKGG*	*2174*	*P15318.2*
1292	*ПАЛОЧКА КОКЛЮША*	*YGGLGDDTL*	*2175*	*P15318.2*
1293	*ПАЛОЧКА КОКЛЮША*	*YGLVQQSHYA*	*2176*	*P15318.2*
1294	*ПАЛОЧКА КОКЛЮША*	*YGYEGDALL*	*2177*	*P15318.2*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1295	ПАЛОЧКА КОКЛЮША	YIPVNPNL	2178	P15318.2
1296	ПАЛОЧКА КОКЛЮША	YSAMIHPGRI	2179	P15318.2
1297	ПАЛОЧКА КОКЛЮША	YSQTGAHAGI	2180	P15318.2
1298	ДИФТЕРИЙНАЯ ПАЛОЧКА	AYNFVESIINLFQVVHNSYN	2181	CAE11230.1
1299	ПАЛОЧКА КОКЛЮША	SGTTIK	2182	BAF35031.1
1300	ПАЛОЧКА КОКЛЮША	RGHTLESAEGRKIFG	2183	AAA22974.1
1301	ПАЛОЧКА КОКЛЮША	AGAMTVRDVAAAADLALQAGDA	2184	AAA22974.1
1302	ПАЛОЧКА КОКЛЮША	AGAMTVRDVAAAADLALQAGDAL	2185	AAA22974.1
1303	ПАЛОЧКА КОКЛЮША	ALAAVLVNPHIFTRIGAAQTSLADGAAGPA	2186	AAA22974.1
1304	ПАЛОЧКА КОКЛЮША	ALSIDSMTALGA	2187	AAA22974.1
1305	ПАЛОЧКА КОКЛЮША	DLSAARGADISGEGR	2188	AAA22974.1
1306	ПАЛОЧКА КОКЛЮША	DQNRYEYIWGLY	2189	AAA22974.1
1307	ПАЛОЧКА КОКЛЮША	DYTVSADAIALA	2190	AAA22974.1
1308	ПАЛОЧКА КОКЛЮША	GPIVVEAGELVSHAGG	2191	AAA22974.1
1309	ПАЛОЧКА КОКЛЮША	GRPEGLKIGAHSATSVSGSFDAL	2192	AAA22974.1
1310	ПАЛОЧКА	ITVTSRGGFDNEGKMESNK	2193	AAA22974.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1311	ПАЛОЧКА КОКЛЮША	LDQNRYEYIWGLYP	2194	AAA22974.1
1312	ПАЛОЧКА КОКЛЮША	LSAARGADISG	2195	AAA22974.1
1313	ПАЛОЧКА КОКЛЮША	NKIRLMGPLQ	2196	AAA22974.1
1314	ПАЛОЧКА КОКЛЮША	NKLGRIRAGEDM	2197	AAA22974.1
1315	ПАЛОЧКА КОКЛЮША	NKLGRIRAGEDMHLDAPRIE	2198	AAA22974.1
1316	ПАЛОЧКА КОКЛЮША	PHLRNTGQVVAG	2199	AAA22974.1
1317	ПАЛОЧКА КОКЛЮША	QVDLHDLSAARGADISG	2200	AAA22974.1
1318	ПАЛОЧКА КОКЛЮША	RDVAAAADLALQ	2201	AAA22974.1
1319	ПАЛОЧКА КОКЛЮША	SAARGADISGEG	2202	AAA22974.1
1320	ПАЛОЧКА КОКЛЮША	TKGEMQIAGKGGGSP	2203	AAA22974.1
1321	ПАЛОЧКА КОКЛЮША	TVSADAIALAAQ	2204	AAA22974.1
1322	ПАЛОЧКА КОКЛЮША	VVAGHDIHI	2205	AAA22974.1
1323	ПАЛОЧКА КОКЛЮША	PSGPNHTKVVQLPKISKNALKANG	2206	CAD12823.1
1324	ВИРУС КРАСНУХИ	LVGATPERPRLRLVDADDPLLRTAPGPGEVWVTPVIGSQAR	2207	BAA19902.1
1325	ВИРУС КРАСНУХИ	QQSRWGLGSPNCHGPDWASPVCQRHSP	2208	BAA19902.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1326	ПАЛОЧКА КОКЛЮША	AGEAMVLVYYESIAYSF	2209	ACI04548.1
1327	ПАЛОЧКА КОКЛЮША	GGVGLASTLWYAESNALSKRLGEL	2210	AAZ74322.1
1328	ПАЛОЧКА КОКЛЮША	GTLVRIAPVIGACMARQA	2211	ACI04548.1
1329	ПАЛОЧКА КОКЛЮША	IRRVTRVYHNGITGETTT	2212	ACI04548.1
1330	ПАЛОЧКА КОКЛЮША	IVKTGERQHGIHIQGSDP	2213	AAZ74322.1
1331	ПАЛОЧКА КОКЛЮША	IVKTGERQHGIHIQGSDPGGVRTA	2214	AAZ74338.1
1332	ПАЛОЧКА КОКЛЮША	LRDTNVTAVPASGAPAAVSVLGAS	2215	AAZ74338.1
1333	ПАЛОЧКА КОКЛЮША	PEAPAPQPPAGRELSAAANAAVNT	2216	AAZ74322.1
1334	ПАЛОЧКА КОКЛЮША	AAADFAHAE	2217	WP_019247158.1
1335	ПАЛОЧКА КОКЛЮША	AAAEVAGAL	2218	WP_019249248.1
1336	ПАЛОЧКА КОКЛЮША	AAESTFESY	2219	WP_019247158.1
1337	ПАЛОЧКА КОКЛЮША	AAGFDPEVQ	2220	WP_019248145.1
1338	ПАЛОЧКА КОКЛЮША	AALGRGHSL	2221	AGS56996.1
1339	ПАЛОЧКА КОКЛЮША	AAMQGAVVH	2222	AGT50936.1
1340	ПАЛОЧКА КОКЛЮША	AAPAAHADW	2223	AGS56996.1
1341	ПАЛОЧКА	AAQATVVQR	2224	WP_0192471

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			*58.1*
1342	*ПАЛОЧКА КОКЛЮША*	*AARVAGDNY*	*2225*	*WP_019249248.1*
1343	*ПАЛОЧКА КОКЛЮША*	*AAVALLNKL*	*2226*	*WP_019249248.1*
1344	*ПАЛОЧКА КОКЛЮША*	*ADDPPATVY*	*2227*	*AAW72734.1*
1345	*ПАЛОЧКА КОКЛЮША*	*AEAGRFKVL*	*2228*	*AGS56996.1*
1346	*ПАЛОЧКА КОКЛЮША*	*AEATQLVTA*	*2229*	*WP_019247158.1*
1347	*ПАЛОЧКА КОКЛЮША*	*AEGGATLGA*	*2230*	*WP_019249248.1*
1348	*ПАЛОЧКА КОКЛЮША*	*AEHGEVSIQ*	*2231*	*WP_019249248.1*
1349	*ПАЛОЧКА КОКЛЮША*	*AEIAFYPKE*	*2232*	*WP_019249248.1*
1350	*ПАЛОЧКА КОКЛЮША*	*AEKVTTPAV*	*2233*	*WP_019247158.1*
1351	*ПАЛОЧКА КОКЛЮША*	*AELQTYLRQ*	*2234*	*1BCP_C*
1352	*ПАЛОЧКА КОКЛЮША*	*AEQSLIEVG*	*2235*	*WP_019249248.1*
1353	*ПАЛОЧКА КОКЛЮША*	*AESNALSKR*	*2236*	*AGS56996.1*
1354	*ПАЛОЧКА КОКЛЮША*	*AESSEAMAA*	*2237*	*AFK26302.1*
1355	*ПАЛОЧКА КОКЛЮША*	*AEVKVGYRA*	*2238*	*WP_019247158.1*
1356	*ПАЛОЧКА КОКЛЮША*	*AEVTDTSPS*	*2239*	*WP_019249248.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1357	ПАЛОЧКА КОКЛЮША	AGKSLKKKN	2240	WP_019247158.1
1358	ПАЛОЧКА КОКЛЮША	AGLAGPSAV	2241	WP_019249248.1
1359	ПАЛОЧКА КОКЛЮША	AIRVGRGAR	2242	AGT50936.1
1360	ПАЛОЧКА КОКЛЮША	ALAAIASAA	2243	WP_019248145.1
1361	ПАЛОЧКА КОКЛЮША	ALADVPYVL	2244	AAA22983.1
1362	ПАЛОЧКА КОКЛЮША	ALANDGTIV	2245	WP_019248658.1
1363	ПАЛОЧКА КОКЛЮША	ALGRGHSLY	2246	AGS56996.1
1364	ПАЛОЧКА КОКЛЮША	ALILAASPV	2247	WP_019248658.1
1365	ПАЛОЧКА КОКЛЮША	ALMLACTGL	2248	AAA22974.1
1366	ПАЛОЧКА КОКЛЮША	AMQGAVVHL	2249	AGT50936.1
1367	ПАЛОЧКА КОКЛЮША	AMTHLSPAL	2250	AAA22983.1
1368	ПАЛОЧКА КОКЛЮША	AMYGKHITL	2251	WP_019249248.1
1369	ПАЛОЧКА КОКЛЮША	ANEANALLW	2252	WP_019249248.1
1370	ПАЛОЧКА КОКЛЮША	APLSITLQA	2253	AGT50936.1
1371	ПАЛОЧКА КОКЛЮША	APNALAWAL	2254	AAA22974.1
1372	ПАЛОЧКА	APPAPKPAP	2255	AGS56996.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1373	ПАЛОЧКА КОКЛЮША	APPGAGFIY	2256	1BCP_C
1374	ПАЛОЧКА КОКЛЮША	APQAAPLSI	2257	AGT50936.1
1375	ПАЛОЧКА КОКЛЮША	APRIENTAK	2258	WP_019249248.1
1376	ПАЛОЧКА КОКЛЮША	AQGKALLYR	2259	AGT50936.1
1377	ПАЛОЧКА КОКЛЮША	AQITSYVGF	2260	WP_019248658.1
1378	ПАЛОЧКА КОКЛЮША	AQLEVRGQR	2261	WP_019249248.1
1379	ПАЛОЧКА КОКЛЮША	AQQLKQADR	2262	WP_019247699.1
1380	ПАЛОЧКА КОКЛЮША	AQVTVAGRY	2263	WP_019249248.1
1381	ПАЛОЧКА КОКЛЮША	ARRSRVRAL	2264	NP_882284.1
1382	ПАЛОЧКА КОКЛЮША	ASPRRARRA	2265	WP_019249248.1
1383	ПАЛОЧКА КОКЛЮША	ASSPDAHVP	2266	AAA22983.1
1384	ПАЛОЧКА КОКЛЮША	ASVSNPGTF	2267	WP_019249248.1
1385	ПАЛОЧКА КОКЛЮША	ATWNFQSTY	2268	WP_019249248.1
1386	ПАЛОЧКА КОКЛЮША	ATYIADSGF	2269	AGS56996.1
1387	ПАЛОЧКА КОКЛЮША	AVAAPAVGA	2270	WP_019249248.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1388	ПАЛОЧКА КОКЛЮША	AVFMQQRPL	2271	AAA22983.1
1389	ПАЛОЧКА КОКЛЮША	AVLVNPHIF	2272	WP_019249248.1
1390	ПАЛОЧКА КОКЛЮША	CFGKDLKRP	2273	AAA22983.1
1391	ПАЛОЧКА КОКЛЮША	CPSSLGNGV	2274	WP_019248145.1
1392	ПАЛОЧКА КОКЛЮША	DAGHEHDTW	2275	AAA22984.1
1393	ПАЛОЧКА КОКЛЮША	DAKHDLTVT	2276	WP_019249248.1
1394	ПАЛОЧКА КОКЛЮША	DASGQHRLW	2277	AGS56996.1
1395	ПАЛОЧКА КОКЛЮША	DATFETYAL	2278	YP_006628018.1
1396	ПАЛОЧКА КОКЛЮША	DATFQTYAL	2279	1BCP_C
1397	ПАЛОЧКА КОКЛЮША	DATLVGAKF	2280	WP_019247158.1
1398	ПАЛОЧКА КОКЛЮША	DDEVDVSGR	2281	AAA22974.1
1399	ПАЛОЧКА КОКЛЮША	DENGKPQTY	2282	WP_019247158.1
1400	ПАЛОЧКА КОКЛЮША	DGPPSRPTT	2283	WP_019247158.1
1401	ПАЛОЧКА КОКЛЮША	DHLTGRSCQ	2284	AAW72734.1
1402	ПАЛОЧКА КОКЛЮША	DNEGKMESN	2285	WP_019249248.1
1403	ПАЛОЧКА	DPPATVYRY	2286	AAW72734.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1404	ПАЛОЧКА КОКЛЮША	EATEGDATL	2287	WP_019247158.1
1405	ПАЛОЧКА КОКЛЮША	EATQQAAGF	2288	WP_019248145.1
1406	ПАЛОЧКА КОКЛЮША	ECSGKQDCP	2289	AAA22983.1
1407	ПАЛОЧКА КОКЛЮША	EGGKLRGKD	2290	AAA22974.1
1408	ПАЛОЧКА КОКЛЮША	EGKMESNKD	2291	WP_019249248.1
1409	ПАЛОЧКА КОКЛЮША	EHRMQEAVE	2292	AAW72734.1
1410	ПАЛОЧКА КОКЛЮША	EKRLDIDDA	2293	WP_019249248.1
1411	ПАЛОЧКА КОКЛЮША	EPASANTLL	2294	AGS56996.1
1412	ПАЛОЧКА КОКЛЮША	EPQAELAVF	2295	AGS56996.1
1413	ПАЛОЧКА КОКЛЮША	EPVKLTLTG	2296	AGT50936.1
1414	ПАЛОЧКА КОКЛЮША	ESAEGRKIF	2297	WP_019249248.1
1415	ПАЛОЧКА КОКЛЮША	ESYSESHNF	2298	WP_019247158.1
1416	ПАЛОЧКА КОКЛЮША	ETFCITTIY	2299	1BCP_C
1417	ПАЛОЧКА КОКЛЮША	EVAGALELS	2300	WP_019249248.1
1418	ПАЛОЧКА КОКЛЮША	EVAKVEVVP	2301	WP_019247158.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1419	ПАЛОЧКА КОКЛЮША	EVDGIIQEF	2302	WP_019249248.1
1420	ПАЛОЧКА КОКЛЮША	EVRADNNFY	2303	WP_019248344.1
1421	ПАЛОЧКА КОКЛЮША	FAILSSTTE	2304	WP_019247158.1
1422	ПАЛОЧКА КОКЛЮША	FAISAYALK	2305	AAA22984.1
1423	ПАЛОЧКА КОКЛЮША	FALYDGTYL	2306	AFK26303.1
1424	ПАЛОЧКА КОКЛЮША	FDTMLGFAI	2307	AAA22984.1
1425	ПАЛОЧКА КОКЛЮША	FEGKPALEL	2308	AAA22983.1
1426	ПАЛОЧКА КОКЛЮША	FELGADHAV	2309	AGS56996.1
1427	ПАЛОЧКА КОКЛЮША	FEPGITTNY	2310	WP_019248658.1
1428	ПАЛОЧКА КОКЛЮША	FETYALTGI	2311	YP_006628018.1
1429	ПАЛОЧКА КОКЛЮША	FIYRETFCI	2312	1BCP_C
1430	ПАЛОЧКА КОКЛЮША	FPTRTTAPG	2313	NP_882284.1
1431	ПАЛОЧКА КОКЛЮША	FQTYALTGI	2314	1BCP_C
1432	ПАЛОЧКА КОКЛЮША	FTHADGWFL	2315	AGS56996.1
1433	ПАЛОЧКА КОКЛЮША	FVRDGQSVI	2316	1BCP_C
1434	ПАЛОЧКА	FVRSGQPVI	2317	YP_00662801

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			8.1
1435	ПАЛОЧКА КОКЛЮША	FVWYVDTVI	2318	WP_019248866.1
1436	ПАЛОЧКА КОКЛЮША	GAASSRQAL	2319	WP_019249248.1
1437	ПАЛОЧКА КОКЛЮША	GAASSYFEY	2320	AAW72734.1
1438	ПАЛОЧКА КОКЛЮША	GAFDLKTTF	2321	AFK26303.1
1439	ПАЛОЧКА КОКЛЮША	GAPAAVSVL	2322	AGS56996.1
1440	ПАЛОЧКА КОКЛЮША	GATRAVDSL	2323	AGS56996.1
1441	ПАЛОЧКА КОКЛЮША	GAVPGGAVP	2324	AGT50936.1
1442	ПАЛОЧКА КОКЛЮША	GEAMVLVYY	2325	AFK26302.1
1443	ПАЛОЧКА КОКЛЮША	GEIALGDAT	2326	WP_019249248.1
1444	ПАЛОЧКА КОКЛЮША	GELMAAQVA	2327	WP_019247158.1
1445	ПАЛОЧКА КОКЛЮША	GGVPGGAVP	2328	AAZ74338.1
1446	ПАЛОЧКА КОКЛЮША	GHEHDTWFD	2329	AAA22984.1
1447	ПАЛОЧКА КОКЛЮША	GIGALKAGA	2330	WP_019249248.1
1448	ПАЛОЧКА КОКЛЮША	GIVIPPKAL	2331	1BCP_C
1449	ПАЛОЧКА КОКЛЮША	GKDLKRPGS	2332	AAA22983.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1450	ПАЛОЧКА КОКЛЮША	GKLPKPVTV	2333	WP_019247158.1
1451	ПАЛОЧКА КОКЛЮША	GKSLKKKNQ	2334	WP_019247158.1
1452	ПАЛОЧКА КОКЛЮША	GLDVQQGTV	2335	WP_019249248.1
1453	ПАЛОЧКА КОКЛЮША	GLTDGVSRI	2336	WP_019249248.1
1454	ПАЛОЧКА КОКЛЮША	GLYPTYTEW	2337	WP_019249248.1
1455	ПАЛОЧКА КОКЛЮША	GLYQTYTEW	2338	YP_006626470.1
1456	ПАЛОЧКА КОКЛЮША	GPPSRPTTP	2339	WP_019247158.1
1457	ПАЛОЧКА КОКЛЮША	GPSAVAAPA	2340	WP_019249248.1
1458	ПАЛОЧКА КОКЛЮША	GVAPTAQQL	2341	WP_019248866.1
1459	ПАЛОЧКА КОКЛЮША	GVGLASTLW	2342	AGS56996.1
1460	ПАЛОЧКА КОКЛЮША	GYEAGFSLG	2343	WP_019247158.1
1461	ПАЛОЧКА КОКЛЮША	HADDGTIVI	2344	WP_019248145.1
1462	ПАЛОЧКА КОКЛЮША	HADWNNQSI	2345	AGS56996.1
1463	ПАЛОЧКА КОКЛЮША	HAEHEKDVR	2346	WP_019247158.1
1464	ПАЛОЧКА КОКЛЮША	HANHYGTRI	2347	WP_019247158.1
1465	ПАЛОЧКА	HAQGKALLY	2348	AGT50936.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1466	ПАЛОЧКА КОКЛЮША	HFIGYIYEV	2349	AAW72734.1
1467	ПАЛОЧКА КОКЛЮША	HLSPALADV	2350	AAA22983.1
1468	ПАЛОЧКА КОКЛЮША	HSLYASYEY	2351	AGS56996.1
1469	ПАЛОЧКА КОКЛЮША	HVRGMLVPV	2352	AAA22974.1
1470	ПАЛОЧКА КОКЛЮША	HVSKEEQYY	2353	YP_006628018.1
1471	ПАЛОЧКА КОКЛЮША	HVTRGWSIF	2354	AFK26303.1
1472	ПАЛОЧКА КОКЛЮША	IADSGFYLD	2355	AGS56996.1
1473	ПАЛОЧКА КОКЛЮША	IAHRTELRG	2356	AGS56996.1
1474	ПАЛОЧКА КОКЛЮША	IENTAKLSG	2357	WP_019249248.1
1475	ПАЛОЧКА КОКЛЮША	IESKISQSV	2358	WP_019249248.1
1476	ПАЛОЧКА КОКЛЮША	IETGGARRF	2359	AGT50936.1
1477	ПАЛОЧКА КОКЛЮША	IIKDAPPGA	2360	1BCP_C
1478	ПАЛОЧКА КОКЛЮША	IIQEFAADL	2361	WP_019249248.1
1479	ПАЛОЧКА КОКЛЮША	ILAGALATY	2362	AAW72734.1
1480	ПАЛОЧКА КОКЛЮША	ILLENPAAE	2363	AGS56996.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1481	ПАЛОЧКА КОКЛЮША	ILPILVLAL	2364	NP_882286.1
1482	ПАЛОЧКА КОКЛЮША	IPFQRALRL	2365	WP_019248145.1
1483	ПАЛОЧКА КОКЛЮША	ISVRVHVSK	2366	YP_006628018.1
1484	ПАЛОЧКА КОКЛЮША	ITNETGKTY	2367	WP_019247158.1
1485	ПАЛОЧКА КОКЛЮША	ITNKRAALI	2368	WP_019249248.1
1486	ПАЛОЧКА КОКЛЮША	ITSYVGFSV	2369	WP_019248658.1
1487	ПАЛОЧКА КОКЛЮША	ITVTSRGGF	2370	WP_019249248.1
1488	ПАЛОЧКА КОКЛЮША	IVIPPKALF	2371	1BCP_C
1489	ПАЛОЧКА КОКЛЮША	IVVEAGELV	2372	WP_019249248.1
1490	ПАЛОЧКА КОКЛЮША	KAAKSVNLM	2373	WP_019247158.1
1491	ПАЛОЧКА КОКЛЮША	KAAPLRRTT	2374	AGS56996.1
1492	ПАЛОЧКА КОКЛЮША	KAGKLSATG	2375	WP_019249248.1
1493	ПАЛОЧКА КОКЛЮША	KAGTIAAPW	2376	WP_019249248.1
1494	ПАЛОЧКА КОКЛЮША	KAKSLTTEI	2377	WP_019249248.1
1495	ПАЛОЧКА КОКЛЮША	KATVTTVQV	2378	WP_019247158.1
1496	ПАЛОЧКА	KDYRDKDGG	2379	WP_019247

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			*58.1*
1497	*ПАЛОЧКА КОКЛЮША*	*KEAATIVAA*	*2380*	*WP_019249248.1*
1498	*ПАЛОЧКА КОКЛЮША*	*KEDVDAAQI*	*2381*	*WP_019248658.1*
1499	*ПАЛОЧКА КОКЛЮША*	*KEVDGIIQE*	*2382*	*WP_019249248.1*
1500	*ПАЛОЧКА КОКЛЮША*	*KGPKLAMPW*	*2383*	*AGS56996.1*
1501	*ПАЛОЧКА КОКЛЮША*	*KLASGGGAV*	*2384*	*WP_019249248.1*
1502	*ПАЛОЧКА КОКЛЮША*	*KLKGKNQEF*	*2385*	*AAA22984.1*
1503	*ПАЛОЧКА КОКЛЮША*	*KLLHHILPI*	*2386*	*NP_882286.1*
1504	*ПАЛОЧКА КОКЛЮША*	*KPAPQPGPQ*	*2387*	*AGS56996.1*
1505	*ПАЛОЧКА КОКЛЮША*	*KPAPTAPPM*	*2388*	*WP_019249248.1*
1506	*ПАЛОЧКА КОКЛЮША*	*KPAVSVKVA*	*2389*	*WP_019249248.1*
1507	*ПАЛОЧКА КОКЛЮША*	*KPDRAARVA*	*2390*	*WP_019249248.1*
1508	*ПАЛОЧКА КОКЛЮША*	*KPLADIAVI*	*2391*	*YP_006626470.1*
1509	*ПАЛОЧКА КОКЛЮША*	*KPLADIAVV*	*2392*	*WP_019249248.1*
1510	*ПАЛОЧКА КОКЛЮША*	*KPLPKPLPV*	*2393*	*WP_019247158.1*
1511	*ПАЛОЧКА КОКЛЮША*	*KQADRDFVW*	*2394*	*WP_019247699.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1512	ПАЛОЧКА КОКЛЮША	KSLPGGKLP	2395	WP_019247158.1
1513	ПАЛОЧКА КОКЛЮША	KSYTLRYLA	2396	WP_019248658.1
1514	ПАЛОЧКА КОКЛЮША	KTNMVVTSV	2397	AAA22983.1
1515	ПАЛОЧКА КОКЛЮША	KVLAPRLYL	2398	AAA22974.1
1516	ПАЛОЧКА КОКЛЮША	KVLSTKTTL	2399	WP_019247158.1
1517	ПАЛОЧКА КОКЛЮША	LAAGAGLTL	2400	WP_019249248.1
1518	ПАЛОЧКА КОКЛЮША	LAAIASAAH	2401	WP_019248145.1
1519	ПАЛОЧКА КОКЛЮША	LAANGNGQW	2402	AGS56996.1
1520	ПАЛОЧКА КОКЛЮША	LAAQVTQRG	2403	WP_019249248.1
1521	ПАЛОЧКА КОКЛЮША	LAARGDGAL	2404	AAA22974.1
1522	ПАЛОЧКА КОКЛЮША	LAAVLVNPH	2405	WP_019249248.1
1523	ПАЛОЧКА КОКЛЮША	LAGSGLFRM	2406	AGS56996.1
1524	ПАЛОЧКА КОКЛЮША	LAKALSAAL	2407	WP_019249248.1
1525	ПАЛОЧКА КОКЛЮША	LALQAGDAL	2408	WP_019249248.1
1526	ПАЛОЧКА КОКЛЮША	LAMPWTFHA	2409	AGS56996.1
1527	ПАЛОЧКА	LANDGTIVI	2410	WP_019248658

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			*58.1*
1528	*ПАЛОЧКА КОКЛЮША*	*LAPTVGVAF*	*2411*	*WP_019247158.1*
1529	*ПАЛОЧКА КОКЛЮША*	*LASDGSVDF*	*2412*	*AGS56996.1*
1530	*ПАЛОЧКА КОКЛЮША*	*LATYQSEYL*	*2413*	*AAW72734.1*
1531	*ПАЛОЧКА КОКЛЮША*	*LAWALMLAC*	*2414*	*AAA22974.1*
1532	*ПАЛОЧКА КОКЛЮША*	*LEAGRRFTH*	*2415*	*AGS56996.1*
1533	*ПАЛОЧКА КОКЛЮША*	*LFTQQGGAY*	*2416*	*1BCP_C*
1534	*ПАЛОЧКА КОКЛЮША*	*LKRPGSSPM*	*2417*	*AAA22983.1*
1535	*ПАЛОЧКА КОКЛЮША*	*LLGSHVARA*	*2418*	*AFK26303.1*
1536	*ПАЛОЧКА КОКЛЮША*	*LLHHILPIL*	*2419*	*NP_882286.1*
1537	*ПАЛОЧКА КОКЛЮША*	*LLNAGGTLL*	*2420*	*WP_019249248.1*
1538	*ПАЛОЧКА КОКЛЮША*	*LMGPLQVNA*	*2421*	*WP_019249248.1*
1539	*ПАЛОЧКА КОКЛЮША*	*LNDSKITMG*	*2422*	*WP_019248145.1*
1540	*ПАЛОЧКА КОКЛЮША*	*LPEPVKLTL*	*2423*	*AGT50936.1*
1541	*ПАЛОЧКА КОКЛЮША*	*LPILVLALL*	*2424*	*NP_882286.1*
1542	*ПАЛОЧКА КОКЛЮША*	*LPKISKNAL*	*2425*	*WP_019248145.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1543	ПАЛОЧКА КОКЛЮША	LPKPVTVKL	2426	WP_019247158.1
1544	ПАЛОЧКА КОКЛЮША	LPLKANPMH	2427	NP_882285.1
1545	ПАЛОЧКА КОКЛЮША	LPPRPVVAE	2428	WP_019247158.1
1546	ПАЛОЧКА КОКЛЮША	LPSIPGTSI	2429	AGS56996.1
1547	ПАЛОЧКА КОКЛЮША	LPTHLYKNF	2430	AAA22984.1
1548	ПАЛОЧКА КОКЛЮША	LPVRGVALV	2431	WP_019249248.1
1549	ПАЛОЧКА КОКЛЮША	LPVSLTALD	2432	WP_019249248.1
1550	ПАЛОЧКА КОКЛЮША	LQGGNKVPV	2433	WP_019249248.1
1551	ПАЛОЧКА КОКЛЮША	LSAALGADW	2434	WP_019249248.1
1552	ПАЛОЧКА КОКЛЮША	LSDAGHEHD	2435	AAA22984.1
1553	ПАЛОЧКА КОКЛЮША	LSGEVQRKG	2436	WP_019249248.1
1554	ПАЛОЧКА КОКЛЮША	LSSPSAITV	2437	WP_019249248.1
1555	ПАЛОЧКА КОКЛЮША	LTWLAILAV	2438	AAW72734.1
1556	ПАЛОЧКА КОКЛЮША	LVFSHVRGM	2439	AAA22974.1
1557	ПАЛОЧКА КОКЛЮША	LVSDAGADL	2440	WP_019249248.1
1558	ПАЛОЧКА	LVYYESIAY	2441	AFK26302.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1559	ПАЛОЧКА КОКЛЮША	MAAESTFES	2442	WP_019247158.1
1560	ПАЛОЧКА КОКЛЮША	MAAGHDATL	2443	WP_019249248.1
1561	ПАЛОЧКА КОКЛЮША	MAAWSERAG	2444	AFK26302.1
1562	ПАЛОЧКА КОКЛЮША	MALGALGAA	2445	AGS56996.1
1563	ПАЛОЧКА КОКЛЮША	MAPVMGACM	2446	ADA85123.1
1564	ПАЛОЧКА КОКЛЮША	MATKGEMQI	2447	WP_019249248.1
1565	ПАЛОЧКА КОКЛЮША	MDAKGGTLL	2448	WP_019249248.1
1566	ПАЛОЧКА КОКЛЮША	MESNKDIVI	2449	WP_019249248.1
1567	ПАЛОЧКА КОКЛЮША	MEVMLRAVF	2450	AAA22983.1
1568	ПАЛОЧКА КОКЛЮША	MEYFKTPLP	2451	WP_019249248.1
1569	ПАЛОЧКА КОКЛЮША	MHTIASILL	2452	AAA22984.1
1570	ПАЛОЧКА КОКЛЮША	MIYMSGLAV	2453	1BCP_C
1571	ПАЛОЧКА КОКЛЮША	MLACTGLPL	2454	AAA22974.1
1572	ПАЛОЧКА КОКЛЮША	MPIDRKTLC	2455	AFK26303.1
1573	ПАЛОЧКА КОКЛЮША	MPKAPELDL	2456	WP_019249248.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1574	ПАЛОЧКА КОКЛЮША	MQRQAGLPL	2457	NP_882285.1
1575	ПАЛОЧКА КОКЛЮША	NALAWALML	2458	AAA22974.1
1576	ПАЛОЧКА КОКЛЮША	NITNKRAAL	2459	WP_019249248.1
1577	ПАЛОЧКА КОКЛЮША	NLMAAESTF	2460	WP_019247158.1
1578	ПАЛОЧКА КОКЛЮША	NNETMSGRQ	2461	WP_019249248.1
1579	ПАЛОЧКА КОКЛЮША	NPGSLIAEV	2462	WP_019249248.1
1580	ПАЛОЧКА КОКЛЮША	NPMHTIASI	2463	NP_882285.1
1581	ПАЛОЧКА КОКЛЮША	NPQTQLSNI	2464	WP_019248145.1
1582	ПАЛОЧКА КОКЛЮША	NPYTSRRSV	2465	AFK26302.1
1583	ПАЛОЧКА КОКЛЮША	PAPTAPPMP	2466	WP_019249248.1
1584	ПАЛОЧКА КОКЛЮША	PASANTLLL	2467	AGS56996.1
1585	ПАЛОЧКА КОКЛЮША	PAVALPRPL	2468	WP_019249248.1
1586	ПАЛОЧКА КОКЛЮША	PDAHVPFCF	2469	AAA22983.1
1587	ПАЛОЧКА КОКЛЮША	PELGAAIRV	2470	AGT50936.1
1588	ПАЛОЧКА КОКЛЮША	PFIIKLKDC	2471	WP_019248145.1
1589	ПАЛОЧКА	PGPQPPQPP	2472	AGS56996.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1590	ПАЛОЧКА КОКЛЮША	PGPQPPQPQ	2473	AAZ74338.1
1591	ПАЛОЧКА КОКЛЮША	PGPTTDYST	2474	WP_019248145.1
1592	ПАЛОЧКА КОКЛЮША	PGTFTAGKD	2475	WP_019249248.1
1593	ПАЛОЧКА КОКЛЮША	PGTPGDLLE	2476	AAA22984.1
1594	ПАЛОЧКА КОКЛЮША	PKPKPKAER	2477	WP_019247158.1
1595	ПАЛОЧКА КОКЛЮША	PKPKPKPKA	2478	WP_019247158.1
1596	ПАЛОЧКА КОКЛЮША	PKPKPKPKP	2479	WP_019247158.1
1597	ПАЛОЧКА КОКЛЮША	PLPPRPVVA	2480	WP_019247158.1
1598	ПАЛОЧКА КОКЛЮША	PPAPKPAPQ	2481	AGS56996.1
1599	ПАЛОЧКА КОКЛЮША	PPKPAPVAK	2482	WP_019247158.1
1600	ПАЛОЧКА КОКЛЮША	PPRPVAAQV	2483	WP_019247158.1
1601	ПАЛОЧКА КОКЛЮША	PPRPVVAEK	2484	WP_019247158.1
1602	ПАЛОЧКА КОКЛЮША	PPSRPTTPP	2485	WP_019247158.1
1603	ПАЛОЧКА КОКЛЮША	PRRARRALR	2486	WP_019249248.1
1604	ПАЛОЧКА КОКЛЮША	QAAPLSITL	2487	AGT50936.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1605	ПАЛОЧКА КОКЛЮША	QADRDFVWY	2488	WP_019247699.1
1606	ПАЛОЧКА КОКЛЮША	QAIVVGKDL	2489	WP_019249248.1
1607	ПАЛОЧКА КОКЛЮША	QALGALKLY	2490	ACI16088.1
1608	ПАЛОЧКА КОКЛЮША	QELALKLKG	2491	AAA22984.1
1609	ПАЛОЧКА КОКЛЮША	QITQHGGPY	2492	NP_882283.1
1610	ПАЛОЧКА КОКЛЮША	QITQHGSPY	2493	AFK26303.1
1611	ПАЛОЧКА КОКЛЮША	QPLPPRPVA	2494	WP_019247158.1
1612	ПАЛОЧКА КОКЛЮША	QPPAGRELS	2495	AGS56996.1
1613	ПАЛОЧКА КОКЛЮША	QQLKQADRD	2496	WP_019247699.1
1614	ПАЛОЧКА КОКЛЮША	QQVQVLQRQ	2497	WP_019247158.1
1615	ПАЛОЧКА КОКЛЮША	QSIVEAPEL	2498	AGT50936.1
1616	ПАЛОЧКА КОКЛЮША	QVGSSNSAF	2499	AAW72734.1
1617	ПАЛОЧКА КОКЛЮША	RAGLSPATW	2500	WP_019249248.1
1618	ПАЛОЧКА КОКЛЮША	RARRALRQD	2501	WP_019249248.1
1619	ПАЛОЧКА КОКЛЮША	RASASRARI	2502	WP_019249248.1
1620	ПАЛОЧКА	RELSAAANA	2503	AGS56996.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1621	*ПАЛОЧКА КОКЛЮША*	*RETFCITTI*	*2504*	*1BCP_C*
1622	*ПАЛОЧКА КОКЛЮША*	*RGFAQRQQL*	*2505*	*AGS56996.1*
1623	*ПАЛОЧКА КОКЛЮША*	*RGSAATFTL*	*2506*	*WP_019248295.1*
1624	*ПАЛОЧКА КОКЛЮША*	*RGWSIFALY*	*2507*	*AFK26303.1*
1625	*ПАЛОЧКА КОКЛЮША*	*RKMLYLIYV*	*2508*	*YP_006628018.1*
1626	*ПАЛОЧКА КОКЛЮША*	*RLRKMLYLI*	*2509*	*YP_006628018.1*
1627	*ПАЛОЧКА КОКЛЮША*	*RPQITDAVT*	*2510*	*WP_019249248.1*
1628	*ПАЛОЧКА КОКЛЮША*	*RPSVNGGRI*	*2511*	*WP_019249248.1*
1629	*ПАЛОЧКА КОКЛЮША*	*RRFTHADGW*	*2512*	*AGS56996.1*
1630	*ПАЛОЧКА КОКЛЮША*	*RSGARATSL*	*2513*	*AAA22974.1*
1631	*ПАЛОЧКА КОКЛЮША*	*RSRVRALAW*	*2514*	*NP_882284.1*
1632	*ПАЛОЧКА КОКЛЮША*	*RSRVRALSW*	*2515*	*YP_006628019.1*
1633	*ПАЛОЧКА КОКЛЮША*	*RTHGVGASL*	*2516*	*AGS56996.1*
1634	*ПАЛОЧКА КОКЛЮША*	*RTRGQARSG*	*2517*	*AAA22974.1*
1635	*ПАЛОЧКА КОКЛЮША*	*RVAPPAVAL*	*2518*	*WP_019249248.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1636	ПАЛОЧКА КОКЛЮША	RVLPEPVKL	2519	AGT50936.1
1637	ПАЛОЧКА КОКЛЮША	RVRALAWLL	2520	NP_882284.1
1638	ПАЛОЧКА КОКЛЮША	RVRALSWLL	2521	YP_006628019.1
1639	ПАЛОЧКА КОКЛЮША	RVTVSGGSL	2522	AGT50936.1
1640	ПАЛОЧКА КОКЛЮША	SEAMAAWSE	2523	AFK26302.1
1641	ПАЛОЧКА КОКЛЮША	SESHNFHAS	2524	WP_019247158.1
1642	ПАЛОЧКА КОКЛЮША	SGEGRVNIG	2525	WP_019249248.1
1643	ПАЛОЧКА КОКЛЮША	SGLAVRVHV	2526	1BCP_C
1644	ПАЛОЧКА КОКЛЮША	SLADISLGA	2527	WP_019249248.1
1645	ПАЛОЧКА КОКЛЮША	SLFAILSST	2528	WP_019247158.1
1646	ПАЛОЧКА КОКЛЮША	SLFAPHGNV	2529	AAZ74338.1
1647	ПАЛОЧКА КОКЛЮША	SLSIDNATW	2530	AGS56996.1
1648	ПАЛОЧКА КОКЛЮША	SPMEVMLRA	2531	AAA22983.1
1649	ПАЛОЧКА КОКЛЮША	SPQPIRATV	2532	WP_019247158.1
1650	ПАЛОЧКА КОКЛЮША	SPRRARRAL	2533	WP_019249248.1
1651	ПАЛОЧКА	SPSRLAGTL	2534	WP_0192492

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			48.1
1652	ПАЛОЧКА КОКЛЮША	SSTPLGSLF	2535	WP_019247158.1
1653	ПАЛОЧКА КОКЛЮША	STYELLDYL	2536	WP_019249248.1
1654	ПАЛОЧКА КОКЛЮША	SVAMKPYEV	2537	AAA22983.1
1655	ПАЛОЧКА КОКЛЮША	SVAPNALAW	2538	AAA22974.1
1656	ПАЛОЧКА КОКЛЮША	SVKVAKKLF	2539	WP_019249248.1
1657	ПАЛОЧКА КОКЛЮША	TAFMSGRSL	2540	AAA22984.1
1658	ПАЛОЧКА КОКЛЮША	TAGATPFDI	2541	WP_019248658.1
1659	ПАЛОЧКА КОКЛЮША	TAPVTSPAW	2542	AAW72734.1
1660	ПАЛОЧКА КОКЛЮША	TARTGWLTW	2543	AAW72734.1
1661	ПАЛОЧКА КОКЛЮША	TEAQGVQVR	2544	WP_019248145.1
1662	ПАЛОЧКА КОКЛЮША	TEVYLEHRM	2545	AAW72734.1
1663	ПАЛОЧКА КОКЛЮША	TFEGKPALE	2546	AAA22983.1
1664	ПАЛОЧКА КОКЛЮША	TFTGKVTNG	2547	WP_019248658.1
1665	ПАЛОЧКА КОКЛЮША	TGDGGGHTD	2548	AGS56996.1
1666	ПАЛОЧКА КОКЛЮША	TLAKALSAA	2549	WP_019249248.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1667	ПАЛОЧКА КОКЛЮША	TLANVGDTW	2550	AGS56996.1
1668	ПАЛОЧКА КОКЛЮША	TLNASNLTL	2551	WP_019249248.1
1669	ПАЛОЧКА КОКЛЮША	TLSSAHGNV	2552	WP_019249248.1
1670	ПАЛОЧКА КОКЛЮША	TPFDIKLKE	2553	WP_019248658.1
1671	ПАЛОЧКА КОКЛЮША	TPFIIKLKD	2554	WP_019248145.1
1672	ПАЛОЧКА КОКЛЮША	TPGWSIYGL	2555	1BCP_C
1673	ПАЛОЧКА КОКЛЮША	TPLGSAATF	2556	AGS56996.1
1674	ПАЛОЧКА КОКЛЮША	TPLGSLFAI	2557	WP_019247158.1
1675	ПАЛОЧКА КОКЛЮША	TPLPVSLTA	2558	WP_019249248.1
1676	ПАЛОЧКА КОКЛЮША	TRQGIMDQY	2559	YP_006626873.1
1677	ПАЛОЧКА КОКЛЮША	TSKQDERNY	2560	WP_019247158.1
1678	ПАЛОЧКА КОКЛЮША	TSPYDGKYW	2561	YP_006628018.1
1679	ПАЛОЧКА КОКЛЮША	TSRRSVASI	2562	AFK26302.1
1680	ПАЛОЧКА КОКЛЮША	TSRTVTMRY	2563	NP_879898.1
1681	ПАЛОЧКА КОКЛЮША	TTEYPNARY	2564	ADA85123.1
1682	ПАЛОЧКА	TTEYSNARY	2565	AFK26302.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1683	ПАЛОЧКА КОКЛЮША	TTLGLEQTF	2566	WP_019247158.1
1684	ПАЛОЧКА КОКЛЮША	TVLAAGAGL	2567	WP_019249248.1
1685	ПАЛОЧКА КОКЛЮША	TVQELALKL	2568	AAA22984.1
1686	ПАЛОЧКА КОКЛЮША	TVVQRNKHW	2569	WP_019247158.1
1687	ПАЛОЧКА КОКЛЮША	VAAAADLAL	2570	WP_019249248.1
1688	ПАЛОЧКА КОКЛЮША	VALASQARW	2571	AGS56996.1
1689	ПАЛОЧКА КОКЛЮША	VAMKPYEVT	2572	AAA22983.1
1690	ПАЛОЧКА КОКЛЮША	VARLVKLQG	2573	WP_019249248.1
1691	ПАЛОЧКА КОКЛЮША	VAVAGGRWH	2574	AGS56996.1
1692	ПАЛОЧКА КОКЛЮША	VEASAITTY	2575	WP_019248145.1
1693	ПАЛОЧКА КОКЛЮША	VEDIGGKNY	2576	WP_019247158.1
1694	ПАЛОЧКА КОКЛЮША	VEVSSPPPV	2577	WP_019247158.1
1695	ПАЛОЧКА КОКЛЮША	VGAGGYEAG	2578	WP_019247158.1
1696	ПАЛОЧКА КОКЛЮША	VGGGEHGRW	2579	WP_019249248.1
1697	ПАЛОЧКА КОКЛЮША	VHVSKEEQY	2580	YP_006628018.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1698	ПАЛОЧКА КОКЛЮША	VIDGQKVLA	2581	AAA22974.1
1699	ПАЛОЧКА КОКЛЮША	VIGACTSPY	2582	YP_006628018.1
1700	ПАЛОЧКА КОКЛЮША	VKLGGVYEA	2583	AAA22974.1
1701	ПАЛОЧКА КОКЛЮША	VLAPRLYLT	2584	AAA22974.1
1702	ПАЛОЧКА КОКЛЮША	VLVKTNMVV	2585	AAA22983.1
1703	ПАЛОЧКА КОКЛЮША	VPASGAPAA	2586	AGS56996.1
1704	ПАЛОЧКА КОКЛЮША	VPFCFGKDL	2587	AAA22983.1
1705	ПАЛОЧКА КОКЛЮША	VPVSEHCTV	2588	AAA22974.1
1706	ПАЛОЧКА КОКЛЮША	VPVTPPKVE	2589	WP_019247158.1
1707	ПАЛОЧКА КОКЛЮША	VRTVSAMEY	2590	WP_019249248.1
1708	ПАЛОЧКА КОКЛЮША	VSGGSLSAP	2591	AGT50936.1
1709	ПАЛОЧКА КОКЛЮША	VSSATKAKG	2592	WP_019248658.1
1710	ПАЛОЧКА КОКЛЮША	VSSPPPVSV	2593	WP_019247158.1
1711	ПАЛОЧКА КОКЛЮША	VSVKVAKKL	2594	WP_019249248.1
1712	ПАЛОЧКА КОКЛЮША	VTMRYLASY	2595	WP_019248145.1
1713	ПАЛОЧКА	VTSVAMKPY	2596	AAA22983.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1714	*ПАЛОЧКА КОКЛЮША*	*VVAEKVTTP*	*2597*	*WP_019247158.1*
1715	*ПАЛОЧКА КОКЛЮША*	*VVDGPPSRP*	*2598*	*WP_019247158.1*
1716	*ПАЛОЧКА КОКЛЮША*	*VVETAQPLP*	*2599*	*WP_019247158.1*
1717	*ПАЛОЧКА КОКЛЮША*	*WLTWLAILA*	*2600*	*AAW72734.1*
1718	*ПАЛОЧКА КОКЛЮША*	*WTFHAGYRY*	*2601*	*AGS56996.1*
1719	*ПАЛОЧКА КОКЛЮША*	*WVMTDNSNV*	*2602*	*AGS56996.1*
1720	*ПАЛОЧКА КОКЛЮША*	*YAEHGEVSI*	*2603*	*WP_019249248.1*
1721	*ПАЛОЧКА КОКЛЮША*	*YAIDGTAAG*	*2604*	*WP_019249248.1*
1722	*ПАЛОЧКА КОКЛЮША*	*YALKSRIAL*	*2605*	*AAA22984.1*
1723	*ПАЛОЧКА КОКЛЮША*	*YATNPQTQL*	*2606*	*WP_019248145.1*
1724	*ПАЛОЧКА КОКЛЮША*	*YDTGDLIAY*	*2607*	*WP_019248658.1*
1725	*ПАЛОЧКА КОКЛЮША*	*YEAGFSLGS*	*2608*	*WP_019247158.1*
1726	*ПАЛОЧКА КОКЛЮША*	*YEDATFETY*	*2609*	*YP_006628018.1*
1727	*ПАЛОЧКА КОКЛЮША*	*YENKSSTPL*	*2610*	*WP_019247158.1*
1728	*ПАЛОЧКА КОКЛЮША*	*YEVTPTRML*	*2611*	*AAA22983.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1729	ПАЛОЧКА КОКЛЮША	YEYIWGLYP	2612	WP_019249248.1
1730	ПАЛОЧКА КОКЛЮША	YEYIWGLYQ	2613	YP_006626470.1
1731	ПАЛОЧКА КОКЛЮША	YEYSKGPKL	2614	AGS56996.1
1732	ПАЛОЧКА КОКЛЮША	YFEPGPTTD	2615	WP_019248145.1
1733	ПАЛОЧКА КОКЛЮША	YIWGLYPTY	2616	WP_019249248.1
1734	ПАЛОЧКА КОКЛЮША	YIWGLYQTY	2617	YP_006626470.1
1735	ПАЛОЧКА КОКЛЮША	YLRQITPGW	2618	1BCP_C
1736	ПАЛОЧКА КОКЛЮША	YMIYMSGLA	2619	1BCP_C
1737	ПАЛОЧКА КОКЛЮША	YPALRAALI	2620	WP_019248658.1
1738	ПАЛОЧКА КОКЛЮША	YPGTPGDLL	2621	AAA22984.1
1739	ПАЛОЧКА КОКЛЮША	YPTYTEWSV	2622	WP_019249248.1
1740	ПАЛОЧКА КОКЛЮША	YQTYTEWSV	2623	YP_006626470.1
1741	ПАЛОЧКА КОКЛЮША	YSTGDLRAY	2624	WP_019248145.1
1742	ПАЛОЧКА КОКЛЮША	YTLRYLASY	2625	WP_019248658.1
1743	ПАЛОЧКА КОКЛЮША	YVLVKTNMV	2626	AAA22983.1
1744	ПАЛОЧКА	YYDYEDATF	2627	YP_00662801

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			8.1
1745	ПАЛОЧКА КОКЛЮША 509	AAFIALYPNSQLAPT	2628	Q7VU05
1746	ПАЛОЧКА КОКЛЮША 509	GGAEYNLALGQRRA	2629	Q7VU04
1747	ПАЛОЧКА КОКЛЮША 509	GGAEYNLALGQRRADA	2630	Q7VU04
1748	ПАЛОЧКА КОКЛЮША 509	IALYPNSQLAPT	2631	Q7VU05
1749	ПАЛОЧКА КОКЛЮША 509	KPDQGEVVAVGPGKKTED	2632	P0A339.1
1750	ПАЛОЧКА КОКЛЮША 509	KPDQGEVVAVGPGKKTEDG	2633	P0A339.1
1751	ПАЛОЧКА КОКЛЮША 509	LAEVLDYHNFVLTQ	2634	Q7VWM1.1
1752	ДИФТЕРИЙНАЯ ПАЛОЧКА	QSIALSSLMVAQAIP	2635	AAV70486.1
1753	ДИФТЕРИЙНАЯ ПАЛОЧКА	SIGVLGYQKTVDHTKVNSKLSLF	2636	AAV70486.1
1754	ПАЛОЧКА КОКЛЮША	AAHADWNNQSIVKT	2637	ABO77783.1
1755	ПАЛОЧКА КОКЛЮША	AALGRG	2638	ABO77783.1
1756	ПАЛОЧКА КОКЛЮША	AALGRGHSLYASYE	2639	ABO77783.1
1757	ПАЛОЧКА КОКЛЮША	AAPLRRTTLAMALG	2640	ABO77783.1
1758	ПАЛОЧКА КОКЛЮША	AAPLSITLQAGAHA	2641	ABO77783.1
1759	ПАЛОЧКА КОКЛЮША	ADAQGDIVATELPS	2642	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1760	ПАЛОЧКА КОКЛЮША	ADSGFYLDATLRAS	2643	ABO77783.1
1761	ПАЛОЧКА КОКЛЮША	AELA	2644	ABO77783.1
1762	ПАЛОЧКА КОКЛЮША	AELAVFRAGGGAYR	2645	ABO77783.1
1763	ПАЛОЧКА КОКЛЮША	AELQFRNGSVTSSG	2646	ABO77783.1
1764	ПАЛОЧКА КОКЛЮША	AGGRWHLGGLAGYT	2647	ABO77783.1
1765	ПАЛОЧКА КОКЛЮША	AGVAAMQGAVVHLQ	2648	ABO77783.1
1766	ПАЛОЧКА КОКЛЮША	AGYTRGDRGFTGDG	2649	ABO77783.1
1767	ПАЛОЧКА КОКЛЮША	ALASQARWTGATRA	2650	ABO77783.1
1768	ПАЛОЧКА КОКЛЮША	AMPWTFHAGYRYSW	2651	ABO77783.1
1769	ПАЛОЧКА КОКЛЮША	AMQGAVVHLQRATIRRGDAP	2652	ABO77783.1
1770	ПАЛОЧКА КОКЛЮША	ANGLRVRDE	2653	ABO77783.1
1771	ПАЛОЧКА КОКЛЮША	ANGLRVRDEGGSSV	2654	ABO77783.1
1772	ПАЛОЧКА КОКЛЮША	ANKDGKVDIGTYRY	2655	ABO77783.1
1773	ПАЛОЧКА КОКЛЮША	APAAVSVLGASELT	2656	ABO77783.1
1774	ПАЛОЧКА КОКЛЮША	APPAPKPAPQPGPQ	2657	ABO77783.1
1775	ПАЛОЧКА	AQGILLENPAAELQ	2658	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1776	ПАЛОЧКА КОКЛЮША	ARWTGATRAVDSLS	2659	ABO77783.1
1777	ПАЛОЧКА КОКЛЮША	ASLEAGRRFTHADG	2660	ABO77783.1
1778	ПАЛОЧКА КОКЛЮША	ASYEYSKGPKLAMP	2661	ABO77783.1
1779	ПАЛОЧКА КОКЛЮША	ATFTLANKD	2662	ABO77783.1
1780	ПАЛОЧКА КОКЛЮША	ATFTLANKDGKVDI	2663	ABO77783.1
1781	ПАЛОЧКА КОКЛЮША	ATRAVDSLSIDNAT	2664	ABO77783.1
1782	ПАЛОЧКА КОКЛЮША	DDDGIALYVAGEQAQ	2665	ABO77783.1
1783	ПАЛОЧКА КОКЛЮША	DDGIALYVAGEQAQ	2666	ABO77783.1
1784	ПАЛОЧКА КОКЛЮША	DGGHITGGRAAGVA	2667	ABO77783.1
1785	ПАЛОЧКА КОКЛЮША	DGIRRFLGTVTVKAGK	2668	ABO77783.1
1786	ПАЛОЧКА КОКЛЮША	DGSVDFQQPAEAGR	2669	ABO77783.1
1787	ПАЛОЧКА КОКЛЮША	DGYAVKGKYRTHGV	2670	ABO77783.1
1788	ПАЛОЧКА КОКЛЮША	DIVATELPSIPGTS	2671	ABO77783.1
1789	ПАЛОЧКА КОКЛЮША	DKLVVMQDASGQHR	2672	ABO77783.1
1790	ПАЛОЧКА КОКЛЮША	DLGLSDKLVVMQDA	2673	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1791	ПАЛОЧКА КОКЛЮША	DNATWVMTDNSNVGA	2674	ABO77783.1
1792	ПАЛОЧКА КОКЛЮША	DNATWVMTDNSNVGALRLA	2675	ABO77783.1
1793	ПАЛОЧКА КОКЛЮША	DNRAGRRFDQKVAG	2676	ABO77783.1
1794	ПАЛОЧКА КОКЛЮША	EAGRFKVLTVNTLA	2677	ABO77783.1
1795	ПАЛОЧКА КОКЛЮША	ELAQSIVEAPELGA	2678	ABO77783.1
1796	ПАЛОЧКА КОКЛЮША	ELGAAIRVGRGARV	2679	ABO77783.1
1797	ПАЛОЧКА КОКЛЮША	ELGADHAVAVAGGR	2680	ABO77783.1
1798	ПАЛОЧКА КОКЛЮША	ELPSIPGTSIGPLD	2681	ABO77783.1
1799	ПАЛОЧКА КОКЛЮША	EPVKLTLTGGADAQ	2682	ABO77783.1
1800	ПАЛОЧКА КОКЛЮША	EQAQASIADSTLQG	2683	ABO77783.1
1801	ПАЛОЧКА КОКЛЮША	ERGANVTVQRSAIV	2684	ABO77783.1
1802	ПАЛОЧКА КОКЛЮША	ERQHGIHIQGSDPG	2685	ABO77783.1
1803	ПАЛОЧКА КОКЛЮША	EVGKRIELAGGRQV	2686	ABO77783.1
1804	ПАЛОЧКА КОКЛЮША	FDGAGTVHTNGIAH	2687	ABO77783.1
1805	ПАЛОЧКА КОКЛЮША	FQQPAEAGRFKVLT	2688	ABO77783.1
1806	ПАЛОЧКА	FRAGGGAYRAANGL	2689	ABO77783.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮШ А			
1807	*ПАЛОЧКА КОКЛЮША*	*GAHAQGKALLYRVL*	*2690*	*ABO77783.1*
1808	*ПАЛОЧКА КОКЛЮША*	*GARVTVSGGSLSAP*	*2691*	*ABO77783.1*
1809	*ПАЛОЧКА КОКЛЮША*	*GAYRAANGLRVRDE*	*2692*	*ABO77783.1*
1810	*ПАЛОЧКА КОКЛЮША*	*GDAPAGGAVPGGAV*	*2693*	*ABO77783.1*
1811	*ПАЛОЧКА КОКЛЮША*	*GGAVPGGAVPGGFG*	*2694*	*ABO77783.1*
1812	*ПАЛОЧКА КОКЛЮША*	*GGAVPGGFGPVLDG*	*2695*	*ABO77783.1*
1813	*ПАЛОЧКА КОКЛЮША*	*GGFGPVLDGWYGVD*	*2696*	*ABO77783.1*
1814	*ПАЛОЧКА КОКЛЮША*	*GGLHIGALQSLQPE*	*2697*	*ABO77783.1*
1815	*ПАЛОЧКА КОКЛЮША*	*GGVQIERGANVTVQ*	*2698*	*ABO77783.1*
1816	*ПАЛОЧКА КОКЛЮША*	*GHSLYASYEYSKGP*	*2699*	*ABO77783.1*
1817	*ПАЛОЧКА КОКЛЮША*	*GHTDSVHVGGYATY*	*2700*	*ABO77783.1*
1818	*ПАЛОЧКА КОКЛЮША*	*GIAHRTELRGTRAE*	*2701*	*ABO77783.1*
1819	*ПАЛОЧКА КОКЛЮША*	*GKALLYRVLPEPVK*	*2702*	*ABO77783.1*
1820	*ПАЛОЧКА КОКЛЮША*	*GLGMAAALGRGHSL*	*2703*	*ABO77783.1*
1821	*ПАЛОЧКА КОКЛЮША*	*GNVIETGGARRFAP*	*2704*	*ABO77783.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1822	ПАЛОЧКА КОКЛЮША	GPLDVALASQARWT	2705	ABO77783.1
1823	ПАЛОЧКА КОКЛЮША	GQHRLWVRN	2706	ABO77783.1
1824	ПАЛОЧКА КОКЛЮША	GRGFAQRQQLDNRA	2707	ABO77783.1
1825	ПАЛОЧКА КОКЛЮША	GRLGLEVGKRIELA	2708	ABO77783.1
1826	ПАЛОЧКА КОКЛЮША	GRQVQPYIKASVLQ	2709	ABO77783.1
1827	ПАЛОЧКА КОКЛЮША	GRRFTHADGWFLEPQAELA	2710	ABO77783.1
1828	ПАЛОЧКА КОКЛЮША	GSEPASANTLLLVQ	2711	ABO77783.1
1829	ПАЛОЧКА КОКЛЮША	GSSVLGRLGLEVGK	2712	ABO77783.1
1830	ПАЛОЧКА КОКЛЮША	GTTIKVSGRQAQGI	2713	ABO77783.1
1831	ПАЛОЧКА КОКЛЮША	GTVTVKAGKLVADH	2714	ABO77783.1
1832	ПАЛОЧКА КОКЛЮША	HAVAVAGGRWHLGG	2715	ABO77783.1
1833	ПАЛОЧКА КОКЛЮША	IELAGGRQVQPYIK	2716	ABO77783.1
1834	ПАЛОЧКА КОКЛЮША	IHIQGSDPGGVRTA	2717	ABO77783.1
1835	ПАЛОЧКА КОКЛЮША	IRRFLGTVTVKAGK	2718	ABO77783.1
1836	ПАЛОЧКА КОКЛЮША	IRVGRGARVTVSGG	2719	ABO77783.1
1837	ПАЛОЧКА	ITLQAGAHA	2720	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1838	ПАЛОЧКА КОКЛЮША	ITLQAGAHAQGKAL	2721	ABO77783.1
1839	ПАЛОЧКА КОКЛЮША	IVEAPELGAAIRVG	2722	ABO77783.1
1840	ПАЛОЧКА КОКЛЮША	IVKTGERQHGIHIQ	2723	ABO77783.1
1841	ПАЛОЧКА КОКЛЮША	KAGKLVADHATLAN	2724	ABO77783.1
1842	ПАЛОЧКА КОКЛЮША	KGKYRTHGVGASLE	2725	ABO77783.1
1843	ПАЛОЧКА КОКЛЮША	KPAPQPGPQPPQPP	2726	ABO77783.1
1844	ПАЛОЧКА КОКЛЮША	KVAGFELGADHAVA	2727	ABO77783.1
1845	ПАЛОЧКА КОКЛЮША	KVAGSDGYAVKGKY	2728	ABO77783.1
1846	ПАЛОЧКА КОКЛЮША	KVDIGTYRYRLAAN	2729	ABO77783.1
1847	ПАЛОЧКА КОКЛЮША	KVLTVNTLAGSGLF	2730	ABO77783.1
1848	ПАЛОЧКА КОКЛЮША	LAANGNGQWSLVGA	2731	ABO77783.1
1849	ПАЛОЧКА КОКЛЮША	LAMPWTFHAGYRYS	2732	ABO77783.1
1850	ПАЛОЧКА КОКЛЮША	LASTLWYAESNALS	2733	ABO77783.1
1851	ПАЛОЧКА КОКЛЮША	LENDFKVAGSDGYA	2734	ABO77783.1
1852	ПАЛОЧКА КОКЛЮША	LENPAAELQFRNGS	2735	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1853	ПАЛОЧКА КОКЛЮША	LGAAPAAHADWNNQ	2736	ABO77783.1
1854	ПАЛОЧКА КОКЛЮША	LGGLAGYTRGDRGFTGDG	2737	ABO77783.1
1855	ПАЛОЧКА КОКЛЮША	LLENP	2738	ABO77783.1
1856	ПАЛОЧКА КОКЛЮША	LLVQTPLGSAATFT	2739	ABO77783.1
1857	ПАЛОЧКА КОКЛЮША	LPPSRVVLRDTNVT	2740	ABO77783.1
1858	ПАЛОЧКА КОКЛЮША	LQPEDLPPS	2741	ABO77783.1
1859	ПАЛОЧКА КОКЛЮША	LQPEDLPPSRVVLR	2742	ABO77783.1
1860	ПАЛОЧКА КОКЛЮША	LRASRLENDFKVAG	2743	ABO77783.1
1861	ПАЛОЧКА КОКЛЮША	LRLASDGSVDFQQP	2744	ABO77783.1
1862	ПАЛОЧКА КОКЛЮША	LSAAANAAVNTGGV	2745	ABO77783.1
1863	ПАЛОЧКА КОКЛЮША	LSAPHGNVIETGGA	2746	ABO77783.1
1864	ПАЛОЧКА КОКЛЮША	LSDDGIRRFLGTVT	2747	ABO77783.1
1865	ПАЛОЧКА КОКЛЮША	LVGAKAPPAPKPAP	2748	ABO77783.1
1866	ПАЛОЧКА КОКЛЮША	LYVAGEQAQASIAD	2749	ABO77783.1
1867	ПАЛОЧКА КОКЛЮША	MALGALGAAPAAHA	2750	ABO77783.1
1868	ПАЛОЧКА	MNMSLSRIVKAAPL	2751	ABO77783.1

№	*ЗАБОЛЕВАНИЕ*	*АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА*	*SEQ ID NO*	*№ ДОСТУПА*
	КОКЛЮША			
1869	*ПАЛОЧКА КОКЛЮША*	*MQDASGQHR*	*2752*	*ABO77783.1*
1870	*ПАЛОЧКА КОКЛЮША*	*MQGAVVHLQRATIR*	*2753*	*ABO77783.1*
1871	*ПАЛОЧКА КОКЛЮША*	*NAAVNTGGVGLAST*	*2754*	*ABO77783.1*
1872	*ПАЛОЧКА КОКЛЮША*	*NALSKRLGELRLNP*	*2755*	*ABO77783.1*
1873	*ПАЛОЧКА КОКЛЮША*	*NGQWSLVGAKAPPA*	*2756*	*ABO77783.1*
1874	*ПАЛОЧКА КОКЛЮША*	*NTLAGSGLFRMNVF*	*2757*	*ABO77783.1*
1875	*ПАЛОЧКА КОКЛЮША*	*PAGRELSAAANAAV*	*2758*	*ABO77783.1*
1876	*ПАЛОЧКА КОКЛЮША*	*PAPQPPAGRELSAA*	*2759*	*ABO77783.1*
1877	*ПАЛОЧКА КОКЛЮША*	*PGPQPPQPPQPQPE*	*2760*	*ABO77783.1*
1878	*ПАЛОЧКА КОКЛЮША*	*PGTSIGPLDVALAS*	*2761*	*ABO77783.1*
1879	*ПАЛОЧКА КОКЛЮША*	*PLGSAATFTLANKD*	*2762*	*ABO77783.1*
1880	*ПАЛОЧКА КОКЛЮША*	*PQPEAPAPQPPAGR*	*2763*	*ABO77783.1*
1881	*ПАЛОЧКА КОКЛЮША*	*PQPPQPQPEAPAPQ*	*2764*	*ABO77783.1*
1882	*ПАЛОЧКА КОКЛЮША*	*PYIKASVLQEFDGA*	*2765*	*ABO77783.1*
1883	*ПАЛОЧКА КОКЛЮША*	*RFAPQAAPLSITLQ*	*2766*	*ABO77783.1*

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1884	ПАЛОЧКА КОКЛЮША	RLGELRLNPDAGGA	2767	ABO77783.1
1885	ПАЛОЧКА КОКЛЮША	RLNPDAGGAWGRGF	2768	ABO77783.1
1886	ПАЛОЧКА КОКЛЮША	RNGSVTSSGQLSDD	2769	ABO77783.1
1887	ПАЛОЧКА КОКЛЮША	RRFDQKVAGFELGA	2770	ABO77783.1
1888	ПАЛОЧКА КОКЛЮША	RTTLAMALGALGAA	2771	ABO77783.1
1889	ПАЛОЧКА КОКЛЮША	SAIVDGGLHIGALQ	2772	ABO77783.1
1890	ПАЛОЧКА КОКЛЮША	SANTLLLVQTPLGS	2773	ABO77783.1
1891	ПАЛОЧКА КОКЛЮША	SDPGGVRTASGTTI	2774	ABO77783.1
1892	ПАЛОЧКА КОКЛЮША	SELTLDGGHITGGR	2775	ABO77783.1
1893	ПАЛОЧКА КОКЛЮША	SGLFRMNVF	2776	ABO77783.1
1894	ПАЛОЧКА КОКЛЮША	SGLFRMNVFADLGL	2777	ABO77783.1
1895	ПАЛОЧКА КОКЛЮША	SGSSVELAQSIVEA	2778	ABO77783.1
1896	ПАЛОЧКА КОКЛЮША	SIADSTLQGAGGVQ	2779	ABO77783.1
1897	ПАЛОЧКА КОКЛЮША	SKGPKLAMPWTFHA	2780	ABO77783.1
1898	ПАЛОЧКА КОКЛЮША	SNVGALRLASDGSV	2781	ABO77783.1
1899	ПАЛОЧКА	SRIVKAAPLRRTTL	2782	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1900	ПАЛОЧКА КОКЛЮША	SVLGASELTLDGGH	2783	ABO77783.1
1901	ПАЛОЧКА КОКЛЮША	SVLQEFDGA	2784	ABO77783.1
1902	ПАЛОЧКА КОКЛЮША	SVLQEFDGAGTVHT	2785	ABO77783.1
1903	ПАЛОЧКА КОКЛЮША	TELR	2786	ABO77783.1
1904	ПАЛОЧКА КОКЛЮША	TELRGTRAELGLGM	2787	ABO77783.1
1905	ПАЛОЧКА КОКЛЮША	TGDGGGHTDSVHVG	2788	ABO77783.1
1906	ПАЛОЧКА КОКЛЮША	TGGARRFAPQAAPL	2789	ABO77783.1
1907	ПАЛОЧКА КОКЛЮША	TGGRAAGVAAMQGA	2790	ABO77783.1
1908	ПАЛОЧКА КОКЛЮША	TGGVGLASTLWYAE	2791	ABO77783.1
1909	ПАЛОЧКА КОКЛЮША	THGVGASLEAGRRF	2792	ABO77783.1
1910	ПАЛОЧКА КОКЛЮША	TIRRGDAPA	2793	ABO77783.1
1911	ПАЛОЧКА КОКЛЮША	TLANVGDTWDDDGI	2794	ABO77783.1
1912	ПАЛОЧКА КОКЛЮША	TLQGAGGVQIERGA	2795	ABO77783.1
1913	ПАЛОЧКА КОКЛЮША	TLTGGADAQGDIVA	2796	ABO77783.1
1914	ПАЛОЧКА КОКЛЮША	TNVTAVPASGAPAA	2797	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1915	ПАЛОЧКА КОКЛЮША	TRAELGLGMAAALG	2798	ABO77783.1
1916	ПАЛОЧКА КОКЛЮША	TSSGQLSDDGIRRF	2799	ABO77783.1
1917	ПАЛОЧКА КОКЛЮША	TVHTNGIAHRTELR	2800	ABO77783.1
1918	ПАЛОЧКА КОКЛЮША	TYRYRLAANGNGQW	2801	ABO77783.1
1919	ПАЛОЧКА КОКЛЮША	VADHATLANVGDTW	2802	ABO77783.1
1920	ПАЛОЧКА КОКЛЮША	VHVGGYATYIADSG	2803	ABO77783.1
1921	ПАЛОЧКА КОКЛЮША	VLDGWYGVD	2804	ABO77783.1
1922	ПАЛОЧКА КОКЛЮША	VPASGAPAAVSVLG	2805	ABO77783.1
1923	ПАЛОЧКА КОКЛЮША	VRDEGGSSVLGRLG	2806	ABO77783.1
1924	ПАЛОЧКА КОКЛЮША	VRTASGTTIKVSGR	2807	ABO77783.1
1925	ПАЛОЧКА КОКЛЮША	VSGGSLSAPHGNVI	2808	ABO77783.1
1926	ПАЛОЧКА КОКЛЮША	VSGRQAQGILLENP	2809	ABO77783.1
1927	ПАЛОЧКА КОКЛЮША	VTVQRSAIVDGGLH	2810	ABO77783.1
1928	ПАЛОЧКА КОКЛЮША	VVLRDTNVTAVPAS	2811	ABO77783.1
1929	ПАЛОЧКА КОКЛЮША	WNNQSIVKTGERQH	2812	ABO77783.1
1930	ПАЛОЧКА	WVRNSGSEPASANT	2813	ABO77783.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	КОКЛЮША			
1931	ПАЛОЧКА КОКЛЮША	WYAESNALSKRLGE	2814	ABO77783.1
1932	ПАЛОЧКА КОКЛЮША	YATYIADSGFYLDA	2815	ABO77783.1
1933	ПАЛОЧКА КОКЛЮША	YGVDVSGSS	2816	ABO77783.1
1934	ПАЛОЧКА КОКЛЮША	YGVDVSGSSVELAQ	2817	ABO77783.1
1935	ПАЛОЧКА КОКЛЮША	YLDATLRASRLEND	2818	ABO77783.1
1936	ПАЛОЧКА КОКЛЮША	YRVLPEPVKLTLTG	2819	ABO77783.1
1937	ПАЛОЧКА КОКЛЮША	VKAQNITNKRAALIEA	2820	AAA22974.1
1938	ПАЛОЧКА КОКЛЮША	YYSNVTATRLLSSTNS	2821	AAA83981.1
1939	ПАЛОЧКА КОКЛЮША	SPNLTDERAAQAGVT	2822	CPP72976.1
1940	МОРБИЛЛИВИРУС КОРИ	SSRASDERAAHLPTS	2823	BAA33867.1
1941	ДИФТЕРИЙНАЯ ПАЛОЧКА	QVVHNSYNRPAYSPG	2824	1007216A
1942	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AEGGEIHEL	2825	AAF85692.1
1943	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AENLISNGIGKY	2826	AAF85698.1
1944	ШТАММ EDMONSTON	AEVDGDVKL	2827	CAB43772.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ-B			
1945	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AIYTAEIHK	2828	AAF85697.1
1946	ШТАММ EDMONSTON ВИРУСА КОРИ-B	APVFHMTNY	2829	CAB43772.1
1947	ШТАММ EDMONSTON ВИРУСА КОРИ-B	APVFHMTNYLEQPVSN	2830	AAR89413.1
1948	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AQRLNEIY	2831	AAF85698.1
1949	ШТАММ EDMONSTON ВИРУСА КОРИ-B	ARVPHAYSL	2832	AAF85698.1
1950	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AVRDLERAM	2833	P03424.1
1951	ШТАММ EDMONSTON ВИРУСА КОРИ-B	AVRDLERAMTTLK	2834	P03424.1
1952	ШТАММ EDMONSTON ВИРУСА КОРИ-B	DALLRLQAM	2835	Q89933.1
1953	ШТАММ EDMONSTON ВИРУСА КОРИ-B	DIKEKVINL	2836	AAF85698.1
1954	ШТАММ EDMONSTON ВИРУСА КОРИ-B	DQGLFKVL	2837	AAF85695.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1955	ШТАММ EDMONSTON ВИРУСА КОРИ-В	DTGVDTRIW	2838	Q9EMA9.1
1956	ШТАММ EDMONSTON ВИРУСА КОРИ-В	EPIGSLAIEEAM	2839	AAF85692.1
1957	ШТАММ EDMONSTON ВИРУСА КОРИ-В	EPIRDALNAM	2840	P69354.1
1958	ШТАММ EDMONSTON ВИРУСА КОРИ-В	FPKLGKTL	2841	AAF85692.1
1959	ШТАММ EDMONSTON ВИРУСА КОРИ-В	FRSVNAVAF	2842	AAF85695.1
1960	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GKIIDNTEQL	2843	AAF85695.1
1961	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GLNEKLVFY	2844	AAF85695.1
1962	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GMYGGTYLVEK	2845	AAC35876.2
1963	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GPPISLERLDVGTN	2846	P69354.1
1964	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GPRQAQVSFL	2847	Q89933.1
1965	ШТАММ	GRLVPQVRVID	2848	AAF85695.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ-В			
1966	ШТАММ EDMONSTON ВИРУСА КОРИ-В	GSAPISMGFR	2849	AAF85692.1
1967	ШТАММ EDMONSTON ВИРУСА КОРИ-В	HILAKSTAL	2850	AAF85698.1
1968	ШТАММ EDMONSTON ВИРУСА КОРИ-В	HYREVNLVY	2851	AAF85698.1
1969	ШТАММ EDMONSTON ВИРУСА КОРИ-В	IPPMKNLAL	2852	AAC35876.2
1970	ШТАММ EDMONSTON ВИРУСА КОРИ-В	IPYQGSGKGVSF	2853	CAB43772.1
1971	ШТАММ EDMONSTON ВИРУСА КОРИ-В	ISKESQHVY	2854	AAF85698.1
1972	ШТАММ EDMONSTON ВИРУСА КОРИ-В	IVSSHFFVY	2855	AAF85698.1
1973	ШТАММ EDMONSTON ВИРУСА КОРИ-В	KEIKETGRLF	2856	AAF85698.1
1974	ШТАММ EDMONSTON ВИРУСА КОРИ-В	KESQHVYYL	2857	AAF85698.1
1975	ШТАММ EDMONSTON	KIIDNTEQL	2858	AAF85695.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	ВИРУСА КОРИ-B			
1976	ШТАММ EDMONSTON ВИРУСА КОРИ-B	KKQINRQN	2859	AAA63285.1
1977	ШТАММ EDMONSTON ВИРУСА КОРИ-B	KKVDTNFIYQ	2860	AAF85698.1
1978	ШТАММ EDMONSTON ВИРУСА КОРИ-B	KLIDGFFPA	2861	AAF85698.1
1979	ШТАММ EDMONSTON ВИРУСА КОРИ-B	KPNLSSKRSEL	2862	BAB39848.1
1980	ШТАММ EDMONSTON ВИРУСА КОРИ-B	KVSPYLFTV	2863	AAR89413.1
1981	ШТАММ EDMONSTON ВИРУСА КОРИ-B	LETRTTNQFL	2864	CAB43772.1
1982	ШТАММ EDMONSTON ВИРУСА КОРИ-B	LLKEATEL	2865	AAF85695.1
1983	ШТАММ EDMONSTON ВИРУСА КОРИ-B	LLKKGNSLY	2866	AAF85698.1
1984	ШТАММ EDMONSTON ВИРУСА КОРИ-B	LPAPIGGMNY	2867	AAF85698.1
1985	ШТАММ EDMONSTON ВИРУСА КОРИ-B	MPEETLHQVM	2868	AAF85698.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
1986	ШТАММ EDMONSTON ВИРУСА КОРИ-В	PTTIRGQFS	2869	CAB43772.1
1987	ШТАММ EDMONSTON ВИРУСА КОРИ-В	QEISRHQALGY	2870	P03424.1
1988	ШТАММ EDMONSTON ВИРУСА КОРИ-В	RITHVDTESY	2871	P69354.1
1989	ШТАММ EDMONSTON ВИРУСА КОРИ-В	RPGLKPDL	2872	P69354.1
1990	ШТАММ EDMONSTON ВИРУСА КОРИ-В	RPIYGLEV	2873	AAF85698.1
1991	ШТАММ EDMONSTON ВИРУСА КОРИ-В	RQAGQEMILAV	2874	P69354.1
1992	ШТАММ EDMONSTON ВИРУСА КОРИ-В	SAVRIATVY	2875	AAF85698.1
1993	ШТАММ EDMONSTON ВИРУСА КОРИ-В	SLMPEETLHQV	2876	AAF85698.1
1994	ШТАММ EDMONSTON ВИРУСА КОРИ-В	SMIDLVTKF	2877	AAF85698.1
1995	ШТАММ EDMONSTON ВИРУСА КОРИ-В	SMLNSQAIDNLRA	2878	P69354.1
1996	ШТАММ	SMYRVFEV	2879	CAB43772.1

№	ЗАБОЛЕВАНИЕ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЭПИТОПА	SEQ ID NO	№ ДОСТУПА
	EDMONSTON ВИРУСА КОРИ-B			
1997	ШТАММ EDMONSTON ВИРУСА КОРИ-B	SQQGMFHAY	2880	AAF85698.1
1998	ШТАММ EDMONSTON ВИРУСА КОРИ-B	TDTPIVYNDRNLLD	2881	Q89933.1
1999	ШТАММ EDMONSTON ВИРУСА КОРИ-B	VIINDDQGLFKV	2882	AAF85695.1
2000	ШТАММ EDMONSTON ВИРУСА КОРИ-B	YESGVRIASL	2883	AAF85698.1
2001	ШТАММ EDMONSTON ВИРУСА КОРИ-B	YLKDKALA	2884	AAF85698.1
2002	ШТАММ EDMONSTON ВИРУСА КОРИ-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

E. Доставка мультилоскутных праймирующих редакторов

[0809] Согласно другому аспекту настоящее изобретение относится к доставке мультилоскутных праймирующих редакторов *in vitro* и *in vivo* с использованием различных стратегий, в том числе в отдельных векторах с использованием расщепленных интеинов, а также стратегии прямой доставки рибонуклеопротеинового комплекса (т.е.

праймированное редактирование в комплексе с РЕгРНК и/или с гРНК второго сайта) с использованием таких способов, как электропорация, использование опосредованных катионными липидами составов и способов индуцированного эндоцитоза с использованием лигандов рецепторов, слитых с рибонуклеопротеиновыми комплексами. Любые такие способы рассматриваются в данном документе.

Обзор вариантов доставки

[0810] Согласно некоторым аспектам настоящее изобретение относится к способам, предусматривающим доставку одного или нескольких полинуклеотидов, кодирующих мультилоскутный праймирующий редактор, как например, один или несколько векторов, как описано в настоящем документе, кодирующих один или несколько компонентов системы мультилоскутного праймированного редактирования, описанной в настоящем документе, один или несколько их транскриптов, и/или один или несколько белков, транскрибируемых из них, в клетку-хозяин. Согласно некоторым аспектам настоящее изобретение относится к способам, получаемым такими способами, и организмам (таким как животные, растения или грибы), содержащим или полученным из таких клеток. Согласно некоторым вариантам осуществления мультилоскутный праймирующий редактор, как описано в настоящем документе, в комбинации с (и необязательно в комплексе с) направляющей последовательностью, доставляется в клетку. Для введения нуклеиновых кислот в клетки млекопитающих или ткани-мишени можно использовать обычные вирусные и невирусные способы переноса генов. Такие способы можно использовать для введения нуклеиновых кислот, кодирующих компоненты системы мультилоскутного праймирующего редактора, в клетки в культуре или в организме-хозяине. Невирусные векторные системы доставки включают ДНК-плазмиды, РНК (например, транскрипт вектора, описанного в настоящем документе), голую нуклеиновую кислоту и нуклеиновую кислоту в комплексе с носителем доставки, таким как липосома. Вирусные векторные системы доставки включают ДНК- и РНК-вирусы, которые после доставки в клетку имеют либо эписомальные, либо интегрированные геномы. Обзор процедур генной терапии см. 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., in Current Topics in Microbiology and Immunology Doerfler and Bihm (eds) (1995), и Yu et al., Gene Therapy 1:13-26 (1994).

[0811] 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 в 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*).

[0812] Получение комплексов липид:нуклеиновая кислота, включающих липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области техники (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), патент США № 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028 и 4946787).

[0813] Применение систем на основе вирусной РНК или ДНК для доставки нуклеиновых кислот основано на продвинутых процессах нацеливания вируса на специфические клетки в организме и направлении полезной нагрузки вируса в ядро. Вирусные векторы могут быть введены напрямую пациентам *(in vivo)* или они могут быть использованы для обработки 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.

[0814] Тропизм вирусов может быть изменен посредством 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 цис-

действующих длинных концевых повторов с показателем упаковывания до 6-10 т.о. чужеродной последовательности. Минимальные цис-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 мышиного лейкоза (MuLV), вирусе лейкоза гиббонов (GaLV), вирусе 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). При применения, 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 относительно простой системе. Векторы на основе адено-ассоциированного вируса («AAV») также могут применяться для трансдукции клеток с использованием нуклеиновых кислот-мишеней, например, при *in vitro* продукции 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 № 4,797,368, 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).

[0815] 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 получения 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 в транс посредством 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 линию 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 техники. См., например, публикацию US20030087817, включенную в настоящий документ посредством ссылки.

[0816] Согласно различным вариантам осуществления конструкции двойного PE (включая расщепленные конструкции) могут быть сконструированы для доставки в один или несколько векторов rAAV. rAAV в рамках любого из предусмотренных в настоящем документе способов и композиций может иметь любой серотип, включая любое производное 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 (т.е. вектор в виде рекомбинантной нуклеиновой кислоты, которая экспрессирует представляющий интерес ген, например, слитый белок целого или расщепленного PE, который переносится rAAV в клетку), которая предназначена для доставки в клетку. rAAV может являться химерным.

[0817] В контексте настоящего изобретения серотип rAAV относится к серотипу каспидных белков рекомбинантного вируса. Неограничивающие примеры производных и псевдотипов включают rAAV2/1, rAAV2/5, rAAV2/8, rAAV2/9, AAV2-AAV3 гибрид, 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, AAV клон 32/83, AAVShH10, AAV2 (Y->F), AAV8 (Y733F), AAV2.15, AAV2.4, AAVM41 и AAVr3.45. Неограничивающий пример производных и псевдотипов, которые имеют химерные белки VP1, представляет собой rAAV2/5-1VP1u, который имеет геном AAV2, капсидный каркас AAV5 и VP1u of AAV1. Другой неограничивающий пример производных и псевдотипов, которые имеют химерные белки VP 1, представляет собой rAAV2/5-8VP1u, rAAV2/9-1VP1u и rAAV2/9-8VP1u.

[0818] Производные/псевдотипы AAV и способы получения таких производных /псевдотипов известны в данной области техники (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 векторов rAAV известны в данной области техники (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).

[0819] Способы получения или упаковывания частиц rAAV известны в данной области техники, 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 включены в настоящий документ посредством ссылки, и плазмиды и наборы доступны от ATCC и Cell Biolabs, Inc.). Например, плазмида, содержащая 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 модифицированную область VP2, как описано в настоящем документе), и транфецированы в рекомбинантные клетки, так что частица rAAV может быть упакована и далее очищена.

[0820] Рекомбинантный AAV может содержать вектор нуклеиновой кислоты, который может содержать как минимум: (а) один или несколько гетерологичных участков нуклеиновой кислоты, содержащих последовательность, кодирующую представляющий интерес белок или пептид или представляющую интерес РНК (например siРНК или микроРНК), и (Ь) один или несколько участков, содержащих последовательности инвертированного концевого повтора (ITR) (например ITR последовательности дикого типа или сконструированные ITR последовательности), фланкирующие один или несколько участков нуклеиновой кислоты (например, гетерологичных участков нуклеиновой кислоты). При этом гетерологичные участки нуклеиновой кислоты, содержащие последовательность, кодирующую представляющий интерес белок или представляющую интерес РНК, обозначаются как представляющие интерес гены.

[0821] Любая из предусмотренных в настоящем документе частиц rAAV может иметь капсидные белки, которые имеют аминокислоты различных серотипов вне участка VPlu. Согласно некоторым вариантам осуществления серотип каркаса белка VP1 отличается от серотипа ITR и/или гена Rep. Согласно некоторым вариантам осуществления серотип каркаса капсидного белка VP 1 частицы является таким же, что и серотип ITR. Согласно некоторым вариантам осуществления серотип каркаса капсидного белка VP 1 частицы является таким же, что и серотип гена Rep. Согласно некоторым вариантам осуществления капсидные белки частиц rAAV содержат аминокислотные мутации, которые приводят к

повышению эффективность трансдукции.

[0822] Согласно некоторым вариантам осуществления вектор нуклеиновой кислоты содержит один или несколько участков, содержащих 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 техники. Неограничивающие примеры последовательностей управления экспрессией включают промоторы, инсулаторы, сайленсеры, элементы ответа, интроны, энхансеры, участки инициации, сигналы терминации, и поли(А)-хвосты. Настоящим изобретением охватывается любая комбинация таких контрольных последовательностей (например, промотор и энхансер).

[0823] Конечные конструкции AAV могут включать последовательность, кодирующую РЕгРНК. Согласно другим вариантам осуществления конструкции AAV могут включать последовательность, кодирующую гидовую РНК для разрыва второго участка. Согласно другим вариантам осуществления конструкции AAV могут включать последовательность, кодирующую гидовую РНК для разрыва второго участка, и последовательность, кодирующую РЕгРНК.

[0824] Согласно различным вариантам осуществления РЕгРНК и гидовые РНК для разрыва второго участка могут быть экспрессированы из подходящего промотора, например, промотора U6 (hU6) человека, промотора U6 (mU6) мыши или другого подходящего промотора. РЕгРНК и гидовые РНК для разрыва второго участка могут регулироваться одними и теми же промоторами или различными промоторами.

[0825] Согласно некоторым вариантам осуществления конструкции rAAV или раскрытые в настоящем документе композиции вводят субъекту энтерально. Согласно некоторым вариантам осуществления конструкции rAAV или раскрытые в настоящем документе композиции вводят субъекту парентерально. Согласно некоторым вариантам осуществления rAAV частицу или раскрытые в настоящем документе 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. Согласно некоторым вариантам осуществления rAAV частицу или раскрытые в настоящем документе композиции вводят субъекту с помощью инъекции в печеночную артерию или воротную вену.

Стратегии на основе вектора расщепленного мультилоскутного PE

[0826] Согласно этому аспекту мультилоскутные праймирующие редакторы могут быть разделены в участке расщепления и представлены в виде двух половин целого/полного праймирующего редактора. Две половины могут быть доставлены в клетки (например, в виде экспрессированных белков или в отдельных векторах экспрессии), при этом при контакте внутри клетки две половины образуют полный праймирующий редактор посредством аутосплайсингового действия интеинов на каждой половине праймирующего редактора. Последовательности расщепленного интеина могут быть встроены в каждую из половин кодируемого праймирующего редактора для обеспечения их транс-сплайсинга внутри клетки и сопутствующего восстановления полного функционирующего PE.

[0827] Эти способы на основе расщепленного интеина обеспечивают преодоление нескольких барьеров при *in vivo* доставке. Например, ДНК, кодирующая праймирующие редакторы, является более крупной, чем предел упаковывания rAAV, и поэтому требует специальных решений. Одно подобное решение заключается в составлении редактора, слитого с парами расщепленного 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 мультилоскутного праймированного редактирования, включая оптимизацию мишеней разрыва второго участка и правильное упаковывание мультилоскутных праймирующих редакторов в вирусные векторы, включая лентивирусы и rAAV.

[0828] Согласно этому аспекту мультилоскутные праймирующие редакторы могут быть разделены в участке расщепления и представлены в виде двух половин целого/полного праймирующего редактора. Две половины могут быть доставлены в клетки (например, в виде экспрессированных белков или в отдельных векторах экспрессии), и при контакте внутри клетки две половины образуют полный праймирующий редактор посредством аутосплайсингового действия интеинов на каждой половине праймирующего редактора. Последовательности расщепленного интеина могут быть встроены в каждую из половин кодируемого праймирующего редактора для обеспечения их транс-сплайсинга внутри клетки и сопутствующего восстановления полного функционирующего PE.

[0829] На Фиг. 66 изображен один вариант осуществления праймирующего редактора, представленного в виде двух половин белка PE, которые соединяются в виде цельного праймирующего редактора посредством аутосплайсингового действия половин расщепленного интеина, расположенных на конце или начале каждой из половин белка праймирующего редактора. Мультилоскутные праймирующие редакторы, описанные в

настоящем документе, также могут быть представлены таким же образом. В контексте настоящего изобретения термин "N-концевая половина PE" обозначает N-концевую половину полного праймирующего редактора, которая включает «N-интеин» на C-конце N-концевой половины PE (т.е. N-концевой экстеин) полного праймирующего редактора. Термин «N-интеин» обозначает N-концевую половину полного полностью сформированного фрагмента расщепленного интеина. В контексте настоящего изобретения термин «C- концевая половина PE» обозначает C-концевую половину полного праймирующего редактора, которая содержит «C-интеин» на N-конце C-концевой половины (т.е. C-концевой экстеин) полного праймирующего редактора. Когда две половины белка, т.е. 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 C-концом N-концевой половины PE и N-концом C-концевой половины PE с образованием полного белка праймированного редактора, содержащего полный домен napDNAbp (например, никазу Cas9) и домен RT. Хотя это не показано на чертежах, праймирующий редактор также может содержать дополнительные последовательности, включающие NLS на N-конце и/или C-конце, а также последовательности аминокислотных линкеров, присоединяющих каждый домен.

[0830] Согласно различным вариантам осуществления мультилоскутные праймирующие редакторы могут быть сконструированы в виде двух половин белка (т.е. N-концевой половины PE и C-концевой половины PE) путем «расщепления» цельного праймирующего редактора в «участке расщепления». Термин «участок расщепления» обозначает месторасположение вставки последовательностей расщепленного интеина (т.е. N-интеина и C-интеина) между двумя смежными аминокислотными остатками в праймирующем редакторе. Более конкретно, термин «участок расщепления» обозначает месторасположение разделения целого праймирующего редактора на две отдельные половины, при этом каждая половина слита в участке расщепления с мотивами либо N-интеина, либо C-интеина. Участок расщепления может представлять собой любое подходящее месторасположение в рамках слитого белка праймирующего редактора, однако, участок расщепления 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 расщепления, являются доступными для достаточного взаимодействия друг с другом, когда одна половина белка контактирует с другой половиной белка внутри клетки.

[0831] Согласно некоторым вариантам осуществления участок расщепления расположен в домене napDNAbp. Согласно другим вариантам осуществления участок расщепления расположен в домене RT. Согласно другим вариантам осуществления участок расщепления расположен в линкере, которые соединяет домен napDNAbp и домен RT.

[0832] Согласно различным вариантам осуществления конструкция участка расщепления требует нахождения участков для расщепления и вставки N- и C-концевых интеинов, оба из которых являются структурно допустимыми для целей упаковывания двух половин в виде доменов праймирующего редактора в два различных генома AAV. Кроме того, остатки интеина, необходимые для транс-сплайсинга, могут быть включены посредством мутирования остатков на N-конце C-концевого экстеина или вставки остатков, которые оставят интеиновый «шрам».

[0833] Иллюстративные конфигурации расщепления расщепленных мультилоскутных (*например*, двухлоскутных или четырехлоскутных) праймирующих редакторов, содержащих либо никазу SpCas9, либо никазу SaCas9, являются следующими:

S. PYOGENES PE, РАСЩЕПЛЕНИЕ ПРИ ЛИНКЕРЕ, СТРУКТУРА N-КОНЦЕВОЙ ЧАСТИ: [N ЭКСТЕИН]-[N ИНТЕИН]
MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIA QLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQ LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQK AQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQT TQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEF VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS

PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGS SGGSCLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCL EDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPNSGGSKRTADGSEFEPKKKRKV(SEQ ID NO: 443)

ОБОЗНАЧЕНИЯ:
NLS (SEQ ID NO: 124, 125)
CAS9 (SEQ ID NO: 445)
ЛИНКЕР (SEQ ID NO: 446)
NPUN ИНТЕИН (SEQ ID NO: 447)

S. PYOGENES PE, РАСЩЕПЛЕНИЕ ПРИ ЛИНКЕРЕ, СТРУКТУРА С-КОНЦЕВОЙ ЧАСТИ: [С ИНТЕИН]-[С ЭКСТЕИН]

MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN***CFNSGSETPGTS ESATPESSGGSSGGSS***TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIP LKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREV NKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT WTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPG FAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQ KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLP DADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYT DSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM ADQAARKAAITETPDTSTLLIENSSP**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)

NPUC ИНТЕИН (SEQ ID NO: 452)
RT (SEQ ID NO: 454)

S. AUREUS PE, РАСЩЕПЛЕНИЕ МЕЖДК ОСТАТКАМИ 740/741, СТРУКТУРА N-КОНЦЕВОЙ ЧАСТИ: [N ЭКСТЕИН]-[N ИНТЕИН]

MKRTADGSEFESPKKKRKV**GKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVEN**
NEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEE
EFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDG
EVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFG
WKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEK
FQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEII
ENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILD
ELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKY
GLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDM
QEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ
YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYA
TRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANA
DFIFKEWKKLDKAKKVMENQMFEEKQAECLSYETEILTVEYGLLPIGKIVEKRIECTVYSVD
NNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRV
DNLPN**SGGS**KRTADGSEFEPKKKRKV(SEQ ID NO: 458)

ОБОЗНАЧЕНИЯ:
NLS (SEQ ID NO: 124, 125)
CAS9 (SEQ ID NO: 460)
ЛИНКЕР (SEQ ID NO: 174)
NPUN ИНТЕИН (SEQ ID NO: 462)

S. AUREUS PE, РАСЩЕПЛЕНИЕ МЕЖДУ ОСТАТКАМИ 740/741, СТРУКТУРА С-КОНЦЕВОЙ ЧАСТИ: [C ИНТЕИН]-[C ЭКСТЕИН]

MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN***CFN**EIETEQEY
KEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDK
DNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYS
KKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTV
KNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLL
NRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIK
KG<u>**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFP*
QAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPL
LPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTS
QPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAAT
SELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTP
RQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTK
PFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQ
PLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLD

ILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELI
ALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIH
*CPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP***SGGS**KRTADGSEFEPKKKRKV(S
EQ 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)
NPUC ИНТЕИН (SEQ ID NO: 452)
RT (SEQ ID NO: 471)

[0834] Согласно различным вариантам осуществления с применением никазы Cas9 (SEQ ID NO: 18, 1368 аминокислот) в качестве примера, расщепление может Однако предпочтительно расщепления могут быть расположены между центральным участком белка, например, в диапазоне 50-1250, или в диапазоне 100-1200, или в диапазоне 150-1150, или в диапазоне 200-1100, или в диапазоне 250-1050, или в диапазоне 300-1000, или в диапазоне 350-950, или в диапазоне 400-900, или в диапазоне 450-850, или в диапазоне 500-800, или в диапазоне 550-750, или в диапазоне 600-700 последовательности SEQ ID NO: 18. Согласно конкретным иллюстративным вариантам осуществления участок расщепления может быть между 740/741, или 801/802, или 1010/1011, или 1041/1042. Согласно другим вариантам осуществления участок расщепления может быть 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 соседних 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 последовательности SpCas9 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% идентичности последовательности с SEQ ID NO: 18, или между любыми двумя соответствующими остатками в варианте или эквиваленте SpCas9 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% идентичности последовательности с любой из последовательностей SEQ ID NO: 19-88.

[0835] Согласно различным вариантам осуществления последовательности расщепленного интеина могут быть сконструированы на основе следующих последовательностей интеина:

НАЗВАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ ЛИГАНД-ЗАВИСИМОГО ИНТЕИНА
2-4 ИНТЕИН:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 472)
3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYTNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 473)
30R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 474)
30R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL

	LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 475)
30R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPIPYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLECAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 476)
37R3-1 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYNPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC ((SEQ ID NO: 477)
37R3-2 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFD QGTRDVIGLRIAGGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 478)
37R3-3 ИНТЕИН	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFD QGTRDVIGLRIAGGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGS GNSLALSLTADQMVSALLDAEPPILYSEYDPTSPFSEASMMGLLTNLADRE LVHMINWAKRVPGFVDLTLHDQAHLLERAWLEILMIGLVWRSMEHPGKL LFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSII LLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQR

	LAQLLLILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAG GSGASRVQAFADALDDKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEE LHTLVAEGVVVHNC (SEQ ID NO: 479)

[0836] Согласно различным другим вариантам осуществления могут применяться следующие последовательности расщепленного интеина:

ИНТЕИН-N	ИНТЕИН-C
NPU-N CLSYETEILTVEYGLLPIGKIVEKRIECTV YSVDNNGNIYTQPVAQWHDRGEQEVFE YCLEDGSLIRATKDHKFMTVDGQMLPID EIFERELDLMRVDNLPNSGGS (SEQ ID NO: 447)	NPU-C IKIATRKYLGKQNVYDIGVERDHNFALKNGFIAS N (SEQ ID NO: 452)

[0837] Согласно различным вариантам осуществления расщепленные интеины могут применяться для раздельной доставки раздельных частей полного слитого белка PE в клетку, которые при экспрессии в клетке восстанавливаются в виде полного слитого белка PE посредством транс-сплайсинг.

[0838] Согласно некоторым вариантам осуществления настоящее изобретение относится к способу доставки PE слитого белка в клетку, предусматривающему:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью расщепленного интеина,

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностию расщепленного интеина,

(c) доставку первого и второго вектора экспрессии в клетку,

где 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 расщепленного интеина.

[0839] Согласно некоторым вариантам осуществления участок расщепления может быть расположен где угодно в слиянии праймирующего редактора, включая домен napDNAbp, линкер или домен обратной транскриптазы.

[0840] Согласно другим вариантам осуществления участок расщепления расположен в домене napDNAbp.

[0841] Согласно другим вариантам осуществления участок расщепления расположен в домене обратной транскриптазы или полимеразы.

[0842] Согласно другим вариантам осуществления участок расщепления находится в линкере.

[0843] Согласно различным вариантам осуществления настоящее изобретение относится к мультилоскутным праймирующим редакторам, содержащим napDNAbp (например, домен Cas9) и обратную транскриптазу, где или оба napDNAbp и/или обратная транскриптаза содержат 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 в настоящем документе посредством ссылки. Кроме того, при использовании структуры в виде расщепленного Cas9, согласно некоторым вариантам осуществления, интеин содержит аминокислотную последовательность, выбранную из группы, состоящей из SEQ ID NO: 8-15, 447, 452, 462 и 472-479.

[0844] Согласно различным вариантам осуществления домены napDNAbp являются доменами napDNAbp меньшего размера по сравнению со каноническим доменом SpCas9 последовательности SEQ ID NO: 18.

[0845] Канонический белок SpCas9 имеет длину 1368 аминокислот и имеет предполагаемую молекулярную массу 158 кДа. Термин «вариант Cas9 небольшого размера», в контексте настоящего изобретения относится к любому варианту вариант Cas9, встречающемуся в природе, сконструированному или полученному иным образом, который составляет менее по меньшей мере 1300 аминокислот или по меньшей мере менее 1290 аминокислот, или менее 1280 аминокислот, или менее 1270 аминокислот, или менее 1260 аминокислот, или менее 1250 аминокислот, или менее 1240 аминокислот, или менее 1230 аминокислот, или менее 1220 аминокислот, или менее 1210 аминокислот, или менее 1200 аминокислот, или менее 1190 аминокислот, или менее 1180 аминокислот, или менее 1170 аминокислот, или менее 1160 аминокислот, или менее 1150 аминокислот, или менее 1140 аминокислот, или менее 1130 аминокислот, или менее 1120 аминокислот, или менее 1110 аминокислот, или менее 1100 аминокислот, или менее 1050 аминокислот, или менее 1000 аминокислот, или менее 950 аминокислот, или менее 900 аминокислот, или менее 850 аминокислот, или менее 800 аминокислот, или менее 750 аминокислот, или менее 700 аминокислот, или менее 650 аминокислот, или менее 600 аминокислот, или менее 550 аминокислот, или менее 500 аминокислот, но по меньшей мере более приблизительно 400 аминокислот и Cas9.

[0846] Согласно одному варианту осуществления, как показано в Примере 20,

описание охватывает следующие конструкции PE на основе расщепленного интеина, которые разделены между остатками 1024 и 1025 канонического 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).

[0847] Сначала показана аминокислотная последовательность SEQ ID NO: 18, которая является следующей, с указанием месторасположения участка расщепления между остатками 1024 («K») и 1025 («S»):

Описание	**Последовательность**	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* M1 SwissProt № доступа Q99ZW2 Дикого типа	**M**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD AKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKL INGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVV GTALIKKYPKLESEFVYGDYKVYDVRKMIA**KS**EQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLK SVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKY	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 1023/1024 в случае, если аминокислотная последовательность не включает Met в положении 1.

	SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADA NLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	1.

[0848] В этой конфигурации аминокислотная последовательность N-концевой половины (аминокислоты 1-1024) является следующей:

[0849] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF
IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNF
KSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPL
SASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIP
YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL
LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD
DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDI
QKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY
DVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS
DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA**K**
(SEQ ID NO: 3877).

[0850] В этой конфигурации аминокислотная последовательность N-концевой половины (аминокислоты 1-1023) (где белок не включает Met в положении 1) является следующей:

[0851] DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG
ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ
LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFK
SNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL
YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV
EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD
KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ
KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG

QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD
VDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT
KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF
RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA**K** (SEQ
ID NO: 3878).

[0852] В этой конфигурации аминокислотная последовательность С-концевой половины (аминокислоты 1024-1368 (или подсчитываемые как аминокислоты 1023-1367 в Cas9 без Met) является следующей:

[0853] **S**EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE
FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE
VLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 3879).

[0854] Как показано в Примере 20, PE2 (который основан на последовательности SpCas9 согласно SEQ Ш NO: 18) был расщеплен в положении 1023/1024 (по сравнению с SEQ ID NO: 18 без Met) на две отдельные конструкции следующим образом:

[0855] SpPE2 расщепление при 1023/1024 N-концевая половина

[0856] MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLG
NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL
SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGY
IDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF
YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKV
TVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF
EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRH
KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNG
RDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY
WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN
DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAK*CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQE*
*VFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN***SGGS***KRTADGSEFEPKKKR*
KV (SEQ ID NO: 3875)

[0857] Ключевое обозначение: NLS, Cas9, мутированные остатки, **Линкер**, *NpuN интеин*, *NpuC интеин*, **RT**

[0858] SpPE2 расщепление при 1023/1024 C-концевая половина

MKRTADGSEFESPKKKRKV*IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN*CFNEIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSA YNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGD**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**TLNIEDEYRLHETSKEPDVSLGS **TWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQ GILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQW YTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRD LADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVK YLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPL TKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTR PDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIAL TQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLP KRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP**SGGS***KRTADGSEF EPKKKRK* (SEQ ID NO: 3876)

[0859] Ключевое обозначение: NLS, Cas9, мутированные остатки, **Линкер**, *NpuN интеин*, *NpuC интеин*, **RT**

[0860] Настоящее изобретение также относится к способам доставки мультилоскутных праймирующих редакторов на основе расщепленных интеинов в клетки и/или обработки клеток мультилоскутными праймирующими редакторами на основе расщепленных интеинов.

[0861] Согласно некоторым вариантам осуществления настоящее изобретение относится к способу доставки PE слитого-белка в клетку, предусматривающему:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью расщепленного интеина,

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностию расщепленного интеина,

(c) доставку первого и второго вектора экспрессии в клетку,

где 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 последовательностей расщепленного интеина.

[0862] Согласно определенным вариантам осуществления N-концевой фрагмент слитого белка PE, слитого с последовательностью первого расщепленного 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% идентичности последовательности с SEQ ID NO: 3875.

[0863] Согласно другим вариантам осуществления C-концевой фрагмент слитого белка PE, слитого с последовательностью первого расщепленного 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% идентичности последовательности с SEQ ID NO: 3876.

[0864] Согласно другим вариантам осуществления настоящее изобретение относится к способу редактирования целевой последовательности ДНК в клетке, предусматривающему:

(a) конструирование первого вектора экспрессии, кодирующего N-концевой фрагмент слитого белка PE, слитого с первой последовательностью расщепленного интеина,

(b) конструирование второго вектора экспрессии, кодирующего C-концевой фрагмент слитого белка PE, слитого со второй последовательностию расщепленного интеина,

(c) доставку первого и второго вектора экспрессии в клетку,

где 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 последовательностей расщепленного интеина.

[0865] Согласно определенным вариантам осуществления N-концевой фрагмент слитого белка PE, слитого с последовательностью первого расщепленного 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% идентичности последовательности с SEQ ID NO: 3875.

[0866] Согласно другим вариантам осуществления C-концевой фрагмент слитого белка PE, слитого с последовательностью первого расщепленного 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% идентичности последовательности с SEQ ID NO: 3876.

Доставка рибонуклеопротеиновых комплексов PE

[0867] Согласно этому аспекту мультилоскутные праймирующие редакторы могут быть доставлены с помощью стратегий невирусной доставки, включающих доставку комплекса мультилоскутного праймированного редактирования с PEгРНК (т.е., 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*).

[0868] Получение комплексов липид:нуклеиновая кислота, включая липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области техники (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).

[0869] Может быть сделана дополнительная ссылка на следующие ссылочные источники, в рамках которых описаны походы для невирусной доставки рибонуклеопротеиновых комплексов, каждый из которых включена в настоящий документ посредством ссылки.

[0870] Chen, Sean, et al. «Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes.» *Journal of Biological Chemistry* (2016): jbc-M116. PubMed

[0871] 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

[0872] 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.

[0873] 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 посредством иРНК

[0874] Другой способ, который может применяться для доставки мультилоскутных праймирующих редакторов и/или PEгРНК в клетки, в которых требуется осуществить геномное редактирование на основе мультилоскутного праймированного редактирования, заключается в использовании способов и технологий доставки на основе информационной РНК (иРНК). Примеры способов доставки на основе иРНК и композиций, которые могут применяться согласно настоящему изобретению, включают, например, PCT7US2014/028330, US8822663B2, NZ700688A, ES2740248T3, EP2755693A4, EP2755986A4, WO2014152940A1, EP3450553B1, BR112016030852A2, и EP3362461A1, каждая из которых включена в настоящий документ посредством ссылки во всей своей полноте. См. включенное в настоящий документ посредством ссылки описание в Kowalski et al., «Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery,» *Mol Therap.*, 2019, 27(4): 710-728.

[0875] В отличие от ДНК вектора, кодирующего мультилоскутные праймирующие редакторы, РНК в качестве агента доставки для мультилоскутных праймирующих 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 праймирующего редактора) и продукты в виде нуклеиновой кислоты (например, PEгРНК). Однако для большей стабильности (например, устойчивости к расщепляющим РНК ферментам в цитоплазме) согласно некоторым вариантам 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 праймирующего 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.

[0876] 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 т.о.) по сравнению с другими типами РНК (малыми интерферирующими РНК [siРНК], ~14 кДа, антисмысловыми олигонуклеотидами [ASO], 4-10 кДа).

[0877] гРНК 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.

[0878] Согласно некоторым вариантам осуществления композиции на основе иРНК согласно настоящему изобретению включают иРНК (кодирующую праймирующий редактор и/или РЕгРНК), средство для транспорта, и необязательно агент, который обеспечивает контакт с клеткой-мишенью и последующее трансфицирование.

[0879] Согласно некоторым вариантам 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 согласно настоящему 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), поли А-хвост, структуру Capí, или гормон роста человека (hGH). Согласно некоторым вариантам осуществления иРНК является модифицированной для снижения иммуногенности иРНК.

[0880] Согласно одному варианту осуществления иРНК мультилоскутного праймирующего редактора в композиции согласно настоящему изобретению может быть составлена в носителе 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, DOD AP, HGT4003, ICE, HGT5000 или HGT5001. Согласно вариантам осуществления носитель для переноса включает холестерин (chol) и/или модифицированные ПЭГ липиды. Согласно некоторым вариантам осуществления носитель для переноса включает DMG-PEG2K. Согласно определенным вариантам 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.

[0881] Настоящее изобретение также относится к композициям и способам, полезным для обеспечения трансфекции клеток-мишеней одной или несколькими РЕ-кодирующими молекулами иРНК. Например, композиции и способы согласно настоящему изобретению предусматривают применение нацеленных лигандов, которые могут повышать аффинность композиции в отношении одной или нескольких клеток-мишеней. Согласно одному варианту осуществления нацеливающий лиганд представляет собой аполипопротеин B или аполипопротеин E, 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 согласно 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, B-клетки, T-клетки, ретикулоциты, лейкоциты, гранулоциты и опухолевые клетки. Однако они не ограничиваются ими.

[0882] Согласно некоторым вариантам осуществления РЕ-кодирующая 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 композициях согласно настоящему изобретению может расщепляться со временем полужизни от 30 минут до нескольких дней. иРНК в композициях согласно настоящему изобретению могут сохранять по меньшей мере некоторую способность подвергаться трансляции, тем самым продуцируя функциональный белок или фермент. Соответственно,

настоящее изобретение относится к композициям, содержащим, и способам, предусматривающим введение иРНК. Согласно некоторым вариантам 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.

[0883] Соответственно, согласно некоторым вариантам 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 контексте настоящего изобретения термин «модификация» и «модифицированные» при их использовании в отношении указанных в настоящем документе 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 контексте настоящего изобретения термина «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 настоящим изобретением 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 согласно 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, M., Nucleic Acids Res 15 (20): 8125-48 (1987)).

[0884] Согласно некоторым вариантам осуществления иРНК, используемые в композициях согласно настоящему изобретению, подвергнуты химической или биологической модификации для придания им большей стабильности. Иллюстративные модификации в иРНК включают устранение основания (например, путем удаления или замещения одного нуклеотида на другой) или модификацию основания, например, химическую модификацию основания. В контексте настоящего изобретения термин «химические модификации» 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).

[0885] Другие подходящие полинуклеотидные модификации, которые могут быть встроены в РЕ-кодирующую иРНК, используемую в композициях согласно настоящему изобретению, включают 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 поли А-хвоста или более длинного поли А-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).

[0886] Согласно некоторым вариантам 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.

[0887] Дополнительные аналоги кэпа включаю 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)).

[0888] Как правило, наличие «хвоста» служит для защиты иРНК от расщепления экзонуклеазой. Предполагается, что поли A- или поли U-хвост стабилизирует природные информационные и синтетические смысловые РНК. Таким образом, в определенных вариантах осуществления длинный поли A- или поли U-хвост может быть добавлен в молекулу иРНК, что, таким образом, делает РНК более стабильной. Поли A- или поли U-хвосты могут быть добавлены с помощью ряда известных в данной области способов. Например, длинные поли A-хвосты могут быть добавлены в синтетическую или транскрибированную in vitro РНК с помощью поли A-полимеразы (Yokoe, et al. Nature Biotechnology. 1996, 14: 1252-1256). Вектор транскрипции также может кодировать длинные поли A-хвосты. Кроме того, поли A-хвосты могут быть добавлены путем транскрипции напрямую из продуктов ПЦР. Поли 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)).

[0889] Как правило, длина поли A- или поли U-хвоста может составлять по меньшей мере приблизительно 10, 50, 100, 200, 300, 400 и по меньшей мере 500 нуклеотидов. Согласно некоторым вариантам осуществления поли A-хвост на 3'-конце иРНК обычно включает приблизительно 10 - 300 нуклеотидов аденозина (например, от приблизительно 10 до 200 нуклеотидов аденозина, приблизительно 10 - 150 нуклеотидов аденозина, приблизительно 10 - 100 нуклеотидов аденозина, приблизительно 20 - 70 нуклеотидов аденозина, или приблизительно 20 - 60 нуклеотидов аденозина). Согласно некоторым

вариантам осуществления иРНК включают структуру в виде 3' поли(С)-хвоста. Подходящий поли С-хвост на 3'-конце иРНК обычно включает приблизительно 10 - 200 нуклеотидов цистеина (например, приблизительно 10 - 150 нуклеотидов цистеина, приблизительно 10 to 100 нуклеотидов цистеина, приблизительно 20 - 70 нуклеотидов цистеина, приблизительно 20 - 60 нуклеотидов цистеина, или приблизительно 10 - 40 нуклеотидов цистеина). Поли С-хвост может быть добавлен к поли А- или поли U-хвосту или может замещать поли А- или поли U-хвост.

[0890] 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 DTT и ионы магния, и подходящую РНК-полимеразу (например, РНК-полимеразу Т3, Т7 или SP6), ДНКазу I, пирофосфатазу, и/или ингибитор РНКазы. Конкретные условия будут варьироваться в соответствии с конкретным вариантом применения.

[0891] Согласно вариантам, осуществления, включающим доставку иРНК, соотношение иРНК, кодирующей слитый белок PE, и PEгРНК может являться важным для эффективного редактирования. Согласно определенным вариантам осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEгРНК составляет 1:1. В определенных других вариантах осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEгРНК составляет 2:1. Согласно другим вариантам осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEгРНК составляет 1:2. Согласно другим вариантам осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEгРНК 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. Согласно другим вариантам осуществления отношение массы иРНК (кодирующей слитый белок PE) к PEгРНК 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.

F. Применение мультилоскутного праймированного редактирования для вставки доменов индуцируемой димеризации

[0892] Мультилоскутные праймирующие редакторы (как представлено на фиг. 90, 95, 97A, 98A), описанные в настоящем документе, 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. Согласно различным аспектам 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. Согласно другим вариантам 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.

[0893] В контексте настоящего изобретения термин «домен димеризации» означает 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, посредством PE, как описано в настоящем документе), а второй домен димеризации (например, посредством PE, как описано в настоящем документе) 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.

[0894] Термин «биспецифический лиганд» или «биспецифический фрагмент» в контексте настоящего изобретения означает лиганд, который связываются с двумя различными лиганд-связывающими доменами. Согласно различным вариантам 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. Согласно определенным вариантам осуществления лиганд представляет собой низкомолекулярное соединение, или пептид, или полипептид. Согласно другим вариантам осуществления лиганд-связывающий фрагмент представляет собой «домен димеризации», который может быть внесен в виде пептидной метки в белок. Согласно различным вариантам осуществления два белка, каждый из которых включает один и тот же или различные домены димеризации, могут быть индуцированы для димеризации посредством связывания каждого домена димеризации с биспецифическим лигандом. Данные молекулы 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.

[0895] Согласно различным аспектам домены димеризации, устанавливаемые посредством мультилоскутного РЕ могут быть одинаковыми или различными.

[0896] Например, домены димеризации могут представлять собой FKBP12, которые имеют следующую аминокислотную последовательность:

FKBP12	MGVQVETISPGDGRTFPKRGQTCVVHYTGML EDGKKFDSSRDRNKPFKFMLGKQEVIRGWEE GVAQMSVGQRAKLTISPDYAYGATGHPGIIPPH ATLVFDVELLKLE (SEQ ID NO: 488)

[0897] Согласно другому примеру домен димеризации может представлять собой мутант FKBP12, упоминаемый как FKBP12-F36V, мктантFKBP12 со встроенным отверстием, которое связывает синтетический имеющий «выступы» миметик FK506 (2, фиг.3)[107]:

FKBP12-F36V	MGVQVETISPGDGRTFPKRGQTCVVHYTGML EDGKKVDSSRDRNKPFKFMLGKQEVIRGWEE GVAQMSVGQRAKLTISPDYAYGATGHPGIIPPH ATLVFDVELLKLE (SEQ ID NO: 489)

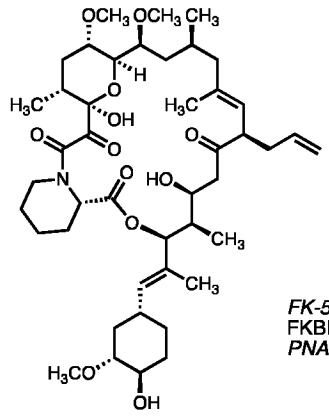
[0898] Согласно другому примеру домен димеризации может представлять собой циклофилин, следующим образом:

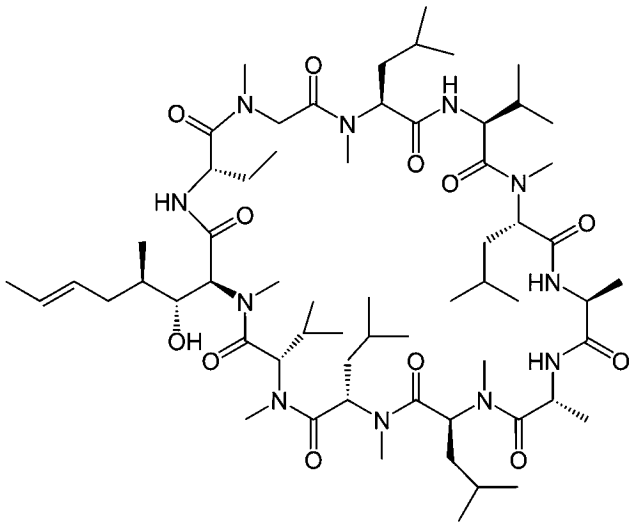
ЦИКЛОФИЛИН	MVNPTVFFDIAVDGEPLGRVSFELF ADKVPKTAENFRALSTGEKGFGYKGSCFH RIIPGFMCQGGDFTRHNGTGGKSIYGEKFE

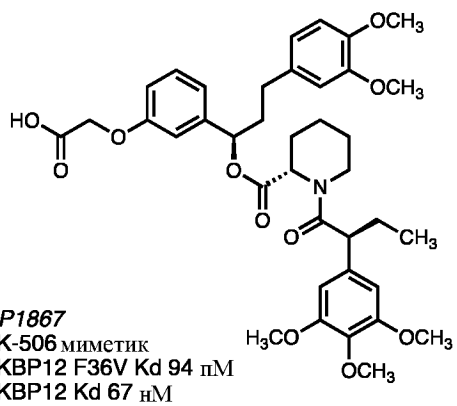
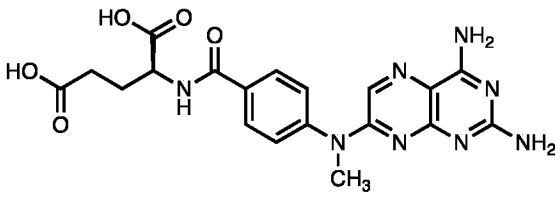
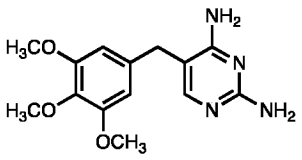
	DENFILKHTGPGILSMANAGPNTNGSQFFI CTAKTEWLDGKHVVFGKVKEGMNIVEA MERFGSRNGKTSKKITIADCGQLE (SEQ ID NO: 490)

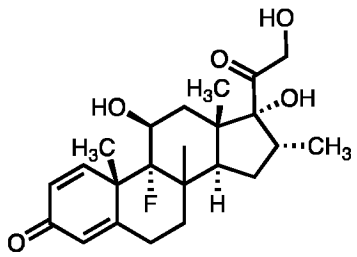
[0899] Согласно различным вариантам осуществления аминокислотные последовательности этих доменов димеризации могут быть изменены для оптимизации связывания или улучшения ортогональности связывания с нативными мишенями. Последовательности нуклеиновой кислоты генов, кодирующих низкомолекулярные связывающиеся белки, могут быть изменены для оптимизации эффективности процесса РЕ, например, за счет снижения вторичной структуры РЕгРНК.

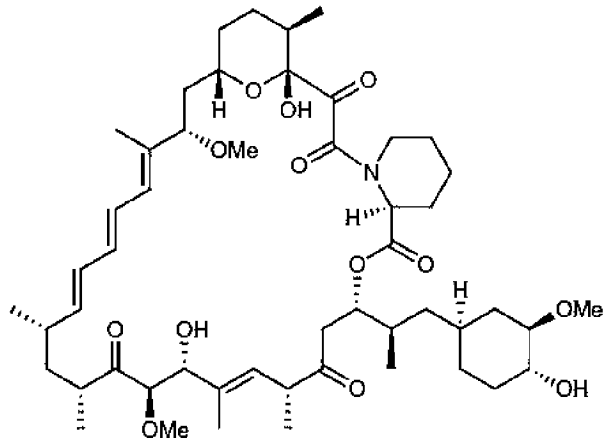
[0900] Другие примеры подходящих доменов димеризации и родственного низкомолекулярного соединения, которое связывается с ними, приведены ниже. Следует обратить внимание, что родственное низкомолекулярное соединение может быть соединено (например, через химический линкер) со вторым низкомолекулярным соединением (которое представляет собой либо такое же соединение, либо другое соединение) с образованием биспецифического лиганда, который может связывать два домена димеризации. В некоторых случаях, например, при использовании FK506 и циклоспорина А, димеризация каждого из них (например, FK506-FK506 или циклоспорин А-циклоспорин А) снижает или устраняет иммуносупрессивную активность мономерных соединений.

МАЛАЯ МОЛЕКУЛА СВЯЗЫВАЕТСЯ С ДОМЕНОМ ДИМЕРИЗАЦИИ – ДИМЕР ЭТИХ МОЛЕКУЛ СОСТАВЛЯЕТ БИСПЕЦИФИЧЕСКИЙ ЛИГАНД, КОТОРЫЙ БУДЕТ СВЯЗЫВАТЬСЯ С ДВУМЯ ДОМЕНАМИ ДИМЕРИЗАЦИИ	ДОМЕН ДИМЕРИЗАЦИИ (ДОМЕНЫ ДИМЕРИЗАЦИИ)
FK506  *FK-506* FKBP12 Kd 0.4 нМ *PNAS* **1990**, *87*, 9231.	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12: MGVQVETISPGDGRTFPKRGQTCVVHYTG MLEDGKKFDSSRDRNKPFKFMLGKQEVIR GWEEGVAQMSVGQRAKLTISPDYAYGATG HPGIIPPHATLVFDVELLKLE (SEQ ID NO: 491) КАЛЬЦИНЕВРИН:

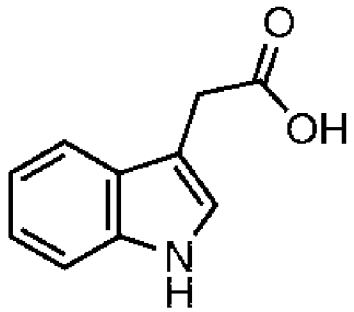
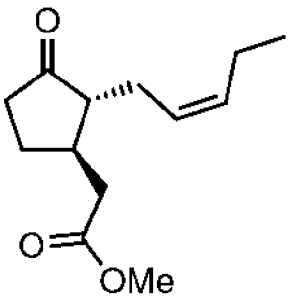
МИШЕНИ: FKBP12 + КАЛЬЦИНЕВРИН	
ЦИКЛОСПОРИН А  МИШЕНИ: ЦИКЛОФИЛИН + КАЛЬЦИНЕВРИН	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЦИКЛОФИЛИНА ЧЕЛОВЕКА А: MVNPTVFFDIAVDGEPLGRVSFELFAD KVPKTAENFRALSTGEKGFGYKGSCFH RIIPGFMCQGGDFTRHNGTGGKSIYGE KFEDENFILKHTGPGILSMANAGPNTN GSQFFICTAKTEWLDGKHVVFGKVKE GMNIVEAMERFGSRNGKTSKKITIADC GQLE (SEQ ID NO: 490) АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ЦИКЛОФИЛИНА ЧЕЛОВЕКА В: MLRLSERNMKVLLAAALIAGSVFFLLL PGPSAADEKKKGPKVTVKVYFDLRIG DEDVGRVIFGLFGKTVPKTVDNFVALA TGEKGFGYKNSKFHRVIKDFMIQGGDF TRGDGTGGKSIYGERFPDENFKLKHYG PGWVSMANAGKDTNGSQFFITTVKTA WLDGKHVVFGKVLEGMEVVRKVEST KTDSRDKPLKDVIIADCGKIEVEKPFAI AKE (SEQ ID NO: 493) АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МЫШИНОГО ЦИКЛОФИЛИНА С: MSPGPRLLLPAVLCLGLGALVSSSGSSG VRKRGPSVTDKVFFDVRIGDKDVGRIV IGLFGNVVPKTVENFVALATGEKGYGY KGSIFHRVIKDFMIQGGDFTARDGTGG MSIYGETFPDENFKLKHYGIGWVSMA NAGPDTNGSQFFITLTKPTWLDGKHVV

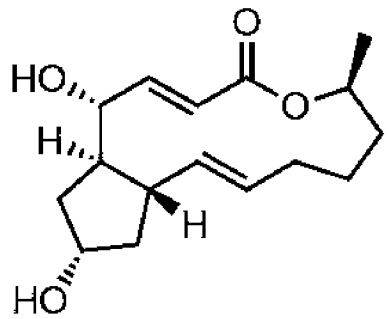
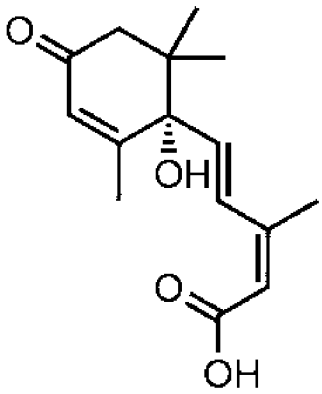
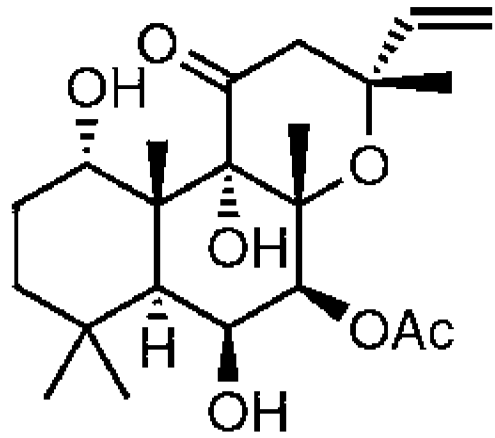
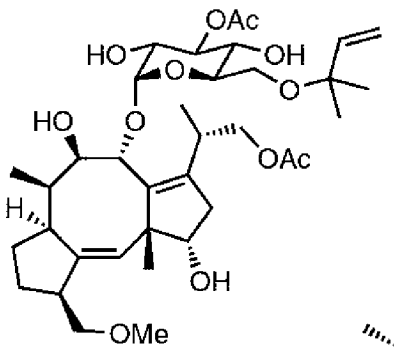
	FGKVLDGMTVVHSIELQATDGHDRPLT DCTIVNSGKIDVKTPFVVEVPDW (SEQ ID NO: 494)
AP1867  *AP1867* FK-506 миметик FKBP12 F36V Kd 94 пМ FKBP12 Kd 67 нМ *PNAS* **1998**, *95*, 10437. МИШЕНЬ (МИШЕНИ): FKBP12	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12-F36V MGVQVETISPGDGRTFPKRGQTCVVHYTG MLEDGKKVDSSRDRNKPFKFMLGKQEVIR GWEEGVAQMSVGQRAKLT ISPDYAYGATGHPGIIPPHATLVFDVELLKL E (SEQ ID NO: 495)
МЕТОТРЕКСАТ  МЕТОТРЕКСАТ Human DHFR Kd < 10 нМ *J. Biol. Chem.* **1988**, *263*, 10304. *E. coli* DHFR Kd 9.5 нМ *PNAS* **2002**, *99*, 13481. МИШЕНЬ (МИШЕНИ): ДИГИДРОФОЛАТРЕДУКТАЗА	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ДИГИДРОФОЛАТРЕДУКТАЗЫ ЧЕЛОВЕКА MVGSLNCIVAVSQNMGIGKNGDLPWPPLR NEFRYFQRMTTTSSVEGKQNLVIMGKKTW FSIPEKNRPLKGRINLVLSRELKEPPQGAHF LSRSLDDALKLTEQPELANKVDMVWIVGG SSVYKEAMNHPGHLKLFVTRIMQDFESDT FFPEIDLEKYKLLPEYPGVLSDVQEEKGIK YKFEVYEKND (SEQ ID NO: 496)
ТРИМЕТОПРИМ  ТРИМЕТОПРИМ *E. coli* DHFR K_I 1.3 нМ *Biochemistry* **1982**, *21*, 5068. МИШЕНЬ (МИШЕНИ): ДИГИДРОФОЛАТРЕДУКТАЗА	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ДИГИДРОФОЛАТРЕДУКТАЗЫ E. COLI MISLIAALAVDRVIGMENAMPWNLPADLA WFKRNTLNKPVIMGRHTWESIGRPLPGRK NIILSSQPGTDDRVTWVKSVDEAIAACGDV PEIMVIGGGRVYEQFLPKAQKLYLTHIDAE

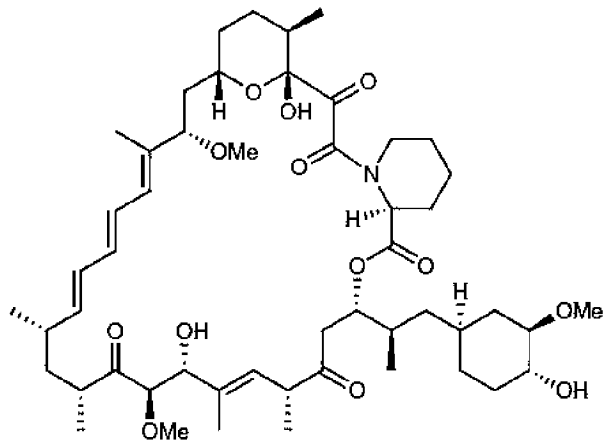
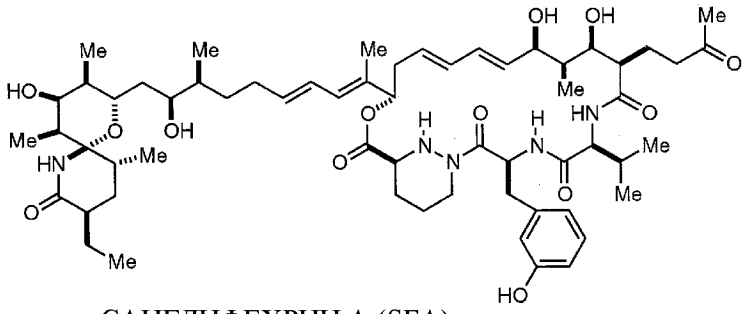
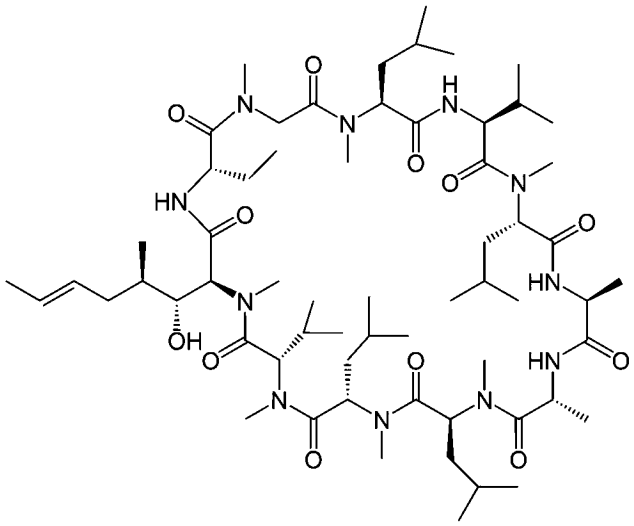
	VEGDTHFPDYEPDDWESVFSEFHDADAQN SHSYCFEILERR (SEQ ID NO: 497)
ДЕКСАМЕТАЗОН  ДЕКСАМЕТАЗОН Human GR Kd 4.6 нМ *Mol. Endocrin.* **1999**, *13*, 1855.	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ГЛЮКОКОРТИКОИДНОГО РЕЦЕПТОРА ЧЕЛОВЕКА MDSKESLTPGREENPSSVLAQERGDVMDF YKTLRGGATVKVSASSPSLAVASQSDSKQR RLLVDFPKGSVSNAQQPDLSKAVSLSMGL YMGETETKVMGNDLGFPQQGQISLSSGET DLKLLEESIANLNRSTSVPENPKSSASTAVS AAPTEKEF PKTHSDVSSEQQHLKGQTGTNGGNVKLYT TDQSTFDILQDLEFSSGSPGKETNESPWRS DLLIDENCLLSPLAGEDDSFLLEGNSNEDC KPLILPDTKPKIKDNGDLVLSSPSNVTLPQV KTEKEDFIELCTPGVIKQEKLGTVYCQASF PGANIIGNKMSAISVHGVSTSGGQMYHYD MNTASLSQQQDQKPIFNVIPPIPVGSENWN RCQGSGDDNLTSLGTLNFPGRTVFSNGYSS PSMRPDVSSPPSSSSTATTGPPPKLCLVCSD EASGCHYGVLTCGSCKVFFKRAVEGQHNY LCAGRNDCIIDKIRRKNCPACRYRKCLQAG MNLEARKTKKKIKGIQQATTGVSQETSEN PGNKTIVPATLPQLTPTLVSLLEVIEPEVLYA GYDSSVPDSTWRIMTTLNMLGGRQVIAAV KWAKAIPGFRNLHLDDQMTLLQYSWMFL MAFALGWRSYRQSSANLLCFAPDLIINEQR MTLPCMYDQCKHMLYVSSELHRLQVSYE EYLCMKTLLLLSSVPKDGLKSQELFDEIRM TYIKELGKAIVKREGNSSQNWQRFYQLTK LLDSMHEVVENLLNYCFQTFLDKTMSIEFP EMLAEIITNQIPKYSNGNIKKLLFHQK (SEQ ID NO: 498)

 РАПАМИЦИН	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ FKBP12: MGVQVETISPGDGRTFPKRGQTCVVHYTG MLEDGKKFDSSRDRNKPFKFMLGKQEVIR GWEEGVAQMSVGQRAKLTISPDYAYGATG HPGIIPPHATLVFDVELLKLE (SEQ ID NO: 491)

[0901] Другие примеры встречающихся в природе бифункциональных молекул и их двойных рецепторов-мишеней являются следующими. Праймированное редактирование может применяться для внесения двойных рецепторов-мишеней в различные белки. После модификации различных белков с помощью PE для включения рецептора бифункциональной молекулы могут быть внесены бифункциональные молекулы, что вызывает димеризацию белков, которые были модифицированы для включения различных доменов димеризации. Примеры пар (1) бифункциональных молекул и (2) их двойных рецепторов-мишеней являются следующими:

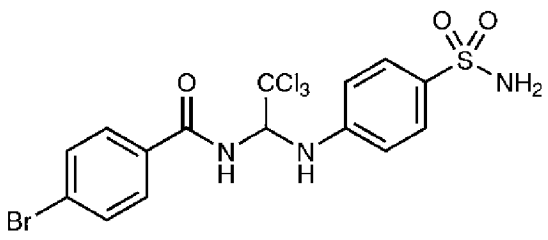
ВСТРЕЧАЮЩИЕСЯ В ПРИРОДЕ БИОФУНКЦИОНАЛЬНЫЕ МОЛЕКУЛЫ	РЕЦЕПТОРЫ-МИШЕНИ БИОФУНКЦИОНАЛЬНЫХ МОЛЕКУЛ
 ауксин	Рецептор-мишень 1: рецептор ауксина Рецептор-мишень 2: TIR1 E3 лигаза
 метилжасмонат	Рецептор-мишень 1: JAZ рецептор Рецептор-мишень 2: Col1 E3 лигаза

Брефельдин А 	рецептор-мишень 1: GBF1 рецептор-мишень 2: GTPаза Arf1p
Абсцизовая кислота 	рецептор-мишень 1: рецептор PYR рецептор-мишень 2: фосфатаза фосфопротеина 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

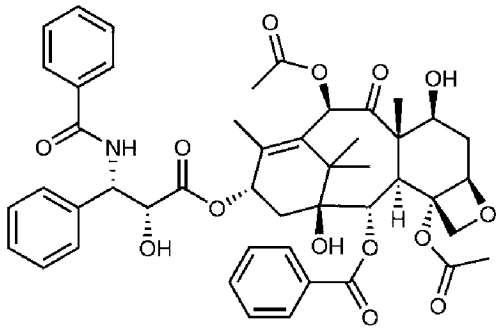
Рапамицин 	рецептор-мишень 1: FKBP12 рецептор-мишень 2: mTOR
Санглифехрин А  САНГЛИФЕХРИН А (SFA)	рецептор-мишень 1: циклофилин рецептор-мишень 2: IMP дегидрогеназа 2
циклоспорин А 	рецептор-мишень 1: циклофилин рецептор-мишень 2: кальциневрин

[0902] Примеры других бифункциональных молекул, которые могут применяться в этом аспекте праймированного редактирования, являются следующими:

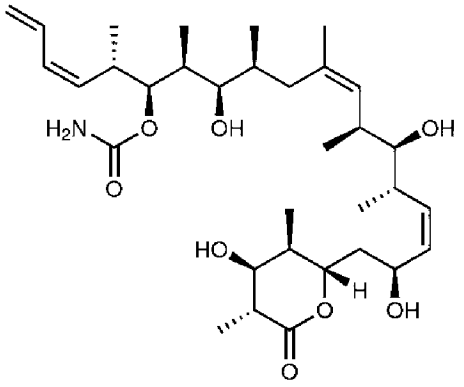
[0903] Синстаб А:



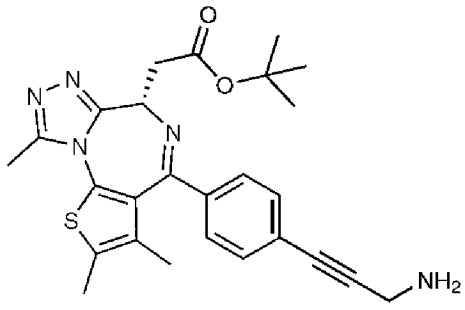
[0904] Паклитаксел:



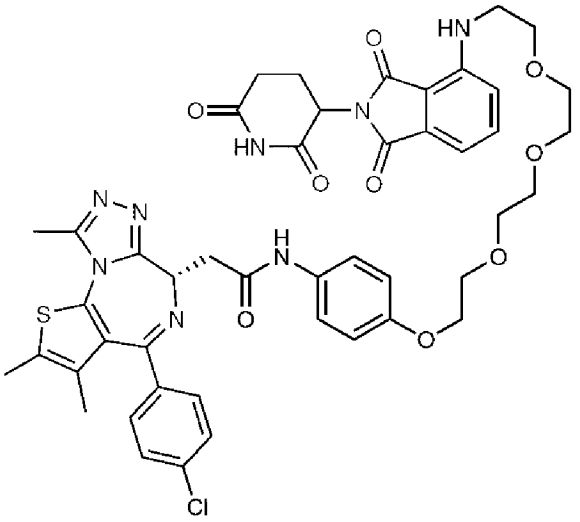
[0905] Дискодермолид:



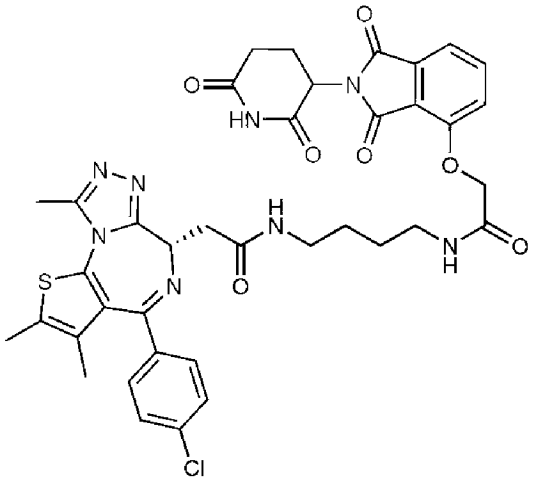
[0906] GNE-0011



[0907] ARV-825 и



[0908] dBET1



[0909] 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 dBETl содержат 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.

[0910] PEгPHK для введения доменов димеризации могут содержать следующие структуры (со ссылкой на фиг. 3D):

5′-[спейсер]-[ядро гРНК]-[плечо удлинения]-3′, где плечо удлинения содержит 5′-[плечо гомологии]-[матрица редактирования]-[участок связывания праймера]-3′, или

5′-[плечо удлинения]-[спейсер]-[ядро гРНК]-3′, где плечо удлинения содержит 5′-[плечо гомологии]-[матрица редактирования]-[участок связывания праймера]-3′, и где при

любой конфигурации «матрица редактирования» содержит нуклеотидную последовательность домена димеризации.

[0911] В одном примере РЕгРНК для вставки домена димеризации FKBP12 на С-конце рецептора инсулина человека является следующей (спейсер подчеркнут, ядро гРНК указано обычным шрифтом, гомология лоскута указана жирным шрифтом, вставка FKBP12 обозначена курсивом, участок отжига обозначен жирным курсивом):

РЕГРНК ДЛЯ ВВЕДЕНИЯ FKBP12 В РЕЦЕПТОР ИНСУЛИНА ЧЕЛОВЕКА	CACGGUAGGCACUGUUAGGAGUUUUAGAG CUAGAAAUAGCAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAAGUGGCACCG AGUCGGUG**CUUGCCUCGGUCCAAUCCUUC** **C***GGAGUGCAGGUGGAAACCAUCUCCCCAGGA* *GACGGGCGCACCUUCCCCAAGCGCGGCCAGA* *CCUGCGUGGUGCACUACACCGGGAUGCUUG* *AAGAUGGAAAGAAAUUUGAUUCCUCCCGGGA* *CAGAAACAAGCCCUUUAAGUUUAUGCUAGGC* *AAGCAGGAGGUGAUCCGAGGCUGGGAAGAAG* *GGGUUGCCCAGAUGAGUGUGGGUCAGAGAG* *CCAAACUGACUAUAUCUCCAGAUUAUGCCUA* *UGGUGCCACUGGGCACCCAGGCAUCAUCCCA* *CCACAUGCCACUCUCGUCUUCGAUGUGGAG* *CUUCUAAAACUGGAA**UAACAGUGCCUACC*** (SEQ ID NO: 499)

[0912] Согласно другому примеру РЕгРНК для вставки домена димеризации FKBP12 в локусе HEK3 (для оптимизации):

РЕГРНК ДЛЯ ВВЕДЕНИЯ FKBP12 В HEK3	GGCCCAGACTGAGCACGTGAGTTTTAGAGC TAGAAATAGCAAGTTAAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGT CGGTGCTGGAGGAAGCAGGGCTTCCTTTCCT CTGCCATCATTCCAGTTTTAGAAGCTCCACA TCGAAGACGAGAGTGGCATGTGGTGGGATG ATGCCTGGGTGCCCAGTGGCACCATAGGCA TAATCTGGAGATATAGTCAGTTTGGCTCTCT GACCCACACTCATCTGGGCAACCCCTTCTTC CCAGCCTCGGATCACCTCCTGCTTGCCTAGC ATAAACTTAAAGGGCTTGTTTCTGTCCCGGG AGGAATCAAATTTCTTTCCATCTTCAAGCAT CCCGGTGTAGTGCACCACGCAGGTCTGGCC

	GCGCTTGGGGAAGGTGCGCCCGTCTCCTGG GGAGATGGTTTCCACCTGCACTCCCGTGCTC AGTCTG (SEQ ID NO: 500)

[0913] Белки-мишени для внесения доменов димеризации не являются конкретно ограниченными, однако их димеризации (после модифицирования с помощью РЕ) в присутствии биспецифического лиганда продуцирует некоторый полезный биологический эффект, например, путь передачи сигнала, сниженную иммунореактивность и т.д. Согласно различным аспектам белки-мишени, подвергаемые димеризации с помощью РЕ-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 т.д. Иллюстративные белки-мишени, для которых может использоваться РЕ для внесения доменов димеризации, включают без ограничения:

СВЯЗАННЫЙ С МЕМБРАНОЙ РЕЦЕПТОР	КИНАЗНЫЙ ДОМЕН СЛИТЫЙ С	ИСПОЛЬЗУЕМЫЙ CID	КАСКАД ПЕРЕДАЧИ СИГНАЛА	ССЫЛОЧНЫЙ ИСТОЧНИК
Т-КЛЕТОЧНЫЙ РЕЦЕПТОР	FKBP12	FK1012 (FK506 DIMER)	СИГНАЛ Т-КЛЕТОЧНОГО РЕЦЕПТОРА	*SCIENCE* 262, 1019–24 (1993) *CHEM. BIOL.* 1, 163–172 (1994).
РЕЦЕПТОР FAS	MURINE CYCLOPHILIN C	CYCLOSPORIN A DIMER	ПУТЬ FAS ДЛЯ АПОПТОЗА	*CHEM. BIOL.* 3, 731–738 (1996).
РЕЦЕПТОР ИНСУЛИНА	FKBP12	FK1012 (FK506 DIMER)	ПЕРЕДАЧА СИГНАЛА ИНСУЛИНА	*CURR. BIOL.* 8, 11–18 (1998).
ПОЛУЧЕННЫЙ ИЗ ТРОМБОЦИТОВ ФАКТОР РОСТА БЕТА (PDGF	FKBP12	FK1012 (FK506 DIMER)	СИГНАЛ ОБРАЗОВАНИЯ МЕЗОДЕРМЫ PDGF	*CURR. BIOL.* 8, 11–18 (1998).
РЕЦЕПТОР ЭРИТРОПОЭТИ	FKBP12	FK1012 (FK506	EPOR-ОПОСРЕДОВ	*PROC. NATL. ACAD. SCI.* 94, 3076–3081 (2002).

НА (EPOR)		DIMER)	АННЫЙ ПРОЛИФЕРАТИВНЫЙ СИГНАЛ	

[0914] Согласно одному аспекту системы мультилоскутного праймирующего редактора, описанные в настоящем документе, могут применяться для введения 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 мультилоскутное праймированное редактирование - CID)», где CID представляет собой биспецифический лиганд, который индуцирует димеризацию белков-мишеней, каждый из которых слит с доменом димеризации, внесенным с помощью мультилоскутного 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 передачу сигнала инсулина. В настоящем документе описан новый способ внесения индуцируемых димеризацией биологических процессов, например, передачи сигнала рецептора, под контролем подходящего низкомолекулярного лекарственного средства (т.е. биспецифического лиганда) путем геномного встраивания генов, кодирующих связывающие малые молекулы белки (т.е. домены димеризации), с помощью мультилоскутного праймированного редактирования.

[0915] Димеризация белков представляет собой повсеместный биологический процесс. Следует отметить, что известно, что гомодимеризация многих связанных с мембраной рецепторов инициирует каскады передачи сигнала, часто со значительными биологическими последствиями. Ряд низкомолекулярных встречающихся в природе продуктов, одобренных для применения в качестве лекарственных средств, 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 стадию передачи сигнала Т-клеточного рецептора.[93] Аналогичным образом, циклоспорин А индуцирует димеризацию циклофилина и кальциневрина, а рапамицин индуцирует димеризацию FKBP и mTOR.[93,94]

[0916] Согласно одному варианту осуществления на основе использования

селективного высокоаффинного связывания FK506:FKBP12 и взаимодействия связывания циклоспорин А:малая молекула циклофилина:белок также были разработаны синтетические химические индукторы димеризации. В одном примере было показано, что малая молекула, состоящая из двух единиц FK506, обозначенная FK1012, оказывает влияние на трансдукцию сигнала, когда цитоплазматические домены сигнальных рецепторов были помечены FKBP12.[95] После этого химические индукторы димеризации (СШ) стали использоваться для управления рядом путей передачи сигнала.[96-103]

[0917] Являясь полезными инструментами для исследования биологических процессов, одной из проблем, связанных с использованием CID для терапевтических применений, является то, что внесение химер FKBP12– или циклофилина–целевого белка пациентам является сложным.

[0918] Настоящее изобретение объединяет множество концепции для создания ранее недоступного терапевтического способа. Первая концепция представляет собой мультилоскутное праймированное редактирование (*например*, двухлоскутное или четырехлоскутное праймированное редактирование), описанное в настоящем документе, 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.

[0919] Били идентифицированы конкретные случаи, при которых химический контроль над димеризацией белка может иметь полезный терапевтический эффект.

[0920] Рецептор инсулина представляет собой гетеротетрамерный трансмембранный белок, который отвечает за связывание инсулина с внеклеточным доменом посредством фосфорилирования цитоплазматического домена киназы.[104] Сконструированный химерный белок, состоящий из компонента локализации на мембране, С-концевого киназного домена рецептора инсулина и трех копий FKBP12, отвечает на FK1012 и инициирует инсулиновый ответ в клеточной культуре. [99] Подобным 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).

[0921] Кроме того, эритропоэтин стимулирует пролиферацию эритроцитов

посредством 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 EK1012-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 праймированного редактирования в клетках пациентов обеспечит FK1012-индуцированный контроль над пролиферацией эритроцитов (эритропоэз). Эта система может применяться для инициирования роста красных кровяных клеток у пациентов с анемией. EK1012-индуцируемый EpoR также может использоваться в качестве *in vivo* маркера отбора для кровяных клеток, которые подверглись *ex vivo* модифицированию.

[0922] В принципе, любая рецепторная тирозинкиназа может представлять собой жизнеспособную мишень для мультилоскутного праймированного редактирования–терапевтического CID. Таблица ниже включает перечень всех рецепторных тирозинкиназ в геноме человека.[110]

Семейство	Рецептор	Синоним	NT доступ	PROT доступ	Хромосома
ALK семейство	ALK	Ki1	NM_004304	NP_004295	2p23
	LTK	TYK1	NM_002344	NP_002335	15q15.1-q21.1
AXL семейство	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
DDR семейство	DDR1	CAK, TRKE, NEP NTRK4, EDDR1, PTK3	NM_013993	NP_001945	6p21.3
	DDR2	TKT, TYRO10, NTRKR3	NM_006182	NP_006173	1q21-q22
EGFR семейство	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
EPH семейство	EPHA1	EPH, EPHT	NM_005232	NP_005223	7q32-q36
	EPHA2	ECK, Sek(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	2q36qter

	EPHA5	HEK7, Ehk(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
FGFR семейство	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
INSR семейство	IGF1R	JTK13	NM_000875	NP_000866	15q25-q26
	INSR	IR	NM_000208	NP_000199	19p13.3-p13.2
	INSRR	IRR	J05046	AAC31759	1q21-q23
MET семейство	MET	HGFR	NM_000245	NP_000236	7q31
	RON	MST1R, CDw136, Fv2(m), STK(m), SEA(ch)	NM_002447	NP_002438	3p21.3
MUSK семейство	MUSK	Nsk2(m), Mlk1(m), Mk2(m)	NM_005592	NP_005583	9q31.3-q32
PDGFR семейство	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
PTK7 семейство	PTK7	CCK4, KLG(ch)	NM_002821	NP_002812	6p21.2-p12.2
RET семейство	RET	MEN2A/B, HSCR1, MTC1	X12949	P07949	10q11.2
ROR семейство	ROR1	NTRKR1	NM_005012	NP_005003	1p32-p31
	ROR2	NTRKR2	NM_004560	NP_004551	
ROS семейство	ROS1	MCF3	NM_002944	NP_002935	6q22
RYK семейство	RYK	Vlk(m), Mrk(m)	S59184	AAB263411	3q22
TIE семейство	TEK	TIE2	NM_000459	NP_000450	9p21
	TIE	TIE1, JTK14	NM_005424	NP_005415	1p34-p33
TRK семейство	NTRK1	TRK, TRKA	NM_002529	NP_002520	1q21-q22
	NTRK2	TRKB	NM_006180	NP_006171	9q22.1
	NTRK3	TRKC	NM_002530	NP_002521	15q25
VEGFR семейство	VEGFR1	FLT1	NM_002019	NP_002010	13q12
	VEGFR2	KDR, FLK1		AAB88005	4q11-q12
	VEGFR3	FLT4, PCL	NM_002020	NP_002011	5q34-q35
AATYK семейство	AATYK	AATK, KIAA0641	NM_004920	NP_004911	17q25.3
	AATYK2	KIAA1079	NM_014916	NP_055731	7q21-q22
	AATYK3				19q13.2-q13.3

[0923] Система CID с мультилоскутным праймирующим редактированием имеет множество преимуществ. Одним из таких преимуществ является то, что она может заменить эндогенные лиганды, которые обычно представляют собой белки, которые создают сложности в производстве, стоимости, доставке, производстве или хранении, на подобные лекарственным средствам малым молекулам, которые можно вводить перорально вместо внутривенного введения или инъекции, являются легко получаемыми из одобренных FDA лекарственных средств (или сами уже являются лекарственными средствами) и не приводят к высоким затратам на производство или хранение, обычно связанными с белковыми лекарственными средствами. Другое преимущество состоит в том, что редактирование само по себе не должно иметь никакого физиологического эффекта. Степень димеризации целевого белка можно контролировать путем дозирования низкомолекулярных CID. Кроме того, димеризация целевого белка легко и быстро

обратима при добавлении мономерной формы CID. Еще одно преимущество заключается в том, что в случаях, когда один лиганд нацелен на множество рецепторов, селективность может быть достигнута путем праймированного редактирования только одного рецептора. Наконец, в зависимости от способа доставки, используемого для праймированного редактирования, также может быть возможно ограничить редактирование локализованной тканью или органом, позволяя индуцируемую активацию рецептора только в определенных областях.

[0924] Если эффективность редактирования достаточно высока при мультилоскутном праймированном редактировании, так что два отдельных события редактирования могут происходить на высоких уровнях, также можно было бы пометить два представляющих интерес белка с помощью разных низкомолекулярных связывающих доменов (таких как FKBP и циклофилин) и индуцировать гетеродимеризацию с низкомолекулярными гетеродимерами (такими как димер FK506-циклоспорин A).

[0925] Слияние FKBP12 или других низкомолекулярных связывающих белков с нативными белками было достигнуто, как правило, путем сверхэкспрессии из плазмиды в культуре ткани. Было продемонстрировано, что последующая химически индуцированная димеризация вызывает фенотипические изменения в клетках, продуцирующих слитые белки.

[0926] Следующие ссылочные источники цитируются в приведенном выше разделе и включены в настоящий документ посредством 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).

G. Применение мультилоскутного праймированного редактирования для включения сайтов-мишеней рекомбиназы

[0927] Согласно другому аспекту мультилоскутное праймированное редактирование (например, как представлено на фиг. 90, 95, 97A, 98A) можно использовать для вставки сайтов рекомбиназы (или «последовательностей распознавания рекомбиназы»)

в желаемый участок генома. Вставка сайтов рекомбиназы обеспечивает программируемое место для осуществления сайт-специфических генетических изменений в геноме. Такие генетические изменения могут включать, например, геномную интеграцию плазмиды, геномную делецию или вставку, хромосомные транслокации и обмен кассетами, среди других генетических изменений. Эти иллюстративные типы генетических изменений проиллюстрированы на фиг. 64(b)-(f). Затем установленные последовательности распознавания рекомбиназы можно использовать для проведения сайт-специфической рекомбинации в этом сайте для осуществления различных результатов рекомбинации, таких как вырезание, интеграция, инверсия или замена фрагментов ДНК. Например, на фиг. 65 проиллюстрирована установка сайта рекомбиназы, который затем можно использовать для интеграции матрицы донора ДНК, содержащей маркер экспрессии GFP. Клетки, содержащие интегрированную систему экспрессии GFP в сайт рекомбиназы, будут флуоресцировать.

[0928] Механизм вставки сайта рекомбиназы в геном аналогичен установке в геном других последовательностей, таких как пептидные/белковые и РНК-метки. Схема, иллюстрирующая установку целевой последовательности рекомбиназы, показана на фиг. 64 (a). Способ начинается с выбора желаемого локуса-мишени, в который будет введена целевая последовательность рекомбиназы. Затем предоставляется слияние праймирующего редактора («RT-Cas9:гРНК»). Согласно настоящему изобретению «гРНК» относится к РЕгРНК, которую можно сконструировать с использованием описанных в настоящем документе принципов. РЕгРНК согласно различным вариантам осуществления будет иметь конструкцию, соответствующую фиг.3D (5'-[~20-нуклеотидный спейсер]-[ядро гРНК]-[плечо удлинения]-3', где плечо удлинения содержит в направлении от 3' к 5' участок связывания праймера («A»), матрицу редактирования («B») и плечо гомологии («C»). Матрица редактирования («B») будет содержать последовательность, соответствующую сайту рекомбиназы, т.е. РЕгРНК одноцепочечной РНК, которая кодирует комплементарную одноцепочечную ДНК, которая является либо смысловой, либо антисмысловой цепью сайта рекомбиназы, и которая встраивается в целевой локус геномной ДНК в ходе способа мультилоскутного праймированного редактирования.

[0929] Согласно различным аспектам настоящее раскрытие относится к применению мультилоскутного РЕ для введения последовательностей распознавания рекомбиназы в важные локусы человеческого или другого генома, которые после воздействия сайт-специфической рекомбиназой (рекомбиназами) будут направлять точные и эффективные геномные модификации (фиг. 64). Согласно различным вариантам осуществления, показанным на фиг. 64, РЕ можно использовать для (b) вставки одной SSR-

мишени для использования в качестве сайта для геномной интеграции матрицы донора ДНК, (c) показано, как можно использовать тандемную вставку целевых участков SSR для удаления части генома, (d) показано, как можно использовать тандемную вставку целевых участков SSR для инвертирования части генома, (e) показано, как вставка двух целевых участков SSR в двух дистальных участках хромосомы может привести к хромосомной транслокации, (f) показано, как вставку двух разных целевых участков SSR в геном можно использовать для замены кассеты из матрицы донора ДНК. Каждый из типов модификаций генома предполагает использование PE для вставки SSR-меток, но этот перечень также не является ограничивающим.

[0930] Мультилоскутное PE-опосредованное введение последовательностей распознавания 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 методик 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 введение последовательностей распознавания рекомбиназы 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 перестроение в трансгенных растениях в мультилоскутное PE-специфичных мишенях может обеспечивать преодоление одного из препятствий в получении сельскохозяйственных растений с улучшенными свойствами[8,9].

[0931] Таблица 6. Примеры генетических заболеваний, связанных с крупномасштабными геномными модификациями, которые могут быть устранены путем установки последовательностей распознавания рекомбиназы на основе мультилоскутного PE.

ЗАБОЛЕВАНИЕ	ПРИЧИНА
ТРИСОМИЯ 17P	ДУБЛИРОВАНИЕ ГЕНОВ
БОЛЕЗНЬ ШАРКО-МАРИ-ТУТА ТИПА I	ДУБЛИРОВАНИЕ ГЕНОВ
СИНДРОМ СМИТА-МАГЕНИСА	УДАЛЕНИЕ ГЕНА

СИНДРОМ УИЛЬЯМСА-БОЙРЕНА	УДАЛЕНИЕ ГЕНА
СИНДРОМ ДЕ ЛЯ ШАПЕЛЯ	ХРОМОСОМНАЯ ТРАНСЛОКАЦИЯ
СИНДРОМ ДАУНА (НЕКОТОРЫЕ ФОРМЫ)	ХРОМОСОМНАЯ ТРАНСЛОКАЦИЯ
ГЕМОФИЛИЯ А	ГЕННАЯ ИНВЕРСИЯ
СИНДРОМ ХАНТЕРА	ГЕННАЯ ИНВЕРСИЯ

[0932] Был охарактеризован ряд членов семейства SSR, а их последовательности распознавания рекомбиназы были 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 встраивания в геном, 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 этих последовательностей распознавания может быть внесена в геном в определенном месте, например, в локус «safe harbor»[41-43], в зависимости от желаемого применения.

[0933] Например, внесение одной мишени распознавания рекомбиназы в геном путем праймированного редактирования будет приводить к интегративной рекомбинации на основе донорской матрицы ДНК (фиг. 64b). Сериновые интегразы, которые функционируют повсеместно в клеток человека, могут являться особенно подходящими для встраивания в ген[44,45].

[0934] Кроме того, внесение двух мишеней распознавания 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). Посредством выбора эндогенных последовательностей, которые уже близко напоминают мишени рекомбиназы, объем редактирования, необходимого для внесения полной мишени рекомбиназы, будет снижен.

[0935] Наконец, было продемонстрировано, что несколько рекомбиназ встраиваются в геном человека или эукариотические геномы в нативно расположенных псевдосайтах[46-64]. Мультилоскутное праймированное редактирование 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.

[0936] В настоящем документе раскрыта общая методология введения последовательностей-мишеней рекомбиназы в эукариотические геномы с помощью мультилоскутного PE, применения которого являются практически неограниченными.

Реакции геномного редактирования предназначены для применения с химерным слиянием «праймирующих редакторов» CRISPR/белка Cas9 и домена обратной транскриптазы, который использует гидовую РНК праймированного редактирования (PEгРНК). Посредством удлинения инструменты на основе Cas9 и пути репарации, направляемой гомологией, (HDR) также могут использоваться для внесения последовательностей распознавания рекомбиназоф с помощью ДНК-матриц путем понижения показателей вставок-делеций с помощью ряда способов[65-67]. Эксперимент для подтверждения данной концепции в культуре клеток человека показан на фиг. 65. Согласно некоторым вариантам осуществления праймирующие редакторы представляют собой мультилоскутные праймирующие редакторы.

[0937] Следующие несколько таблиц цитируются в вышеприведенном описании, относящемся к PE-направленной установке последовательностей распознавания рекомбиназы, и представляют перечень иллюстративных рекомбиназ, которые можно использовать, и родственные им последовательности распознавания рекомбиназы, которые могут быть установлены мультилоскутным PE.

Таблица 7. Тирозинрекомбиназы и целевые последовательности SSR

Рекомбиназа	**Последовательность распознавания рекомбиназой**	**наз-вание**
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	GAGTTTCATTAAGGAATAACTAATTCCCTAATGAAACTC (SEQ ID NO: 522)	*loxB*
B3	GGTTGCTTAAGAATAAGTAATTCTTAAGCAACC (SEQ ID NO: 523)	*loxB3*
Kw	ACGAAAAATGGTAAGGAATAGACCATTCCTTACCATTTTTGGT (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	AAACGATATCAGACATTTGTCTGATAATGCTTCATTATCAGACAAAT GTCTGATATCGTTT (SEQ ID NO: 530)	*loxK*

Рекомбиназа	**Последовательность распознавания рекомбиназой**	**наз-вание**
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 целевые последовательности SSR

Рекомбиназа	**Последовательность распознавания рекомбиназой, левая**	**Последовательность распознавания рекомбиназой, правая**
Bxb1	GGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACC (SEQ ID NO: 536)	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCAT (SEQ ID NO: 537)
phiC31	GTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGG (SEQ ID NO: 538)	TGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCC (SEQ ID NO: 539)
R4	TGTTCCCCAAAGCGATACCACTTGAAGCAGTGGTACTGCTTGTGGGTACA (SEQ ID NO: 540)	GCATGTTCCCCAAAGCGATACCACTTGAAGCAGTGGTACTGCTTGTGGGTACACTCTGCGGGTG (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	TTTCGGATCAAGCTATGAAGGACGCAAAGAGGGAACTAAA (SEQ ID NO: 553)

Рекомбиназа	**Последовательность распознавания рекомбиназой, левая**	**Последовательность распознавания рекомбиназой, правая**
	NO: 552)	
phiRV1	GTAGTGTATCTCACAGGTCCACGG TTGGCCGTGGACTGCTGAAGAAC ATTCC (SEQ ID NO: 554)	GAAGGTGTTGGTGCGGGGTTGGCCGT GGTCGAGGTGGGGT (SEQ ID NO: 555)
phi370.1	AAAAAAATACAGCGTTTTTCATGT ACAACTATACTAGTTGTAGTGCCT AAAA (SEQ ID NO: 556)	TTGTAAAGGAGACTGATAATGGCATG TACAACTATACTCGTCGGTAAAAAGG CA (SEQ ID NO: 557)
TG1	TCCAGCCCAACAGTGTTAGTCTTT GCTCTTACCCAGTTGGGCGGGA (SEQ ID NO: 558)	GATCAGCTCCGCGGGCAAGACCTTTCT CCTTCACGGGGTGGAAGGTC (SEQ ID NO: 559)
WB	CTAGTTTTAAAGTTGGTTATTAGT TACTGTGATATTTATCACGGTACC CAATAACCAATGAAT (SEQ ID NO: 560)	CGGAAGGTAGCGTCAACGATAGGTGT AACTGTCGTGTTTGTAACGGTACTTCC AACAGCTGGCGCCGCCAC (SEQ ID NO: 561)
BL3	CAATGAAAAACTAGGCATGTAGA AGTTGTTTGT (SEQ ID NO: 562)	TTTCCACAGACAACTCACGTGGAGGT AGTCAC (SEQ ID NO: 563)
SprA	TGTAGTAAGTATCTTAATATACAG CTTTATCTGTTTTTTAAGATACTTA CTACTTT (SEQ ID NO: 564)	CACCCATTGTGTTCACAGGAGATACA GCTTTATCTGTACTGATATTAATGACA TGCTG (SEQ ID NO: 565)
phiJoe	AGTTGTGGCCATGTGTCCATCTGG GGGCAGATGGAGACGGGGTCACA (SEQ ID NO: 566)	ATCTGGATGTGGGTGTCCATCTGCGGG CAGACGCCGCAGTCGAAGCACGG (SEQ ID NO: 567)
--	--	ACCTTGATCTCGGTGTCCATCGCCGGG CAGACGCCGCAGTCGAAGCACGG (SEQ ID NO: 568)
phiK38	CCCTAATACGCAAGTCGATAACTC TCCTGGGAGCGTTGACAACTTGCG CACCCTGATCTG (SEQ ID NO: 569)	GAGCGCCGGATCAGGGAGTGGACGGC CTGGGAGCGCTACACGCTGTGGCTGC GGTCGGTGC (SEQ ID NO: 570)
Int2	GCTCATGTATGTGTCTACGCGAGA TTCTCGCCCGAGAACTTCTGCAAG GCACTGCTCTTGGCT (SEQ ID NO: 571)	GGACGGCGCAGAAGGGGAGTAGCTCT TCGCCGGACCGTCGACATACTGCTCAG CTCGTC (SEQ ID NO: 572)
Int3	ATGGATAAAAAAATACAGCGTTT TTCATGTACAACTATACTAGTTGT AGTGCCTAAATAATGCTT (SEQ ID	GTTTGTAAAGGAGACTGATAATGGCA TGTACAACTATACTCGTCGGTAAAAA GGCATCTTAT (SEQ ID NO: 574)

Рекомбиназа	Последовательность распознавания рекомбиназой, левая	Последовательность распознавания рекомбиназой, правая
	NO: 573)	
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 (SEQ ID NO: 585)	ATGGATTTTGCAGATTCCCAGATGCCCCTACAGAAAGAGGTACAAAACATTTATTGGAATTAATT (SEQ ID NO: 586)
Int12	TTTTTGTATGTTAGTTGTGTCACTGGGTAGACCTAAATAGTGACACAACTGCTATTAAAATTTAA (SEQ ID NO: 587)	GTTCGTGGTAACTATGGGTGGTACAGGTGCCACATTAGTTGTACCATTTATGTTTATGTGGTTAAC (SEQ ID NO: 588)
Int13	CAATAACGGTTGTATTTGTAGAACTTGACCAGTTGTTTTAGTAACATAAATACAACTCCGAATA (SEQ ID NO: 589)	GCATACATTGTTGTTGTTTTTCCAGATCCAGTTGGTCCTGTAAATATAAGCAATCCATGTGAGT (SEQ ID NO: 590)
LI	GTTTAGTATCTCGTTATCTCTCGTTGGAGGGAGAAGAAACGGGATACC	TAACTTTTTCGGATCGAGTTATGATGGACGTAAAGAGGGAACAAAGCATCTA

Рекомбиназа	**Последовательность распознавания рекомбиназой, левая**	**Последовательность распознавания рекомбиназой, правая**
	AAAA (SEQ ID NO: 591)	(SEQ ID NO: 592)
Peaches	TAGTTTCCAATGTTACAGGAACTG CTGGCAGAATCCAACACATTGGA AGTCG (SEQ ID NO: 593)	CGGTCTCCATCGGGATCTGCTGATCGA GCAGCATGCCGACCA (SEQ ID NO: 594)
Bxz2	TAACCGCAAGTGTACATCCCTCGG CTGGCCGAGACAAGTACAGTTGC GACAG (SEQ ID NO: 595)	CGGTCTCCATCGGGATCTGCTGATCGA GCAGCATGCCGACCA (SEQ ID NO: 596)
SV1	ATGTGGTCCTTTAGATCCACTGAC GTGGGTCAGTGTCTCTAAAGGACT CGCG (SEQ ID NO: 597)	CATCAGGGCGGTCAGGCCGTAGATGT GGAAGAAACGGCAGCACGGCGAGGA CG (SEQ ID NO: 598)

Таблица 9. Сериновые резольвазы и целевые последовательности SSR

Резольваза	**Последовательность распознавания рекомбиназой (левая)**	**Последовательность распознавания рекомбиназой (правая)**
Gin	CGTTTCCTGTAAACCGAGGTTTTGGA TAAACA (SEQ ID NO: 599)	CGTTTCCTGTAAACCGAGGTTTTGGA TAATGG (SEQ ID NO: 600)
Cin	GAGTTCTCTTAAACCAAGGTTTAGGA TTGAAA (SEQ ID NO: 601)	GAGTTCTCTTAAACCAAGGTATTGGA TAACAG (SEQ ID NO: 602)
Hin	TGGTTCTTGAAAACCAAGGTTTTTGA TAAAGC (SEQ ID NO: 603)	AAATTTTCCTTTTTGGAAGGTTTTTGA TAACCA (SEQ ID NO: 604)
Min	GCCTTCCCCTAAACCAACGTTTTTAT GCCGCC (SEQ ID NO: 605)	GCCTTCCCCCAAACCAAGGTAATCAA GAACGC (SEQ ID NO: 606)
Sin	TTGTGAAATTTGGGTACACCCTAATC ATACAA (SEQ ID NO: 607)	CGTATGATTAGGGTGTATATTAATTT (SEQ ID NO: 608)

Таблица 10. Тирозининтегразы и целевые последовательности

Интеграза	***attP***	***attB***
HK022	CAAATGATTTTATTTTGACTAATAATGAC CTACTTACATTAATTTACTGATAATTAAA GAGATTTTAAATATACAACTTATTCACCT AAAGGATGACAAAA (SEQ ID NO: 609)	GCACTTTAGGTGAAAAAGGTT (SEQ ID NO: 610)
	TAACATTAATCACTTAAAAATCATCGCAT TACACTAATCTGTGGTTAAATGATAGACT ACATAATGCGACAAAACGCAACATATCC AGTCACTATGAATCAACTACTTAGATAGT	

	ATTAGTGACCT (SEQ ID NO: 611)	
P22	CTAAGTGGTTTGGGACAAAAATGGGACA TACAAATCTTTGCATCGGTTTGCAAGGCT TTGCATGTCTTTCGAAGATGGGACGTGTG AGCGCAGGTATGACGTGGTATGTGTTGA CTTAAAAGGTAGTTCTTATAATTCGTAAT GCGAAGGTCGTAGGTTCGACTCCTATTAT CGGCACCAGTTAAATCAAATACTTACGT ATTATTCGTGCCTTCCTTATTTTTACTGTG GGACATATTTGGGACAGAAGTACCAAAA A (SEQ ID NO: 612)	GCAGCGCATTCGTAATGCGAAGG TCGT (SEQ ID NO: 613)
L5	GCGATCCCCATCCGCGACGTGCCAACTA GGTCTCCTCTCGTCGTGAACAAGGCTACC GGGTTGCAACTCCTGTGCAACTCTCAGGC TTCAACGCGCTTCTACGACCTGCAATTTC TTTCCACTTAGAGGATGCAGCCGAGAGG GGTAAAAACCTATCTTGACCGGCCCATA TGTGGTCGGCAGACACCCATTCTTCCAAA CTAGCTACGCGGGTTCGATTCCCGTCGCC CGCTCCGCTGGTCAGAGGGTGTTTTCGCC CTCTGGCCATTTTTCTTTCCAGGGGTCTG CAACTCTTGTGCGACTCTTCTGACCTGGG CATACGCGGTTGCAACGCATCCCTGATCT GGCTACTTTCGATGCTGACAAACGAATA GAGCCCCCCGCCTGCGCGAACAGACGAG GGGCATTCACA (SEQ ID NO: 614)	GAGCGGGCGACGGGAATCGAACC CGCGTAGCTAGTTTGGAAGA (SEQ ID NO: 615)

[0938] Согласно различным другим аспектам настоящее изобретение относится к способам применения мультилоскутного PE для внесения одной или нескольких последовательностей распознавания рекомбиназой и их применению для сайт-специфической рекомбинации.

[0939] Согласно некоторым вариантам осуществления сайт-специфическая рекомбинация может обеспечивать получение ряда результатов рекомбинации, например, эксцизии, встраивания, инверсии, или замены фрагментов ДНК.

[0940] Согласно некоторым вариантам осуществления способы полезны для внесения рекомбинации одного или нескольких участков или между одним или несколькими участками одной или нескольких молекул нуклеиновых кислот (например,

ДНК). Согласно другим вариантам осуществления способы полезны для внесения рекомбинации одного или нескольких участков или между одним или несколькими участками в одной молекуле нуклеиновой кислоты (например, ДНК).

[0941] Согласно некоторым вариантам осуществления настоящее изобретение относится к способу встраивания донорской матрицы ДНК с помощью сайт-специфической рекомбинации, предусматривающему: (a) внесение последовательности распознавания рекомбиназой в геномный локус с помощью мультилоскутного праймированного редактирования, (b) контакт геномного локуса с донорской матрицей ДНК, которая также содержит последовательность распознавания рекомбиназой, в присутствии рекомбиназы.

[0942] Согласно другим вариантам осуществления настоящее изобретение относится к способу удаления участка генома с помощью сайт-специфической рекомбинации, предусматривающему: (a) внесение пары последовательностей распознавания рекомбиназой в геномный локус с помощью мультилоскутного праймированного редактирования, (b) контакт геномного локуса с рекомбиназой, тем самым обеспечивая катализ удаления участка генома между парой последовательностей распознавания рекомбиназой.

[0943] Согласно другим вариантам осуществления настоящее изобретение относится к способу инверсии участка генома с помощью сайт-специфической рекомбинации, предусматривающему: (a) внесение пары последовательностей распознавания рекомбиназой в геномный локус с помощью мультилоскутного праймированного редактирования, (b) контакт геномного локуса с рекомбиназой, тем самым обеспечивая катализ инверсии участка генома между парой последовательностей распознавания рекомбиназой.

[0944] Согласно другим вариантам осуществления, настоящее изобретение относится к способу индуцирования хромосомной транслокации между первым участком генома и вторым участком генома, предусматривающему: (a)внесение первой последовательности распознавания рекомбиназой в первый геномный локус с помощью праймированного редактирования, (b) внесение второй последовательности распознавания рекомбиназой во второй геномный локус с помощью мультилоскутного праймированного редактирования, (c) контакт первого и второго геномного локуса с рекомбиназой, тем самым обеспечивая хромосомную транслокацию первого и второго геномного локуса.

[0945] Согласно другим вариантам осуществления настоящее изобретение относится к способу индуцирования обмена кассетами между геномным локусом и донорской ДНК, включающей кассету, предусматривающему: (a) внесение первой последовательности распознавания рекомбиназой в первый геномный локус с помощью

мультилоскутного праймированного редактирования, (b) внесение второй последовательности распознавания рекомбиназой во второй геномный локус с помощью мультилоскутного праймированного редактирования, (c) контакт первого и второго геномного локуса с донорской ДНК, включающей кассету, которая фланкируется первой и второй последовательностию распознавания рекомбиназой, и рекомбиназой, тем самым обеспечивая катализ обмена фланкированного геномного локуса и кассеты в донорской ДНК.

[0946] Согласно различным вариантам осуществления, включающих вставку более чем одной последовательности распознавания рекомбиназой в геном, последовательности распознавания рекомбиназой могут являться одинаковыми или различающимися. Согласно некоторым вариантам осуществления последовательности распознавания рекомбиназой являются одинаковыми. Согласно другим вариантам осуществления данные последовательности распознавания рекомбиназой являются различающимися.

[0947] Согласно различным вариантам осуществления рекомбиназа может представлять собой тирозиновую рекомбиназу, например, Cre, Dre, Vcre, Sere, Flp, B2, B3, Kw, R, TD1-40, Vika, Nigri, Panto, Kd, Fre, Cre(ALSHG), Tre, Brecl или Cre-R3M3, как показано в Таблице 7. Согласно таким вариантам осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 7, которая соответствует используемой рекомбиназе.

[0948] Согласно различным другим вариантам 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. Согласно таким вариантам осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 8, которая соответствует используемой рекомбиназе.

[0949] Согласно другим вариантам 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. Согласно таким вариантам осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 8, которая соответствует используемой рекомбиназе.

[0950] Согласно другим вариантам осуществления рекомбиназа может представлять собой сериновую резольвазу, например, Gin, Cin, Hin, Min или Sin, как показано в Таблице 9. Согласно таким вариантам осуществления последовательность распознавания

рекомбиназой может представлять собой RRS из Таблицы 9, которая соответствует используемой рекомбиназе.

[0951] Согласно различным другим вариантам осуществления рекомбиназа может представлять собой тирозиновую интегразу, например, HK022, P22, или L5, как показано в Таблице 10. Согласно таким вариантам осуществления последовательность распознавания рекомбиназой может представлять собой RRS из Таблицы 10, которая соответствует используемой рекомбиназе.

[0952] Согласно некоторым вариантам осуществления любой из способов сайт-специфической рекомбинации с использованием мультилоскутного PE может быть реализован *in vivo* или *in vitro.* Согласно некоторым вариантам осуществления любой из способов сайт-специфической 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 в настоящем документе субъект (например, представляющий собой человека субъект). Описанные в настоящем документе 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.

[8] Варианты осуществления мультилоскутных праймирующих редакторов, описанных в настоящем документе

[0953] Следующие неограничивающие варианты осуществления описывают аспекты двойных праймирующих редакторов и их применения, как описано в настоящем документе.

1. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, где система содержит первый комплекс праймированного редактирования и второй комплекс праймированного редактирования, где каждый из первого и второго комплексов праймированного редактирования содержит (1) праймирующий редактор, содержащий (i) программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и (ii) полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и (2) PEгРНК, содержащую спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера, где матрица для синтеза ДНК PEгРНК первого комплекса праймированного редактирования кодирует первую последовательность одноцепочечной ДНК, и матрица для синтеза ДНК PEгРНК второго комплекса праймированного

редактирования кодирует вторую последовательность одноцепочечной ДНК, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть, по сравнению с последовательностью ДНК в целевом участке, подлежащем редактированию, который вставляется в целевой участок, подлежащий редактированию.

2. Система согласно варианту осуществления 1, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

3. Система согласно варианту осуществления 1 или 2, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

4. Система согласно варианту осуществления 1 или 2, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат разные праймирующие редакторы.

5. Система согласно любому из вариантов осуществления 1-4, где napDNAbp представляет собой домен Cas9 или его вариант.

6. Система согласно любому из вариантов осуществления 1-5, где napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, никазный домен Cas9 или их вариант.

7. Система согласно любому из вариантов осуществления 1-4, где napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

8. Система согласно любому из вариантов осуществления 1-7, где napDNAbp обладает никазной активностью.

9. Система согласно любому из вариантов осуществления 1-4, где napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

10. Система согласно любому из вариантов осуществления 1-9, где полипептид,

обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

11. Система согласно любому из вариантов осуществления 1-10, где полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

12. Система согласно любому из вариантов осуществления 1-11, где праймирующий редактор дополнительно содержит линкер, который соединяет napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

13. Система согласно варианту осуществления 12, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

14. Система согласно варианту осуществления 12 или 13, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

15. Система согласно любому из вариантов осуществления 1-14, где РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

16. Система согласно любому из вариантов осуществления 1-15, где спейсерная последовательность РЕгРНК первого комплекса праймированного редактирования связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

17. Система согласно любому из вариантов осуществления 1-16, где спейсерная последовательность РЕгРНК второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности

двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

18. Система согласно варианту осуществления 17, где связывание спейсерной последовательности РЕгРНК первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM, на каждой цепи посредством napDNAbp каждого из первого и второго комплексов праймированного редактирования.

19. Система согласно варианту осуществления 18, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

20. Система согласно варианту осуществления 19, где непрерывная область между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по

меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

21. Система согласно варианту осуществления 19, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

22. Система согласно любому из вариантов осуществления 1-21, где матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по

меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

23. Система согласно любому из вариантов осуществления 1-22, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

24. Система согласно любому из вариантов осуществления 1-22, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

25. Система согласно любому из вариантов осуществления 1-24, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

26. Система согласно варианту осуществления 25, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине желаемой последовательности ДНК и ее комплемента.

27. Система согласно варианту осуществления 25, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть.

28. Система согласно любому из вариантов осуществления 1-27, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

29. Система согласно варианту осуществления 21, где замещение непрерывной

области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

30. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, причем указанная система содержит:

(a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первая РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

(b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторая РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, и вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, и где первая цепь и вторая цепь последовательности двухцепочечной ДНК комплементарны друг другу.

31. Система согласно варианту осуществления 30, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

32. Система согласно варианту осуществления 30, где первая и вторая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

33. Система согласно варианту осуществления 30, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть, которая вставляется в целевой участок, подлежащий редактированию.

34. Система согласно варианту осуществления 30, где первая последовательность одноцепочечной ДНК и/или вторая последовательность одноцепочечной ДНК не имеют гомологии последовательности по сравнению с последовательностью ДНК в целевом участке.

35. Система согласно варианту осуществления 30, где первая РЕгРНК дополнительно содержит первую спейсерную последовательность, первое ядро гРНК и первый участок связывания праймера, и вторая РЕгРНК дополнительно содержит вторую спейсерную последовательность, второе ядро гРНК и второй участок связывания праймера.

36. Система согласно любому из вариантов осуществления 30-35, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

37. Система согласно любому из вариантов осуществления 30-35, где первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, являются одинаковыми.

38. Система согласно любому из вариантов осуществления 30-37, где первый и/или второй napDNAbp представляет собой домен Cas9 или его вариант.

39. Система согласно любому из вариантов осуществления 30-38, где первый и/или второй napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, никазный домен Cas9 или их вариант.

40. Система согласно любому из вариантов осуществления 30-39, где первый и/или второй napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

41. Система согласно любому из вариантов осуществления 30-40, где napDNAbp обладает никазной активностью.

42. Система согласно любому из вариантов осуществления 30-41, где первый и/или второй napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

43. Система согласно любому из вариантов осуществления 30-42, где первый

и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

44. Система согласно любому из вариантов осуществления 30-43, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

45. Система согласно любому из вариантов осуществления 30-44, где первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

46. Система согласно варианту осуществления 45, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

47. Система согласно варианту осуществления 45 или 46, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

48. Система согласно любому из вариантов осуществления 30-47, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-18 и 223-244, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

49. Система согласно любому из вариантов осуществления 30-48, где первый спейсерная последовательность связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

50. Система согласно варианту осуществления 49, где вторая спейсерная

последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

51. Система согласно варианту осуществления 50, где связывание спейсерной последовательности первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM на каждой цепи посредством первого и/или второго napDNAbp каждого из первого и второго комплексов праймированного редактирования.

52. Система согласно варианту осуществления 51, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

53. Система согласно варианту осуществления 52, где непрерывная область между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

54. Система согласно варианту осуществления 53, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

55. Система согласно любому из вариантов осуществления 30-54, где первая и/или вторая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотида, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по

меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

56. Система согласно любому из вариантов осуществления 30-55, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

57. Система согласно любому из вариантов осуществления 30-55, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

58. Система согласно любому из вариантов осуществления 30-57, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

59. Система согласно любому из вариантов осуществления 30-58, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

60. Система согласно любому из вариантов осуществления 30-59, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс.

61. Система согласно любому из вариантов осуществления 30-60, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

62. Система согласно варианту осуществления 54, где замещение непрерывной области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции

или замене ДНК в целевом участке.

63. Полинуклеотид, кодирующий систему согласно любому из вариантов осуществления 1-62.

64. Полинуклеотид, кодирующий первый и второй праймирующие редакторы согласно любому из вариантов осуществления 1-62.

65. Полинуклеотид, кодирующий первую и/или вторую РЕгРНК согласно любому из вариантов осуществления 1-29.

66. Полинуклеотид, кодирующий первую и/или вторую РЕгРНК согласно любому из вариантов осуществления 30-62.

67. Полинуклеотид, кодирующий первую и вторую РЕгРНК согласно любому из вариантов осуществления 1-29.

68. Полинуклеотид, кодирующий первую и вторую РЕгРНК согласно любому из вариантов осуществления 30-62.

69. Вектор, содержащий полинуклеотид согласно любому из вариантов осуществления 63-68.

70. Клетка, содержащая полинуклеотид согласно любому из вариантов осуществления 63-69 или вектор согласно варианту осуществления 64.

71. Фармацевтическая композиция, содержащая полинуклеотид согласно любому из вариантов осуществления 63-68, вектор согласно варианту осуществления 69, или клетку согласно варианту осуществления 70, и фармацевтическое вспомогательное вещество.

72. Набор, содержащий полинуклеотид согласно любому из вариантов осуществления 63-69 или вектор согласно варианту осуществления 70 и инструкции по одновременному редактированию обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию.

73. Способ одновременного редактирования первой и второй комплементарных цепей последовательности двухцепочечной ДНК в целевом участке, причем указанный способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем указанная пара содержит:

a. первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности геномной

ДНК против хода транскрипции целевого участка,

b. второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок, (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв в последовательности, комплементарной первому участку связывания, и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв в последовательности, комплементарной второму участку связывания, и последующую полимеризацию второй последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного вторым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, содержащий редактирование, и

где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

74. Способ согласно варианту осуществления 73, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы

75. Способ согласно варианту осуществления 73, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

76. Способ согласно варианту осуществления 73, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат разные праймирующие редакторы.

77. Способ согласно любому из вариантов осуществления 73-76, где первый и/или второй napDNAbp представляет собой домен Cas9 или его вариант.

78. Способ согласно любому из вариантов осуществления 73-77, где первый и/или второй napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, никазный домен Cas9 или их вариант.

79. Способ согласно любому из вариантов осуществления 73-76, где первый и/или второй napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

80. Способ согласно любому из вариантов осуществления 73-79, где первый и/или второй napDNAbp обладает никазной активностью.

81. Способ согласно любому из вариантов осуществления 73-80, где первый и/или второй napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

82. Способ согласно любому из вариантов осуществления 73-81, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

83. Способ согласно любому из вариантов осуществления 73-82, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

84. Способ согласно любому из вариантов осуществления 73-83, где первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

85. Способ согласно варианту осуществления 84, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

86. Способ согласно варианту осуществления 84 или 85, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

87. Способ согласно любому из вариантов осуществления 73-86, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

88. Способ согласно любому из вариантов осуществления 73-87, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

89. Способ согласно любому из вариантов осуществления 73-88, где первая спейсерная последовательность первого комплекса праймированного редактирования связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

90. Способ согласно любому из вариантов осуществления 73-89, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

91. Способ согласно варианту осуществления 90, где связывание первой и второй спейсерных последовательностей первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого первого и второго комплекса праймированного редактирования.

92. Способ согласно варианту осуществления 91, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

93. Способ согласно варианту осуществления 92, где непрерывная область

между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по

меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

94. Способ согласно варианту осуществления 92 или 93, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

95. Способ согласно любому из вариантов осуществления 73-94, где первая и/или вторая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

96. Способ согласно любому из вариантов осуществления 73-95, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

97. Способ согласно любому из вариантов осуществления 73-95, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

98. Способ согласно любому из вариантов осуществления 73-97, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей

мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

99. Способ согласно любому из вариантов осуществления 73-98, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по

меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине желаемой последовательности ДНК и ее комплемента.

100. Способ согласно любому из вариантов осуществления 73-99, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс.

101. Способ согласно любому из вариантов осуществления 73-100, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

102. Способ согласно варианту осуществления 94, где замещение непрерывной области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

103. Способ согласно любому из вариантов осуществления 73-102, где дуплекс содержит редактирование, содержащее один или несколько сайтов рекомбиназы.

104. Способ согласно варианту осуществления 103, дополнительно предусматривающий предоставление рекомбиназы для стимуляции опосредованного рекомбиназой редактирования последовательности ДНК в целевом участке, подлежащем редактированию.

105. Пара РЕгРНК для применения в двойном праймированном редактировании, где пара содержит:

a. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции целевого участка, подлежащего

редактированию, где первая РЕгРНК содержит первую спейсерную последовательность, первое ядро гРНК, первую матрицу для синтеза ДНК, и первый участок связывания праймера, где первая матрица для синтеза ДНК кодирует первую последовательность одноцепочечной ДНК, содержащую отредактированную часть,

b. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где вторая РЕгРНК содержит вторую спейсерную последовательность, второе ядро гРНК, вторую матрицу для синтеза ДНК, и второй участок связывания праймера, где вторая матрица для синтеза ДНК кодирует вторую последовательность одноцепочечной ДНК, содержащую отредактированную часть, где вторая последовательность одноцепочечной ДНК содержит область комплементарности с первой последовательностью одноцепочечной ДНК.

106. Пара РЕгРНК согласно варианту осуществления 105, где первая РЕгРНК содержит структуру:

5′-[первая спейсерная последовательность гРНК]-[первое ядро гРНК]-[первая матрица для синтеза ДНК]-[первый участок связывания праймера]-3′, и

где вторая РЕгРНК содержит структуру:

5′-[вторая спейсерная последовательность гРНК]-[второе ядро гРНК]-[вторая матрица для синтеза ДНК]-[второй участок связывания праймера]-3′.

107. Пара РЕгРНК согласно варианту осуществления 105 или 106, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, который интегрируется в последовательность двухцепочечной ДНК в целевом участке, подлежащем редактированию.

108. Пара РЕгРНК согласно варианту осуществления 107, где интеграция дуплекса приводит к вставке, делеции или замене ДНК в целевом участке.

109. Пара РЕгРНК согласно варианту осуществления 108, где вставка составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по

меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

110. Пара РЕгРНК согласно варианту осуществления 109, где делеция составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по

меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

111. Пара РЕгРНК согласно варианту осуществления 110, где замена составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по

меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

112. Полинуклеотид, кодирующий пару РЕгРНК согласно любому из вариантов осуществления 105-111.

113. Вектор, содержащий полинуклеотид согласно варианту осуществления 112 для экспрессии пары РЕгРНК в клетке.

114. Клетка, содержащая вектор согласно варианту осуществления 113.

115. Фармацевтическая композиция, содержащая пару РЕгРНК согласно любому из вариантов осуществления 105-111, вектор согласно варианту осуществления 113 или клетку согласно варианту осуществления 114 и фармацевтическое вспомогательное вещество.

116. Пара комплексов праймированного редактирования, причем указанная пара содержит:

a. первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и первый полипептид,

обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первой целевой последовательностью на цепи последовательности геномной ДНК против хода транскрипции от целевого участка,

b. второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со второй целевой последовательностью на второй цепи последовательности геномной ДНК по ходу транскрипции от целевого участка.

117. Пара комплексов праймированного редактирования согласно варианту осуществления 116, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

118. Полинуклеотид, кодирующий пару комплексов праймированного редактирования согласно варианту осуществления 116 или 117.

119. Вектор, содержащий полинуклеотид согласно варианту осуществления 118, для экспрессии пары комплексов праймированного редактирования в клетке.

120. Клетка, содержащая вектор согласно варианту осуществления 119.

121. Фармацевтическая композиция, содержащая пару комплексов праймированного редактирования согласно варианту осуществления 116 или 117, вектор согласно варианту осуществления 119 или клетку согласно варианту осуществления 120 и фармацевтическое вспомогательное вещество.

122. Система согласно любому из вариантов осуществления 1-62, или способ согласно любому из вариантов осуществления 68-98, где интеграция дуплекса, содержащего отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

123. Система или способ согласно варианту осуществления 122, где вставка представляет собой пептидную метку.

124. Система или способ согласно варианту осуществления 123, 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.

125. Система или способ согласно варианту осуществления 123, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из последовательности SEQ ID NO: 1-6, 245-249, 252-262, 264-273, 275-276, 281, 278-288 и 622.

126. Система или способ согласно варианту осуществления 122, где вставка представляет собой метку иммуноэпитопа.

127. Система или способ согласно варианту осуществления 126, где метка иммуноэпитопа выбрана из группы, состоящей из столбнячного токсина (SEQ ID NO: 396), мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630), иммуноэпитопа паротита 1 (SEQ ID NO: 400), иммуноэпитопа паротита 2 (SEQ ID NO: 402), иммуноэпитоп паротита 3 (SEQ ID NO: 404), вируса краснухи (SEQ ID NO: 406), гемагглютинина (SEQ ID NO: 408), нейраминидазы (SEQ ID NO: 410), TAP1 (SEQ ID NO: 412), TAP2 (SEQ ID NO: 414), эпитопов гемагглютинина к HLA класса I (SEQ ID NO: 416), эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418), эпитопов гемагглютинина к HLA класса II (SEQ ID NO: 420), эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422), эпитопов гемагглютинина к HLA класса I и класса II, H5N1-связанного, (SEQ ID NO: 424) и эпитопов нейраминидазы к HLA класса I и класса II, H5N1-связанного, (SEQ ID NO: 426).

128. Система или способ согласно варианту осуществления 122, где вставка представляет собой домен димеризации.

129. Система или способ согласно варианту осуществления 128, где домен димеризации представляет собой домен связывания малой молекулы FKBP12 последовательности SEQ ID NO: 488, FKBP12-F36V последовательности SEQ ID NO: 489 или циклофилин последовательности SEQ ID NO: 490 или 493-494.

130. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, причем система содержит:

a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

b) второй комплекс праймированного редактирования, содержащий:

iii. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

iv. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

c) третий комплекс праймированного редактирования, содержащий:

iii. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

iv. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

d) четвертый комплекс праймированного редактирования, содержащий:

iii. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

iv. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, и вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, третья РЕгРНК содержит третью матрицу для синтеза ДНК, кодирующую третью последовательность одноцепочечной ДНК, и четвертая РЕгРНК содержит четвертую матрицу для синтеза ДНК, кодирующую четвертую последовательность одноцепочечной ДНК,

где первая и третья последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где вторая и четвертая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

131. Система для редактирования одной или нескольких последовательностей

двухцепочечной ДНК, причем система содержит:

e) первый комплекс праймированного редактирования, содержащий:

iii. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

iv. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем редактированию,

f) второй комплекс праймированного редактирования, содержащий:

v. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp), и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

vi. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем редактированию,

g) третий комплекс праймированного редактирования, содержащий:

v. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

vi. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с первым участком связывания на первой цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

h) четвертый комплекс праймированного редактирования, содержащий:

v. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

vi. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается со вторым участком связывания на второй цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, третья РЕгРНК содержит третью матрицу для синтеза ДНК, кодирующую третью последовательность одноцепочечной ДНК, и четвертая РЕгРНК содержит четвертую

матрицу для синтеза ДНК, кодирующую четвертую последовательность одноцепочечной ДНК,

где первая и третья последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где вторая и четвертая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

132. Система согласно варианту осуществления 130 или 131, где праймирующий редактор первого комплекса праймированного редактирования, второй комплекс праймированного редактирования, третьего комплекса праймированного редактирования и/или четвертого комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы

133. Система согласно любому из вариантов осуществления 129-132, где

(i) первая РЕгРНК дополнительно содержит первую спейсерную последовательность, первое ядро гРНК и первый участок связывания праймера,

(ii) вторая РЕгРНК дополнительно содержит вторую спейсерную последовательность, второе ядро гРНК и второй участок связывания праймера,

(iii) третья РЕгРНК дополнительно содержит третью спейсерную последовательность, третье ядро гРНК и третий участок связывания праймера, и

(iv) четвертая РЕгРНК содержит четвертую спейсерную последовательность, четвертое ядро гРНК, четвертую матрицу для синтеза ДНК и четвертый участок связывания праймера.

134. Система согласно варианту осуществления 131, где область комплементарности первой и третьей последовательности одноцепочечной ДНК и/или область комплементарности второй и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

135. Система согласно варианту осуществления 134, где первая спейсерная последовательность связывается с первым участком связывания первой последовательности двухцепочечной ДНК против хода транскрипции от первого целевого участка, подлежащего редактированию, и вторая спейсерная последовательность связывается со вторым участком связывания на первой последовательности двухцепочечной ДНК по ходу транскрипции от первого целевого участка, подлежащего редактированию.

136. Система согласно варианту осуществления 134 или 135, где третья спейсерная последовательность связывается с первым участком связывания второй последовательности двухцепочечной ДНК против хода транскрипции от второго целевого

участка, подлежащего редактированию, и четвертая спейсерная последовательность связывается со вторым участком связывания на второй последовательности двухцепочечной ДНК по ходу транскрипции от второго целевого участка, подлежащего редактированию.

137. Система согласно любому из вариантов осуществления 134-136, где первый и второй участки связывания определяют оба конца первой непрерывной области первой двухцепочечной ДНК.

138. Система согласно любому из вариантов осуществления 134-137, где третий и четвертый участки связывания определяют оба конца второй непрерывной области второй двухцепочечной ДНК.

139. Система согласно варианту осуществления 138, где непрерывная область между первым и вторым участками связывания или третьим и четвертым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по

меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

140. Система согласно любому из вариантов осуществления 134-140, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК и/или область комплементарности второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

141. Система согласно варианту осуществления 140, где дуплекс встраивается в первую двухцепочечную ДНК и/или вторую двухцепочечную ДНК.

142. Система согласно любому из вариантов осуществления 134-141, где система приводит к замещению первой непрерывной области второй непрерывной областью в первой двухцепочечной ДНК.

143. Система согласно любому из вариантов осуществления 134-141, где система

приводит к замене первой непрерывной области и второй непрерывной области между первой двухцепочечной ДНК и второй двухцепочечной ДНК.

144. Система согласно любому из вариантов осуществления 130-143, где первый комплекс праймированного редактирования, второй комплекс праймированного редактирования, третий комплекс праймированного редактирования и четвертый комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

145. Система согласно любому из вариантов осуществления 130-144, где первый, второй, третий и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы являются одинаковыми.

146. Система согласно любому из вариантов осуществления 130-145, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 или его вариант.

147. Система согласно любому из вариантов осуществления 130-146, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, или никазный домен Cas9, или их вариант.

148. Система согласно любому из вариантов осуществления 130-147, где первый, второй, третий и/или четвертый napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c и Argonaute.

149. Система согласно любому из вариантов осуществления 130-148, где первый, второй, третий и/или четвертый napDNAbp имеет никазную активность.

150. Система согласно любому из вариантов осуществления 130-149, где первый, второй, третий и/или четвертый napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 2-65, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 2-65.

151. Система согласно любому из вариантов осуществления 130-150, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

152. Система согласно любому из вариантов осуществления 130-151, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110.

153. Система согласно любому из вариантов осуществления 130-152, где

(i) первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp, и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(ii) второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(iii) третий праймирующий редактор дополнительно содержит линкер, который соединяет третий napDNAbp и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или

(iv) четвертый праймирующий редактор дополнительно содержит линкер, который соединяет четвертый napDNAbp и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

154. Система согласно варианту осуществления 153, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 119-128, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 119-128.

155. Система согласно варианту осуществления 153 или 154, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

156. Система согласно любому из вариантов осуществления 130-155, где первая, вторая, третья и/или четвертая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 192-203, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 192-203.

157. Система согласно любому из вариантов осуществления 130-156, где первая спейсерная последовательность связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

158. Система согласно любому из вариантов осуществления 130-157, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего

редактированию.

159. Система согласно любому из вариантов осуществления 130-158, где третья спейсерная последовательность связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

160. Система согласно любому из вариантов осуществления 130-159, где четвертая спейсерная последовательность четвертого комплекса праймированного редактирования связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

161. Система согласно варианту осуществления 160, где связывание спейсерной последовательности первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM на каждой цепи, посредством первого и/или второго napDNAbp каждого из первого и второго комплексов праймированного редактирования.

162. Система согласно варианту осуществления 161, где связывание спейсерной последовательности третьего и четвертого комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM, на каждой цепи посредством третьего и/или четвертого napDNAbp каждого третьего и четвертого комплексов праймированного редактирования.

163. Система согласно варианту осуществления 160 или 162, где первый и третий участки связывания, или второй и четвертый участки связывания определяют два конца непрерывной области двухцепочечной ДНК.

164. Система согласно варианту осуществления 163, где непрерывная область между первым и третьим участками связывания или вторым и четвертым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар

нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по

меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

165. Система согласно любому из вариантов осуществления 130-164, где область комплементарности первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК, и/или область комплементарности второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

166. Система согласно варианту осуществления 164 или 165, где непрерывная область замещается дуплексом.

167. Система согласно варианту осуществления 164 или 165, где дуплекс встраивается в двухцепочечную ДНК в целевом участке.

168. Система согласно варианту осуществления согласно любому из вариантов осуществления 130-167, где первая, вторая, третья и/или четвертая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по

меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

169. Система согласно любому из вариантов осуществления 130-168, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК, и четвертая последовательность одноцепочечной ДНК имеют одинаковые длины.

170. Система согласно любому из вариантов осуществления 130-168, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК, и четвертая последовательность одноцепочечной ДНК имеют разные длины.

171. Система согласно любому из вариантов осуществления 130-170, где область комплементарности первой, второй, третьей и/или четвертой последовательности

одноцепочечной ДНК составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК, или второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК.

172. Система согласно любому из вариантов осуществления 130-171, где область комплементарности первой, второй, третьей и/или четвертой последовательности одноцепочечной ДНК составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по

меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК, и второй последовательности одноцепочечной ДНК, и четвертой последовательности одноцепочечной ДНК.

173. Система согласно любому из вариантов осуществления 130-172, где первая, вторая, третья и четвертая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

174. Система согласно варианту осуществления 167, где включение дуплекса, содержащего отредактированную часть приводит к вставке, делеции или замене ДНК в целевом участке.

175. Полинуклеотид, кодирующий первый, второй, третий и четвертый праймирующие редакторы согласно любому из вариантов осуществления 130-174.

176. Полинуклеотид, кодирующий первую, вторую, третью и четвертую РЕгРНК согласно любому из вариантов осуществления 130-174.

177. Вектор, содержащий полинуклеотид согласно варианту осуществления 175 или 176.

178. Клетка, содержащая полинуклеотид согласно варианту осуществления 175 или 176 или вектор согласно варианту осуществления 177.

179. Фармацевтическая композиция, содержащая полинуклеотид согласно варианту осуществления 175 или 176, вектор согласно варианту осуществления 177 или клетку согласно варианту осуществления 178 и фармацевтическое вспомогательное вещество.

180. Набор, содержащий полинуклеотид согласно варианту осуществления 175 или 176, или вектор согласно варианту осуществления 177 и инструкции по одновременному редактированию обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию.

181. Способ одновременного редактирования первой и второй комплементарных цепей последовательности двухцепочечной ДНК в целевом участке, причем способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем пара содержит:

(a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первой целевой последовательностью на первой цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую PEгРНК), которая связывается со второй целевой последовательностью на второй цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью PEгРНК), которая связывается с третьей целевой последовательностью на первой цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

(d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий второй программируемый

нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертой целевой последовательностью на второй цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв на первой целевой последовательности и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв на второй целевой последовательности и последующую полимеризацию второй последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного вторым разрывом,

где третий комплекс праймированного редактирования вызывает третий разрыв на третьей целевой последовательности и последующую полимеризацию третьей последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного третьим разрывом,

где четвертый комплекс праймированного редактирования вызывает четвертый разрыв на четвертой целевой последовательности и последующую полимеризацию четвертой последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного четвертым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК, и

где третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

182. Способ согласно варианту осуществления 181, где праймирующий редактор первого комплекса праймированного редактирования, второй комплекс праймированного редактирования, третьего комплекса праймированного редактирования и/или четвертого комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

183. Способ согласно варианту осуществления 181 или 182, где первый комплекс праймированного редактирования, второй комплекс праймированного редактирования, третий комплекс праймированного редактирования и четвертый комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

184. Способ согласно варианту осуществления 181 или 182, где первый, второй, третий и/или четвертый комплекс праймированного редактирования содержат разные праймирующие редакторы.

185. Способ согласно любому из вариантов осуществления 181-184, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 или его вариант.

186. Способ согласно любому из вариантов осуществления 181-185, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью или никазный домен Cas9, или их вариант.

187. Способ согласно любому из вариантов осуществления 181-184, где первый, второй, третий и/или четвертый napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c.

188. Способ согласно любому из вариантов осуществления 181-187, где первый, второй, третий и/или четвертый napDNAbp обладает никазной активностью.

189. Способ согласно любому из вариантов осуществления 181-188, где первый, второй, третий и/или четвертый napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 2-65, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 2-65.

190. Способ согласно любому из вариантов осуществления 181-189, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

191. Способ согласно любому из вариантов осуществления 181-190, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110.

192. Способ согласно любому из вариантов осуществления 181-191, где

(i) первый праймирующий редактор дополнительно содержит линкер, который

соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(ii) второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(iii) третий праймирующий редактор дополнительно содержит линкер, который соединяет третий napDNAbp и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или

(iv) четвертый праймирующий редактор дополнительно содержит линкер, который соединяет четвертый napDNAbp и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

193. Способ согласно варианту осуществления 192, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 119-128, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 119-128.

194. Способ согласно варианту осуществления 192 или 193, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

195. Способ согласно любому из вариантов осуществления 181-191, где первая, вторая, третья и/или четвертая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 192-203, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 192-203.

196. Способ согласно любому из вариантов осуществления 181-195, где первая спейсерная последовательность первого комплекса праймированного редактирования связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

197. Способ согласно любому из вариантов осуществления 181-196, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

198. Способ согласно любому из вариантов осуществления 181-197, где третья спейсерная последовательность третьего комплекса праймированного редактирования связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

199. Способ согласно любому из вариантов осуществления 181-198, где четвертая спейсерная последовательность четвертого комплекса праймированного редактирования связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

200. Способ согласно варианту осуществления 197, где связывание первой и второй спейсерных последовательностей первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого первого и второго комплекса праймированного редактирования.

201. Способ согласно варианту осуществления 199, где связывание третьей и четвертой спейсерных последовательностей третьего и четвертого комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого третьего и четвертого комплекса праймированного редактирования.

202. Способ согласно варианту осуществления 200 или 201, где первый и второй участки связывания, и/или третий и четвертый участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

203. Способ согласно варианту осуществления 202, где непрерывная область между первым и вторым участками связывания, и/или непрерывная область между третьим и четвертым участками связывания, составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей

мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98

нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

204. Способ согласно варианту осуществления 202 или 203, где непрерывная область замещается дуплексом.

205. Способ согласно любому из вариантов осуществления 181-204, где первая, вторая, третья и/или четвертая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64

нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

206. Способ согласно любому из вариантов осуществления 181-205, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК имеют одинаковые длины.

207. Способ согласно любому из вариантов осуществления 181-206, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной, и четвертая последовательность одноцепочечной ДНК имеют разные длины.

208. Способ согласно любому из вариантов осуществления 181-207, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК, или третьей последовательности

одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК.

209. Способ согласно любому из вариантов осуществления 181-208, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по

меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

210. Способ согласно любому из вариантов осуществления 181-209, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК, и/или третьей последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

211. Способ согласно любому из вариантов осуществления 181-210, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК, представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

212. Способ согласно варианту осуществления 204, где замещение непрерывной области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

213. Способ согласно любому из вариантов осуществления 181-212, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК находятся на концах целевой последовательности ДНК, и где третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на противоположных концах той же целевой последовательности ДНК.

214. Способ согласно варианту осуществления 213, где целевая последовательность ДНК является инвертированной.

215. Способ согласно любому из вариантов осуществления 181-214 дополнительно предусматривает предоставление донора кольцевой ДНК.

216. Способ согласно варианту осуществления 215, где первая последовательность одноцепочечной ДНК и третья последовательность одноцепочечной

ДНК находятся на противоположных концах целевой последовательности ДНК, и где вторая последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на доноре кольцевой ДНК.

217. Способ согласно варианту осуществления 216, где часть донора кольцевой ДНК между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК замещает целевую последовательность между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК.

218. Способ согласно любому из вариантов осуществления 181-214, где первая последовательность одноцепочечной ДНК и третья последовательность одноцепочечной ДНК находятся на первой молекуле нуклеиновой кислоты, и где вторая последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на второй молекуле нуклеиновой кислоты.

219. Способ согласно варианту осуществления 218, где часть первой молекулы нуклеиновой кислоты между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК встраивается во вторую молекулу нуклеиновой кислоты, и где часть второй молекулы нуклеиновой кислоты между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК встраивается в первую молекулу нуклеиновой кислоты.

220. Способ согласно варианту осуществления 218 или 219, где первая молекула нуклеиновой кислоты представляет собой первую хромосому, и вторая молекула нуклеиновой кислоты представляет собой вторую хромосому.

221. Набор РЕгРНК для применения в мультилоскутном праймированном редактировании, где РЕгРНК содержат:

(a) первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию, где первая РЕгРНК содержит первую спейсерную последовательность, первое ядро гРНК, первую матрицу для синтеза ДНК и первый участок связывания праймера, где первая матрица для синтеза ДНК кодирует первую последовательность одноцепочечной ДНК,

(b) вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию, где вторая РЕгРНК содержит вторую спейсерную последовательность,

второе ядро гРНК, вторую матрица для синтеза ДНК и второй участок связывания праймера, где вторая матрица для синтеза ДНК кодирует вторую последовательность одноцепочечной ДНК, содержащую отредактированную часть, где вторая последовательность одноцепочечной ДНК комплементарна первой последовательности одноцепочечной ДНК.

(c) третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где третья РЕгРНК содержит третью спейсерную последовательность, третье ядро гРНК, третью матрицу для синтеза ДНК и третий участок связывания праймера, где третья матрица для синтеза ДНК кодирует третью последовательность одноцепочечной ДНК,

(d) четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где четвертая РЕгРНК содержит четвертую спейсерную последовательность, четвертое ядро гРНК, четвертую матрица для синтеза ДНК и четвертый участок связывания праймера, где четвертая матрица для синтеза ДНК кодирует четвертую последовательность одноцепочечной ДНК, где четвертая последовательность одноцепочечной ДНК содержит область комплементарности с третьей последовательностью одноцепочечной ДНК.

222. Набор РЕгРНК согласно варианту осуществления 221, где

первая РЕгРНК содержит структуру:

5′-[первая спейсерная последовательность гРНК]-[первое ядро гРНК]-[первая матрица для синтеза ДНК]-[первый участок связывания праймера]-3′,

вторая РЕгРНК содержит структуру:

5′-[вторая спейсерная последовательность гРНК]-[второе ядро гРНК]-[вторая матрица для синтеза ДНК]-[второй участок связывания праймера]-3′,

третья РЕгРНК содержит структуру:

5′-[третья спейсерная последовательность гРНК]-[третье ядро гРНК]-[третья матрица для синтеза ДНК]-[третий участок связывания праймера]-3′, и

четвертая РЕгРНК содержит структуру:

5′-[четвертая спейсерная последовательность гРНК]-[четвертое ядро гРНК]-[четвертая матрица для синтеза ДНК]-[четвертый участок связывания праймера]-3′.

223. Набор РЕгРНК согласно варианту осуществления 221 или 222, где первая

последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК образуют дуплекс, который интегрируется в последовательность двухцепочечной ДНК в целевом участке, подлежащем редактированию.

224. Набор РЕгРНК согласно любому из вариантов осуществления 221-223, где интеграция дуплекса приводит к вставке, делеции или замене ДНК в целевом участке.

225. Набор РЕгРНК согласно варианту осуществления 224, где вставка составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по

меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

226. Набор РЕгРНК согласно варианту осуществления 224, где делеция составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 на 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

227. Набор РЕгРНК согласно варианту осуществления 224, где замещение составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106

нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

228. Полинуклеотид, кодирующий набор РЕгРНК согласно любому из вариантов осуществления 221-227.

229. Вектор, содержащий полинуклеотид согласно варианту осуществления 228 для экспрессии множества РЕгРНК в клетке.

230. Клетка, содержащая вектор согласно варианту осуществления 229.

231. Фармацевтическая композиция, содержащая множество РЕгРНК согласно любому из вариантов осуществления 221-227, вектор согласно варианту осуществления 229 или клетку согласно варианту осуществления 230, и фармацевтическое вспомогательное вещество.

232. Способ модификации генома в клетке, предусматривающий:

(i) приведение генома в контакт с системой двухлоскутного праймированного редактирования, причем указанная система содержит первый комплекс праймирующего редактора/РЕгРНК, способный устанавливать первый 3'-лоскут нуклеиновой кислоты в первом участке разрыва на первой цепи целевого участка, и второй комплекс праймирующего редактора/РЕгРНК, способный устанавливать второй 3'-лоскут нуклеиновой кислоты во втором участке разрыва на второй цепи целевого участка,

где первый 3'-лоскут нуклеиновой кислоты и второй 3'-лоскут нуклеиновой кислоты представляют собой обратно комплементарные последовательности, которые образуют дуплекс, и

где дуплекс содержит первый сайт рекомбиназы, и

где клетка устанавливает дуплекс, содержащий первый сайт рекомбиназы, в целевом участке вместо эндогенного дуплекса, расположенного между первым и вторым участками разрыва,

(ii) контакт первого сайта рекомбиназы с донорной нуклеиновой кислотой, содержащей второй сайт рекомбиназы, и

(iii) контакт генома с рекомбиназой, способной к рекомбинации между первым и вторым сайтами рекомбиназы, тем самым интегрируя донорную нуклеиновую кислоту в целевой участок в геноме.

[9] Способы

Способы мультилоскутного праймированного редактирования

[0954] Согласно одному аспекту настоящее изобретение относится к способам одновременного редактирования первой и второй комплементарных цепей

последовательности двухцепочечной ДНК в целевом участке, причем указанный способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем указанная пара содержит:

a. первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первым участком связывания на первой цепи последовательности геномной ДНК против хода транскрипции целевого участка,

b. второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую PEгРНК), которая связывается со вторым участком связывания на второй цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв в последовательности, комплементарной первому участку связывания, и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв в последовательности, комплементарной второму участку связывания, и последующую полимеризацию второй последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного вторым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, содержащий редактирование, и

где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

[0955] Согласно другому аспекту настоящее изобретение относится к способам одновременного редактирования первой и второй комплементарных цепей последовательности двухцепочечной ДНК в целевом участке, причем способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем пара содержит:

(a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первой целевой последовательностью на первой цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую PEгРНК), которая связывается со второй целевой последовательностью на второй цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью PEгРНК), которая связывается с третьей целевой последовательностью на первой цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

(d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую PEгРНК), которая связывается с четвертой целевой последовательностью на второй цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв на первой целевой последовательности и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв на второй целевой последовательности и последующую полимеризацию второй

последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного вторым разрывом,

где третий комплекс праймированного редактирования вызывает третий разрыв на третьей целевой последовательности и последующую полимеризацию третьей последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного третьим разрывом,

где четвертый комплекс праймированного редактирования вызывает четвертый разрыв на четвертой целевой последовательности и последующую полимеризацию четвертой последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного четвертым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК, и

где третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

[0956] В настоящем описании описываются системы мультилоскутного праймированного редактирования (включая, например, систему двухлоскутного праймированного редактирования и систему четырехлоскутного праймированного редактирования), которые решают проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в подлежащую редактированию комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. В системе двухлоскутного праймированного редактирования две РЕгРНК используются для нацеливания на противоположные цепи участка генома и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 90**). В системе четырехлоскутного праймированного редактирования используются четыре РЕгРНК, которые управляют синтезом четырех 3'-лоскутов, два из которых комплементарны друг другу, а два других комплементарны друг другу (**фиг. 95, фиг. 96А** и **фиг. 98А**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная цепь. Поскольку обе цепи дуплекса синтезируются как редактируемая ДНК, система

мультилоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК нужно только вырезать парные 5'-лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные цепи ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое содержит минимальную геномную гомологию. Нуклеазоактивные версии мультилоскутных праймированных редакторов, которые разрезают обе цепи ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК.

[0957] Подобно классическому праймированному редактированию, мультилоскутное праймированное редактирование является универсальным и точным способом редактирования генома, который непосредственно записывает новую генетическую информацию в указанный участок ДНК с использованием программируемого нуклеиновой кислотой ДНК-связывающего белка («napDNAbp»), работающего в сочетании с полимеразой (т.е. в форме слитого белка или иным образом представленного *в транс* с napDNAbp), где система праймированного редактирования запрограммирована с помощью гидовой РНК праймированного редактирования (PE) («PEгРНК»), которая одновременно определяет целевой сайт и матрицы для синтеза желаемое редактирование в виде замещающей цепи ДНК посредством удлинения (либо ДНК, либо РНК), сконструированного на гидовой РНК (например, на 5'- или 3'-конце или во внутренней части гидовой РНК). Цепь замещения, содержащая желаемое редактирование (например, замену одного нуклеотидного основания), имеет ту же последовательность, что и эндогенная цепь целевого участка, подлежащего редактированию (за исключением того, что она включает желаемое редактирование). С помощью механизмов репарации и/или репликации ДНК эндогенная цепь целевого участка заменяется вновь синтезированной замещающей цепью, содержащей желаемое редактирование. В некоторых случаях праймированное редактирование можно рассматривать как технологию редактирования генома «поиск и замена», поскольку двойные праймирующие редакторы, как описано в настоящем документе, не только ищут и находят желаемый целевой сайт для редактирования, но в то же время кодируют замещающую цепь, содержащую желаемое редактирование, которое устанавливается вместо соответствующей цепи эндогенной ДНК целевого участка.

[0958] В двухлоскутном праймированном редактировании последовательность

двухцепочечной ДНК контактирует в целевом участке с первым и вторым комплексами праймированного редактирования. Каждый содержит слитый белок и РЕгРНК. Согласно некоторым вариантам осуществления каждый слитый белок содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы (например, обратная транскриптаза), и каждая РЕгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера. Каждая матрица для синтеза ДНК может кодировать последовательность одноцепочечной ДНК, которая может содержать отредактированную часть одного или нескольких нуклеотидов. Две кодируемые одноцепочечные последовательности ДНК могут быть комплементарны друг другу и образовывать дуплекс, который может интегрироваться в редактируемый целевой участок. Различные элементы комплексов праймированного редактора (например, слитые белки, napDNAbp, полимераза, РЕгРНК и т.д.) могут включать любой из вариантов осуществления систем, раскрытых в настоящем документе.

[0959] При четырехлоскутном праймированном редактировании последовательность двухцепочечной ДНК контактирует в целевом участке с первым, вторым, третьим и четвертым комплексами праймированного редактирования. Каждый комплекс содержит слитый белок и РЕгРНК. Согласно некоторым вариантам осуществления каждый слитый белок содержит программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы (например, обратная транскриптаза), и каждая РЕгРНК содержит спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК, и участок связывания праймера. Каждая матрица для синтеза ДНК кодирует последовательность одноцепочечной ДНК. Две кодируемые одноцепочечные последовательности ДНК могут быть комплементарны друг другу и образовывать дуплекс, который может интегрироваться в редактируемый целевой участок. Различные элементы комплексов праймированного редактора (например, слитые белки, napDNAbp, полимераза, РЕгРНК и т.д.) могут включать любой из вариантов осуществления систем, раскрытых в настоящем документе.

[0960] Способы мультилоскутного праймированного редактирования согласно настоящему изобретению могут быть полезны для множества применений. Например, их можно использовать для облегчения инверсии последовательности ДНК-мишени. В настоящей заявке первая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК первой РЕгРНК, и вторая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК второй РЕгРНК, находятся на противоположных концах последовательности ДНК-мишени, и третья последовательность одноцепочечной

ДНК, кодируемая матрицей для синтеза ДНК третьей РЕгРНК, и четвертая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК четвертой РЕгРНК, находятся на противоположных концах одной и той же последовательности ДНК-мишени.

[0961] Согласно некоторым вариантам осуществления способы мультилоскутного праймированного редактирования согласно настоящему изобретению предусматривают дополнительно предоставление донора кольцевой ДНК, часть которого может быть интегрирована в двухцепочечную нуклеиновую кислоту в целевом участке. В настоящей заявке первая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК первой РЕгРНК, и третья последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК третьей РЕгРНК, находятся на противоположных концах последовательности ДНК-мишени, и вторая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК второй РЕгРНК, и четвертая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК четвертой РЕгРНК, находятся на доноре кольцевой ДНК. Часть донора кольцевой ДНК между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК может образовывать дуплекс, который заменяет последовательность ДНК-мишени между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной Д ДНК.

[0962] Согласно другому применению способы мультилоскутного праймированного редактирования согласно настоящему изобретению позволяют транслокацию последовательности ДНК-мишени из первой молекулы нуклеиновой кислоты (*например*, первой хромосомы) во вторую молекулу нуклеиновой кислоты (*например*, вторую хромосому). В настоящей заявке первая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК первой РЕгРНК, и третья последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК третьей РЕгРНК, находятся на первой молекуле нуклеиновой кислоты, а вторая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК второй РЕгРНК, и четвертая последовательность одноцепочечной ДНК, кодируемая матрицей для синтеза ДНК четвертой РЕгРНК, находятся на второй молекуле нуклеиновой кислоты. Часть первой молекулы нуклеиновой кислоты между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК может быть включена во вторую молекулу нуклеиновой кислоты, а часть второй молекулы нуклеиновой кислоты между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК включена в первую молекулу нуклеиновой кислоты.

Другие способы

[0963] Настоящее изобретение относится к способам лечения субъекта, у которого диагностировано заболевание, связанное с точечной мутацией или вызванное ею, или другими мутациями (например, делеция, вставка, инверсия, дупликация и т.д.), которые можно скорректировать с помощью системы мультилоскутного праймированного редактирования согласно настоящему изобретению, в качестве примера, но без ограничения прионной болезни (например, пример 5 в настоящем документе), болезни с экспансией тринуклеотидных повторов (например, пример 3 в настоящем документе) или синдрома дефицита CDKL5 (CDD) (например, пример 23 в настоящем документе).

[0964] Практически любой генетический дефект, вызывающий заболевание, может быть устранен с помощью мультилоскутного праймированного редактирования, которое включает выбор подходящего слитого белка праймирующего редактора (включая napDNAbp и полимеразу (например, обратную транскриптазу) и конструирование подходящей PEгРНК, предназначенной для (a) нацеливания на соответствующую целевую ДНК, содержащую участок редактирования, и (b) предоставления матрицы для синтеза одной цепи ДНК от 3'-конца участка разрыва, которая включает желаемое редактирование, которая смещает и заменяет эндогенную цепь сразу по ходу транскрипции от участка разрыва Мультилоскутное праймированное редактирование можно применять без ограничения для (a) внесения корректирующих мутацию изменений в нуклеотидную последовательность, (b) установки белковых и РНК-меток, (в) установки иммуноэпитопов на представляющие интерес белки, (d) установки индуцируемых доменов димеризации в белках, (e) установки или удаления последовательности, чтобы изменить эту активность биомолекулы, (f) установки сайтов-мишеней рекомбиназы, чтобы направлять определенные генетические изменения, и (g) мутагенеза целевой последовательности с использованием RT сниженной точности.

[0965] Способ лечения нарушения может предусматривать в качестве ранней стадии конструирование подходящего слитого белка PEгРНК и праймирующего редактора в соответствии с описанными в настоящем документе способами, которые предусматривают ряд аспектов, которые могут быть приняты во внимание, такие как:

(a) целевая последовательность, т.е. нуклеотидная последовательность, в которой одна или несколько модификаций нуклеотидных должны быть установлены праймирующим редактором,

(b) расположение участка разреза в целевой последовательности, т.е. конкретное положение нуклеотидного основания, в котором праймирующий редактор будет

индуцировать одноцепочечный разрыв с созданием 3'-концевой последовательности праймера RT на одной стороне разрыва и 5'-концового эндогенного лоскута на другой стороне разрыва (который в конечном итоге удаляется с помощью FEN1 или ее эквивалента и заменяется 3'-лоскутом оцДНК. Участок разреза аналогичен «положению редактирования», поскольку он создает 3'-концевую последовательность праймера RT, которая удлиняется посредством RT в ходе РНК-зависимой полимеризации ДНК с созданием 3'-лоскута одноцепочечной ДНК, содержащего желаемое редактирование, которое затем заменяет 5'-лоскут эндогенной ДНК в целевой последовательности.

(c) доступные последовательности PAM (включая канонические участки PAM SpCas9, а также неканонические участки PAM, распознаваемые вариантами и эквивалентами Cas9 с удлиненными или отличающимися специфичностями PAM),

(d) расстояние между доступными последовательностями PAM и расположение участка разреза в целевой последовательности,

(e) конкретный Cas9, вариант Cas9 или эквивалент Cas9 используемого праймирующего редактора),

(f) последовательность и длина участка связывания праймера,

(g) последовательность и длина матрицы редактирования,

(h) последовательность и длина плеча гомологии,

(i) спейсерная последовательность и длина и

(j) последовательность ядра.

[0966] Подходящая РЕгРНК и, необязательно, руководство по конструированию вносящей разрыв сгРНК для разрыва второго участка могут быть разработаны с помощью следующего иллюстративного пошагового набора инструкций, который учитывает один или несколько из приведенных выше аспектов. Стадии ссылаются на примеры, показанные на фиг. 70A-70I.

1. **Определение целевой последовательности и редактирование**. Получите последовательность целевого участка ДНК (~ 200 п.о.), сосредоточенную вокруг местоположения желаемого редактирования (точечная мутация, вставка, удаление или их комбинация). См. фиг. 70A.

2. **Положение целевых PAM**. Определите PAM в непосредственной близости от желаемого места редактирования. PAM можно идентифицировать на любой цепи ДНК, проксимально к желаемому месту редактирования. В то время как PAM, близкие к положению редактирования, являются предпочтительными (т. е. сайт разрыва находится менее чем в пределах 30 нуклеотидов от положения редактирования или менее чем в пределах 29 нуклеотидов, 28 нуклеотидов, 27 нуклеотидов, 26 нуклеотидов, 25

нуклеотидов, 24 нуклеотидов, 23 нуклеотидов, 22 нуклеотидов, 21 нуклеотидов, 20 нуклеотидов, 19 нуклеотидов, 18 нуклеотидов, 17 нуклеотидов, 16 нуклеотидов, 15 нуклеотидов, 14 нуклеотидов, 13 нуклеотидов, 12 нуклеотидов, 11 нуклеотидов, 10 нуклеотидов, 9 нуклеотидов, 8 нуклеотидов, 7 нуклеотидов, 6 нуклеотидов, 5 нуклеотидов, 4 нуклеотидов, 3 нуклеотидов или 2 нуклеотидов от положения редактирования до участка разрыва), можно установить редактирования с помощью протоспейсеров и PAM, которые помещают разрыв на ≥ 30 нуклеотидов от положения редактирования. См. фиг. 70.

3. **Положение участков разрыва**. Для каждого рассматриваемого PAM определите соответствующий участок разрыва и на какой цепи он находится. Для никазы Sp Cas9 H840A расщепление происходит в цепи, содержащей PAM, между 3-м и 4-м основаниями 5' по отношению к NGG PAM. Все отредактированные нуклеотиды должны располагаться на 3'-конце участка разрыва, поэтому соответствующие PAM должны размещать участок разрыва на 5'-конце целевого редактирования на цепи, содержащей PAM. В примере, показанном ниже, есть два возможных PAM. Для простоты оставшиеся стадии демонстрируют конструкцию PEгРНК с использованием только PAM 1. См. фиг. 70C.

4. **Конструирование спейсерной последовательности**. Протоспейсер Sp Cas9 соответствует 20 нуклеотидам 5' по отношению к NGG PAM на цепи, содержащей PAM. Для эффективной инициации транскрипции Pol III требуется, чтобы G был первым транскрибируемым нуклеотидом. Если первым нуклеотидом протоспейсера является G, последовательность спейсера для PEгРНК представляет собой просто последовательность протоспейсера. Если первый нуклеотид протоспейсера не представляет собой G, последовательность спейсера PEгРНК представляет собой G, за которой следует последовательность протоспейсера. См. фиг. 70D.

5. **Конструирование участка связывания праймера (PBS)**. Используя исходную последовательность аллеля, определите праймер ДНК на цепи, содержащей PAM. 3'-конец ДНК-праймера представляет собой нуклеотид непосредственно перед участком разрыва (т.е. 4-е основание 5' по отношению к NGG PAM для Sp Cas9). В качестве общего принципа конструирования для использования с PE2 и PE3 участок связывания праймера PEгРНК (PBS), содержащий от 12 до 13 нуклеотидов, комплементарных ДНК-праймеру, можно использовать для последовательностей, которые содержат ~ 40-60% GC. Для последовательностей с низким содержанием GC следует тестировать более длинные (от 14 до 15 нуклеотидов) PBS. Для последовательностей с более высоким содержанием GC следует тестировать более короткие (от 8 до 11 нуклеотидов) PBS. Оптимальные последовательности PBS должны быть определены эмпирически, независимо от

содержания GC. Чтобы сконструировать последовательность PBS длины *p*, возьмите обратный комплемент первых *p* нуклеотидов 5' от участка разрыва в цепи, содержащей PAM, используя последовательность *исходного* аллеля. См. фиг. 70E.

6. **Конструирование матрицы RT (или матрицы для синтеза ДНК)**. Матрица RT (или матрица для синтеза ДНК, где полимераза не является обратной транскриптазой) кодирует сконструированное редактирование и гомологию с последовательностью, соседней с редактированием. Согласно одному варианту осуществления эти области соответствуют матрице для синтеза ДНК на фиг. 3D и фиг. 3E, где матрица для синтеза ДНК содержит «матрицу редактирования» и «плечо гомологии». Оптимальная длина матрицы RT зависит от целевого участка. Для коротких редактирований (положения от +1 до +6) рекомендуется протестировать короткую (от 9 до 12 нуклеотидов), среднюю (от 13 до 16 нуклеотидов) и длинную (от 17 до 20 нуклеотидов) матрицу RT. Для редактирований дальнего диапазона (положения +7 и далее) рекомендуется использовать матрицы RT, которые простираются как минимум на 5 нуклеотидов (предпочтительно на 10 или более нуклеотидов) от положения редактирования, чтобы обеспечить достаточную гомологию 3' лоскута ДНК. Для редактирований дальнего диапазона следует скринировать несколько матриц RT, чтобы определить функциональные конструкции. Для более крупных вставок и делеций (≥5 нуклеотидов) рекомендуется включение большей 3'-гомологии (~ 20 нуклеотидов или более) в матрицу RT. Эффективность редактирования обычно снижается, когда матрица RT кодирует синтез G в качестве последнего нуклеотида в продукте обратной транскрипции ДНК (соответствует C в матрице RT PEгРНК). Поскольку многие матрицы RT поддерживают эффективное праймированное редактирование, при разработке матриц RT рекомендуется избегать использования G в качестве конечного синтезированного нуклеотида. Чтобы сконструировать последовательность матрицы RT длины *r*, используйте *желаемую* последовательность аллеля и возьмите обратный комплемент первых нуклеотидов *r* 3' участка разрыва в цепи, которая первоначально содержала PAM. Обратите внимание, что по сравнению с редактированиями SNP, редактирования вставкой или делецией с использованием матриц RT той же длины не будет содержать идентичной гомологии. См. фиг. 70F.

7. **Сборка полной последовательности PEгРНК**. Соедините компоненты PEгРНК в следующем порядке (от 5 до 3): спейсер, каркас, матрица RT и PBS. См. фиг. 70G.

8. **Конструирование вносящих разрыв сгРНК для PE3**. Идентифицируйте PAM в не редактируемой цепи до и после редактирования. Оптимальные положения разрыва сильно зависят от локуса и должны быть определены эмпирически. В целом, разрывы, расположенные на расстоянии от 40 до 90 нуклеотидов в 5' положении от

РЕгРНК-индуцированного разрыва, приводят к более высокому выходу редактирования и меньшему количеству вставок-делеций. Вносящая разрыв сгРНК имеет спейсерную последовательность, которая соответствует 20-нуклеотидному протоспейсеру в *исходном* аллеле, с добавлением 5'-G, если протоспейсер не начинается с G. См. фиг. 70H.

9. **Конструирование вносящих разрыв сгРНК PE3b**. Если PAM существует в комплементарной цепи и соответствующий ему протоспейсер перекрывается с последовательностью, предназначенной для редактирования, это редактирование может быть кандидатом для системы PE3b. В системе PE3b спейсерная последовательность вносящей разрыв сгРНК соответствует последовательности желаемого редактируемого аллеля, но не исходного аллеля. Система PE3b работает эффективно, когда отредактированный нуклеотид (нуклеотиды) попадает в область затравки (~ 10 нуклеотидов, примыкающую к PAM) протоспейсера вносящей затравку сгРНК. Это предотвращает разрыв комплементарной цепи до установки отредактированной цепи, предотвращая конкуренцию между РЕгРНК и сгРНК за связывание ДНК-мишени. PE3b также позволяет избежать образования одновременных разрывов на обеих цепях, что значительно снижает образование дефектов при сохранении высокой эффективности редактирования. сгРНК PE3b должны иметь спейсерную последовательность, которая соответствует 20-нуклеотидному протоспейсеру в *желаемом* аллеле, с добавлением 5'G, если это необходимо. См. фиг. 70I.

[0967] Приведенный выше постадийный способ конструирования подходящей РЕгРНК и сгРНК с разрывом второго сайта никоим образом не является ограничивающим. В описании рассматриваются варианты вышеописанного постадийного способа, которые могут быть получены специалистом в данной области техники.

[0968] После того, как подходящие РЕгРНК и слитый белок PE выбраны/сконструированы, их можно вводить с помощью подходящей методики, такой как векторная трансфекция (в которой один или несколько векторов содержат ДНК, кодирующую РЕгРНК и слитый белок PE, и которые экспрессируются внутри клетки после трансфекции векторами), прямая доставка слитого белка PE в комплексе с РЕгРНК (например, доставка RNP) в формате доставки (например, липидные частицы, наночастицы) или с помощью системы доставки на основе иРНК. Такие способы описаны в настоящем описании, а также можно использовать любой известный способ.

[0969] РЕгРНК и слитый белок PE (или обозначаемые вместе комплекс PE) могут доставлять в клетку в терапевтически эффективном количестве, так что при контакте с представляющей интерес ДНК-мишенью в ней устанавливается желаемое редактирование.

[0970] Любое заболевание, по-видимому, поддается лечению такими способами,

если осуществима доставка в соответствующие клетки. Специалист в данной области техники сможет выбрать и/или отобрать методику доставки PE в соответствии с предполагаемой целью и предполагаемыми клетками-мишенями.

[0971] Например, согласно некоторым вариантам осуществления настоящее изобретение относится к способу, который предусматривает введение субъекту, страдающему таким заболеванием, как рак, связанный с точечной мутацией, как описано выше, эффективного количества описанной в настоящем документе системы мультилоскутного праймированного редактирования, которая корректирует точечную мутацию или вводит деактивирующую мутацию в ген, ассоциированный с заболеванием, посредством направляемой гомологией репарации, в присутствии молекулы донорной ДНК, содержащей желаемое генетическое изменение. Согласно некоторым вариантам осуществления предложен способ, который предусматривает введение субъекту, страдающему таким заболеванием, как рак, связанный с точечной мутацией, как описано выше, эффективного количества описанной в настоящем документе системы мультилоскутного праймированного редактирования, которая корректирует точечную мутацию или вводит деактивирующую мутацию в ген, ассоциированный с заболеванием. Согласно некоторым вариантам осуществления заболевание представляет собой пролиферативное заболевание. Согласно некоторым вариантам осуществления заболевание представляет собой генетическое заболевание. Согласно некоторым вариантам осуществления заболевание представляет собой неопластическое заболевание. Согласно некоторым вариантам осуществления заболевание представляет собой метаболическое заболевание. Согласно некоторым вариантам осуществления заболевание представляет собой лизосомную болезнь накопления. Другие заболевания, которые можно лечить путем коррекции точечной мутации или введения деактивирующей мутации в ген, ассоциированный с заболеванием, известны специалистам в данной области техники, и описание не ограничивается в этом отношении.

[0972] Настоящее изобретение относится к способам лечения дополнительных заболеваний или нарушений, например, заболеваний или нарушений, которые связаны или вызваны точечной мутацией, которая может быть скорректирована с помощью редактирования генов, опосредованного TPRT. Некоторые такие заболевания описаны в настоящем документе, и дополнительные подходящие заболевания, которые можно лечить с помощью стратегий и слитых белков, предложенных в настоящем документе, будут очевидны специалистам в данной области на основании настоящего раскрытия. Типичные подходящие заболевания и нарушения перечислены ниже. Понятно, что нумерация конкретных положений или остатков в соответствующих последовательностях зависит от

конкретного белка и используемой схемы нумерации. Нумерация может различаться, например, у предшественников зрелого белка и самого зрелого белка, и различия в последовательностях от вида к виду могут влиять на нумерацию. Специалист в данной области техники сможет идентифицировать соответствующий остаток в любом гомологичном белке и в соответствующей кодирующей нуклеиновой кислоте способами, хорошо известными в данной области техники, например, путем выравнивания последовательностей и определения гомологичных остатков. Иллюстративные подходящие заболевания и нарушения включают без ограничения следующие: 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 Аарского, синдром Ааза, ахондрогенез 2 типа, ахроматопсия 2 и 7, приобретенный синдром удлиненного интервала QT, акрокаллозальный синдром по типу Шницеля, 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 X-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, аутосомнодоминантная 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 аутосомнорецессивной гипогидротической эктодермальной дисплазии, эктодермальная дисплазия lib, гипогидротическая / волос / зубов, аутосомно-рецессивная, аутосомнорецессивная 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 (аутосомнодоминантный), 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 Берна-Маккауна, 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 IB, 2B2, 2C, 2F, 21, 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, X-сцепленная рецессивная и 2 X-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 1-2-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 cl, 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, X-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, ID, 1G, 1H, 1J, IK, 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, Al 1 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 1 A, 1B, 1C, IE, 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 аутосомнодоминантная 2, ишемическая болезнь сердца, гиперальфалипопротеинемия 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, аутосомнодоминантная, Х-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 Ш и IVF, синдром Дежерина-Соттаса, аутосомнодоминантный, дефицит дендритных клеток, моноцитов, В-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, IAA, 1C, IG, IBB, 1DD, IFF, 1HH, II, IKK, IN, IS, 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, Х-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 аутосомнодоминантный, 3, врожденный дискератоз, аутосомно-рецессивный, 1, 3, 4 и 5, врожденный дискератоз Х-сцепленный, дискинезия, семейная, с миокимией лица, дисплазминогенемия, дистония 2 (скрученность, аутосомно-рецессивная), 3 (скрученность, Х-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 Т-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 ПА и ПВ, глутаровая ацидурия, тип 1, дефицит глутатионсинтетазы, болезнь накопления гликогена 0 (мышцы), II (взрослая форма), IXa2,1Xc, тип 1 A, 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 типаIIB иIIA, 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 11c 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, аутосомнодоминантный, 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 Нойака, 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 (X-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 A или B, атипичная, мандибуло-лицевой дизостоз по типу Тричера-Коллинза, аутосомно-рецессивный, дефицит маннозо-связывающего белка, заболевание мочи типа «кленовый сироп» 1A и 3 типа, синдром Мардена-Уокера, синдром Марфана, синдром Маринеско-Шегрена, синдром Марцольфа, диабет зрелого возраста у молодых людей типа 1, типа 2, типа И, 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 2a, 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 Миеси 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 гуанинoксидазы 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, аутосомнодоминантная 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, Х-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 Нейл-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 1С, дефект транспорта нейтральной 1 аминокислоты, болезнь накопления нейтральных липидов с миопатией, синдром нейтрофильного иммунодефицита, синдром Николаидеса-Барайцера, болезнь Ниманна-Пика типа Cl, С2, типа А и типа С1, 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 1В, 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 аутосомнодоминантный 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 (аутосомнорецессивное раннее начало) и 9, частичный альбинизм, частичный дефицит гипоксантин-гуанинфосфорибозилтрансферазы, узорчатая дистрофия пигментного эпителия сетчатки,

PC-Кба, 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, Х-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 1 А, 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 El-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, аутосомнорецессивный, с поздней нейросенсорной тугоухостью или с гемолитической анемией, почечный канальцевый ацидоз, проксимальный, с глазными аномалиями и умственной отсталостью, дистрофия конуса сетчатки 3В, 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 Шар леву а-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, аутосомнорецессивная, 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, IB, ID, 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, отонебнодигитальный синдром, тип I, X-сцепленный тяжелый комбинированный иммунодефицит, синдром Циммермана-Лабанда и синдром Циммермана-Лабанда 2, и зонулярная пылевидная катаракта 3.

[0973] Целевая нуклеотидная последовательность 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.

[0974] Таким образом, согласно некоторым аспектам дезаминирование мутантного C приводит к изменению аминокислоты, кодируемой мутантным кодоном, что в некоторых случаях может приводить к экспрессии аминокислоты дикого типа. Согласно другим аспектам дезаминирование мутантного A приводит к изменению аминокислоты,

кодируемой мутантным кодоном, что в некоторых случаях может приводить к экспрессии аминокислоты дикого типа.

[0975] Способы, описанные в настоящем документе, предусматривающие контактирование клетки с композицией или rAAV частицей, могут осуществляться *in vitro, ex vivo,* или *in vivo*. Согласно определенным вариантам осуществления стадия контактирования осуществляется в субъекте. Согласно определенным вариантам осуществления у субъекта было диагностировано заболевание, нарушение или состояние.

[0976] Согласно некоторым вариантам осуществления способы, раскрытые в настоящем документе, предусматривают контактирование клетки млекопитающего с композицией или rAAV частицей. Согласно конкретным вариантам осуществления способы предусматривают контактирования клетки сетчатки, кортикальной клетки или церебеллярной клетки.

[0977] Расщепленный белок Cas9 или расщепленный праймирующий редактор, доставляемый с применением способов, описанных в настоящем документе, предпочтительно имеет сравнимую активность по сравнению с исходным белком Cas9 или праймирующим редактором (*m.e.*, неразделенным белком, доставляемым в клетку или экспрессируемым в клетке в целом). Например, расщепленный белок Cas9 расщепленный праймирующий 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%) от активности исходного белка Cas9 или праймирующего редактора. Согласно некоторым вариантам осуществления расщепленный белок Cas9 расщепленный праймирующий редактор праймирующий редактор является более активным (например, 2-кратно, 5-кратно, 10-кратно, 100-кратно, 1000-кратно или более), чем исходный белок Cas9 или праймирующий редактор.

[0978] Описанные в настоящем документе композиции могут быть введены субъекту, нуждающемуся в этом, в терапевтически эффективном количестве для лечения и/или профилактики заболевания или нарушения, которым страдает субъект. Любое заболевание или нарушение, которое может быть подвергнуто лечению и/или профилактике с помощью технологии геномного редактирования на основе CRISPR/Cas9, может быть подвергнуто лечению с помощью расщепленного белка Cas9 или расщепленного мультилоскутного праймирующего редактора, описанных в настоящем документе Следует понимать, что если нуклеотидные последовательности, кодирующие расщепленный белок Cas9 или мультилоскутный праймирующий редактор, далее не кодируют гРНК, то может быть введен отдельный вектор нуклеиновой кислоты,

кодирующей гРНК, наряду с описанными в настоящем документе композициями.

[0979] Иллюстративные подходящие заболевания, нарушения или состояния включают без ограничения заболевание или нарушение, выбранное из группы, состоящей из: муковисцидоза, фенилкетонурии, эпидермолитического гиперкератоза (ЕНК), хронической обструктивной болезни легких (COPD), 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).

[0980] Согласно некоторым вариантам осуществления заболевание, нарушение или состояние ассоциировано с точечной мутацией в гене *NPC,* гене *DNMT1,* гене *PCSK9,* или гене *TMC1.* Согласно определенным вариантам осуществления точечная мутация представляет собой мутацию T3182C в *NPC,* которая приводит к аминокислотному замещению I1061T.

[0981] Согласно определенным вариантам 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.

[0982] Согласно некоторым вариантам осуществления заболевание, нарушение или состояние ассоциировано с точечной мутацией, которая приводит к образованию стоп-кодона, например, преждевременного стоп-кодона в кодирующем участке гена.

[0983] Дополнительные примеры заболеваний, нарушений 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, and 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, and Ali *et al.*, *Hematol.* **2014,** 93: 381-384, эпидермолитический гиперкератоз (ЕНК) - 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. номер доступа P04264 в базе данных UNIPROT по ссылке 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, также см. номер доступа 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. номер доступа 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) – 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, hereditary lymphedema – *например*, наследственную лимфедему - например, мутацию гистидин на аргинин в положении 1035 или гомологичном остатке в VEGFR3 тирозинкиназе (мутация A>G mutation), 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 аß (мутация A>G) – см., *например*, Kumar *et al.*, *J. Biol. Chem.* **1999,** 274: 24137-24141. Полное содержание всех ссылочных источников и записей в базах данных включено в настоящий документ посредством ссылки.

Заболевание, связанное с экспансией тринуклеотидных повторов

[0984] 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 праймированного редактирования.

[0985] Участок против хода транскрипции от участка повтора может быть разрезан 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 лоскутом ведет к сокращению повторяющегося аллеля. Термин «сокращенный» обозначает укорачивание длины участка нуклеотидного повтора, что тем самым приводит к репарации участка тринуклеотидного повтора.

[0986] Системы мультилоскутного праймированного редактирования (*например*, система двухлоскутного и четырехлоскутного праймированного редактирования), описанные в настоящем документе, могут 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 доминантноотрицательным образом (например, при полиглутаминовых заболеваниях).

[0987] Без ограничения 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. Мультилоскутное праймированное редактирование может применяться для снижения или устранения данных участков экспансии триплетов посредством делеции одного или нескольких триплетов в виде нежелательных повторяющихся кодонов. Согласно варианту осуществления этого применения на фиг.23 приведена схема конструкции РЕгРНК для сокращения или уменьшения последовательностей тринуклеотидных повторов с помощью праймированного редактирования.

[0988] Мультилоскутное праймированное редактирование может быть реализовано для сокращения участков экспансии триплетов посредством разрыва участка против хода транскрипции от участка в виде триплетного повтора с помощью праймирующего редактора, содержащего РЕгРНК, которая подходящим образом нацелена на участок

разрыва. Праймирующий редактор затем обеспечивает синтез новой цепи ДНК (оцДНК-лоскута) на основе РЕгРНК в качестве матрицы (т.е. ее матрицы редактирования), которая кодирует здоровое количество триплетных повторов (которое зависит от конкретного гена и заболевания). Новая синтезированная цепь оцДНК, содержащая здоровую последовательность триплетных повторов, также синтезируется таким образом, чтобы включать короткий участок гомологии (т.е. 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-лоскутом ведет к сокращению повторяющегося аллеля.

[0989] 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).

[0990] 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).

[0991] Неограничивающие примеры белков, ассоциированных с заболеваниями, связанными с экспансией тринуклеотидных повторов, которые могут быть скорректированы с помощью праймированного 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 эндонуклеазу лоскутов 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 (El A-связывающий белок p400), GIGYF2 (белок GYF 2, взаимодействующий с GRB10), OGG1 (8-оксогуанин-ДНК-гликозилазу), STC1 (станниокальцин 1), CNDP1 (карнозиндипептидазу 1 (металлопептидазу семейства M20)), C1 Oorf2 (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 Spl), 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 (BCE2-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 Dl), 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-8-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.

[0992] Согласно определенному аспекту настоящее изобретение относится к способам 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 X-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, которая включена в настоящий документ посредством 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 нарушение (HPE), врожденный синдром центральной гиповентиляции (CCHS), ARX-несиндромную X-сцепленную умственную отсталость (XLMR) и окулофарингеальную мышечную дистрофию (OPMD). 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.

[0993] Согласно одному варианту осуществления на фиг.23 изображен способ лечения заболевания, связанного с тринуклеотидными повторами. В целом, данный подход предусматривает применение геномного редактирования на основе TPRT (т.е. праймированного редактирования) в комбинации с удлиненной гРНК, которая включает участок, который кодирует требуемую и здоровую замещающую последовательность тринуклеотидного повтора, которая должна заместить эндогенную связанную с заболеванием последовательность тринуклеотидного повтора посредством механизма процесса праймированного редактирования. Схема примера конструкции гРНК для сокращения последовательностей тринуклеотидных повторов и сокращение тринуклеотидных повторов с помощью редактирования генома на основе TPRT (т.е. праймированного редактирования) показаны на фиг.23.

Прионное заболевание

[0994] Мультилоскутное праймированное редактирование может применяться для профилактики 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.

[0995] 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. Мультилоскутное праймированное редактирование может применяться для внесения защитных точечных мутаций в прионный белок для предотвращения или остановки прогрессирования любого из данных прионных заболеваний.

[0996] 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.

[0997] 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.

[0998] 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.

[0999] 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.

[1000] 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.

[1001] В контексте настоящего изобретения 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 в настоящем документе 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.

[1002] В целом, и без ограничения теорией, прионные заболевания вызываются неправильным фолдингом прионных белков. При таких заболеваниях, которые часто называются в настоящем документе 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, иB_n является ß-обогащенным и мультимерным.

[1003] 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 которой включено в настоящий документ посредством ссылки.

[1004] PRNP (NCBI RefSeq No. NP_000302.1 (SEQ ID NO: 291)) - прионный белок человека — кодируется геном длиной 16 т.о., расположенным на хромосоме 20 (4686151-4701588). Он содержит два экзона, при этом экзон 2 несет открытую рамку считывания, которая кодирует бело PrP в эндоплазматическую сеть (ER), в то время как CTF

отщепляется GPI-якорем (сигнальный пептид [GPI-SP] гликозилфосфатидилинозитол [GPI]). GP 1-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 ß-листов, трех a-спиралей, и C-концевого участка (который содержит GPI-якорь Расщепление PrP приводит к получению гликопротеина длиной 208 AA, присоединенного к клеточной мембране.

[1005] Последовательность из 253 аминокислот PRNP (NP_000302.1) является следующей:

MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQ GGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPK TNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQVY YRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYER ESQAYYQRGSSMVLFSSPPVILLISFLIFLIVG (SEQ ID NO: 291).

[1006] Последовательность из 253 аминокислот PRNP (NP_000302.1) кодируется нуклеотидной последовательностью (NCBI ссыл. Seq No. NM_000311.5, «прионный белок homo sapiens (PRNP), транскриптный вариант 1, иРНК), является следующей:

GCGAACCTTGGCTGCTGGATGCTGGTTCTCTTTGTGGCCACATGGAGTGACCT GGGCCTCTGCAAGAAGCGCCCGAAGCCTGGAGGATGGAACACTGGGGGCAGCCGAT ACCCGGGGCAGGGCAGCCCTGGAGGCAACCGCTACCCACCTCAGGGCGGTGGTGGC TGGGGGCAGCCTCATGGTGGTGGCTGGGGGCAGCCTCATGGTGGTGGCTGGGGGCA GCCCCATGGTGGTGGCTGGGGACAGCCTCATGGTGGTGGCTGGGGTCAAGGAGGTG GCACCCACAGTCAGTGGAACAAGCCGAGTAAGCCAAAAACCAACATGAAGCACATG GCTGGTGCTGCAGCAGCTGGGGCAGTGGTGGGGGGCCTTGGCGGCTACATGCTGGG AAGTGCCATGAGCAGGCCCATCATACATTTCGGCAGTGACTATGAGGACCGTTACTAT CGTGAAAACATGCACCGTTACCCCAACCAAGTGTACTACAGGCCCATGGATGAGTAC AGCAACCAGAACAACTTTGTGCACGACTGCGTCAATATCACAATCAAGCAGCACACG GTCACCACAACCACCAAGGGGGAGAACTTCACCGAGACCGACGTTAAGATGATGGA GCGCGTGGTTGAGCAGATGTGTATCACCCAGTACGAGAGGGAATCTCAGGCCTATTAC CAGAGAGGATCGAGCATGGTCCTCTTCTCCTCTCCACCTGTGATCCTCCTGATCTCTTT CCTCATCTTCCTGATAGTGGGATGAGGAAGGTCTTCCTGTTTTCACCATCTTTCTAATC TTTTTCCAGCTTGAGGGAGGCGGTATCCACCTGCAGCCCTTTTAGTGGTGGTGTCTCA CTCTTTCTTCTCTCTTTGTCCCGGATAGGCTAATCAATACCCTTGGCACTGATGGGCAC TGGAAAACATAGAGTAGACCTGAGATGCTGGTCAAGCCCCCTTTGATTGAGTTCATCA TGAGCCGTTGCTAATGCCAGGCCAGTAAAAGTATAACAGCAAATAACCATTGGTTAAT

CTGGACTTATTTTTGGACTTAGTGCAACAGGTTGAGGCTAAAACAAATCTCAGAACA
GTCTGAAATACCTTTGCCTGGATACCTCTGGCTCCTTCAGCAGCTAGAGCTCAGTATA
CTAATGCCCTATCTTAGTAGAGATTTCATAGCTATTTAGAGATATTTTCCATTTTAAGAA
AACCCGACAACATTTCTGCCAGGTTTGTTAGGAGGCCACATGATACTTATTCAAAAAA
ATCCTAGAGATTCTTAGCTCTTGGGATGCAGGCTCAGCCCGCTGGAGCATGAGCTCTG
TGTGTACCGAGAACTGGGGTGATGTTTTACTTTTCACAGTATGGGCTACACAGCAGCT
GTTCAACAAGAGTAAATATTGTCACAACACTGAACCTCTGGCTAGAGGACATATTCAC
AGTGAACATAACTGTAACATATATGAAAGGCTTCTGGGACTTGAAATCAAATGTTTGG
GAATGGTGCCCTTGGAGGCAACCTCCCATTTTAGATGTTTAAAGGACCCTATATGTGG
CATTCCTTTCTTTAAACTATAGGTAATTAAGGCAGCTGAAAAGTAAATTGCCTTCTAGA
CACTGAAGGCAAATCTCCTTTGTCCATTTACCTGGAAACCAGAATGATTTTGACATAC
AGGAGAGCTGCAGTTGTGAAAGCACCATCATCATAGAGGATGATGTAATTAAAAAAT
GGTCAGTGTGCAAAGAAAAGAACTGCTTGCATTTCTTTATTTCTGTCTCATAATTGTC
AAAAACCAGAATTAGGTCAAGTTCATAGTTTCTGTAATTGGCTTTTGAATCAAAGAAT
AGGGAGACAATCTAAAAAATATCTTAGGTTGGAGATGACAGAAATATGATTGATTTGA
AGTGGAAAAAGAAATTCTGTTAATGTTAATTAAAGTAAAATTATTCCCTGAATTGTTTG
ATATTGTCACCTAGCAGATATGTATTACTTTTCTGCAATGTTATTATTGGCTTGCACTTT
GTGAGTATTCTATGTAAAAATATATATGTATATAAAATATATATTGCATAGGACAGACTTA
GGAGTTTTGTTTAGAGCAGTTAACATCTGAAGTGTCTAATGCATTAACTTTTGTAAGG
TACTGAATACTTAATATGTGGGAAACCCTTTTGCGTGGTCCTTAGGCTTACAATGTGCA
CTGAATCGTTTCATGTAAGAATCCAAAGTGGACACCATTAACAGGTCTTTGAAATATG
CATGTACTTTATATTTTCTATATTTGTAACTTTGCATGTTCTTGTTTTGTTATATAAAAAA
ATTGTAAATGTTTAATATCTGACTGAAATTAAACGAGCGAAGATGAGCACCA (SEQ ID NO: 292)

[1007] Участки мутации относительно PRNP (NP_000302.1), которые связаны с CJD и FFI, являются следующими. Данные мутации могут быть удалены или внесены с помощью мультилоскутных праймирующих редакторов, раскрытых в настоящем документе.

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТА PRNP, СВЯЗАННАЯ С ПРИОННЫМ ЗАБОЛЕВАНИЕМ CJD (СМ. ТАБЛИЦУ 1 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 ДЛЯ PRNP NP_000302.1)
D178N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHNCVNITIKQHTVTTTTKGENFTETDVKMMERV

	VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3940)
T188K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHKVTTTTKGENFTETDVKMMER VVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3941)
E196K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGKNFTETDVKMMER VVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3942)
E196A	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGANFTETDVKMMER VVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 296)
E200K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTKTDVKMMER VVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3943)
E200G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTGTDVKMMER VVEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 298)
V203I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDIKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3944)
R208H	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMEH VVEQMCITQYERESQAYYQRGSSMVLFSSPPV(SEQ ID NO: 3945)
V210I	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS

	KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV IEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3946)
E211Q	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VQQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3947)
M232R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSRVLFSSPPV (SEQ ID NO: 3948)

[1008] Участки мутации относительно PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с GSS, сообщены и являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТА PRNP, СВЯЗАННАЯ С ПРИОННЫМ ЗАБОЛЕВАНИЕМ GSS (СМ. ТАБЛИЦУ 2 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 ДЛЯ PRNP NP_000302.1)
P102L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKLS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3949)
P105L	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KLKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3950)
A117V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAVAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHR YPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3951)
G131V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLVSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV

	VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3952)
V176G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFGHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3953)
H187R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQRTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3954)
	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 487)
F198S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENSTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3955)
D202N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETNVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3956)
Q212P	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEPMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3957)
Q217R	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITRYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3958)
M232T	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS

	KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSTVLFSSPPV (SEQ ID NO: 3959)

[1009] Участки мутации относительно PRNP (NP_000302.1) (SEQ ID NO: 291), которые связаны с возможной защитной природой против прионного заболевания, являются следующими:

МУТАЦИЯ	АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ МУТАНТА PRNP, СВЯЗАННОГО С ЗАЩИТНОЙ ПРИРОДОЙ ПРОИВ ПРИОННОГО ЗАБОЛЕВАНИЯ (СМ. ТАБЛИЦУ 4 В BAGYNSZKY ET AL., 2018) (ОТНОСИТЕЛЬНО SEQ ID NO: 291 ДЛЯ PRNP NP_000302.1)
G127S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGSYMLGSAMSRPIIHFGSDYEDRYYRENMHR YPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3960)
G127V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGVYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3961)
M129V	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYVLGSAMSRPIIHFGSDYEDRYYRENMHR YPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3962)
D167G	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMGEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3963)
D167N	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMNEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3964)
N171S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS

	KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSSQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3965)
E219K	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYKRESQAYYQRGSSMVLFSSPPV (SEQ ID NO: 3966)
P238S	MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQG GGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPS KPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMH RYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERV VEQMCITQYERESQAYYQRGSSMVLFSSSPV (SEQ ID NO: 3967)

[1010] Таким образом, согласно различным вариантам осуществления мультилоскутное праймированное 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, мультилоскутное праймированное редактирование можно применять 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). Согласно другим вариантам осуществления мультилоскутное праймированное редактирование можно применять 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 применением праймированного редактирования риск развития прионного заболевания может быть снижен или устранен.

[1011] Согласно другим вариантам осуществления мультилоскутное праймированное редактирование может применяться для внесения защитной мутации в PRNP, которая связана с защитным эффектом против одного или нескольких прионных заболеваний. Например, мультилоскутное праймированное 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). Согласно другим вариантам 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).

[1012] Согласно конкретным вариантам осуществления мультилоскутное

праймированное редактирование может применяться для внесения защитной мутации G127V в PRNP, как показано на фиг.27 и описано в Примере 5.

[1013] Согласно другому варианту осуществления праймированное редактирование может применяться для внесения защитной мутации E219K в PRNP.

[1014] 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. Праймированное 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 праймированного редактирования или мультилоскутного праймированного редактирования для снижения или устранения экспрессии *PRNP* путем внесения в ген раннего стоп-кодона.

[1015] С помощью принципов, описанных в настоящем документе, для конструирования РЕгРНК, могут быть сконструированы походящие РЕгРНК для внесения требуемых защитных мутаций или для удаления ассоциированных с прионным заболеванием мутаций из PRNP. Например, приведенный ниже перечень РЕгРНК может применяться для внесения защитного аллеля G127V и защитного аллеля E219K в PRNP человека, а также защитного аллеля G127V в PRNP различных животных.

[10] Фармацевтические композиции

[1016] Согласно другому аспекту настоящее изобретение относится к фармацевтическим композициям, содержащим любой из различных компонентов систем мультилоскутного праймированного редактирования, описанной в настоящем документе (например, включая без ограничения napDNAbp, обратные транскриптазы, слитые белки (например, содержащие napDNAbp и обратные транскриптаза), удлиненные гидовые РНК, и комплексы, содержащие слитые белки и удлиненные гидовые РНК, а также вспомогательные элементы, такие как компоненты для разрыва второй цепи и нуклеазы для удаления 5'-эндогенного лоскута ДНК для помощи в направлении процесса мультилоскутного праймированного редактирования в сторону образования

отредактированного продукта).

[1017] Термин «фармацевтическая композиция» в контексте настоящего изобретения относится к композиции, составленной для фармацевтического применения. Согласно некоторым вариантам осуществления фармацевтическая композиция дополнительно содержит фармацевтически приемлемый носитель. Согласно некоторым вариантам осуществления фармацевтическая композиция содержит дополнительные агенты (например, для специфичной доставки, увеличения периода полужизни, или другие терапевтические соединения).

[1018] В контексте настоящего изобретения 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 в настоящем документе взаимозаменяемым образом.

[1019] Согласно некоторым вариантам осуществления фармацевтическая композиция составлена для доставки субъекту, например, для редактирования генов. Подходящие способы введения описанной в настоящем документе 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 интрацеребровентрикулярное введение.

[1020] Согласно некоторым вариантам осуществления описанную в настоящем документе фармацевтическую композицию вводят местно в участок заболевания (например, участок с опухолью). Согласно некоторым вариантам осуществления описанную в настоящем документе 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.

[1021] Согласно другим вариантам осуществления фармацевтическая композиция, описанная в настоящем документе, доставляется в системе с контролируемым высвобождением. Согласно одному варианту 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). Согласно другому варианту 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. See also 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.

[1022] Согласно некоторым вариантам осуществления фармацевтическая

композиция составлена в соответствии со стандартными процедурами в виде композиции, адаптированной для внутривенного или подкожного введения субъекту, например, человеку. Согласно некоторым вариантам 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 предоставлена ампула со стерильной водой или солевым раствором для инъекций, так что ингредиенты могут быть смешаны перед введением.

[1023] 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.

[1024] Фармацевтическая композиция может содержаться в липидной частице или везикуле, 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 (PEG) 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 везикул. Получение подобных липидных частиц является хорошо известным. См., например, патенты США 4880635, 4906477, 4911928, 4917951, 4920016 и 4921757, каждый из которых включен в настоящий документ посредством ссылки.

[1025] Описанная в настоящем документе фармацевтическая композиция может

быть, например, введена или упакована в виде единичной дозы. Термин «единичная доза» при использовании в отношении фармацевтической композиции согласно настоящему изобретению обозначает физически отдельные единицы, являющиеся подходящими в качестве единичной дозы для субъекта, при этом каждая единица содержит предварительно определенное количество активного вещества, рассчитанное для получения требуемого терапевтического эффекта, совместно с требуемым разбавителем, т.е. носителем или средой.

[1026] Кроме того, фармацевтическая композиция может быть представлена в виде фармацевтического набора, содержащего (a) контейнер, содержащий соединение согласно настоящему 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 согласно настоящему 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.

[1027] Согласно другому аспекту предусмотрено изделие, содержащее материалы, которые могут применяться для лечения описанных выше заболеваний. Согласно некоторым вариантам 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. Согласно некоторым вариантам осуществления контейнер содержит композицию, которая является эффективной для лечения описанного в настоящем документе 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. Согласно некоторым вариантам 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.

[11] Наборы, клетки, векторы и доставки

Наборы

[1028] Композиции согласно настоящему изобретению могут быть собраны в наборы. Согласно некоторым вариантам осуществления набор содержит векторы нуклеиновой кислоты для экспрессии описанных в настоящем документе мультилоскутных праймирующих редакторов Согласно другим вариантам осуществления набор дополнительно включает подходящие гидовые нуклеотидные последовательности (например, РЕгРНК и гРНК для второго участка) или векторы нуклеиновой кислоты для экспрессии таких гидовых нуклеотидных последовательностей для нацеливания белка Cas9 или праймирующего редактора на желаемую целевую последовательность.

[1029] Описанный в настоящем документе набор может содержать один или несколько контейнеров, включающих компоненты для осуществления описанных в настоящем документе способов и необязательно инструкции по применению. Любой из описанных в настоящем документе наборов может дополнительно содержать 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.

[1030] Согласно некоторым вариантам осуществления наборы могут необязательно включать инструкции и/или материалы относительно применения представленных компонентов. В контексте настоящего изобретения «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 животному. В контексте настоящего изобретения «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 в настоящем документе.

[1031] Наборы могут содержать любой один или несколько описанных в настоящем документе 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.

[1032] 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, содержащим конструкцию нуклеиновой кислоты, содержащую нуклеотидную последовательность, кодирующую различные компоненты описанных в настоящем документе систем мультилоскутного праймированного редактирования (*например*, системы двухлоскутного праймированного редактирования и четырехлоскутного праймированного редактирования), описанные в настоящем документе (например, включая без ограничения napDNAbp, обратные транскриптазы, полимеразы,

слитые белки (например, включающие napDNAbp и обратные транскриптазы (или, более широко, полимеразы), удлиненные гидовые РНК, и комплексы, содержащие слитые белки и удлиненные гидовые РНК, а также вспомогательные элементы, такие как компоненты для разрыва второй цепи (например, гРНК для разрыва второй цепи) и нуклеазы для удаления 5'-эндогенного лоскута ДНК для направления мультилоскутного праймированного редактирования в сторону образования отредактированного продукта). Согласно некоторым вариантам осуществления нуклеотидная последовательность (последовательности) содержит гетерологичный промотор (или более одного промотора), который управляет экспрессией компонентов мультилоскутной системы праймированного редактирования.

[1033] Согласно другим аспектам настоящее изобретение относится к наборам, содержащим одну или несколько конструкций нуклеиновой кислоты, кодирующих различные компоненты систем мультилоскутного праймированного редактирования, описанных в настоящем документе, например, содержащих нуклеотидную последовательность, кодирующую компоненты системы мультилоскутного праймированного редактирования, способную модифицировать целевую последовательность ДНК. Согласно некоторым вариантам осуществления нуклеотидная последовательность содержит гетерологичный промотор, который обеспечивает экспрессию компонентов системы мультилоскутного праймированного редактирования.

[1034] Согласно некоторым аспектам настоящее изобретение относится к наборам, содержащим конструкцию нуклеиновой кислоты, содержащую (a) нуклеотидную последовательность, кодирующую napDNAbp (*например,* домен Cas9), слитый с обратной транскриптазой, и (b) гетерологичный промотор, который обеспечивает экспрессию последовательности (a).

Клетки

[1035] Клетки, которые могут содержать любую из описанных в настоящем документе композиций, включают прокариотические клетки и эукариотические клетки. Описанные в настоящем документе способы предусматривают доставку белка Cas9 или мультилоскутного праймирующего редактора в эукариотическую клетку (например, клетку млекопитающего, такую как клетка человека). Согласно некоторым вариантам осуществления клетка представлена *in vitro* (например, является культивируемой клеткой). Согласно некоторым вариантам осуществления клетка представлена *in vivo* (например, располагается в субъекте, таком как представляющий собой человека субъект). Согласно некоторым вариантам осуществления клетка представлена *ex vivo* (например, выделена из субъекта и может быть введена обратно тому же или другому субъекту).

[1036] 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). Согласно некоторым вариантам осуществления rAAV векторы доставляют в эмбриональные клетки почки человека (HEK) (например, клетки HEK 293 или HEK 293T). Согласно некоторым вариантам 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» означает 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» означает 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 в настоящем документе 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).

[1037] 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, High Five клетки, HL-60, HMEC, HT-29, HUVEC, J558L клетки, Jurkat, JY клетки, K562 клетки, 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, RBL клетки, RenCa, RIN-5F, RMA/RMAS, S2, Saos-2 клетки, Sf21, Sf9, SiHa, SKBR3, SKOV-3, T-47D, T2, T84, THP1, U373, U87, U937, VCaP, WM39, WT-49, X63, YAC-1 и YAR клетки.

[1038] Согласно некоторым аспектам настоящее изобретение относится к клеткам, содержащим любую из конструкций, раскрытых в настоящем документе. Согласно некоторым вариантам осуществления клетка-хозяин транзиентно или нетранзиентно трансфицируется одним или несколькими описанными в настоящем документе векторами. Согласно некоторым вариантам осуществления клетка трансфицируется в естественном виде, в котором она находится в субъекте. Согласно некоторым вариантам осуществления клетку, которая подвергается трансфекции, забирают от субъекта. Согласно некоторым вариантам 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 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, фибробласты мыши 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A 172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 клетки, 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 клетки, K562 клетки, 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, OPCN/OPCT клеточные линии, Peer, PNT-1A/PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2

клетки, Sf-9, SkBr3, T2, T-47D, T84, THP1 клеточные линии, U373, U87, U937, VCaP, Vero клетки, WM39, WT-49, X63, YAC-1, YAR и их трансгенные варианты.

[1039] Линии клеток доступны из ряда источников, известных специалистам в данной области (см., например, Американскую коллекцию типовых культур (American Type Culture Collection, ATCC) (Manassus, Va.)). Согласно некоторым вариантам осуществления клетка, трансфицированная одним или несколькими описанными в настоящем документе векторами, применяется для получения новой линии клеток, включающих одну или несколько последовательностей из вектора. Согласно некоторым вариантам осуществления клетка, транзиентно трансфицированная компонентами системы CRISPR, как описано в настоящем документе, (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. Согласно некоторым вариантам осуществления клетки транзиентно или нетранзиентно трансфицируют одним или несколькими описанными в настоящем документе векторами или линии клеток, полученные из подобных клеток, применяют при анализе одного или нескольких тестируемых соединений.

Векторы

[1040] Согласно некоторым аспектам настоящее изобретение относится к применению рекомбинантных вирусных векторов (например, адено-ассоциированных вирусных векторов, аденовирусных векторов, или векторов на основе вируса простого герпеса) для доставки описанных в настоящем документе мультилоскутных праймирующих редакторов или их компонентов, например, расщепленного белка Cas9 или расщепленных мультилоскутных праймирующих редакторов, в клетку. В случае подхода на основе расщепленого PE N-концевая часть слитого белка PE и C-концевая часть 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 мультилоскутные праймирующие редакторы превышают лимит упаковывания различных вирусных векторов, например, rAAV (~4,9 т.о.).

[1041] Таким образом, согласно одному варианту осуществления настоящее изобретение охватывает векторы, способные доставлять расщепленные слитые белки мультилоскутного праймирующего редактора, или их расщепленные компоненты. Согласно некоторым вариантам осуществления настоящее изобретение относится к композиции для доставки расщепленного слитого белка Cas9 или расщепленного

праймирующего редактора в клетку (например, клетку млекопитающего, клетку человека). Согласно некоторым вариантам осуществления композиция согласно настоящему изобретению содержит: (i) частицу первого рекомбинантного адено-ассоциированного вируса (rAAV), содержащую первую нуклеотидную последовательность, кодирующую N-концевую часть белка Cas9 или праймирующий редактор, слитый на его C-конце с интеином-N, и (ii) вторую частицу рекомбинантного адено-ассоциированного вируса (rAAV), содержащую вторую нуклеотидную последовательность, кодирующую интеин-C, слитый с N-концом C-концевой части белка Cas9 или праймирующий редактор. rAAV частицы согласно настоящему изобретению включают rAAV вектор (т.е. рекомбинантный геном rAAV), инкапсулированный в белках капсида вируса.

[1042] Согласно некоторым вариантам осуществления вектор rAAV содержит: (1) а гетерологичный участок нуклеиновой кислоты, содержащий первую или вторую нуклеотидную последовательность, кодирующую N-концевую часть или C-концевую часть расщепленного белка Cas9 или расщепленного мультилоскутного праймирующего редактора в любой форме, как описано в настоящем документе, (2) одну или несколько нуклеотидных последовательностей, содержащих последовательность, которая обеспечивает экспрессию гетерологичного участка нуклеиновой кислоты (например, промотора), и (3) один или несколько участков нуклеиновой кислоты, содержащих последовательность, которая обеспечивает встраивание гетерологичного участка нуклеиновой кислоты (необязательно с одним или несколькими участками нуклеиновой кислоты, содержащими последовательность, которая обеспечивает экспрессию), в геном клетки. Согласно некоторым вариантам осуществления вирусные последовательности, которые обеспечивают встраивание, содержат последовательности инвертированных концевых повторов (ITR). Согласно некоторым вариантам осуществления первая или вторая нуклеотидная последовательность, кодирующая N-концевую часть или C-концевую часть расщепленного белка Cas9 или расщепленного мультилоскутного праймирующего редактора, фланкируется на каждой стороне последовательностью ITR. Согласно некоторым вариантам осуществления вектор нуклеиновой кислоты дополнительно содержит участок, кодирующий белок AAV Rep, как описано в настоящем документе, 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. Согласно некоторым вариантам осуществления последовательности ITR получены из AAV2 или AAV6.

[1043] Таким образом, согласно некоторым вариантам осуществления описанные в

настоящем документе rAAV частицы содержат, по меньшей мере, одну rAAV2 частицу, rAAV6 частицу, rAAV8 частицу, rPHP.B частицу, rPHP.eB частицу, или rAAV9 частицу, или их вариант. Согласно определенным вариантам осуществления описанные rAAV частицы представляют собой rPHP.B частицы, rPHP.eB частицы, rAAV9 частицы.

[1044] Последовательности ITR и плазмиды, содержащие последовательности ITR, известны в данной области техники и коммерчески доступны (см., *например*, продукты и услуги, доступные от Vector Biolabs, Philadelphia, PA, Cellbiolabs, San Diego, CA, Agilent Technologies, Santa Clara, Ca, and Addgene, Cambridge, MA, and 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, and 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 включены в настоящий документ посредством ссылки).

[1045] Согласно некоторым вариантам осуществления rAAV вектор согласно настоящему изобретению содержит один или несколько регуляторных элементов для контролирования экспрессии гетерологичного участка нуклеиновой кислоты (например, промоторы, терминаторы транскрипции, и/или другие регуляторные элементы). Согласно некоторым вариантам 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 расщепленного белка Cas9 или расщепленного мультилоскутного праймирующего редактора. Согласно некоторым вариантам осуществления терминатор транскрипции, используемый в настоящем изобретении, представляет собой терминатор транскрипции bGH. Согласно некоторым вариантам осуществления rAAV вектор дополнительно включает посттранскрипционный регуляторный элемент вируса гепатита сурка (WPRE). Согласно определенным вариантам осуществления WPRE представляет собой усеченную последовательность WPRE, например, "W3". Согласно некоторым вариантам осуществления WPRE вставлен 5'

относительно терминатора транскрипции. При транскрипции такие последовательности создают третичную структуру, которая повышает экспрессию, особенно из вирусных векторов.

[1046] Согласно некоторым вариантам осуществления используемые согласно настоящему изобретению векторы могут кодировать слитые белки PE или любые их компоненты (например, napDNAbp, линкеры или полимеразы). Кроме того, применяемые в настоящем документе векторы могут кодировать PEгРНК и/или вспомогательную гРНК для разрыва второй цепи. Векторы могут быть способны обеспечивать экспрессию одной или нескольких кодирующих последовательностей в клетке. Согласно некоторым вариантам осуществления клетка может представлять собой прокариотическую клетку, например, бактериальную клетку. Согласно некоторым вариантам осуществления клетка может представлять собой эукариотическую клетку, например, клетку дрожжей, растения, насекомого или млекопитающего. Согласно некоторым вариантам осуществления эукариотическая клетка может представлять собой клетку млекопитающего. Согласно некоторым вариантам осуществления эукариотическая клетка может представлять собой клетку грызуна. Согласно некоторым вариантам осуществления эукариотическая клетка может представлять собой клетку человека. Подходящие промоторы для обеспечения экспрессии в различных типах клеток известны в данной области. Согласно некоторым вариантам 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.

[1047] Согласно некоторым вариантам осуществления промоторы, которые могут применяться в рамках векторов праймирующего редактора, могут являться конститутивными, индуцируемыми, или специфичными в отношении ткани. Согласно некоторым вариантам 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. Согласно некоторым вариантам осуществления промотор может представлять собой промотор CMV.

Согласно некоторым вариантам осуществления промотор может представлять собой усеченный промотор CMV. В других вариантах осуществления промотор может представлять собой промотор EFla. Согласно некоторым вариантам 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. Согласно некоторым вариантам осуществления индуцируемый промотор может представлять собой промотор, имеющий низкий базальный (неиндуцированный) уровень экспрессии, например, промотор Tet-On® (Clontech). Согласно некоторым вариантам осуществления промотор может представлять собой специфический в отношении ткани промотор. Согласно некоторым вариантам 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-ß, промотор Mb, промотор Nphsl, промотор OG-2, промотор SP-B, промотор SYN1, и промотор WASP.

[1048] Согласно некоторым вариантам осуществления векторы праймирующего редактора (например, включая любые векторы, кодирующие слитый белок праймирующего редактора и/или РЕгРНК, и/или вспомогательные гРНК для разрыва второй цепи) могут содержать 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. Согласно некоторым вариантам осуществления индуцируемый промотор может представлять собой промотор, имеющий низкий базальный (неиндуцированный) уровень экспрессии, например, промотор Tet-On® (Clontech).

[1049] Согласно дополнительным вариантам осуществления векторы праймирующего редактора (например, включая любые векторы, кодирующие слитый белок праймирующего редактора и/или РЕгРНК, и/или вспомогательные гРНК для разрыва второй цепи) могут содержать 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 GF AP, промотор GPIIb, промотор ICAM- 2, промотор INF-ß, промотор Mb, промотор Nphsl, промотор OG-2, промотор SP-B, промотор SYN1, и промотор WASP.

[1050] Согласно некоторым вариантам осуществления нуклеотидная последовательность, кодирующая РЕгРНК (или любые гидовые РНК, используемые в связи с мультилоскутным праймированным редактированием), может быть функционально связана, по меньшей мере, с одной последовательностью контроля транскрипции или трансляции. Согласно некоторым вариантам осуществления нуклеотидная последовательность, кодирующая гидовую РНК может быть функционально связана по меньшей мере с одним промотором. Согласно некоторым вариантам осуществления промотор может распознаваться РНК-полимеразой III (Pol III). Неограничивающие примеры Pol III промоторов включают промоторы U6, HI и тРНК. Согласно некоторым вариантам осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана с промотором U6 мыши или человека. В других вариантах осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана с промотором HI мыши или человек. Согласно некоторым вариантам осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана с тРНК промотором мыши или человека. В вариантах осуществления с более чем одной гидовой РНК промоторы, применяемые для обеспечения экспрессии, могут являться одинаковыми или различающимися. Согласно некоторым вариантам осуществления нуклеотид, кодирующий сгРНК гидовой РНК, и нуклеотид, кодирующий tracrРНК гидовой РНК, могут быть внесены на одном векторе. Согласно некоторым вариантам осуществления нуклеотид, кодирующий сгРНК, и нуклеотид, кодирующий tracrРНК, могут быть управляемы с помощью одного промотора. Согласно некоторым вариантам осуществления сгРНК и tracrРНК могут быть транскрибированы в один транскрипт. Например, сгРНК и tracrРНК могут быть процессированы из одного транскрипта для образования гидовой РНК в виде двойной молекулы. В качестве альтернативы, сгРНК и tracrРНК могут быть транскрибированы в гидовую РНК в виде одной молекулы.

[1051] Согласно некоторым вариантам осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть расположена на одном векторе, содержащем нуклеотидную последовательность, кодирующую слитый белок РЕ. Согласно некоторым вариантам осуществления экспрессия гидовой РНК и слитого белка РЕ может быть обеспечена посредством их соответствующих промоторов. Согласно некоторым вариантам осуществления экспрессия гидовой РНК может быть обеспечена

посредством того же промотора, который обеспечивает экспрессию слитого белка PE. Согласно некоторым вариантам осуществления гидовая РНК и транскрипт слитого белка PE могут содержаться в одном транскрипте. Например, гидовая РНК может располагаться в нетранслируемом участке (UTR) транскрипта белка Cas9. Согласно некоторым вариантам осуществления гидовая РНК может располагаться в 5' UTR транскрипта слитого белка PE. Согласно другим вариантам осуществления направляющая РНК может располагаться в 3' UTR транскрипта слитого белка PE. Согласно некоторым вариантам осуществления внутриклеточный период полужизни транскрипта слитого белка PE может быть снижен посредством включения гидовой РНК в его 3' UTR с укорачиванием, тем самым, длины его 3' UTR. Согласно дополнительным вариантам осуществления гидовая РНК может располагаться в интроне транскрипта слитого белка PE. Согласно некоторым вариантам осуществления в интрон, в котором располагается гидовая РНК, могут быть добавлены подходящие участки сплайсинга, так что гидовая РНК подвергается сплайсингу из транскрипта правильным образом. Согласно некоторым вариантам осуществления экспрессия белка Cas9 и гидовой РНК в непосредственной близости на одном векторе может обеспечивать более эффективное образование комплекса CRISPR.

[1052] Векторная система мультилоскутного праймирующего редактора может содержать один вектор, или два вектора, или три вектора, или четыре вектора, или пять векторов, или более. Согласно некоторым вариантам осуществления векторная система может содержать один единственный вектор, который кодирует как слитый белок PE, так и PEгРНК. Согласно другим вариантам осуществления векторная система может содержать два вектора, при этом один вектор кодирует слитый белок PE, а другой кодирует PEгРНК. Согласно дополнительным вариантам осуществления векторная система может содержать три вектора, при этом третий вектор кодирует гРНК для разрыва второй цепи, применяемую в рамках описанных в настоящем документе способов.

[1053] Согласно некоторым вариантам осуществления композиция, содержащая rAAV частицу (в любой предусмотренной в настоящем документе форме), дополнительно содержит фармацевтически приемлемый носитель. Согласно некоторым вариантам осуществления композиция содержит подходящие фармацевтические носители для введения субъектам, представляющим собой человека или животных.

[1054] Некоторые примеры веществ, 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 в настоящем документе взаимозаменяемым образом.

Способы доставки

[1055] Согласно некоторым аспектам настоящее изобретение относится к способам, предусматривающим доставку одного или нескольких полинуклеотидов, кодирующих мультилоскутные праймирующие редакторы, описанные в настоящем документе, как например, одного или нескольких векторов, как описано в настоящем документе, одного или нескольких их транскриптов, и/или одного или нескольких транскрибированных из них белков в клетку-хозяин. Согласно некоторым аспектам настоящее изобретение дополнительно предусматривает клетки, полученные с помощью таких способов, и организмы (такие как животные, растения или грибы), состоящие или полученные из таких клеток. Согласно некоторым вариантам осуществления редактор оснований, как описано в настоящем документе, в комбинации (и необязательно в виде комплекса) с гидовой последовательностью доставляется в клетку.

[1056] Иллюстративные стратегии доставки описаны в настоящем документе и включают стратегии на основе векторов, доставку на основе рибонуклеопротеинового комплекса PE, и доставку PE с помощью способов на основе гРНК.

[1057] Согласно некоторым вариантам осуществления способ доставки согласно настоящему изобретению предусматривает нуклеофекцию, микроинъекцию, биолистику,

вирусомы, липосомы, иммунолипосомы, конъюгаты поликатион или липид:нуклеиновая кислота, голую ДНК, искусственные вирионы, и агент-энхансированный захват ДНК.

[1058] Иллюстративные способы 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.

[1059] Получение комплексов липид:нуклеиновая кислота, включая липосомы-мишени, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области техники (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 № 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028 и 4,946,787).

[1060] Согласно другим вариантам осуществления способ доставки и вектор согласно настоящему изобретению представляет собой RNP комплекс. RNP доставка слитых белков заметно повышает специфичность в плане ДНК при редактировании оснований. 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 при 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 включен в настоящий документ посредством ссылки.

[1061] Дополнительные способы доставки нуклеиновых кислот в клетки известны

специалистам в данной области. См., например, публикацию US 2003/0087817, приведенную в настоящем документе посредством ссылки.

[1062] Другие аспекты настоящего изобретения относятся к способам доставки конструкций мультилоскутных праймирующих редакторов в клетку с образованием полного и функционального праймирующего редактора в клетке. Например, согласно некоторым вариантам осуществления клетка контактирует с описанной в настоящем документе композицией (например, композициями, содержащими нуклеотидные последовательности, кодирующие расщепленный Cas9 или расщепленный праймирующий редактор или AAV частицы, содержащие векторы нуклеиновой кислоты, содержащие такие нуклеотидные последовательности). Согласно некоторым вариантам осуществления контактирование приводит к доставке таких нуклеотидных последовательностей в клетку, при этом N-концевая часть белка Cas9 или праймирующего редактора и C-концевая часть белка Cas9 или праймирующего редактора экспрессируются в клетке и соединяются с образованием полного белка Cas9 или полного праймирующего редактора.

[1063] Очевидно, что любая предусмотренная в настоящим изобретением rAAV частица, молекула нуклеиновой кислоты или композиция может быть внесена в клетку любым подходящим способом, либо стабильно, или транзиентно. Согласно некоторым вариантам осуществления описанные белки могут быть трансфицированы в клетку. Согласно некоторым вариантам 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. Согласно некоторым вариантам осуществления клетки, экспрессирующие разделенный белок или содержащие разделенный белок, могут быть трансдуцированы или трансфицированы одной или несколькими последовательностями гидовой РНК, например, при доставке разделенного белка Cas9 (например, nCas9). Согласно некоторым вариантам 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.

[1064] Согласно определенным вариантам осуществления композиции согласно настоящему изобретению содержат липид и/или полимер. Согласно определенным

вариантам осуществления липид и/или полимер является катионным. Получение подобных липидных частиц является хорошо известным. См., например, патент США № 4880635, 4906477, 4911928, 4917951, 4920016, 4921757 и 9737604, каждый из которых включен в настоящий документ посредством ссылки.

[1065] Последовательность гидовой РНК может составлять 15-100 нуклеотидов в длину и включать последовательность из по меньшей мере 10, по меньшей мере 15 или по меньшей мере 20 непрерывных нуклеотидов, которая комплементарна целевой нуклеотидной последовательности. Гидовая РНК может содержать 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 непрерывных нуклеотидов, которая комплементарна целевой нуклеотидной последовательности. Гидовая РНК может составлять 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 в длину.

[1066] Согласно некоторым вариантам осуществления целевая нуклеотидная последовательность представляет собой последовательность ДНК в геноме, например, эукариотическом геноме. Согласно определенным вариантам осуществления целевая нуклеотидная последовательность расположена в геноме млекопитающего (например, человека.

[1067] Композиции согласно настоящему 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 вещества, рассчитанное для продуцирования требуемого терапевтического эффекта, совместно с требуемым разбавителем, т.е. носителем или индифферентным веществом.

[1068] Лечение заболевания или нарушения включает замедление развития или прогрессирования заболевания, или снижение тяжести заболевания. Лечение заболевания не обязательно требует результатов в виде излечения.

[1069] В контексте настоящего изобретения «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.

[1070] «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.

[1071] В контексте настоящего изобретения термин «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 сайта заболевания.

[1072] Предполагается, что специалист в данной области техники 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 в настоящем документе публикации приведены в качестве ссылок для указанных в настоящем документе целей или описываемого объекта.

Примеры

ПРИМЕР 1. ПРАЙМИРОВАННОЕ РЕДАКТИРОВАНИЕ (PE) ДЛЯ ВСТАВКИ ТОЧНЫХ НУКЛЕОТИДНЫХ ИЗМЕНЕНИЙ В ГЕНОМ

[1073] Цель состоит в том, чтобы разработать трансформационную технологию редактирования генома для точной и общей установки изменений одиночных нуклеотидов в геномах млекопитающих. Эта технология позволила бы исследователям изучить эффекты однонуклеотидных вариаций практически в любом гене млекопитающих и потенциально

позволила бы проводить терапевтические вмешательства для коррекции патогенных точечных мутаций у пациентов-людей.

[1074] Адоптация системы кластеризованных регулярных промежуточных коротких палиндромных повторов (CRISPR) для редактирования генома произвело революцию 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 двухцепочечных 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 п.о. от основания-мишени. Таким образом, преодоление данных ограничений значительно расширит варианты основного исследовательского и терапевтического использования геномного редактирования.

[1075] Здесь предлагается разработать новый подход точного редактирования, который обеспечивает множество преимуществ редактирования оснований, а именно устранение двухцепочечных 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 в настоящем документе конструкции гидовая РНК CRISPR (гРНК) будет сконструирована таким образом, чтобы нести матрицу, кодирующую синтез мутагенной цепи ДНК, для выполнения ассоциированным ферментом

в виде обратной транскриптазы (RT). Разрезаемый нуклеазой (Cas9) CRISPR целевой участок ДНК будет служить в качестве праймера для обратной транскрипции, обеспечивая прямое встраивание любого требуемого нуклеотидного изменения.

Раздел 1

[1076] Реализация обратной транскрипции мутагенных цепей ДНК с гидовой 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 гРНК 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 разорванные цепи ДНК в комплексах в виде связанных Cas9:гРНК действительно могут праймировать обратную транскрипцию с использованием связанной гРНК в качестве матрицы (фермент RT в транс). Далее будут определены различные линкеры гРНК, участки связывания праймера, и матрицы для синтеза для определения оптимальных правил конструирования *in vitro.* Затем будут оценены *in vitro* различные ферменты RT, действующие в транс или в виде продуктов слияния с Cas9. Наконец будут идентифицированы конструкции модифицированной гРНК, 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

[1077] Установление праймированного редактирования в клетках человека. Основываясь на механизмах процессинга и репарации ДНК, предполагается, что мутагенные цепи ДНК (одноцепочечные лоскуты) можно использовать для прямого и эффективного редактирования целевых нуклеотидов. В качестве одобрения предварительных исследований осуществимость этой стратегии была установлена путем демонстрации редактирования модельными плазмидными субстратами, содержащими мутагенные лоскуты. Параллельно с целью 1 результаты восстановления будут дополнительно оцениваться путем систематического изменения длины мутагенного лоскута, состава последовательности, идентичности целевого нуклеотида и 3'-конца. Также будут протестированы небольшие вставки и делеции от 1 до 3 нуклеотидов. Параллельно и на основе цели 1 будут оцениваться конструкции Cas9-RT, включая слитые белки и стратегии нековалентного рекрутирования. Конструкции Cas9-RT и удлиненные гРНК будут протестированы на клеточное редактирование в нескольких целевых участках генома

человека, а затем будут оптимизированы для достижения высокой эффективности. В случае успеха эта цель немедленно установит редактирование генома TPRT (т.е. праймированное редактирование) для приложений фундаментальной науки.

Раздел 3

[1078] Достижение сайт-специфического редактирования патогенных мутаций в культивируемых клетках человека. Потенциальная универсальность этой технологии может позволить редактировать трансверсионные мутации и вставки, которые в настоящее время не поддаются коррекции с помощью BE. Руководствуясь результатами целей 1 и 2, патогенные трансверсионные мутации будут нацелены на культивируемые клетки человека, включая мутацию, лежащую в основе серповидно-клеточной анемии в бета-глобине (для исправления требуется трансверсия A•T в T•A) и наиболее распространенные мутации болезни Вильсона в ATP7B (для исправления требуется трансверсия G•C в T•A). Также будет изучена коррекция небольших инсерционных и делеционных мутаций, включая 3-нуклеотидную делецию ΔF508 в *CFTR*, которая вызывает муковисциидоз. В случае успеха это заложит основу для разработки мощных терапевтических подходов к этим важным заболеваниям человека.

Подход

[1079] Цель состоит в том, чтобы разработать стратегию редактирования генома, которая напрямую устанавливает точечные мутации в целевых геномных сайтах. На этапе разработки технологии усилия будут сосредоточены на конструировании белков и РНК, чтобы включить функциональность TPRT в систему CRISPR/Cas. Анализы *in* vitro будут использоваться для тщательного изучения функции каждого этапа TPRT, построенного с нуля (Цель 1). Второе основное направление будет оценивать результаты редактирования в клетках млекопитающих с использованием комбинации модельных субстратов и сконструированных систем CRISPR/Cas (Цель 2). Наконец, на этапе применения технология будет использоваться для исправления мутаций, которые не поддавались редактированию генома другими методами (Цель 3).

[1080] Общая схема редактирования показана на фиг. 1A-1B. Никазы Cas9 содержат инактивирующие мутации в домене нуклеазы HNH (Spy Cas9 H840A или N863A), ограничивающие расщепление ДНК цепью, содержащей PAM (нецелевая цепь). Годовые РНК (гРНК) сконструированы таким образом, чтобы содержать матрицу для обратной транскрипции (схемы подробно описаны на слайде 5). Показано 5'-плечо удлинения гРНК, но также могут быть реализованы 3'-удлинения. Никаза Cas9 слита с ферментом обратной транскриптазы (RT) либо через C-конец, либо через N-конец. Комплекс гРНК:Cas9-RT нацеливается на представляющую интерес область ДНК и образует R-петлю после

смещения нецелевой цепи. Cas9 разрывает нецелевую цепь ДНК. Высвобождение цепи с разрывом обнажает свободный 3'-OH-конец, способный праймировать обратную транскрипцию с использованием расширенной гРНК в качестве матрицы. Эта реакция синтеза ДНК осуществляется слитым ферментом RT. Матрица гРНК кодирует последовательность ДНК, гомологичную исходному дуплексу ДНК, за исключением нуклеотида, предназначенного для редактирования. Продукт обратной транскрипции представляет собой лоскут одноцепочечной ДНК, который кодирует желаемое редактирование. Этот лоскут, который содержит свободный 3'-конец, может уравновешиваться с соседней цепью ДНК, в результате чего образуются 5'-лоскуты. Предполагается, что последний вид служит эффективным субстратом для FEN1 (эндонуклеаза 1 лоскута), фермента, который естественным образом вырезает 5'-лоскуты из фрагментов Окадзаки во время синтеза отстающей цепи ДНК и удаляет 5'-лоскуты после синтеза смещения цепи, который происходит во время эксцизионной репарации длинного основания. Лигирование разорванной ДНК дает несовпадающую пару оснований. Это промежуточное соединение может либо подвергнуться реверсии к исходной паре оснований, либо преобразованию в желаемую отредактированную пару оснований с помощью процессов репарации ошибок спаривания (MMR). В качестве альтернативы, полуконсервативная репликация ДНК может дать по одной копии для каждой реверсии и редактирования.

ПРИМЕР 2. ПРАЙМИРОВАННОЕ РЕДАКТИРОВАНИЕ: ОЧЕНЬ УНИВЕРСАЛЬНОЕ И ТОЧНОЕ РЕДАКТИРОВАНИЕ ГЕНОМА ПОИСКА И ЗАМЕНЫ В КЛЕТКАХ ЧЕЛОВЕКА БЕЗ ДВУХЦЕПОЧНЫХ РАЗРЫВОВ ДНК

[1081] Текущие методы редактирования генома могут разрушать, удалять или вставлять целевые гены с сопутствующими побочными продуктами двухцепочечных разрывов ДНК с использованием программируемых нуклеаз и устанавливать четыре точечные мутации транзиции в целевых локусах с помощью редакторов оснований. Небольшие вставки, небольшие делеции и восемь точечных мутаций трансверсии, однако, в совокупности представляют собой наиболее патогенные генетические варианты, но не могут быть эффективно исправлены без избытка побочных продуктов в большинстве типов клеток. В настоящем документе описывается праймированное редактирование, чрезвычайно универсальный и точный способ редактирования генома, который непосредственно записывает новую генетическую информацию в указанный участок ДНК с использованием каталитически поврежденного Cas9, слитого со сконструированной обратной транскриптазой, запрограммированной с помощью сконструированной гидовой

РНК праймированного редактирования (РЕгРНК), которая одновременно указывает целевой участок и кодирует желаемое редактирование. Было выполнено более 175 различных изменений в клетках человека, чтобы установить, что праймированное редактирование может эффективно выполнять целевые вставки, делеции, все 12 возможных типов точечных мутаций и их комбинации (обычно 20–60%, до 77% в несортированных клетках) и с низким содержанием побочных продуктов (обычно 1–10%), не требуя двухцепочечных разрывов или донорских матриц ДНК. Праймированное редактирование применяли в клетках человека для исправления первичных генетических причин серповидноклеточной анемии (требующей трансверсии A•T в T•A в *HBB*) и болезни Тея-Сакса (требующей делеции 4 оснований в *HEXA*), в обоих случаях эффективно возвращает патогенные геномные аллели к дикому типу с минимальными побочными продуктами. Праймированное редактирование также использовалось для создания клеточных линий человека с этими патогенными трансверсиями *HBB* и мутациями вставки *HEXA*, для установки мутации G127V в *PRNP*, которая придает устойчивость к прионной болезни (требующей трансверсии G•C-в-T•A), и для эффективной вставки метки His6, метки эпитопа FLAG и расширенного сайта *LoxP* в локусы-мишени в клетках человека. Праймированное редактирование предлагает преимущества 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 поиска и замены, который требует трех различных событий спаривания оснований, праймированное редактирование гораздо менее склонно к нецелевой модификации ДНК в известных нецелевых сайтах Cas9, чем Cas9. Праймированное редактирование существенно расширяет масштабы и возможности редактирования генома и в принципе может исправить ~ 89% известных патогенных генетических вариантов человека.

[1082] Возможность вносить практически любые целенаправленные изменения в геном любой живой клетки или организма — давняя цель науки о жизни. Несмотря на быстрый прогресс в технологиях редактирования генома, большинство из >75 000 известных генетических вариантов человека, связанных с заболеваниями[111], не могут быть скорректированы или внедрены в большинство терапевтически значимых клеток (фиг. 38A). Программируемые нуклеазы, такие как CRISPR-Cas9, делают двухцепочечные разрывы ДНК (DSB), которые могут разрушать гены, индуцируя смеси вставок и делеций (инделей) в сайтах-мишенях[112-114]. Нуклеазы также могут быть использованы для удаления генов-мишеней[115,116] или вставки экзогенных генов[117-119] посредством независимых от гомологии процессов. Двухцепочечные разрывы ДНК, однако, также связаны с нежелательными последствиями, включая сложные смеси продуктов, транслокации[120] и

активацию p53[121,122]. Более того, подавляющее большинство патогенных аллелей отличаются от своих непатогенных аналогов небольшими вставками, делециями или заменами оснований, для исправления которых требуются гораздо более точные технологии редактирования (фиг. 38A). Направленная на гомологию репарация (HDR), стимулируемая нуклеазно-индуцированными DSB [123] широко используется для установления множества точных изменений ДНК. HDR, однако, опирается на экзогенные матрицы репарации донорской ДНК, обычно генерирует избыток побочных продуктов инделей от репарации концевых соединений DSB и неэффективна в большинстве терапевтически релевантных типов клеток (Т-клетки и некоторые стволовые клетки являются важными исключениями)[124,125]. В то время как повышение эффективности и точности DSB-опосредованного редактирования генома остается в центре многообещающих усилий[126-130], эти проблемы требуют изучения альтернативных стратегий точного редактирования генома.

[1083] Редактирование оснований может эффективно установить или исправить четыре типа переходных мутаций (C в T, G в A, A в G и T в C), не требуя DSB в самых разных типах клеток и организмах, включая млекопитающих[128–131], но в настоящее время не может достичь ни одной из восьми мутаций трансверсии (C в A, C в G, G в C, G в T, A в C, A в T, T в A и T в G), таких как мутация T• A-в-A•T, необходимая для прямой коррекции наиболее распространенной причины серповидно-клеточной анемии (*HBB* E6V)[132]. Кроме того, не было сообщено ни о каком способе без DSB, выполняющем целевые делеции, такие как удаление дупликации 4 оснований, которая вызывает болезнь Тея-Сакса (*HEXA* 1278+TATC)[133], или целевых вставок, таких как точные вставки 3 оснований, необходимые для непосредственной коррекции наиболее частой причины муковисцидоза (*CFTR* ΔF508)[134]. Нацеленные точечные мутации трансверсии, инсерции и делеции, таким образом, трудно установить или корректировать эффективно и без избыточных побочных продуктов в большинстве типов клеток, даже если они в совокупности составляют большинство известных патогенных аллелей (фиг. 38A).

[1084] Описанная в настоящем документе разработка праймированного редактирования, новой технологии редактирования генома «поиск и замена», которая обеспечивает целевые вставки, делеции и все 12 возможных преобразований оснований в целевых локусах в клетках человека без использования двухцепочечного разрыва ДНК или матрицы донорской ДНК. Праймирующие редакторы, первоначально представленные PE1, используют обратную транскриптазу, слитую с программируемой никазой, и расширенную гидовую РНК праймированного редактирования (PEгРНК) для прямого копирования генетической информации из удлинения на PEгРНК в целевой геномный локус.

Праймирующие редактор второго поколения (PE2) использует сконструированную обратную транскриптазу для существенного повышения эффективности редактирования с минимальным (обычно <2%) образованием инделей, в то время как система PE3 третьего поколения добавляет вторую гидовую РНК для разрыва неотредактированной цепи, тем самым способствуя замене неотредактированной цепи и дополнительно повышая эффективность редактирования, как правило, приблизительно до 20-50% в клетках человека с образованием инделей приблизительно 1-10%. PE3 предлагает гораздо меньше побочных продуктов и более высокую или аналогичную эффективность по сравнению с оптимизированным HDR, инициируемым нуклеазой Cas9, и предлагает дополнительные сильные и слабые стороны по сравнению с редакторами оснований текущего поколения.

[1085] PE3 применяли в геномных локусах клеток HEK293T человека для достижения эффективного превращения *HBB E6V в HBB* дикого типа, делеции встроенного TATC для восстановления *HEXA 1278+TATC в HEXA* дикого типа, установки в *PRNP* мутации G127V, которая придает устойчивость к прионному заболеванию[135] (требуется трансверсия G•C-в-T•A), и целенаправленной вставки метки His_6 (18 п.о.), эпитопной метки FLAG (24 п.о.) и удлиненного сайт *LoxP* для Cre-опосредованной рекомбинации (44 п.о.). Праймированное редактирование также было успешным в трех других клеточных линиях человека, а также в постмитотических праймированных кортикальных нейронах мыши с различной эффективностью. Из-за высокой степени гибкости в отношении расстояния между начальным разрывом и местом редактирования праймированное редактирование существенно не ограничено требованием PAM Cas9 и, в принципе, может быть нацелено на подавляющее большинство геномных локусов. Нецелевое праймированное редактирование встречается намного реже, чем нецелевое редактирование Cas9 в известных нецелевых локусах Cas9, вероятно, из-за необходимости трех различных событий спаривания оснований ДНК для того, чтобы продуктивное праймированное редактирование имело место. Обеспечивая точные целевые вставки, делеции и все 12 возможных классов точечных мутаций в самых разных геномных локусах без необходимости использования DSB или матриц донорской ДНК, праймированное редактирование может продвинуть вперед изучение и коррекцию многих вариантов генов.

Результаты

Стратегия переноса информации из удлиненной гидовой РНК в целевой локус ДНК

[1086] Cas9 нацеливается на ДНК, используя гидовую РНК, содержащую спейсерную последовательность, которая гибридизуется с сайтом ДНК-мишени[112-114,136,137]. Цель состояла в том, чтобы сконструировать годовые РНК, чтобы указать ДНК-мишень, как в естественных системах CRISPR[138,139], а также включить новую генетическую

информацию, которая заменяет соответствующие нуклеотиды ДНК в целевом локусе. Прямой перенос генетической информации из удлиненной гидовой РНК в указанный участок ДНК с последующей заменой исходной неотредактированной ДНК, в принципе, может обеспечить общие средства установки целевых изменений последовательности ДНК в живых клетках, независимо от DSB или донорской матрицы ДНК. Чтобы добиться этой прямой передачи информации, цель состояла в том, чтобы использовать геномную ДНК, надрезанную в целевом сайте, чтобы обнажить 3'-гидроксильную группу, чтобы инициировать обратную транскрипцию генетической информации из удлинения на сконструированной гидовой РНК (далее называемой гидовой РНК для праймированного редактирования или РЕгРНК) непосредственно в сайт-мишень (фиг. 38A).

[1087] Эти начальные этапы разрыва и обратной транскрипции, которые напоминают механизмы, используемые некоторыми естественными мобильными генетическими элементами[140], приводят к разветвленному промежуточному продукту с двумя перекрывающимися одноцепочечными лоскутами ДНК на одной цепи: 5'-лоскут, который содержит неотредактированную последовательность ДНК, и лоскут 3', который содержит отредактированную последовательность, скопированную с РЕгРНК (фиг. 38B). Для успешного редактирования это разветвленное промежуточное звено должно быть разрешено таким образом, чтобы отредактированный 3- лоскут заменял неотредактированный 5- лоскут. В то время как гибридизация 5'-лоскута с неотредактированной цепью, вероятно, термодинамически благоприятна, поскольку отредактированный 3'-лоскут может образовывать меньше пар оснований с неотредактированной цепью, 5'-лоскуты являются предпочтительным субстратом для структурно-специфических эндонуклеаз, таких как FEN1[141], которая вырезает 5' лоскуты, образующиеся во время синтеза отстающей цепи ДНК и эксцизионной репарации оснований длинного участка. Было высказано предположение, что преимущественное удаление 5'-лоскута и лигирование 3'-лоскута может стимулировать включение отредактированной цепи ДНК, создавая гетеродуплексную ДНК, содержащую одну отредактированную цепь и одну неотредактированную цепь (фиг. 38B).

[1088] Постоянная установка редактирования может возникнуть в результате последующей репарации ДНК, которая устраняет несоответствие между двумя цепями ДНК таким образом, что информация в отредактированной цепи копируется в комплементарную цепь ДНК (фиг. 38C). Основываясь на аналогичной стратегии, разработанной для максимизации эффективности редактирования оснований ДНК[131-133], предполагалось, что разрыв неотредактированной цепи ДНК, достаточно далеко от места исходного разрыва, чтобы минимизировать образование двухцепочечного разрыва, может

привести к репарации ДНК, чтобы предпочтительно заменить неотредактированную нить.

Валидация стадий праймированного редактирования in vitro и в дрожжевых клетках

[1089] После расщепления PAM-содержащей цепи ДНК доменом RuvC нуклеазы Cas9, PAM-дистальный фрагмент этой цепи может диссоциировать от стабильных комплексов Cas9:сгРНК:ДНК[143]. Было высказано предположение, что 3'-конец этой высвобожденной цепи может быть достаточно доступным для праймированной полимеризации ДНК. Руководящие усилия по инженерии РНК[144-146] и кристаллические структуры комплексов Cas9:сгРНК:ДНК[147-149] позволяют предположить, что 5'- и 3'-концы сгРНК могут быть удлинены без отмены активности Cas9:сгРНК. РЕгРНК были сконструированы путем удлинения сгРНК для включения двух критических компонентов: участка связывания праймеров (PBS), который позволяет 3'-концу цепи ДНК гибридизоваться с РЕгРНК, и матрицы обратной транскриптазы (RT), содержащей желаемое редактирование, которое будет непосредственно копироваться в сайт геномной ДНК, поскольку 3'-конец цепи ДНК с разрывом удлиняется через матрицу РНК полимеразой (фиг. 38C).

[1090] Эти гипотезы были проверены *in vitro* с использованием очищенного белка Cas9 S. pyogenes. Ряд потенциальных РЕгРНК был сконструирован путем удлинения сгРНК на любом конце последовательностью PBS (от 5 до 6 нуклеотидов) и матрицей RT (от 7 до 22 нуклеотидов). Было подтверждено, что 5'-удлиненные РЕгРНК направляют связывание Cas9 с ДНК-мишенью, и что как 5'-удлиненные РЕгРНК, так и 3'-удлиненные РЕгРНК поддерживают Cas9-опосредованное разрезание мишеней *in vitro* и активность расщепления ДНК в клетках млекопитающих (фиг. 44A-44C). Эти потенциальные конструкции РЕгРНК были протестированы с использованием субстратов двухцепочечной ДНК, предварительно меченных 5'-Cy5, каталитически мертвого Cas9 (dCas9) и коммерческого варианта обратной транскриптазы вируса мышиного лейкоза Молони (M-MLV) (фиг. 44D). Когда присутствовали все компоненты, наблюдалось эффективное превращение флуоресцентно меченной цепи ДНК в более длинные продукты ДНК с подвижностью в геле, согласующиеся с обратной транскрипцией вдоль матрицы RT (фиг. 38D, фиг. 44D-44E). Продукты желаемой длины были сформированы либо с 5'-удлиненной, либо с 3'-удлиненной РЕгРНК (фиг. 38D-38E). Исключение dCas9 привело к образованию продуктов трансляции разрыва, полученных в результате опосредованной обратной транскриптазой полимеризации ДНК на матрице ДНК без переноса информации РЕгРНК (фиг. 38D). Никаких продуктов полимеризации ДНК не наблюдалось, когда РЕгРНК была заменена обычной сгРНК, что подтверждает необходимость матричных компонентов PBS и RT для РЕгРНК (фиг. 38D). Эти результаты демонстрируют, что плавление ДНК,

опосредованное Cas9, обнажает одноцепочечную R-петлю, которая, если ее надрезать, способна инициировать обратную транскрипцию либо с 5'-удлиненной, либо с 3'-удлиненной РЕгРНК.

[1091] Затем субстраты дцДНК без разрывов тестировали с помощью никазы Cas9 (мутант H840A), которая разрывает исключительно PAM-содержащую цепь[112]. В этих реакциях 5'-удлиненные РЕгРНК неэффективно генерировали продукты обратной транскрипции, возможно, из-за нарушенной активности никазы Cas9 (фиг. 44F). Тем не менее, 3'-удлиненные РЕгРНК обеспечивают надежный разрыв Cas9 и эффективную обратную транскрипцию (фиг. 38E). Использование 3'-удлиненных РЕгРНК генерировало только один кажущийся продукт, несмотря на принципиальную возможность окончания обратной транскрипции внутри РЕгРНК. Секвенирование ДНК продуктов реакций с никазой Cas9, RT и 3'-удлиненными РЕгРНК показало, что полная последовательность матрицы RT подвергалась обратной транскрипции в ДНК-субстрат (фиг. 44G). Эти эксперименты установили, что 3'-удлиненные РЕгРНК могут служить матрицей для обратной транскрипции новых цепей ДНК, сохраняя при этом способность направлять активность никазы Cas9.

[1092] Для оценки результатов репарации ДНК эукариотических клеток 3'-лоскутов, полученных с помощью запрограммированной РЕгРНК обратной транскрипции *in vitro,* разрыва ДНК и обратной транскрипции с использованием РЕгРНК, никазы Cas9 и RT *in vitro* на субстратах репортерных плазмид, а затем продукты реакции трансформировали в клетки дрожжей (S. cerevisiae) (фиг. 45A). Обнадеживает то, что когда плазмиды редактировали *in vitro* с помощью 3'-удлиненных РЕгРНК, кодирующих трансверсию T•A-в-A•T, которая корректирует преждевременный стоп-кодон, 37% трансформантов дрожжей экспрессировали белки GFP и mCherry (фиг. 38F, фиг. 45C). В соответствии с результатами на фиг. 38E и ФИГ. 44F, реакции редактирования, проведенные in vitro с 5'-удлиненными РЕгРНК, дали меньше двойных положительных колоний GFP и mCherry (9%), чем при использовании 3'-удлиненных РЕгРНК (фиг. 38F и фиг. 45D). Продуктивное редактирование также наблюдалось с использованием 3'-удлиненных РЕгРНК, которые вставляют один нуклеотид (15% дважды положительных трансформантов) или удаляют один нуклеотид (29% дважды положительных трансформантов) для исправления мутаций сдвига рамки считывания (фиг. 38F и фиг. 45E-45F). Секвенирование ДНК отредактированных плазмид, выделенных из двойных положительных колоний дрожжей, подтвердило, что закодированное трансверсионное редактирование произошло в нужном положении последовательности (фиг. 45G). Эти результаты демонстрируют, что репарация ДНК в эукариотических клетках может разрешать 3'-лоскуты ДНК, возникающие в

результате праймированного редактирования, для включения точных изменений ДНК, включая трансверсии, вставки и делеции.

Конструирование праймирующего редактора 1 (PE1)

[1093] В результате положительных, полученных *in vitro* и в дрожжах, была разработана основная система редактирования с минимальным количеством компонентов, способных редактировать геномную ДНК в клетках млекопитающих. Было высказано предположение, что 3'-удлиненные РЕгРНК (далее называемые просто РЕгРНК, фиг. 39A) и прямые слияния Cas9 H840A с обратной транскриптазой через гибкий линкер могут составлять функциональную двухкомпонентную систему праймированного редактирования. Клетки HEK293T (иммортализованная почка эмбриона человека) трансфицировали одной плазмидой, кодирующей слияние обратной транскриптазы M-MLV дикого типа с любым концом никазы Cas9 H840A, а также второй плазмидой, кодирующей РЕгРНК. Первоначальные попытки не привели к обнаруживаемому преобразованию T•A в A•T в целевом локусе HEK3.

[1094] Однако удлинение PBS в РЕгРНК до 8-15 оснований (фиг. 39A) привело к детектируемому редактированию T•A-в-A•T в целевом сайте HEK3 (фиг. 39B) с более высокой эффективностью для праймированного редактора. конструкции, в которых ОТ была слита с С-концом никазы Cas9 (максимальная конверсия T•A в A•T 3,7% при длине PBS в диапазоне от 8 до 15 нуклеотидов) по сравнению со слияниями N-концевой ОТ и никазы Cas9 (1,3% максимальное преобразование T•A в A•T) (фиг. 39B, все данные о клетках млекопитающих в данном документе представляют значения для всей обработанной популяции клеток, без отбора или сортировки, если не указано иное). Эти результаты позволяют предположить, что M-MLV RT дикого типа, слитая с Cas9, требует более длинных последовательностей PBS для редактирования генома в клетках человека по сравнению с тем, что требуется *in vitro* при использовании коммерческого варианта M-MLV RT, поставляемого в транс. Эта обратная транскриптаза M-MLV первого поколения дикого типа, слитая с С-концом никазы Cas9 H840A, была обозначена как PE1.

[1095] Была протестирована способность PE1 точно вводить точечные мутации трансверсии в четыре дополнительных сайта-мишени генома, специфических для РЕгРНК (фиг. 39C). Подобно редактированию локуса HEK3, эффективность этих геномных сайтов зависела от длины PBS, при этом максимальная эффективность редактирования находилась в диапазоне 0,7-5,5% (фиг. 39C). Индели из PE1 были низкими, в среднем 0,2±0,1% для пяти сайтов в условиях, которые максимизировали эффективность редактирования каждого сайта (фиг. 46A). PE1 также был способен устанавливать целевые вставки и делеции,

примерами которых являются однонуклеотидная делеция (эффективность 4,0%), однонуклеотидная вставка (9,7%) и трехнуклеотидная вставка (17%) в локусе HEK3 (фиг. 39C). Эти результаты подтверждают способность PE1 напрямую устанавливать целевые трансверсии, вставки и делеции, не требуя двухцепочечных разрывов ДНК или матриц ДНК.

[1096] Конструирование праймирующего редактора 2 (PE2)

[1097] В то время как PE1 может устанавливать различные модификации в нескольких локусах в клетках HEK293T, эффективность редактирования обычно была низкой (обычно ≤5%) (фиг. 39C). Было высказано предположение, что разработка обратной транскриптазы в PE1 может повысить эффективность синтеза ДНК в рамках уникальных конформационных ограничений комплекса праймированного редактирования, что приведет к более высоким результатам редактирования генома. Ранее сообщалось о мутациях M-MLV RT, которые повышают термостабильность фермента[150,151], процессивность[150] и аффинность субстрата гетеродуплекса ДНК:РНК[152] и инактивируют активность RNaseH[153] . Было сконструировано 19 вариантов PE1, содержащих различные мутации обратной транскриптазы, для оценки их эффективности праймированного редактирования в клетках человека.

[1098] Во-первых, была исследована серия вариантов RT M-MLV, которые ранее возникли в результате лабораторной эволюции на предмет их способности поддерживать обратную транскрипцию при повышенных температурах[150]. Последовательное введение трех из этих аминокислотных замен (D200N, L603W и T330P) в M-MLV RT, далее обозначаемое как M3, привело к среднему 6,8-кратному увеличению эффективности редактирования трансверсии и вставки в пяти геномных локусах в клетках HEK293T. по сравнению с PE1 (фиг. 47A-47S).

[1099] Затем, в комбинации с M3, были протестированы дополнительные мутации обратной транскриптазы, которые, как ранее было показано, усиливают связывание с комплексом матрица:PBS, ферментативную процессивность и термостабильность152. Среди 14 дополнительных проанализированных мутантов вариант с заменами T306K и W313F, в дополнение к мутациям M3, повысил эффективность редактирования дополнительно в 1,3–3,0 раза по сравнению с M3 для шести трансверсий или инсерционных редактирований в пяти геномных сайтах в клетках человека (фиг. 47A-47S). Этот пентамутант обратной транскриптазы M-MLV, включенный в конструкцию PE1 (Cas9 H840A–M-MLV RT (D200N L603W T330P T306K W313F)), далее обозначается как PE2.

[1100] PE2 устанавливает однонуклеотидные трансверсии, вставки и делеционные

мутации со значительно более высокой эффективностью, чем PE1 (фиг. 39C), и совместим с более короткими последовательностями PBS PEгРНК (фиг. 39C), что согласуется с повышенной способностью продуктивно задействовать временные комплексы геномная ДНК: PBS. В среднем, PE2 приводил к повышению эффективности точечной мутации праймированного редактирования в 1,6–5,1 раза по сравнению с PE1 (фиг. 39C), а в некоторых случаях значительно повышал эффективность редактирования до 46 раз (фиг. 47F и фиг. 47I). PE2 также более эффективно, чем PE1, осуществляет целенаправленные вставки и делеции, достигая целевой вставки эпитопной метки FLAG длиной 24 п.о. в локус HEK3 с эффективностью 4,5%, что в 15 раз превышает эффективность установки этой вставки с PE1 (фиг. 47D) и опосредовал делецию 1 п.о. в HEK3 с эффективностью 8,6%, что в 2,1 раза выше, чем у PE1 (фиг. 39C). Эти результаты подтверждают, что PE2 является более эффективным праймирующим редактором, чем PE1.

[1101] Оптимизация признаков PEгРНК

[1102] Взаимосвязь между конструкцией PEгРНК и эффективностью праймированного редактирования систематически исследовали в пяти геномных локусах в клетках HEK293T с PE2 (фиг. 39C). В целом, сайты праймирования с более низким содержанием GC требовали более длинных последовательностей PBS (EMX1 и RNF2, содержащие 40% и 30% GC, соответственно, в первых 10 нуклеотидах перед разрывом), тогда как сайты с более высоким содержанием GC поддерживали праймированное редактирование с помощью более короткие последовательности PBS (HEK4 и FANCF, содержащие 80% и 60% содержания GC, соответственно, в первых 10 нуклеотидах перед разрывом) (фиг. 39C), что соответствует энергетическим требованиям для гибридизации цепи ДНК с разрывом и PBS PEгРНК. Никакая длина PBS или уровень содержания GC не были строго предсказуемы для эффективности праймированного редактирования, и другие факторы, такие как вторичная структура в праймере ДНК или удлинение PEгРНК, также могут влиять на активность редактирования. Рекомендуется начинать с длины PBS ~ 13 нуклеотидов для типичной целевой последовательности и исследовать разные длины PBS, если последовательность отклоняется от содержания GC ~ 40-60%. При необходимости оптимальные последовательности PBS следует определять эмпирически.

[1103] Затем были изучены детерминанты производительности матричной части RT PEгРНК. PEгРНК с RT-матрицами длиной от 10 до 20 нуклеотидов систематически оценивали в пяти геномных целевых сайтах с использованием PE2 (фиг. 39D), а с более длинными RT-матрицами до 31 нуклеотида в трех геномных сайтах (фиг. 48A-48C). Как и в случае с длиной PBS, длину матрицы RT также можно варьировать, чтобы

максимизировать эффективность праймированного редактирования, хотя в целом многие матрицы RT длиной ≥10 нуклеотидов поддерживают более эффективное праймированное редактирование (фиг. 39D). Поскольку для некоторых целевых сайтов предпочтительны более длинные матрицы RT (>15 нуклеотидов) для достижения более высокой эффективности редактирования (FANCF, EMX1), в то время как для других локусов предпочтительны короткие матрицы RT (HEK3, HEK4) (РИС. 39D), рекомендуется использовать как короткие, так и длинные матрицы RT для тестирования при оптимизации РЕгРНК, начиная с ~10-16 нуклеотидов.

[1104] Важно отметить, что матрицы RT, которые помещают C в качестве нуклеотида рядом с концевой шпилькой каркаса сгРНК, обычно приводили к более низкой эффективности редактирования по сравнению с другими РЕгРНК с матрицами RT аналогичной длины (фиг. 48A-48C). Основываясь на структуре сгРНК, связанных с Cas9[148,149] , было высказано предположение, что присутствие C в качестве первого нуклеотида 3'-удлинения канонической сгРНК может разрушить укладку каркаса сгРНК путем спаривания с G81, нуклеотидом, который изначально образует pi-круг с Tyr 1356 в Cas9 и неканонической парой оснований с сгРНК A68. Поскольку многие длины RT-матриц поддерживают праймированное редактирование, рекомендуется выбирать РЕгРНК, в которых первое основание 3'-удлинения (последнее обратно транскрибируемое основание RT-матрицы) не является C.

[1105] Конструирование систем праймирующего редактора 3 (PE3 и PE3b)

[1106] Хотя PE2 может передавать генетическую информацию от РЕгРНК к локусу-мишени более эффективно, чем PE1, способ, которым клетка разрешает результирующую гетеродуплексную ДНК, созданную одной отредактированной цепью и одной неотредактированной цепью, определяет, будет ли редактирование долговечным. Предыдущая разработка редактирования оснований столкнулась с аналогичной проблемой, поскольку исходный продукт дезаминирования цитозина или аденина представляет собой гетеродуплексную ДНК, содержащую одну отредактированную и одну неотредактированную цепь. Чтобы повысить эффективность редактирования оснований, никаза Cas9 D10A использовалась для введения разрыва в нередактируемую цепь и для направления репарации ДНК в эту цепь с использованием отредактированной цепи в качестве матрицы[129,130,142]. Чтобы использовать этот принцип для повышения эффективности праймированного редактирования, была испытана аналогичная стратегия разрыва неотредактированной цепи с использованием никазы Cas9 H840A, уже присутствующей в PE2, и простой сгРНК для индукции предпочтительной замены

неотредактированной цепи клеткой (фиг. 40A). Поскольку редактируемая цепочка ДНК также была надрезана для инициации праймированного редактирования, различные местоположения запрограммированных сгРНК надрезов были протестированы на неотредактированной цепи, чтобы свести к минимуму образование двухцепочечных разрывов ДНК, которые приводят к вставкам инделям.

[1107] Эта стратегия PE3 была впервые протестирована на пяти геномных сайтах в клетках HEK293T путем скрининга сгРНК, которые индуцируют разрыв, расположенный на расстоянии от 14 до 116 оснований от сайта РЕгРНК-индуцированного разрыва, либо 5', либо 3' от PAM. В четырех из пяти протестированных сайтов разрыв неотредактированной цепи увеличивал количество продуктов праймированного редактирования, не содержащих инделей, по сравнению с системой PE2 в 1,5-4,2 раза, до 55% (фиг. 40B). В то время как оптимальное положение разрыва варьировалось в зависимости от геномного сайта, разрывы, расположенные на 3' от PAM (положительные расстояния на фиг. 40B) приблизительно в 40–90 п.о. от РЕгРНК-индуцированного разрыва, обычно приводили к благоприятному увеличению эффективности праймированного редактирования (в среднем 41%) без избыточного образования инделей (в среднем 6,8% инделей для сгРНК, что приводит к самой высокой эффективности редактирования для каждого из пяти протестированных сайтов) (фиг. 40B). Как и ожидалось, в некоторых местах размещение неотредактированного разрыва цепи в пределах 40 п.о. от РЕгРНК-индуцированного разрыва приводило к значительному увеличению образования инделей до 22% (фиг. 40B), предположительно из-за образования двойного разрыва цепи из-за надрезания обеих цепей близко друг к другу. В других сайтах, однако, разрывы на расстоянии всего 14 п.о. от РЕгРНК-индуцированного разрыва давали только 5% инделей (фиг. 40B), указывая на то, что локус-зависимые факторы контролируют превращение проксимальных двойных разрывов в двухцепочечные разрывы ДНК. В одном протестированном сайте (HEK4) разрывы комплементарных цепей либо не давали результата, либо приводили к уровням инделей, которые превышали эффективность редактирования (до 26%), даже при размещении на расстоянии> 70 п.о. от разрыва, индуцированного РЕгРНК, что согласуется с необычным склонностью отредактированной цепи в этом месте к разрыву клеткой или к неэффективному лигированию. Рекомендуется начинать с разрывов нередактируемых цепей приблизительно в 50 п.о. от РЕгРНК-опосредованного разрыва и тестировать альтернативные местоположения разрывов, если частоты инделей превышают допустимые уровни.

[1108] Эта модель того, как надрезы комплементарных цепей улучшают эффективность праймированного редактирования (фиг. 40A), предсказала, что надрезание

неотредактированной цепи только *после* разрешения лоскута редактируемой цепи может свести к минимуму наличие одновременных надрезов, уменьшая частоту двухцепочечных разрывов, которые приводят к инделям. Для достижения временного контроля над разрывом неотредактированной цепи были разработаны сгРНК со спейсерными последовательностями, которые соответствуют отредактированной цепи, но не исходному аллелю. При использовании этой стратегии, именуемой далее PE3b, несоответствия между спейсером и неотредактированным аллелем должны препятствовать разрыву сгРНК до тех пор, пока не произойдет событие редактирования на цепи PAM. Этот подход PE3b был протестирован с пятью различными изменениями в трех геномных сайтах в клетках HEK293T и сравнивался с результатами, достигнутыми с системами PE2 и PE3. Во всех случаях PE3b ассоциировался со значительно более низкими уровнями инделей по сравнению с PE3 (от 3,5 до 30 раз, в среднем в 12 раз ниже, или 0,85%), без какого-либо очевидного снижения общей эффективности редактирования по сравнению с PE3 (фиг. 40C). Таким образом, когда редактирование происходило внутри второго протоспейсера, система PE3b могла уменьшать количество инделей, при этом повышая эффективность редактирования по сравнению с PE2, часто до уровней, сходных с уровнями PE3 (фиг. 40C).

[1109] В совокупности эти результаты показали, что системы PE3 (Cas9-никаза-оптимизированная обратная транскриптаза + РЕгРНК + сгРНК) улучшали эффективность редактирования приблизительно в 3 раза по сравнению с PE2 (фиг. 40B-40C). PE3 сопровождался более широким диапазоном инделей, чем PE2, как и ожидалось, учитывая дополнительную активность PE3 в отношении разрывов. Использование PE3 рекомендуется при приоритете эффективности основного редактирования. Когда минимизация инделей имеет решающее значение, PE2 предлагает приблизительно в 10 раз более низкие частоты инделей. Когда можно использовать сгРНК, которая распознает установленное редактирование, для разрыва неотредактированной цепи, система PE3b может достигать уровней редактирования, подобных PE3, при значительном снижении образования инделей.

[1110] Чтобы продемонстрировать возможности и универсальность праймированного редактирования с помощью PE3, применили установку всех возможных одиночных нуклеотидных замен в положениях от +1 до +8 (считая первое основание 3′ РЕгРНК-индуцированного разрыва как положение +1) целевого участка HEK3 с использованием PE3 и РЕгРНК с матрицами RT из 10 нуклеотидов (фиг. 41A). В совокупности эти 24 различных редактирования охватывают все четыре переходные мутации и все восемь трансверсионных мутаций и происходят с эффективностью редактирования (без вставок) в среднем 33 ± 7,9% (в диапазоне от 14% до 48%) со средним

значением 7,5 ± 1,8% инделей.

[1111] Важно отметить, что матрицы RT большого расстояния также могут способствовать эффективному праймированному редактированию с помощью PE3. Например, при использовании PE3 с матрицей RT из 34 нуклеотидов точечные мутации были установлены в положениях +12, +14, +17, +20, +23, +24, +26, +30 и +33 (от 12 до 33 оснований из PEгРНК-индуцированного разрыва) в локусе HEK3 со средней эффективностью 36±8,7% и 8,6±2,0% инделей (фиг. 41B). Хотя редактирования за пределами положения +10 в других локусах не предпринималось, другие матрицы RT ≥30 нуклеотидов в трех альтернативных сайтах также поддерживают эффективное редактирование (фиг. 48A-C). Жизнеспособность длинных матриц RT позволила эффективно редактировать десятки нуклеотидов из начального участка разрыва. Поскольку NGG PAM на любой цепи ДНК происходит в среднем каждые ~ 8 п.о., что намного меньше максимальных расстояний между редактированием и PAM, которые поддерживают эффективное праймированное редактирование, праймированное редактирование существенно не ограничено доступностью соседней последовательности PAM, по сравнению с другими методами точного редактирования генома[125,142,154]. Учитывая предполагаемую взаимосвязь между вторичной структурой РНК и эффективностью праймированного редактирования, при конструировании PEгРНК для редактирования длинных диапазонов целесообразно тестировать матрицы RT различной длины и, при необходимости, составы последовательностей (например, синонимичные кодоны) для оптимизации эффективности редактирования.

[1112] Для дальнейшего тестирования объема и ограничений системы PE3 для введения точечных мутаций транзиции и трансверсии были протестированы 72 дополнительных редактирования, охватывающих все 12 возможных типов точечных мутаций в шести дополнительных геномных сайтах-мишенях (фиг. 41C-41H). В целом эффективность редактирования без инделей составила в среднем 25±14%, а образование инделей — в среднем 8,3±7,5%. Поскольку матрица RT PEгРНК включала последовательность PAM, праймированное редактирование могло вызвать изменения в последовательности PAM. В этих случаях наблюдали более высокую 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). Это увеличение эффективности и снижение образования инделей для редактирования PAM может быть результатом неспособности никазы Cas9 повторно связываться и разрывать отредактированную цепь до восстановления комплементарной цепи. Поскольку основное редактирование поддерживает комбинированное редактирование без явной потери

эффективности редактирования, рекомендуется по возможности редактировать PAM в дополнение к другим желаемым изменениям.

[1113] Затем было выполнено 14 целевых небольших вставок и 14 целевых небольших делеций в семи геномных сайтах с использованием PE3 (фиг. 41I). Целевые вставки из 1 п.о. протекали со средней эффективностью 32±9,8%, в то время как вставки из 3 п.о. устанавливались со средней эффективностью 39±16%. Целенаправленные делеции 1 и 3 п.о. также были эффективны со средним выходом 29±14% и 32±11%, соответственно. Образование инделей (за пределами целевой вставки или делеции) в среднем составило 6,8 ± 5,4%. Поскольку вставки и делеции, введенные между положениями +1 и +6, изменяют положение или структуру PAM, было высказано предположение, что редактирование вставки и делеции в этом диапазоне обычно более эффективны из-за неспособности никазы Cas9 повторно связываться и разрывать отредактированную цепь ДНК перед восстановлением комплементарной цепи, подобно точечным мутациям, которые редактируют PAM.

[1114] PE3 также тестировали на его способность опосредовать более крупные точные делеции размером от 5 п.о. до 80 п.о. в сайте HEK3 (фиг. 41J). Очень высокая эффективность редактирования (от 52 до 78%) наблюдалась для делеций 5, 10 и 15 п.о. при использовании 13-нуклеотидного PBS и матрицы RT, которая содержала 29, 24 или 19 п.о. гомологии с целевым локусом, соответственно. Использование матрицы RT из 26 нуклеотидов поддерживало более крупную делецию 25 п.о. с эффективностью 72 ± 4,2%, в то время как матрица RT из 20 нуклеотидов обеспечивала делецию 80 п.о. с эффективностью 52 ± 3,8%. Эти целевые делеции сопровождались частотами инделей, составляющими в среднем 11±4,8% (фиг. 41J).

[1115] Наконец, была проверена способность PE3 опосредовать 12 комбинаций множественных редактирований в одном и том же целевом локусе, состоящих из вставок и делеций, вставок и точечных мутаций, делеций и точечных мутаций или двух точечных мутаций в трех геномных сайтах. Эти комбинированные редактирования были очень эффективными, составив в среднем 55% целевого редактирования с 6,4% инделей (фиг. 41K), и продемонстрировали способность праймированного редактирования создавать комбинации точных вставок, делеций и точечных мутаций в отдельных целевых сайтах с высокой эффективностью и низкими частотами инделей.

[1116] Вместе примеры на фиг. 41A-41K представляют 156 различных транзиций, трансверсий, вставок, делеций и комбинированных редактирований в семи геномных локусах человека. Эти результаты подтверждают универсальность, точность и гибкость праймированного редактирования.

[1117] Праймированное редактирование по сравнение с редактированием оснований

[1118] Редакторы цитидиновых оснований (CBE) и редакторы адениновых оснований (ABE) текущего поколения могут устанавливать переходные мутации C•G-в-T•A и переходные мутации A•T-в-G•C с высокой эффективностью и низкими инделями[129,130,142]. Применение редактирования оснований может быть ограничено наличием нескольких цитидиновых или адениновых оснований в пределах окна действия редактирования основания (обычно шириной ~5 п.о.), что дает нежелательные сопутствующие редактирования[129,130,142,155] или отсутствие расположения PAM в пределах 15 ± 2 нуклеотида от целевого нуклеотида[142,156]. Ожидалось, что праймированное редактирование может быть особенно полезным для точной установки переходных мутаций без сопутствующего редактирования или когда отсутствие подходящим образом расположенных PAM исключает благоприятное расположение целевого нуклеотида в пределах окна активности CBE или ABE.

[1119] Праймированное редактирование и редактирование основания цитозина сравнивали путем редактирования трех геномных локусов, которые содержат несколько целевых цитидинов, в окне редактирования канонического основания (положения протоспейсера 4-8, считая PAM как положения 21-23) с использованием оптимизированных CBE[157] без активности никазы (BE2max) или с никазной активностью (BE4max), или с использованием аналогичных систем праймированного редактирования PE2 и PE3. Среди девяти общих цитозинов-мишеней в окнах редактирования оснований трех сайтов BE4max давал в 2,2 раза более высокую среднюю общую конверсию C•G-в-T•A, чем PE3 для оснований в центре окна редактирования оснований (положения протоспейсера 5-7, фиг. 42A). Аналогично, BE2max без разрывов превосходил PE2 в среднем в 1,4 раза на этих хорошо расположенных основаниях (фиг. 42A). Однако PE3 превзошел BE4max в 2,7 раза, а PE2 превзошел BE2max в 2,0 раза для цитозинов за пределами центра основного окна редактирования (среднее редактирование 40 ± 17% для PE3 против 15 ± 18% для BE4max и 22±11% для PE2 против 11±13% для BE2max). В целом частота инделей для PE2 была очень низкой (в среднем 0,86 ± 0,47%), а для PE3 была аналогична или немного выше, чем у BE4max (диапазон BE4max: от 2,5% до 14%, диапазон PE3: от 2,5% до 21%) (фиг. 42B).

[1120] При сравнении эффективности редактирования основания с праймированным редактированием для установки точных редактирований C•G-в-T•A (без какого-либо сопутствующего редактирования) эффективность праймированного редактирования значительно превышала эффективность редактирования оснований на вышеуказанных сайтах, которые, как и большинство сайтов геномной ДНК содержат множество цитозинов в пределах окна редактирования оснований ~5 п.о. (фиг. 42C). В этих

сайтах, таких как EMX1, который содержит цитозины в положениях протоспейсера C5, C6 и C7, BE4max генерировал несколько продуктов, содержащих только конверсию одной целевой пары оснований без дополнительных редактирований. Напротив, праймированное редактирование в этом сайте можно использовать для выборочной установки редактирования C•G-в-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 редактирования одного основания или двух оснований (то есть редактирования, которые не изменяют какие-либо другие близлежащие основания) были намного эффективнее с PE3 или PE2, чем с BE4max или BE2 соответственно, в то время как редактирование трех оснований C•G-в-T•A было более эффективным с BE4max (фиг. 42C), отражая склонность редакторов оснований редактировать все целевые основания в окне активности. В совокупности эти результаты показывают, что редакторы оснований цитозина могут обеспечивать более высокий уровень редактирования в оптимально расположенных целевых основаниях, чем PE2 или PE3, но праймированное редактирование может превзойти редактирование оснований в неоптимально расположенных целевых основаниях и может редактироваться с гораздо большей точностью с множеством редактируемых оснований.

[1121] Редактирование A•T в G•C сравнивали в двух геномных локусах с помощью оптимизированного ABE без разрыва (ABEmax[152] с dCas9 вместо никазы Cas9, далее именуемой ABEdmax) по сравнению с PE2 и с помощью оптимизированного редактора оснований аденина с разрывом ABEmax по сравнению с PE3. На сайте, который содержит два целевых аденина в окне редактирования оснований (HEK3), ABE были более эффективны, чем PE2 или PE3, для преобразования A5, но PE3 был более эффективным для преобразования A8, который находится на краю окна редактирования ABEmax (фиг. 42D). При сравнении эффективности точного редактирования, при котором преобразуется только один аденин, PE3 превосходит ABEmax как при A5, так и при A8 (фиг. 42E). В целом, ABE производили гораздо меньше инделей в HEK3, чем праймирующие редакторы (0,19 ± 0,02% для ABEdmax по сравнению с 1,5 ± 0,46% для PE2 и 0,53 ± 0,16% для ABEmax по сравнению с 11 ± 2,3% для PE3, фиг. 42F). В FANCF, в котором в окне редактирования оснований присутствует только один A, ABE2 и ABEmax превзошли свои основные аналоги редактирования по общему преобразованию целевых пар оснований в 1,8–2,9 раза, причем практически все отредактированные продукты как от редактирования оснований, так и от праймированного редактирования, содержали только точное редактирование (фиг. 42D-42E). Как и в случае с сайтом HEK3, ABE продуцировали гораздо меньше инделей на сайте FANCF (фиг. 42F).

[1122] В совокупности эти результаты показывают, что редактирование оснований

и праймированное редактирование предлагают дополнительные сильные и слабые стороны для создания целевых мутаций транзиции. В случаях, когда в окне редактирования оснований присутствует один целевой нуклеотид, или когда сопутствующее редактирование допустимо, существующие редакторы оснований обычно более эффективны и приводят к меньшему количеству инделей, чем праймирующие редакторы. Когда присутствует несколько цитозинов или аденинов и сопутствующее редактирование нежелательно, или когда целевые основания плохо расположены для редактирования оснований по сравнению с доступными PAM, праймирующие редакторы имеют существенные преимущества.

[1123] Нецелевое праймированное редактирование

[1124] Чтобы привести к продуктивному редактированию, праймированное редактирование требует комплементарности локуса-мишени: спейсера PEгРНК для связывания домена Cas9, комплементарности локуса-мишени: PEгРНК-PBS для инициации обратной транскрипции, праймированной PEгРНК, и комплементарности локуса-мишени: продукта обратной транскриптазы для разрешения лоскута. Было высказано предположение, что эти три различных требования к гибридизации ДНК могут свести к минимуму нецелевое праймированное редактирование по сравнению с другими методами редактирования генома. Чтобы проверить эту возможность, клетки HEK293T обрабатывали PE3 или PE2 и 16 общими PEгРНК, предназначенными для нацеливания на четыре целевых геномных локуса, с Cas9 и четырьмя соответствующими сгРНК, нацеленными на одни и те же протоспейсеры, или с Cas9 и теми же 16 PEгРНК. Эти четыре локуса-мишени были выбраны, потому что каждый имеет по меньшей мере четыре хорошо охарактеризованных нецелевых сайта, для которых известно, что Cas9 и соответствующая целевая сгРНК в клетках HEK293T вызывают существенную нецелевую модификацию[118,159]. После лечения были секвенированы четыре целевых локуса и четыре основных известных нецелевых сайта Cas9 для каждого целевого спейсера, всего 16 нецелевых сайтов (таблица 1).

[1125] В соответствии с предыдущими исследованиями[118], Cas9 и четыре нацеленные сгРНК модифицировали все 16 ранее описанных нецелевых локусов (фиг. 42G). Эффективность нецелевой модификации Cas9 среди четырех нецелевых сайтов для локуса-мишени HEK3 в среднем составила 16%. Cas9 и сгРНК, нацеленные на HEK4, привели в среднем к 60% модификации четырех проверенных известных нецелевых сайтов. Аналогично, нецелевые сайты для EMX1 и FANCF модифицировались с помощью Cas9:сгРНК со средней частотой 48% и 4,3%, соответственно (фиг. 42G). Было отмечено, что PEгРНК с нуклеазой Cas9 модифицировали сайты-мишени с аналогичной (в 1-1,5 раза

ниже) эффективностью в среднем по сравнению с сгРНК, в то время как РЕгРНК с нуклеазой Cas9 модифицировали сайты-мишени приблизительно с в 4 раза более низкой средней эффективностью, чем сгРНК.

[1126] Неожиданно, PE3 или PE2 с теми же 16 протестированными РЕгРНК, содержащими эти четыре целевых спейсера, приводили к гораздо меньшему нецелевому редактированию (фиг. 42H). Из 16 сайтов, которые, как известно, подвергались нецелевому редактированию с помощью Cas9+сгРНК, PE3+РЕгРНКs или PE2+РЕгРНК приводили к обнаруживаемому нецелевому праймированному редактированию только в 3 из 16 нецелевых сайтов, и только 1 из 16 демонстрировал эффективность нецелевого редактирования ≥1% (фиг. 42H). Среднее нецелевое праймированное редактирование для РЕгРНК, нацеленных на HEK3, HEK4, EMX1 и FANCF, в этих 16 известных нецелевых сайтах Cas9 составляло <0,1%, <2,2±5,2%, <0,1% и <0,13±0,11%, соответственно. (фиг. 42H). Примечательно, что в нецелевом сайте 3 HEK4, который Cas9+РЕгРНК1 редактирует с эффективностью 97 %, PE2+РЕгРНК1 приводит лишь к 0,7 % нецелевому редактированию, несмотря на то, что используется одна и та же спейсерная последовательность, демонстрируя, как два дополнительных события гибридизации ДНК, необходимые для праймированной редактирование по сравнению с редактированием Cas9, могут значительно уменьшить нецелевое редактирование. Взятые вместе, эти результаты предполагают, что PE3 и РЕгРНК индуцируют гораздо меньшее нецелевое редактирование ДНК в клетках человека, чем Cas9 и сгРНК, которые нацелены на одни и те же протоспейсеры.

[1127] Обратная транскрипция 3'-удлиненных РЕгРНК, в принципе, может происходить в каркас гидовой РНК. Если полученный 3'-лоскут, несмотря на отсутствие комплементарности на его 3'-конце с неотредактированной цепью ДНК, включается в целевой локус, результатом является вставка нуклеотидов каркаса РЕгРНК, что способствует частоте инделей. Проанализировали данные секвенирования из 66 экспериментов по редактированию, опосредованному PE3, в четырех локусах в клетках HEK293T и наблюдали вставку каркаса РЕгРНК с низкой частотой, в среднем 1,7 ± 1,5% от общего количества вставок любого количества нуклеотидов каркаса РЕгРНК (фиг. 56A-56D). Предполагается, что недоступность каркаса гидовой РНК для обратной транскриптазы из-за связывания домена Cas9, а также клеточного вырезания во время разрешения лоскута несовпадающего 3'-конца 3'-лоскута, возникающая в результате обратной транскрипции каркаса РЕгРНК, сводит к минимуму продукты, которые включают нуклеотиды каркаса РЕгРНК. Хотя такие события являются редкими, будущие усилия по созданию РЕгРНК или белков праймирующего редактора, которые минимизируют

включение каркаса PEгРНК, могут еще больше снизить частоту инделей.

[1128] Деаминазы в некоторых редакторах основания могут действовать независимым от Cas9 способом, что приводит к низкого уровня, но широко распространенному нецелевому редактированию ДНК среди CBE первого поколения (но не ABE)[160–162] и нецелевому редактированию РНК среди CBE первого поколения и ABE[163-165] , хотя более новые варианты CBE и ABE со сконструированными дезаминазами значительно снижают Cas9-независимое нецелевое редактирование ДНК и РНК[163-165]. В праймирующих редакторах отсутствуют ферменты модификации оснований, такие как деаминазы, и, следовательно, они не обладают присущей им способностью модифицировать основания ДНК или РНК независимым от Cas9 способом.

[1129] В то время как домен обратной транскриптазы в праймирующих редакторах, в принципе, мог процессировать надлежащим образом праймированные матрицы РНК или ДНК в клетках, было отмечено, что ретротранспозоны, такие как в семействе LINE-1[166], эндогенные ретровирусы[167,168] и теломераза человека, все обеспечивали активные эндогенные обратные транскриптазы человек. Их естественное присутствие в клетках человека предполагает, что активность обратной транскриптазы сама по себе не является существенно токсичной. Действительно, не наблюдалось никаких PE3-зависимых различий в жизнеспособности клеток HEK293T по сравнению с контрольными клетками, экспрессирующими dCas9, Cas9 H840A никазу или PE2 с мутациями R110S+K103L (PE2-dRT), которые инактивируют обратную транскриптазу[169] (фиг. 49A-49B).

[1130] Несмотря на приведенные выше данные и анализы, необходимы дополнительные исследования для оценки нецелевого праймированного редактирования объективным, полногеномным образом, а также для характеристики степени, в которой варианты обратной транскриптазы в праймирующих редакторах или промежуточных продуктах праймированного редактирования могут воздействовать на клетки.

[1131] Праймированное редактирование патогенных мутаций трансверсии, вставки и делеции в клетках человека

[1132] Была проверена способность PE3 непосредственно устанавливать или корректировать в клетках человека трансверсию, небольшие вставки и небольшие делеционные мутации, вызывающие генетические заболевания. Серповидноклеточная анемия чаще всего вызывается мутацией трансверсии A•T в T•A в HBB, что приводит к мутации Glu6→Val в бета-глобине. Обработка гемопоэтических стволовых клеток *ex vivo* нуклеазой Cas9 и матрицей донорной ДНК для HDR с последующим обогащением отредактированных клеток, трансплантацией и приживлением является многообещающей

потенциальной стратегией лечения серповидно-клеточной анемии[170]. Однако этот подход по-прежнему генерирует много побочных продуктов, содержащих индели, в дополнение к правильно отредактированному аллелю HBB[170-171]. Хотя редакторы оснований обычно производят гораздо меньше инделей, в настоящее время они не могут произвести мутацию трансверсии T•A в A•T, необходимую для непосредственного восстановления нормальной последовательности HBB.

[1133] PE3 использовали для установки мутации HBB E6V в клетках HEK293T с эффективностью 44% и 4,8% инделей (фиг. 43A). Из смеси клеток, обработанных PE3, мы выделили шесть клеточных линий HBB293T, гомозиготных (триплоидных) по HBB E6V аллелю (фиг. 53A-53D), демонстрируя способность праймированного редактирования генерировать клеточные линии человека с патогенными мутациями. Чтобы скорректировать аллель HBB E6V до HBB дикого типа, гомозиготные клетки HBB E6V HEK293T обрабатывали с помощью PE3 и PEгРНК, запрограммированными на прямое возвращение мутации HBB E6V в HBB дикого типа. Всего было протестировано 14 конструкций PEгРНК. Через три дня секвенирование ДНК показало, что все 14 PEгРНК в сочетании с PE3 обеспечивают эффективную коррекцию HBB E6V в HBB дикого типа (≥26 % HBB дикого типа без инделей), а уровни инделей в среднем составляли 2,8±0,70% (фиг. 50A). Наилучшая PEгРНК приводила к 52% коррекции HBB E6V до дикого типа с 2,4% инделей (фиг. 43A). Введение молчащей мутации, которая модифицирует PAM, распознаваемый PEгРНК, слегка улучшила эффективность редактирования и чистоту продукта, корректирование до 58% с инделями 1,4% (фиг. 43A). Эти результаты показывают, что праймированное редактирование может установить и исправить патогенную мутацию точки трансверсии в линии клеток человека с высокой эффективностью и минимальным количеством побочных продуктов.

[1134] Болезнь Тея-Сакса чаще всего вызывается вставкой 4 п.о. в ген HEXA (HEXA 1278+TATC)[136]. PE3 использовали для установки этой вставки из 4 п.о. в клетки HEK293T с эффективностью 31% и 0,8% инделей (фиг. 43B), и выделяли две клеточные линии HEK293T, гомозиготные по аллелю HEXA 1278+TATC (фиг. 53A-53D). Эти клетки использовали для тестирования 43 PEгРНК и трех вносящих разрыв сгРНК с системами PE3 или PE3b для коррекции патогенной вставки в HEXA (фиг. 50B) либо путем полной реверсии к аллелю дикого типа, либо путем смещенной делеции 4 п.о., что разрушает PAM и устанавливает молчащую мутацию. Девятнадцать из 43 протестированных PEгРНК привели к редактированию ≥20%. Идеальная коррекция HEXA дикого типа с помощью PE3 или PE3b и лучшей PEгРНК наблюдали с одинаковой средней эффективностью (30% для PE3 против 33% для PE3b), но система PE3b сопровождалась в 5,3 раза меньшим

количеством инделей (1,7% для PE3 по сравнению с 0,32% для PE3b) (фиг. 43B и фиг. 50B). Эти результаты демонстрируют способность праймированного редактирования делать точные небольшие вставки и делеции, которые эффективно устанавливают или корректируют патогенный аллель в клетках млекопитающих с минимальным количеством побочных продуктов.

[1135] Наконец, была протестирована установка защитного SNP в PRNP, ген, кодирующий прионовый белок человека (PrP). Неправильный фолдинг PrP вызывает прогрессирующую и фатальную нейродегенеративную прионную болезнь, которое может возникать спонтанно, в результате наследственных доминантных мутаций в гене PRNP или при воздействии неправильно свернутого PrP[172]. Встречающийся в природе мутантный аллель PRNP G127V придает устойчивость к прионным заболеваниям у людей [138] и мышей[173]. PE3 использовали для установки G127V в аллель PRNP человека в клетках HEK293T, что требует трансверсии G•C-в-T•A. Четыре PEгРНК и три вносящих разрыв сгРНК оценивали с помощью системы PE3. После трех дней воздействия наиболее эффективных PE3 и PEгРНК секвенирование ДНК показало эффективность установки мутации G127V 53±11% и уровни инделей 1,7±0,7% (фиг. 43C). Взятые вместе, эти результаты подтверждают способность праймированного редактирования в клетках человека устанавливать или корректировать мутацию трансверсию, вставку или делецию, которые вызывают или придают устойчивость к болезням, эффективно и с минимальным количеством побочных продуктов.

[1136] Праймированное редактирование в различных линиях клеток человека и первичных нейронах мыши

[1137] Затем праймированное редактирование тестировали на его способность редактировать эндогенные сайты в трех дополнительных линиях клеток человека. В клетках K562 (лейкемический костный мозг) PE3 использовали для выполнения трансверсионных изменений в сайтах HEK3, EMX1 и FANCF, а также для вставки 18-п.о. метки 6xHis в HEK3. Средняя эффективность редактирования 15-30% наблюдалась для каждого из этих четырех PE3-опосредованных редактирований с вставками в среднем 0,85-2,2% (фиг. 43A). В клетках U2OS (остеосаркома) были установлены трансверсионные мутации в HEK3 и FANCF, а также вставка 3 п.о. и вставка метки 6xHis в HEK3 с эффективностью редактирования 7,9-22%, что превышало образование инделей в 10-76 раз (фиг. 43A). Наконец, в клетках HeLa (рак шейки матки) выполняли вставку 3 п.о. в HEK3 со средней эффективностью 12% и 1,3% инделей (фиг. 43A). В совокупности эти данные показывают, что несколько клеточных линий помимо клеток HEK293T поддерживают

праймированное редактирование, хотя эффективность редактирования зависит от типа клеток и, как правило, менее эффективна, чем в клетках HEK293T. Отношения редактирование:индель оставались высокими во всех протестированных линиях клеток человека.

[1138] Чтобы определить, возможно ли праймированное редактирование в постмитотических, терминально дифференцированных первичных клетках, первичные кортикальные нейроны, собранные у мышей E18.5, трансдуцировали с помощью двойной системы лентивирусной доставки расщепленного PE3, в которой сплайсинг расщепленного интеина[203] восстанавливает белок PE2 из N- концевой и C-концевой половины, каждая из которых доставляется из отдельного вируса. Чтобы ограничить редактирование постмитотических нейронов, промотор synapsin человека, который является высокоспецифичным для зрелых нейронов[204], использовали для управления экспрессией обоих белковых компонентов PE2. GFP был слит через саморасщепляющийся пептид P2A[205] с N-концевой половиной PE2. Ядра нейронов выделяли через две недели после двойной вирусной трансдукции и секвенировали непосредственно или сортировали по экспрессии GFP перед секвенированием. Наблюдали среднее праймированное редактирование 7,1±1,2% для установки трансверсии в локусе DNMT1 со средним числом инделей 0,58±0,14% в отсортированных ядрах (фиг. 43D). Нуклеаза Cas9 в той же двойной лентивирусной системе с расщеплением белка приводила к 31 ± 5,5% инделей среди отсортированных ядер кортикальных нейронов (фиг. 43D). Эти данные показывают, что постмитотические, терминально дифференцированные первичные клетки могут поддерживать праймированное редактирование и, таким образом, установили, что праймированное редактирование не требует репликации клеток.

[1139] Праймированное редактирование по сравнению с Cas9-инициированным HDR

[1140] Эффективность PE3 сравнивали с эффективностью оптимизированным Cas9-инициированным HDR[128,125] в митотических клеточных линиях, которые поддерживают HDR[128]. Клетки HEK293T, HeLa, K562 и U2OS обрабатывали нуклеазой Cas9, сгРНК и матрицей донорных олигонуклеотидов оцДНК, разработанной для установки различных трансверсий и вставок (фиг. 43E-43G и фиг. 51A-51F). HDR, инициированное Cas9, во всех случаях успешно устанавливало желаемое редактирование, но с гораздо более высоким уровнем побочных продуктов (преимущественно вставок), как и ожидалось от процедур, вызывающих двухцепочечные разрывы. Используя PE3 в клетках HEK293T, установка и коррекция *HBB* E6V происходили со средней эффективностью редактирования 42% и 58%

со средними инделями 2,6% и 1,4%, соответственно (фиг. 43E и фиг. 43G). Напротив, те же правки с нуклеазой Cas9 и матрицей HDR приводили к средней эффективности редактирования 5,2% и 6,7% со средней частотой инделей 79% и 51% (фиг. 43E и фиг. 43G). Точно так же PE3 установил *PRNP* G127V с эффективностью 53% и 1,7% инделей, тогда как инициированное Cas9 HDR установило эту мутацию с эффективностью 6,9% и 53% инделей (фиг. 43E и фиг. 43G). Таким образом, соотношение редактирование:индели для установки *HBB* E6V, коррекции HBB E6V и установки *PRNP* G127V в среднем было в 270 раз выше для PE3, чем для Cas9-инициированного HDR.

[1141] Сравнение между PE3 и HDR в линиях клеток человека, отличных от HEK293T, показало аналогичные результаты, хотя и с более низкой эффективностью редактирования PE3. Например, в клетках K562 PE3-опосредованная вставка 3 п.о. в HEK3 происходила с эффективностью 25% и 2,8% инделей по сравнению с 17% редактированием и 72% инделями для HDR, индуцированного Cas9, что дает 40-кратное преимущество в соотношении редактирование:индели в пользу PE3 (фиг. 43F-43G). В клетках U2OS PE3 выполнил эту вставку из 3 п.о. с эффективностью 22% и 2,2% инделей, в то время как инициированный Cas9 HDR привел к 15% редактированию с 74% инделей, т.е. в 49 раз меньшее соотношение редактирование:индели (фиг. 43F-43G). В клетках HeLa PE3 произвел эту вставку с эффективностью 12% и инделями 1,3% по сравнению с редактированием 3,0% и инделями 69% для инициированного Cas9 HDR, что представляет собой 210-кратную разницу в отношении редактирование:индели (фиг. 43F-43G). В совокупности эти данные указывают на то, что HDR обычно приводит к аналогичной или более низкой эффективности редактирования и гораздо более высоким инделям, чем PE3, в четырех протестированных клеточных линиях (фиг. 51A-51F).

[1142] Обсуждение и будущие направления

[1143] Возможность вставлять последовательности ДНК с точностью до одного нуклеотида является особенно важной способностью редактирования. Например, PE3 использовали для точной вставки в локус HEK3 в клетках HEK293T метки His_6 (18 п.о., средняя эффективность 65%), эпитопной метки FLAG (24 п.о., средняя эффективность 18%) и удлиненного сайта *LoxP* (44 п.о., средняя эффективность 23%), который является нативным субстратом для рекомбиназы Cre. Средние индели для этих примеров находились в диапазоне от 3,0% до 5,9% (фиг. 43H). Предполагается, что многие биотехнологические, синтетические биологические и терапевтические применения возникнут из-за способности эффективно и точно вводить новые последовательности ДНК в представляющие интерес целевые участки в живых клетках.

[1144] В совокупности, описанные в настоящем документе эксперименты по праймированному редактированию установили 18 вставок длиной до 44 п.о., 22 делеции длиной до 80 п.о., 113 точечных мутаций, включая 77 трансверсий, и 18 комбинированных правок в 12 эндогенных локусах геномов человека и мыши в местах от 3 п.о. против хода транскрипции до 29 п.о. по ходу транскрипции от начала PAM без явных разрывов двухцепочечной ДНК. Эти результаты подтверждают, что праймированное редактирование является чрезвычайно универсальным способом редактирования генома. Поскольку подавляющее большинство (85-99%) вставок, делеций, инделей и дупликаций в ClinVar имеют размер ≤30 п.о. (фиг. 52A-52D), в принципе праймированное редактирование может исправить до ~89% из 75122 известных в настоящее время патогенных генетических вариантов человека в ClinVar (транзиции, трансверсии, вставки, делеции, индели и дупликации на фиг. 38A) с дополнительным потенциалом для облегчения заболеваний, вызванных увеличением или потерей числа копий.

[1145] Важно отметить, что для любого желаемого редактирования гибкость праймированного редактирования предлагает множество возможных вариантов расположения РЕгРНК-индуцированных разрывов, положений сгРНК-индуцированных вторых разрывов, длин PBS, длин матрицы RT и того, какую цепь редактировать первой, как подробно продемонстрировано в настоящем документе. Эта гибкость, которая контрастирует с более ограниченными вариантами, обычно доступными для других способов точного редактирования генома[125,142,154], позволяет оптимизировать эффективность редактирования, чистоту продукта, специфичность ДНК или другие параметры в соответствии с потребностями данного применения, как показано на фиг. 50A-50B, в которых тестирование 14 и 43 РЕгРНК, охватывающих ряд основных стратегий редактирования, оптимизировало коррекцию патогенных аллелей HBB и HEXA, соответственно.

[1146] Требуется много дополнительных исследований, чтобы лучше понять и улучшить основное редактирование. Могут потребоваться дополнительные модификации систем праймированного редактора для расширения их совместимости с другими типами клеток, такими как постмитотические клетки. Взаимодействие праймированного редактирования с вирусными и невирусными стратегиями доставки *in vitro* и *in vivo* необходимо для полного изучения потенциала праймированного редактирования для обеспечения широкого спектра применений, включая изучение и лечение генетических заболеваний. Однако, обеспечивая высокоточные целевые транзиции, трансверсии, небольшие вставки и небольшие делеции в геномах клеток млекопитающих, не требуя двухцепочечных разрывов или HDR, праймированное редактирование обеспечивает новую

возможность «поиска и замены», которая существенно расширяет область применения. редактирования генома.

Способы

[1147] Общие способы

[1148] 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 сгРНК, конструировали путем лигирования подвергнутых отжигу нуклеотидов в BsmBI-расщепляемый акцепторный вектор. Плазмиды, экспрессирующие РЕгРНК, конструировали на основе сборки согласно методу Гибсона или сборки Golden Gate с использованием модифицированной акцепторной плазмиды (см. дополнительное описание "сборки Golden Gate"). Последовательности используемых в согласно настоящему изобретению конструкций сгРНК и РЕгРНК 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 Pure Yield (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).

[1149] In vitro биологические анализы

[1150] РЕгРНК и сгРНК 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 пикированию субстрата и трех нуклеотидов (Cy5-AVA023, AVA025 и AVA026, 1:1.1:1.1) для предварительно подвергнутого пикированию

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 KC1; 5% глицерин; 0,2 мМ ЭДТА, pH 8,0; 3 мМ MgC12; 0,5 мМ dNTP микс; 5 мМ ДТТ). dCas9 или Cas9 H840A (конечная концентрация 5 мкМ) и сгРНК или РЕгРНК (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 К (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.

[1151] Проводили анализы сдвига электрофоретической подвижности в 1x связывающем буфере (1x буфер расщепления +10 мкг/мл гепарин) с использованием предварительно инкубированных комплексов dCas9: сгРНК или dCas9:РЕгРНК (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.

[1152] 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 праймированного удлинения с использованием фрагмента Klenow (New England Biolabs) с применением праймера AVA134 (A-присоединенные продукты) или AVA135 (G-присоединенные продукты) (Таблицы 2A-2C). Удлинение амплифицировали с помощью ПЦР в рамках 10 циклов с использованием

праймеров AVAHO и AVA122, затем секвенировали с AVA037 с использованием способа Сангера (Таблицы 2A-2C).

[1153] Анализы на основе флуоресцентного репортера дрожжей

[1154] Двойные флуоресцентные репортерные плазмиды, содержащие находящийся в рамке считывания стоп-кодон, +1 сдвиг рамки считывания, или -1 сдвиг рамки считывания, подвергали *in vitro* реакциям праймированного редактирования с использованием 5'-удлиненной PEгРНК или 3'-удлиненной PEгРНК, 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 (8C(глюкоза), L-). Сигналы флуоресценции GFP и mCherry визуализировали из колоний с использованием биомолекулярного сканера Typhoon FLA 7000.

[1155] Общие условия культивирования клеток млекопитающих

[1156] 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% CO2. Линии клеток были проверены их соответствующими поставщиками и протестированы на отсутствие микоплазмы.

[1157] Протокол трансфекции культуры ткани HEK293T и подготовка геномной ДНК

[1158] 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 PEгРНК, 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 lx 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 К (ThermoFisher Scientific)) напрямую в каждую лунку планшета для культивирования ткани. Смесь геномной ДНК инкубировали при 37 °С в течение от 1 до 2 часов с последующим этапом в виде деактивации ферментов при 80 °С 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 сгРНК, 50 нг (1,51 пмоль) 100 нуклеотидной 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.

[1159] Высокопроизводительное секвенирование ДНК образцов геномной ДНК

[1160] 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 °С в течение 2 минут, затем 30 циклов из [98 °С в течение 10 секунд, 61 °С в течение 20 секунд, и 72 °С в течение 30 секунд], и далее конечная стадия при 72 °С 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 °С в течение 2 минут, затем 12 циклов из [98 °С в течение 10 секунд, 61 °С в течение 20 секунд, и 72 °С в течение 30 секунд], и далее конечная стадия при 72 °С 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.

[1161] 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 «discardindelreads» («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 (e flag) 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.

[1162] Нуклеофекции клеток U2OS, K562 и HeLa

[1163] 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 праймирующего редактора, 200 нг плазмиды экспрессии PEгРНК, и 83 нг пикирующей 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 sgPHK, и 200 пмоль (6,6 мкг) 100 нукоетидной 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 $5x10^5$ 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 $x10^5$ 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 $2x10^5$ 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.

[1164] Выделение геномной ДНК для экспериментов HDR

[1165] 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.

[1166] Сравнение между PE2, PE3, BE2, BE4max, ABEdmax и ABEmax

[1167] Клетки HEK293T высевали на 48-луночные покрытые поли-O-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 сгРНК, 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.

[1168] Определение активности PE3 в известных нецелевых участках Cas9

[1169] 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; последовательности праймеров 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.41 A-41K; последовательности PEгРНК и пикирующей сгРНК перечислены в Таблицах 3A-3R. После ПНР 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 PEгРНК или сгРНК, 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 праймирующих последовательностей из Таблицы 4 и секвенировали на Illumina MiSeq.

[1170] 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 PEгРНК, 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 PEгРНК изменением, использовали t-тест (неспаренный, односторонний, a = 0.5) для определения того, встречались ли данные варианты со значительно более высокими уровнями по сравнению с образцами, обработанными PEгРНК, 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.

[1171] Получение линии клеток HEK293T, содержащих мутацию HBB E6V, с помощью Cas9-инициируемого HDR

[1172] 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.

[1173] Получение линии клеток HEK293T, содержащих мутацию HBB E6V, с использованием PE3

[1174] 2,5 x 10^4 клеток HEK293T, выращенных при отсутствии антибиотиков, высевали в 48-луночные покрытые поли-О-лизином планшеты (Coming). 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 PEгРНК, и 83 нг плазмиды sgPHK. 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.53А-53В. 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*.

[1175] Получение линии клеток HEK293T, содержащих HEXA 1278 + вставку TATC, с применением PE3

[1176] Клетки HEK293T, содержащие HEXA 1278+TATC аллель, получали в соответствии с протоколом, описанным выше для создания линии клеток HBB E6V; последовательности PEгРНК и сгРНК перечислены в Таблицах 2А-2С в соответствии с подзаголовками на Фиг.43А-43Н. 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.

[1177] Анализы жизнеспособности клеток

[1178] Клетки HEK293T высевали в 48-луночные планшеты и трансфицировали с

приблизительно 70% конфлюентностью 750 нг плазмиды для редактирования (PE3, PE3 R11OS K103L, никаза Cas9 H840A или dCas9), 250 нг нацеленной на HEK3 плазмиды PEгРНК, 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 Ml000 Pro (Tecan) с временем интеграции, соответствующим 1 секунде.

[1179] Получение лентивируса

[1180] 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-PEгРНК 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) вносящая разрыв сгРНК 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.

[1181] Иссечение и культивирование первичных кортикальных нейронов мышей

[1182] Диссоциированные кортикальные культуры El 8.5 собирали от мышей C57BL/6 с известным сроком беременности (Charles River). Зародышей собирали из беременных мышей после умерщвления с помощью CCE с последующим

обезглавливанием. Кортикальные лоскуты иссекали в ледяном Hibernate-E с добавлением пенициллина/стрептомицина (Life Technologies). После промывания ледяным Hibemate-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.

[1183] Праймированное редактирование в первичных нейронах и выделение ядер

[1184] В DIV 1 добавляли 15 мкл лентивируса при 10:10:1 соотношении N-концевой:С-концевой:вводящей разрыв сгРНК. 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 центрифигурировали 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 IxPBS, 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) при проточной цитометрии Broad Institute. Геномную ДНК очищали в соответствии с инструкциями Agencourt DNAdvance производителя.

[1185] РНК-секвенирование и анализ данных

[1186] Клетки HEK293T котрансфицировали нацеленными на PRNP или нацеленными на HEXA РЕгРНК 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 п.о. 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.

[1187] Анализ ClinVar

[1188] 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 Ш 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.

[1189] 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.

[1190] Доступность данных

[1191] Данные высокопроизводительного секвенирования внесены в базу данных архива считываний секвенирования NCBI. Плазмиды, кодирующие векторы экспрессии PEI, PE2/PE3, и PEгРНК, будут доступны в Addgene.

[1192] Доступность кода

[1193] Скрипт, использованный для количественного анализа вставки каркаса PEгРНК, представлен на фиг. 60A-60B.

[1194] Дополнительная информация: Таблицы и последовательности

[1195] Таблица 1: Активность праймирующих 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 % праймированного 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 сгРНК и PEгРНК приведены в Таблицах 3 A-3R под заголовком, соответствующим Фиг.42A-42H. Все значения представляют собой среднее значение трех независимых биологических повторов.

PE

сайт \ PEгРНК	HEK3 (PE3)				HEK4 (PE2)				EMX1 (PE3)				FANCF (PE3)			
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
нецелевая		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)		56.8 (9.3)	32.4 (16.7)	42.8 (13.6)	47.6 (12.0)
нецелевая 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)
нецелевая 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)
нецелевая 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)
нецелевая 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

сайт \ PEгРНК	HEK3					HEK4					EMX1					FANCF				
	sgRNA	1	2	3	4	sgRNA	1	2	3	4	sgRNA	1	2	3	4	sgRNA	1	2	3	4
нецелевая	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
нецелевая 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
нецелевая 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
нецелевая 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
нецелевая 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

сайт \ PEгРНК	HEK3					HEK4					EMX1					FANCF				
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
нецелевая 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
нецелевая 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
нецелевая 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
нецелевая 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

сайт \ PEгРНК	HEK3					HEK4					EMX1					FANCF				
	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4	---	1	2	3	4
нецелевая 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
нецелевая 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
нецелевая 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
нецелевая 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

[1196] Таблицы 2A-2C: Последовательности олигонуклеотидов, РЕгРНК и сгРНК, применяемые для экспериментов *in vitro*.

[1197] Таблица 2A. Олигонуклеотиды ДНК

ОЛИГОНУКЛЕОТИДЫ	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	CTCTGGAGGATCTAGCGGAGTTTTTTTTTTTTTTTTT (SEQ ID NO: 381)
AVA135	CTCTGGAGGATCTAGCGGAGCCCCCCCCCCCCC (SEQ ID NO: 382)

[1198] Таблица 2B. 5′-удлиненные РЕгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	ПОСЛЕДОВАТЕЛЬНОСТЬ 5′ УДЛИНЕНИЯ	ДЛИНА ЛИНКЕРА (НТ)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
РЕГРНК 1	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 383)	GGCTAACCGTGCCATTTGATCAGGTCA (SEQ ID NO: 429)	15	5	7
РЕГРНК 2	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 384)	GGCTAACCGTGCAAATTAACAAACTAA(SEQ ID NO: 430)	15	5	7
РЕГРНК 3	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 385)	GGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 431)	15	5	8
РЕГРНК 4	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 386)	GGTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 432)	15	5	15
РЕГРНК 5	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 387)	GGCTTCCTTTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 433)	15	5	22

5′-РЕГРНК_RT_7_A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 388)	GGCTAACCGTGCCATTTGATCAGGTCA(SEQ ID NO: 434)	15	5	7
5′-РЕГРНК_RT_7_B	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 389)	GGCTAACCGTGCAAATTAACAAACTAA(SEQ ID NO: 435)	15	5	7
5′-РЕГРНК_RT_8	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 390)	GGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 436)	15	5	8
5′-РЕГРНК_RT_15	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 391)	GGTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 437)	15	5	15
5′-РЕГРНК_RT_22	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 392)	GGCTTCCTTTCCTCTGCCATCTCGTGCAAATTAACAAACTAA(SEQ ID NO: 438)	15	5	22

[1199] Таблица 2C. 3′-удлиненные РЕгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	УДЛИНЕНИЕ 3′ ПОСЛЕДОВАТЕЛЬНОСТИ	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НИ)
3′-РЕГРНК_10	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 506)	TCTGCCATCTCGTGCTC (SEQ ID NO: 439)	7	10
3′-РЕГРНК_YEAST_TTOA	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 507)	TCTGCCATCTCGTGCTC(SEQ ID NO: 440)	7	10
3′-РЕГРНК_YEAST_+1AINS	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 508)	TCTGCCATCATCGTGCTC(SEQ ID NO: 441)	7	11
3′-РЕГРНК_YEAST_+1TDEL	GGCCCAGACTGAGCACGTGA(SEQ ID NO: 509)	TCTGCCATCCGTGCTC(SEQ ID NO: 442)	7	9

[1200] Таблицы 3A-3R: Последовательности РЕгРНК и сгРНК, применяемые в экспериментах на клетках млекопитающих. Все последовательности показаны в 5′ - 3′

направлении. Для конструкции РЕгРНК, спейсерные последовательности, перечисленные ниже, добавляли к 5′ концу каркаса сгРНК, и 3′ удлинения, перечисленные ниже, содержащие участок связывания праймера и матрицу RT, добавляли к 3′ концу каркаса сгРНК. Последовательность каркаса сгРНК является следующей GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGCACCGAGTCGGTGC (SEQ ID NO: 131).

[1201] Таблица 3А. Фиг. 39A-39D РЕгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NOS: 2890-2996)	3′ УДЛИНЕНИЕ (SEQ ID NOS:2997-3103)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
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-C_11	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTC	11	10
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_16	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGATTC	16	17
FANCF_2C_17	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCAAGGTGCTGCAGAAGGGATTCC	17	17
RNF2_2C_9	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACT	9	11
RNF2_2C_10	GTCATCTTAGTCATTACCTG	GAACACCTCATGTAATGACTA	10	11
RNF2_2C_11	GTCATCTTAGTCATT	GAACACCTCATGTAATGAC	11	11

	ACCTG	TAA		
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_11	GGCCCAGACTGAGCACGTGA	CTCTGCCATCTCGTGCTCAGTCTG	13	11
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_1	GAGTCCGAGCAGAA	TGTGATGGGAGCCCTTGTT	13	17

7	GAAGAA	CTTCTGCTCGG		
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_13	GGAATCCCTTCTGCAGCACC	AAGCGATCAAGGTGCTGCAGAAGGGA	13	13
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_12	GGCACTGCGGCTGGAGGTGG	CTTTAACCCCAACCTCCAGCCGC	11	12
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_1	GGCACTGCGGCTGG	CTCCGCTTTAACCCCAACC	11	18

8	AGGTGG	TCCAGCCGC		
HEK4_2D_19	GGCACTGCGGCTGGAGGTGG	CTCCGCTTTAACCCCAACCTCCAGCCGC	11	19

[1202] Таблица 3Вю Фиг. 40A-40C РЕгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3104-3112)	3′ УДЛИНЕНИЕ (SEQ ID NOS: 3113-3121)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
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

[1203] Таблица 3С. Фиг. 40A-40C последовательности вносящей разрыв сгРНК

ВНОСЯЩАЯ РАЗРЫВ СГРНК	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**

[1204] Таблица 3D. Фиг. 41A-41K PEгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3154-3304)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3305-3455)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
HEK3_4A_1TTOA	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_4A_1TTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATCGCGTGCTCAGTCTG	13	10
HEK3_4A_1	GGCCCAGACTGAGCA	TCTGCCATCCCGTGCTCAGT	13	10

TTOG	CGTGA	CTG		
HEK3_4A_2 GTOA	GGCCCAGACTGAGCA CGTGA	TCTGCCATTACGTGCTCAGT CTG	13	10
HEK3_4A_2 GTOC	GGCCCAGACTGAGCA CGTGA	TCTGCCATGACGTGCTCAGT CTG	13	10
HEK3_4A_2 GTOT	GGCCCAGACTGAGCA CGTGA	TCTGCCATAACGTGCTCAGT CTG	13	10
HEK3_4A_3 ATOC	GGCCCAGACTGAGCA CGTGA	TCTGCCAGCACGTGCTCAG TCTG	13	10
HEK3_4A_3 ATOG	GGCCCAGACTGAGCA CGTGA	TCTGCCACCACGTGCTCAG TCTG	13	10
HEK3_4A_3 ATOT	GGCCCAGACTGAGCA CGTGA	TCTGCCAACACGTGCTCAG TCTG	13	10
HEK3_4A_4 TTOA	GGCCCAGACTGAGCA CGTGA	TCTGCCTTCACGTGCTCAGT CTG	13	10
HEK3_4A_4 TTOC	GGCCCAGACTGAGCA CGTGA	TCTGCCGTCACGTGCTCAG TCTG	13	10
HEK3_4A_4 TTOG	GGCCCAGACTGAGCA CGTGA	TCTGCCCTCACGTGCTCAGT CTG	13	10
HEK3_4A_5 GTOA	GGCCCAGACTGAGCA CGTGA	TCTGCTATCACGTGCTCAGT CTG	13	10
HEK3_4A_5 GTOC	GGCCCAGACTGAGCA CGTGA	TCTGCGATCACGTGCTCAGT CTG	13	10
HEK3_4A_5 GTOT	GGCCCAGACTGAGCA CGTGA	TCTGCAATCACGTGCTCAGT CTG	13	10
HEK3_4A_6 GTOA	GGCCCAGACTGAGCA CGTGA	TCTGTCATCACGTGCTCAGT CTG	13	10
HEK3_4A_6 GTOC	GGCCCAGACTGAGCA CGTGA	TCTGGCATCACGTGCTCAGT CTG	13	10
HEK3_4A_6 GTOT	GGCCCAGACTGAGCA CGTGA	TCTGACATCACGTGCTCAGT CTG	13	10
HEK3_4A_7 CTOA	GGCCCAGACTGAGCA CGTGA	TCTTCCATCACGTGCTCAGT CTG	13	10
HEK3_4A_7 CTOG	GGCCCAGACTGAGCA CGTGA	TCTCCCATCACGTGCTCAGT CTG	13	10
HEK3_4A_7 CTOT	GGCCCAGACTGAGCA CGTGA	TCTACCATCACGTGCTCAGT CTG	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_17GTOC	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTGCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_20GTOC	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGGTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_23CTOG	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCACGGCTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_24TTOA	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCTGGGCTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_26CTOG	GGCCCAGACTGAGCACGTGA	TGGAGGAACCAGGGCTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_30CTOG	GGCCCAGACTGAGCACGTGA	TGGACGAAGCAGGGCTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
HEK3_4B_33CTOG	GGCCCAGACTGAGCACGTGA	TCGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACGTGCTCAGTCTG	13	34
RNF2_4C_1CTOA	GTCATCTTAGTCATTACCTG	AACGAACACCTCATGTAATGACTAAGATG	15	14

RNF2_4C_1CTOG	GTCATCTTAGTCATTACCTG	AACGAACACCTCACGTAATGACTAAGATG	15	14
RNF2_4C_1CTOT	GTCATCTTAGTCATTACCTG	AACGAACACCTCAAGTAATGACTAAGATG	15	14
RNF2_4C_2TTOA	GTCATCTTAGTCATTACCTG	AACGAACACCTCTGGTAATGACTAAGATG	15	14
RNF2_4C_2TTOG	GTCATCTTAGTCATTACCTG	AACGAACACCTCCGGTAATGACTAAGATG	15	14
RNF2_4C_3GTOC	GTCATCTTAGTCATTACCTG	AACGAACACCTGAGGTAATGACTAAGATG	15	14
RNF2_4C_4ATOC	GTCATCTTAGTCATTACCTG	AACGAACACCGCAGGTAATGACTAAGATG	15	14
RNF2_4C_4ATOT	GTCATCTTAGTCATTACCTG	AACGAACACCACAGGTAATGACTAAGATG	15	14
RNF2_4C_4ATOG	GTCATCTTAGTCATTACCTG	AACGAACACCCCAGGTAATGACTAAGATG	15	14
RNF2_4C_5GTOT	GTCATCTTAGTCATTACCTG	AACGAACACATCAGGTAATGACTAAGATG	15	14
RNF2_4C_6GTOA	GTCATCTTAGTCATTACCTG	AACGAACATCTCAGGTAATGACTAAGATG	15	14
RNF2_4C_7TTOC	GTCATCTTAGTCATTACCTG	AACGAACGCCTCAGGTAATGACTAAGATG	15	14
FANCF_4D_1ATOG	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT	14	17
FANCF_4D_1ATOT	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGGAGCTGCAGAAGGGAT	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_	GGAATCCCTTCTGCA	GGAAAAGCGATCTAGGTGC	14	17

5GTOA	GCACC	TGCAGAAGGGAT		
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_	GATGTCTGCAGGCCA	AATGTGCCATCTGGAGGCC	13	22

6GTOC	GATGA	TCATCTGGCCTGCAGA		
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_D	GGCCCAGACTGAGCA	TGGAGGAAGCAGGGCTTCC	13	19

EL1-15	CGTGA	CGTGCTCAGTCTG		
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_3-5AGGDEL	GATTCCTGGTGCCAGAAACA	TCCCGTCACCGTTTCTGGCACCAGG	13	12
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

[1205]

[1206] Таблица 3E. Фиг. 41A-41K вносящая разрыв сгРНК

ВНОСЯЩАЯ РАЗРЫВ СГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NOS: 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

[1207]

[1208] Таблица 3F: Фиг. 42A-42H PEгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NOS: 3464-3478)	3′ УДЛИНЕНИЕ (SEQ ID NOS: 3479-3493)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
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

[1209] Таблица 3G. Фиг. 42A-42H вносящая разрыв сгРНК

ВНОСЯЩАЯ РАЗРЫВ СГРНК	ВОЗМОЖНАЯ СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NOS: 616-618)	ВОЗМОЖНАЯ СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ(SEQ ID NOS: 619-621)
HEK3_5A-F_+90	GTCAACCAGTATCCCGGTGC	GTCAACCAGTATCCCGGTGC
FANCF_5A-F_+48	GGGGTCCCAGGTGCTGACGT	GATGTACAGAGAGCCCAGGGC
EMX1_5A-F_+57	GATGTACAGAGAGCCCAGGGC	GGGGTCCCAGGTGCTGACGT

[1210] Таблица 3H. Фиг. 42A-42H сгРНК редактирования оснований

СГРНК 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: 457)

[1211] Таблица 3I. Фиг. 42A-42H нацеленная сгРНК

НАЦЕЛЕННАЯ СГРНК	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)

[1212] Таблица 3J. Фиг. 42A-42H нацеленная сгРНК

НАЦЕЛЕННАЯ СГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 663-677)	3′ УДЛИНЕНИЕ (SEQ ID NO: 678-692)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
HEK3_5G-H_PEГРНК_1	GGCCCAGACTGAGCACGTGA	TCTGCCATCTCGTGCTCAGTCTG	13	10
HEK3_5G-	GGCCCAGACTGAGCAC	TCTGCCATCAAAGCGTG	13	13

H_РЕГРНК_2	GTGA	CTCAGTCTG		
HEK3_5G-H_РЕГРНК_3	GGCCCAGACTGAGCACGTGA	TCTGCCATCCGTGCTCAGTCTG	13	9
HEK3_5G-H_РЕГРНК_4	GGCCCAGACTGAGCACGTGA	TCTGCGATCACGTGCTCAGTCTG	13	10
HEK4_5G-H_РЕГРНК_1	GGCACTGCGGCTGGAGGTGG	TTAACGCCCACCTCCAGCC	9	10
HEK4_5G-H_РЕГРНК_2	GGCACTGCGGCTGGAGGTGG	TTAACCCCCCCCTCCAGCC	9	10
HEK4_5G-H_РЕГРНК_3	GGCACTGCGGCTGGAGGTGG	TTAACCCCTTACACCTCCAGCC	9	13
HEK4_5G-H_РЕГРНК_4	GGCACTGCGGCTGGAGGTGG	TTAACCCCCCCTCCAGCC	9	9
EMX1_5G-H_РЕГРНК_1	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCACTTCTTCTTCTGCTCGGA	14	16
EMX1_5G-H_РЕГРНК_2	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA	14	16
EMX1_5G-H_РЕГРНК_3	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA	14	19
EMX1_5G-H_РЕГРНК_4	GAGTCCGAGCAGAAGAAGAA	GTGATGGGAGTTCTTCTTCTGCTCGGA	14	13
FANCF_5G-H_РЕГРНК_1	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT	14	17
FANCF_5G-H_РЕГРНК_2	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT	14	17
FANCF_5G-H_РЕГРНК_3	GGAATCCCTTCTGCAGCACC	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT	14	20

[1213] Таблица 3К. Фиг. 49A-49BРЕгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3494-3521)	3′ УДЛИНЕНИЕ (SEQ ID NO: 3522-3540)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
HEK3_6A_2GTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATGACGTGCTCAGTCTG	13	10
HEK3_6A_2GTOC	GGCCCAGACTGAGCACGTGA			

EMX1_6A_3G TOC	GAGTCCGAGCAG AAGAAGAA	ATGGGAGCCCTTGTTCTT CTGCTCGG	13	13
EMX1_6A_3G TOC	GAGTCCGAGCAG AAGAAGAA			
FANCF_6A_5G TOT	GGAATCCCTTCTG CAGCACC	AAAAGCGATCAAGGTGC TGCAGAAGGGA	13	15
FANCF_6A_5G TOT	GGAATCCCTTCTG CAGCACC			
HEK3_6A_1HI S6INS	GGCCCAGACTGA GCACGTGA	TGGAGGAAGCAGGGCTT CCTTTCCTCTGCCATCAAT GATGGTGATGATGGTGCG TGCTCAGTCTG	13	52
HEK3_6A_1HI S6INS	GGCCCAGACTGA GCACGTGA			
HEK3_6A_5GT OT	GGCCCAGACTGA GCACGTGA	TCTGCAATCACGTGCTCA GTCTG	13	10
HEK3_6A_5GT OT	GGCCCAGACTGA GCACGTGA			
HEK3_6A_1CT TINS	GGCCCAGACTGA GCACGTGA	TCTGCCATCAAAGCGTGC TCAGTCTG	13	10
HEK3_6A_1CT TINS	GGCCCAGACTGA GCACGTGA			
*HBB*_6B_INSA LL	GCATGGTGCACCT GACTCCTG	AGACTTCTCCACAGGAGT CAGGTGCAC	13	14
*HBB*_6B_INSA LL	GCATGGTGCACCT GACTCCTG			
*HBB*_6B_COR RECT	GCATGGTGCACCT GACTCCTG	AGACTTCTCCTCAGGAGT CAGGTGCAC	13	14
*HBB*_6B_COR RECT	GCATGGTGCACCT GACTCCTG			
*HBB*_6B_COR RECT_W_SILE NT	GCATGGTGCACCT GACTCCTG	AGACTTCTCTTCAGGAGT CAGGTGCAC	13	14
*HBB*_6B_COR RECT_W_SILE NT	GCATGGTGCACCT GACTCCTG			
*HEXA*_6B_INS	GTACCTGAACCGT	AGTCAGGGCCATAGGATA	12	14

TALL	ATATCCTA	GATATACGGTTC		
*HEXA*_6B_CORRECT	GATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTG	10	21
*HEXA*_6B_CORRECT_W_SILENT	GATCCTTCCAGTCAGGGCCAT	GTACCTGAACCGTATATCTTATGGCCCTGACT	9	27
*PRNP*_6C	GCAGTGGTGGGGGGCCTTGG	ATGTAGACGCCAAGGCCCCCCACC	12	12
HEK3_6E-G_1TTOG	GGCCCAGACTGAGCACGTGA	TCTGCCATCCCGTGCTCAGTCTG	13	10
HEK3_6E-G_1CTTINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	10
RNF2_6E-G_1CTOG	GTCATCTTAGTCATTACCTG	AACGAACACCTCACGTAATGACTAAGATG	15	14
*HBB*_6E-G_4ATOT	GCATGGTGCACCTGACTCCTG	AGACTTCTCCACAGGAGTCAGGTGCAC	13	14
HEK3_6H_1HIS6INS	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATGATGGTGATGATGGTGCGTGCTCAGTCTG	13	52
HEK3_6H_1FLAGINS	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACTTATCGTCGTCATCCTTGTAATCCGTGCTCAGTCTG	13	58

[1214] Таблица 3L. Фиг. 47A-74D PEгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3541-3547)	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′ УДЛИНЕНИЯ (SEQ ID NO: 3549-3556)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
HEK3_ED4B_1TDEL	GGCCCAGACTGAGCACGTGA	TCTGCCATCCGTGCTCAGTCTG	13	9
HEK3_ED4B_1AINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCATCGTGCTCAGTCTG	13	11
HEK3_ED4B_1CTTINS	GGCCCAGACTGAGCACGTGA	TCTGCCATCAAAGCGTGCTCAGTCTG	13	13

HEK3_ED4C_2GTOC	GGCCCAGACTGAGCACGTGA	TCTGCCATGACGTGCTCAGTCTG	13	10
HEK3_ED4D_1FLAGINS	GGCCCAGACTGAGCACGTGA	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACTTATCGTCGTCATCCTTGTAATCCGTGCTCAGTCTG	13	58
RNF2_ED4E_1CTOA	GTCATCTTAGTCATTACCTG	AACGAACACCTCATGTAATGACTAAGATG	15	14
EMX1_ED4F_1GTOC	GAGTCCGAGCAGAAGAAGAA	ATGGGAGCCCTTGTTCTTCTGCTCGG	13	13
*HBB*_ED4G_2TTOA	GTAACGGCAGACTTCTCCTC	ATCTGACTCCTGTGGAGAAGTCTGCC	12	14

[1215] Таблица 3M: Фиг. 48A-48C PEгPHK

PEГPHK	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 3557-3627)	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′ УДЛИНЕНИЯ (SEQ ID NO: 3628-3698)	ДЛИНА PBS (HT)	ДЛИНА МАТРИЦЫ RT (HT)
VEGFA_ED5A_31	GATGTCTGCAGGCCAGATGA	CCCTCTGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	31
VEGFA_ED5A_30	GATGTCTGCAGGCCAGATGA	CCTCTGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	30
VEGFA_ED5A_29	GATGTCTGCAGGCCAGATGA	CTCTGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	29
VEGFA_ED5A_28	GATGTCTGCAGGCCAGATGA	TCTGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	28
VEGFA_ED5A_27	GATGTCTGCAGGCCAGATGA	CTGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	27
VEGFA_ED5A_26	GATGTCTGCAGGCCAGATGA	TGACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	26
VEGFA_ED5A_25	GATGTCTGCAGGCCAGATGA	GACAATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA	13	25
VEGFA_ED5	GATGTCTGCAGGCC	ACAATGTGCCATCTGGAGCACT	13	24

A_24	AGATGA	CATCTGGCCTGCAGA		
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_ED5B_31	GATTCCTGGTGCCAGAAACA	AGGACTAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	31
DNMT1_ED5B_30	GATTCCTGGTGCCAGAAACA	GGACTAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	30
DNMT1_ED5B_29	GATTCCTGGTGCCAGAAACA	GACTAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	29
DNMT1_ED5B_28	GATTCCTGGTGCCAGAAACA	ACTAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	28
DNMT1_ED5B_27	GATTCCTGGTGCCAGAAACA	CTAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	27
DNMT1_ED5B_26	GATTCCTGGTGCCAGAAACA	TAGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	26
DNMT1_ED5B_25	GATTCCTGGTGCCAGAAACA	AGTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	25
DNMT1_ED5B_24	GATTCCTGGTGCCAGAAACA	GTTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	24
DNMT1_ED5B_23	GATTCCTGGTGCCAGAAACA	TTCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	23
DNMT1_ED5B_22	GATTCCTGGTGCCAGAAACA	TCTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	22
DNMT1_ED5B_21	GATTCCTGGTGCCAGAAACA	CTGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	21
DNMT1_ED5B_20	GATTCCTGGTGCCAGAAACA	TGCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	20
DNMT1_ED5B_19	GATTCCTGGTGCCAGAAACA	GCCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	19
DNMT1_ED5B_18	GATTCCTGGTGCCAGAAACA	CCCTCCCGTCACCACTGTTTCTGGCACCAGG	13	18
DNMT1_ED5	GATTCCTGGTGCCA	CCTCCCGTCACCACTGTTTCTG	13	17

B_17	GAAACA	GCACCAGG		
DNMT1_ED5 B_16	GATTCCTGGTGCCA GAAACA	CTCCCGTCACCACTGTTTCTGG CACCAGG	13	16
DNMT1_ED5 B_15	GATTCCTGGTGCCA GAAACA	TCCCGTCACCACTGTTTCTGGC ACCAGG	13	15
DNMT1_ED5 B_14	GATTCCTGGTGCCA GAAACA	CCCGTCACCACTGTTTCTGGCA CCAGG	13	14
DNMT1_ED5 B_13	GATTCCTGGTGCCA GAAACA	CCGTCACCACTGTTTCTGGCAC CAGG	13	13
DNMT1_ED5 B_12	GATTCCTGGTGCCA GAAACA	CGTCACCACTGTTTCTGGCACC AGG	13	12
DNMT1_ED5 B_11	GATTCCTGGTGCCA GAAACA	GTCACCACTGTTTCTGGCACCA GG	13	11
DNMT1_ED5 B_10	GATTCCTGGTGCCA GAAACA	TCACCACTGTTTCTGGCACCAG G	13	10
DNMT1_ED5 B_9	GATTCCTGGTGCCA GAAACA	CACCACTGTTTCTGGCACCAGG	13	9
DNMT1_ED5 B_8	GATTCCTGGTGCCA GAAACA	ACCACTGTTTCTGGCACCAGG	13	8
RUNX1_ED5 C_31	GCATTTTCAGGAGG AAGCGA	AATGACTCAAATATGCTGTCTG AAGCAATCGCTTCCTCCTGAAA AT	15	31
RUNX1_ED5 C_30	GCATTTTCAGGAGG AAGCGA	ATGACTCAAATATGCTGTCTGA AGCAATCGCTTCCTCCTGAAAA T	15	30
RUNX1_ED5 C_29	GCATTTTCAGGAGG AAGCGA	TGACTCAAATATGCTGTCTGAA GCAATCGCTTCCTCCTGAAAAT	15	29
RUNX1_ED5 C_28	GCATTTTCAGGAGG AAGCGA	GACTCAAATATGCTGTCTGAAG CAATCGCTTCCTCCTGAAAAT	15	28
RUNX1_ED5 C_27	GCATTTTCAGGAGG AAGCGA	ACTCAAATATGCTGTCTGAAGC AATCGCTTCCTCCTGAAAAT	15	27

RUNX1_ED5C_26	GCATTTTCAGGAGGAAGCGA	CTCAAATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	26
RUNX1_ED5C_25	GCATTTTCAGGAGGAAGCGA	TCAAATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	25
RUNX1_ED5C_24	GCATTTTCAGGAGGAAGCGA	CAAATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	24
RUNX1_ED5C_23	GCATTTTCAGGAGGAAGCGA	AAATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	15	23
RUNX1_ED5C_22	GCATTTTCAGGAGGAAGCGA	AATATGCTGTCTGAAGCAATCGCTTCCTCCTGAAAAT	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_ED5	GCATTTTCAGGAGG	TGAAGCAATCGCTTCCTCCTGA	15	11

C_11	AAGCGA	AAAT		
RUNX1_ED5C_10	GCATTTTCAGGAGGAAGCGA	GAAGCAATCGCTTCCTCCTGAAAAT	15	10
RUNX1_ED5C_9	GCATTTTCAGGAGGAAGCGA	AAGCAATCGCTTCCTCCTGAAAAT	15	9

[1216] Таблица 3N. Фиг. 48A-48C PEгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	ПОСЛЕДОВАТЕЛЬНОСТЬ 3′ УДЛИНЕНИЯ	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
HEK3_ED6_5GTOA	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 393)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 394)	13	10

[1217] Таблица 3O. Фиг. 48A-48C вносящая разрыв сгРНК

ВНОСЯЩАЯ РАЗРЫВ СГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
HEK3_ED6_+63	GCACATACTAGCCCCTGTCT (SEQ ID NO: 395)

[1218] Таблица 3P. Фиг. 50A-50B PEгРНК

РЕГРНК	СПЕЙСЕР (SEQ ID NO: 3699-3754)	3′ УДЛИНЕНИЕ (5′ - 3′)(SEQ ID NO: 3755-3810)	ДЛИНА PBS (NT)	ДЛИНА МАТРИЦЫ RT (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	GCATGGTGCACCTGA	GCAGACTTCTCCTCAGGAGTCAGGTG	13	13

5.5	CTCCTG	CAC		
HBB 5.6	GCATGGTGCACCTGA CTCCTG	GACTTCTCCTCAGGAGTCAGGTGCAC	13	12
HBB 5.7	GCATGGTGCACCTGA CTCCTG	ACTTCTCCTCAGGAGTCAGGTGCAC	13	21
HBB 5.8	GCATGGTGCACCTGA CTCCTG	TAACGGCAGACTTCTCCTCAGGAGTCA GGTGCA	12	19
HBB 5.9	GCATGGTGCACCTGA CTCCTG	ACGGCAGACTTCTCCTCAGGAGTCAG GTGCA	12	17
HBB 5.10	GCATGGTGCACCTGA CTCCTG	GGCAGACTTCTCCTCAGGAGTCAGGT GCA	12	16
HBB 5.11	GCATGGTGCACCTGA CTCCTG	GCAGACTTCTCCTCAGGAGTCAGGTG CA	12	13
HBB 5.12	GCATGGTGCACCTGA CTCCTG	GACTTCTCCTCAGGAGTCAGGTGCA	12	12
HBB 5.13	GCATGGTGCACCTGA CTCCTG	ACTTCTCCTCAGGAGTCAGGTGCA	12	14
HEXA S 1	ATCCTTCCAGTCAGG GCCAT	ATATCTTATGGCCCTGACTGGAA	13	14
HEXA S 2	ATCCTTCCAGTCAGG GCCAT	TATATCTTATGGCCCTGACTGGAA	13	15
HEXA S 3	ATCCTTCCAGTCAGG GCCAT	GTATATCTTATGGCCCTGACTGGAA	13	16
HEXA S 4	ATCCTTCCAGTCAGG GCCAT	ACCGTATATCTTATGGCCCTGACTGGA A	13	19
HEXA S 5	ATCCTTCCAGTCAGG GCCAT	AACCGTATATCTTATGGCCCTGACTGG AA	13	20
HEXA S 6	ATCCTTCCAGTCAGG GCCAT	GAACCGTATATCTTATGGCCCTGACTG GAA	13	21
HEXA S 7	ATCCTTCCAGTCAGG GCCAT	TGAACCGTATATCTTATGGCCCTGACTG GAA	13	22
HEXA S 8	ATCCTTCCAGTCAGG GCCAT	ATATCTTATGGCCCTGACT	9	14
HEXA S 9	ATCCTTCCAGTCAGG GCCAT	TATATCTTATGGCCCTGACT	9	15
HEXA S 10	ATCCTTCCAGTCAGG GCCAT	GTATATCTTATGGCCCTGACT	9	16

HEXA S 11	ATCCTTCCAGTCAGGGCCAT	ACCGTATATCTTATGGCCCTGACT	9	19
HEXA S 12	ATCCTTCCAGTCAGGGCCAT	AACCGTATATCTTATGGCCCTGACT	9	20
HEXA S 13	ATCCTTCCAGTCAGGGCCAT	GAACCGTATATCTTATGGCCCTGACT	9	21
HEXA S 14	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACT	9	22
HEXA S 15	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGAC	8	22
HEXA S 16	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTG	10	22
HEXA S 17	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTGG	11	22
HEXA S 18	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTGGA	12	22
HEXA S 19	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTGGAA	13	22
HEXA S 20	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTGGAAG	14	22
HEXA S 21	ATCCTTCCAGTCAGGGCCAT	TGAACCGTATATCTTATGGCCCTGACTGGAAGG	15	22
HEXA S 22	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCTTATGGCCCTGACT	9	25
HEXA S 23	ATCCTTCCAGTCAGGGCCAT	TACCTGAACCGTATATCTTATGGCCCTGACT	9	26
HEXA S 24	ATCCTTCCAGTCAGGGCCAT	GTACCTGAACCGTATATCTTATGGCCCTGACT	9	27
HEXA S 25	ATCCTTCCAGTCAGGGCCAT	GGTACCTGAACCGTATATCTTATGGCCCTGACT	9	28
HEXA S 26	ATCCTTCCAGTCAGGGCCAT	TGGTACCTGAACCGTATATCTTATGGCCCTGACT	9	29
HEXA 5	ATCCTTCCAGTCAGGGCCAT	ACCTGAACCGTATATCCTATGGCCCTGACTGGAA	13	21
HEXA 6	ATCCTTCCAGTCAGGGCCAT	ACCGTATATCCTATGGCCCTGACTGGAA	13	15
HEXA	ATCCTTCCAGTCAGG	ACCTGAACCGTATATCCTATGGCCCTG	15	21

7	GCCAT	ACTGGAAGG		
HEXA 8	ATCCTTCCAGTCAGG GCCAT	ACCTGAACCGTATATCCTATGGCCCTG ACTGGAAG	14	21
HEXA 9	ATCCTTCCAGTCAGG GCCAT	ACCTGAACCGTATATCCTATGGCCCTG ACTGGA	12	21
HEXA 10	ATCCTTCCAGTCAGG GCCAT	ACCTGAACCGTATATCCTATGGCCCTG ACTGG	11	21
HEXA 11	ATCCTTCCAGTCAGG GCCAT	ACCTGAACCGTATATCCTATGGCCCTG ACTG	10	21
HEXA 12	ATCCTTCCAGTCAGG GCCAT	AACCGTATATCCTATGGCCCTGACTGG AA	13	16
HEXA 13	ATCCTTCCAGTCAGG GCCAT	TGAACCGTATATCCTATGGCCCTGACTG GAA	13	18
HEXA 14	ATCCTTCCAGTCAGG GCCAT	TACCTGAACCGTATATCCTATGGCCCTG ACTGGAA	13	22
HEXA 15	ATCCTTCCAGTCAGG GCCAT	TGGTACCTGAACCGTATATCCTATGGCC CTGACTGGAA	13	25
HEXA 16	ATCCTTCCAGTCAGG GCCAT	GTACCTGAACCGTATATCCTATGGCCCT GACTGGAA	13	23
HEXA 17	ATCCTTCCAGTCAGG GCCAT	AACCGTATATCCTATGGCCCTGACTG	10	16
HEXA 18	ATCCTTCCAGTCAGG GCCAT	TGAACCGTATATCCTATGGCCCTGACTG	10	18
HEXA 19	ATCCTTCCAGTCAGG GCCAT	TACCTGAACCGTATATCCTATGGCCCTG ACTG	10	22
HEXA 20	ATCCTTCCAGTCAGG GCCAT	TGGTACCTGAACCGTATATCCTATGGCC CTGACTG	10	25

[1219] Таблица 3Q. Фиг. 50A-50B вносящая разрыв сгРНК

ВНОСЯЩАЯ РАЗРЫВ СГРНК	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)

[1220] Таблица 3R. Фиг. 51A-51G PEгРНК

РЕГРНК	СПЕЙСЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NO: 632-640)	3′ УДЛИНЕНИЕ (SEQ ID NO: 641-649)	ДЛИНА PBS (НТ)	ДЛИНА МАТРИЦЫ RT (НТ)
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

[1221] Таблица 4. Последовательности праймеров, используемых для амплификации геномной ДНК клеток млекопитающих и HTS[181].

ОПИСАНИЕ	ПОСЛЕДОВАТЕЛЬНОСТЬ (SEQ ID NOS: 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	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTTGGTTGGGGTAACGGTG
HEXA FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCATACAGGTGTGGCGAGAGG
HEXA REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCAGCCTCCTTTGGTTAGCA
RUNX1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCACAAACAAGACAGGGAACTG
RUNX1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGATGTAGGGCTAGAGGGGTG
VEGFA FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACTTGGTGCCAAATTCTTCTCC
VEGFA REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTAAAGAGGGAATGGGCTTTGGA
DNMT FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCACAACAG

	CTTCATGTCAGCC
DNMT REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGTTAATGTTTCCTGATGGTCC
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCCCTGTTGACCTGGAGAA
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTACTTGCCCTGACCA
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTGGTGTTGACAGGGAGCAA
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGATGTGGGCAGAAGGG
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAGAGGGAACAGAAGGGCT
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCCAAAGGCCCAAGAACCT
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCTAGCACTTTGGAAGGTCG
HEK3 НЕЦЕЛЕВОЙ УЧАСТОК 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCATCTTAATCTGCTCAGCC
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGCATGGCTTCTGAGACTCA
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTCCCTTGCACTCCCTGTCTTT
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTGGCAATGGAGGCATTGG
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAAGAGGCTGCCCATGAGAG
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGTCTGAGGCTCGAATCCTG
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGTGGCCTCCATATCCCTG
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTCCACCAGAACTCAGCCC
HEK4 НЕЦЕЛЕВОЙ УЧАСТОК 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTCGGTTCCTCCACAACAC
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTGGGGAGATTTGCATCTGTGGAGG

EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTTTTATACCATCTTGGGGTTACAG
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAATGTGCTTCAACCCATCACGGC
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCATGAATTTGTGATGGATGCAGTCTG
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAGAAGGAGGTGCAGGAGCTAGAC
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCATCCCGACCTTCATCCCTCCTGG
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTAGTTCTGACATTCCTCCTGAGGG
EMX1 НЕЦЕЛЕВОЙ УЧАСТОК 4 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCAAACAAGGTGCAGATACAGCA
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 1 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCGGGCAGTGGCGTCTTAGTCG
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 1 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCTGGGTTTGGTTGGCTGCTC
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 2 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTCCTTGCCGCCCAGCCGGTC
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 2 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGGGGAAGAGGCGAGGACAC
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 3 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCAGTGTTTCCCATCCCCAACAC
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 3 REV	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAATGGATCCCCCCCTAGAGCTC
FANCF НЕЦЕЛЕВОЙ УЧАСТОК 4 FWD	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGGCCCACAGGTCCTTCTGGA

[1222] Таблица 5. Последовательности 100-мерных одноцепочечных ДНК-олигонуклеотидных донорных матриц, использованных в экспериментах HDR и при создании клеточной линии *HBB* E6V HEK293T. Олигонуклеотиды имеют длину 100-103 нуклеотида с плечами гомологии, центрированными вокруг сайта редактирования. Олигонуклеотиды были получены от Integrated DNA Technologies, очищены с помощью PAGE.

HEK3 +3 A TO T	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGTTGGCAGAGGAAAGGAAGCCCTGCTTCCTCCAGAG

	GGCGTCGCAGGAC (SEQ ID NO: 717)
HEK3 +3 A - T, +5,6 GG - TT	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTG AGCACGTGTTTTCAGAGGAAAGGAAGCCCTGCTTCCTCCAGAG GGCGTCGCAGGAC (SEQ ID NO: 718)
HEK3 +1 T - G	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTG AGCACGGGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCAGA GGGCGTCGCAGGAC (SEQ ID NO: 719)
HEK3 +3 A - C	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTG AGCACGTGCTGGCAGAGGAAAGGAAGCCCTGCTTCCTCCAGA GGGCGTCGCAGGAC (SEQ ID NO: 720)
HEK3 +1 CTT ВСТАВКА	GCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTG AGCACGCTTTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA GAGGGCGTCGCAGGAC (SEQ ID NO: 721)
RNF2 +1 C - A	CCCAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTAGTCA TTACATGAGGTGTTCGTTGTAACTCATATAAACTGAGTTCCCATG TTTTGCTTAA (SEQ ID NO: 722)
RNF2 +1 C - G	CCCAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTAGTCA TTACGTGAGGTGTTCGTTGTAACTCATATAAACTGAGTTCCCATG TTTTGCTTAA (SEQ ID NO: 723)
RNF2 +1 GTA ВСТАВКА	CAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTAGTCATT ACGTACTGAGGTGTTCGTTGTAACTCATATAAACTGAGTTCCCAT GTTTTGCTTA (SEQ ID NO: 724)
УСТАНОВКА *HBB* E6V (ТАКЖЕ ИСПОЛЬЗУЕТСЯ ДЛЯ СОЗДАНИЯ КЛЕТОЧНОЙ ЛИНИИ *HBB* E6V HEK293T)	ACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGGCAG ACTTCTCCACAGGAGTCAGATGCACCATGGTGTCTGTTTGAGGT TGCTAGTGAACAC (SEQ ID NO: 725)
HBB E6V КОРРЕКЦИЯ ПРОТОСПЕЙСЕР A	ACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGGCAG ACTTCTCCTCAGGAGTCAGGTGCACCATGGTGTCTGTTTGAGGT TGCTAGTGAACAC (SEQ ID NO: 726)
HBB E6V КОРРЕКЦИЯ ПРОТОСПЕЙСЕР B	GTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGAC TCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG AACGTGGATGAAGT (SEQ ID NO: 727)
HBB E6V КОРРЕЦИЯ ПРОТОСПЕЙСЕР B, МОЛЧАЩАЯ МУТАЦИЯ PAM =	GTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGAC TCCTGATGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG AACGTGGATGAAGT (SEQ ID NO: 728)

PRNP G127V	CACATGGCTGGTGCTGCAGCAGCTGGGGCAGTGGTGGGGGGCC TTGGCGTCTACATGCTGGGAAGTGCCATGAGCAGGCCCATCATA CATTTCGGCAGTG (SEQ ID NO: 729)

[1223] Дополнительные последовательности

[1224] Последовательности двойных флуоресцентных репортерных плазмид дрожжей, используемых согласно настоящему изобретению

p425-GFP_stop_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGATGG
TGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTACTTACGGTAA
ATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATGGCCAACCTTAGTCACTA
CTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGATCATATGAAACAACATGACTTTTTC
AAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTATTTTTTTCAAAGATGACGGTAAC
TACAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAA
GGTATTGATTTTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTC
ACAATGTTTACATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGAC
ACAACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGA
TGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTATCCAAAGATCCA
AACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTGCTGCTGGTATTACCCATGGT
ATGGATGAATTGTACAAAGCTAGCAACCTGGGTCAATCCTTGG**<u>GGCCCAGACTGAGCACGT</u>**
<u>GA</u>TGGCAGAGCACAGGAGACGTC*ATGGTTTCAAAAGGTGAAGAAGATAATATGGCTATTATTAAAG*
AATTTATGAGATTTAAAGTTCATATGGAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAAG
GTGAAGGTAGACCATATGAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACCA
TTTGCTTGGGATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTGATA
TTCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAATTTTGAAGATG
GTGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTATTTATAAAGTTAAATTGAG
AGGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAAAACTATGGGTTGGGAAGCTTCATCAG
AAAGAATGTATCCAGAAGATGGTGCTTTAAAAGGTGAAATTAAACAAAGATTGAAATTAAAAGATGGT
GGTCATTATGATGCTGAAGTTAAAACTACTTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTAT
AATGTTAATATTAAATTGGATATTACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAGC
TGAAGGTAGACATTCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATT
*GATTGA*GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTTGCATT
TAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAATGACATTTT
CGATTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCT
TTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT
TATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC
CTAATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA
CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG

GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT
ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTG
TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA
GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTA
CGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG
CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATT
GTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG
CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAAC
GATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC
CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA
ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAC
CGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC
GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATAT
TATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAA
TAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTG
CTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGC
TGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGCT
TCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCT
GCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTT
CTTTTTTGTTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAGATTAC
TTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTT
AAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACT

TCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCG
ATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATT
CTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAA
TGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTA
AAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGG
AGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATAC
TTTTGAGCAATGTTTGTGGAAGCGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCG
TTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTC
CTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCG
CTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCTATATCTG
CGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGT
ACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTC
CATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTC
AATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCA
TTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTC
GGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACTACGTC
GTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCATTATGGGAAATGCTTC
AAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCATTGAGTGTTTTTTATTTGTTGTAT
TTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAATTAGGAATCGTAGTTTCATGATTTTCTGTT
ACACCTAACTTTTTGTGTGGTGCCCTCCTCCTTGTCAATATTAATGTTAAAGTGCAATTCTTTTT
CCTTATCACGTTGAGCCATTAGTATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCT
CCTTCTTGATAAATGTATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTT
GTAATTTCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAA
GAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGG
TACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATC
TTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCAGCAGACAA
GATAGTGGCGATAGGGTCAACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTC
GTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAA
CCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTG
ATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATT
GATGTTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAAT
CTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGT
AGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTAT
CCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTC
TCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAAG
AGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTT

AGTAAACCTTGTTCAGGTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAAC
AAAACGGCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGC
ATCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATC
AGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGG
CAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGACATTAGAATGGTATATCCTTGAAA
TATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACC
ACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGC
ATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTT
TCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAAT
TTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAA
AAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTT
AACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTCGGCCAAACAACCAATTAC
TTGTTGAGAAATAGAGTATAATTATCCTATAAATATAACGTTTTTGAACACACATGAACAAGGA
AGTACAGGACAATTGATTTTGAAGAGAATGTGGATTTTGATGTAATTGTTGGGATTCCATTTTT
AATAAGGCAATAATATTAGGTATGTGGATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGT
GTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATAT
TTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCG
GCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGA
ACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAA
GCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGA
ACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTG
TAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCG
TCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAG
GGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTAT
AGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAAT
TCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATAAAAAACACGCTTTTTCAGTTCGAG
TTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAAC
TTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGG
CGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACT
AAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAA
ATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACA
GGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGT
AAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCTTCTATTAC
CTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATT
ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATT
TCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACT

TAGTTTCGAATAAACACACATAAACAAACAAAGAATTC (SEQ ID NO: 730)

p425-GFP_+1fs_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGATGG
TGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTACTTACGGTAA
ATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATGGCCAACCTTAGTCACTA
CTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGATCATATGAAACAACATGACTTTTTC
AAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTATTTTTTTCAAAGATGACGGTAAC
TACAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAA
GGTATTGATTTTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTC
ACAATGTTTACATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGAC
ACAACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGA
TGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTATCCAAAGATCCA
AACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTGCTGCTGGTATTACCCATGGT
ATGGATGAATTGTACAAAGCTAGCAAACCTGGGTCAATCCTTGG**<u>GGCCCAGACTGAGCACG</u>**
<u>TGA</u>TGGCAGAGCACAGGAGACGTC*ATGGTTTCAAAAGGTGAAGAAGATAATATGGCTATTATTAAA*
GAATTTATGAGATTTAAAGTTCATATGGAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAA
GGTGAAGGTAGACCATATGAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACC
ATTTGCTTGGGATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTGA
TATTCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAATTTTGAAGAT
GGTGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTATTTATAAAGTTAAATTGA
GAGGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAAAACTATGGGTTGGGAAGCTTCATCA
GAAAGAATGTATCCAGAAGATGGTGCTTTAAAAGGTGAAATTAAACAAAGATTGAAATTAAAAGATGG
TGGTCATTATGATGCTGAAGTTAAAACTACTTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTA
TAATGTTAATATTAAATTGGATATTACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAG
CTGAAGGTAGACATTCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATT
*GATTGA*GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTTGCATT
TAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAATGACATTTT
CGATTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCT
TTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT
TATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC
CTAATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA
CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT
ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC

CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTG
TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA
GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTA
CGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG
CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATT
GTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG
CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAAC
GATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC
CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA
ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAC
CGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC
GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATAT
TATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAA
TAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTG
CTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGC
TGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGCT
TCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCT
GCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTT
CTTTTTTGTTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAGATTAC
TTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTT
AAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACT
TCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCG
ATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATT
CTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAA
TGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTA
AAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGG

AGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATAC
TTTTGAGCAATGTTTGTGGAAGCGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCG
TTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTC
CTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCG
CTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCTATATCTG
CGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGT
ACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTC
CATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTC
AATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCA
TTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTC
GGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACTACGTC
GTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCATTATGGGAAATGCTTC
AAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCATTGAGTGTTTTTTATTTGTTGTAT
TTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAATTAGGAATCGTAGTTTCATGATTTTCTGTT
ACACCTAACTTTTTGTGTGGTGCCCTCCTCCTTGTCAATATTAATGTTAAAGTGCAATTCTTTTT
CCTTATCACGTTGAGCCATTAGTATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCT
CCTTCTTGATAAATGTATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTT
GTAATTTCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAA
GAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGG
TACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATC
TTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCAGCAGACAA
GATAGTGGCGATAGGGTCAACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTC
GTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAA
CCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTG
ATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATT
GATGTTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAAT
CTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGT
AGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTAT
CCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTC
TCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAAG
AGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTT
AGTAAACCTTGTTCAGGTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAAC
AAAACGGCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGC
ATCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATC
AGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGG
CAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGACATTAGAATGGTATATCCTTGAAA

TATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACC
ACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGC
ATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTT
TCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAAT
TTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAA
AAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTT
AACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTCGGCCAAACAACCAATTAC
TTGTTGAGAAATAGAGTATAATTATCCTATAAATATAACGTTTTTGAACACACATGAACAAGGA
AGTACAGGACAATTGATTTTGAAGAGAATGTGGATTTTGATGTAATTGTTGGGATTCCATTTTT
AATAAGGCAATAATATTAGGTATGTGGATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGT
GTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATAT
TTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCG
GCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGA
ACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAA
GCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGA
ACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTG
TAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCG
TCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAG
GGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTAT
AGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAAT
TCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATAAAAAACACGCTTTTTCAGTTCGAG
TTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAAC
TTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGG
CGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACT
AAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAA
ATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACA
GGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGT
AAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCTTCTATTAC
CTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATT
ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATT
TCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACT
TAGTTTCGAATAAACACACATAAACAAACAAAGAATTC (SEQ ID NO: 731)

p425-GFP_-1fs_mCherry:

ATGTCTAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGATGG
TGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTACTTACGGTAA

ATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATGGCCAACCTTAGTCACTA
CTTTCGGTTATGGTGTTCAATGTTTTGCTAGATACCCAGATCATATGAAACAACATGACTTTTTC
AAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTATTTTTTTCAAAGATGACGGTAAC
TACAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAA
GGTATTGATTTTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTC
ACAATGTTTACATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGAC
ACAACATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGA
TGGTCCAGTCTTGTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTATCCAAAGATCCA
AACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTTACTGCTGCTGGTATTACCCATGGT
ATGGATGAATTGTACAAAGCTAGCAAACCTGGGTCAATCCTTGG**<u>GGCCCAGACTGAGCACG</u>**
<u>TGA</u>TGGCAGAGCACAGGACGTC*ATGGTTTCAAAAGGTGAAGAAGATAATATGGCTATTATTAAAGA*
ATTTATGAGATTTAAAGTTCATATGGAAGGTTCAGTTAATGGTCATGAATTTGAAATTGAAGGTGAAGG
TGAAGGTAGACCATATGAAGGTACTCAAACTGCTAAATTGAAAGTTACTAAAGGTGGTCCATTACCAT
TTGCTTGGGATATTTTGTCACCACAATTTATGTATGGTTCAAAAGCTTATGTTAAACATCCAGCTGATAT
TCCAGATTATTTAAAATTGTCATTTCCAGAAGGTTTTAAATGGGAAAGAGTTATGAATTTTGAAGATGG
TGGTGTTGTTACTGTTACTCAAGATTCATCATTACAAGATGGTGAATTTATTTATAAAGTTAAATTGAGA
GGTACTAATTTTCCATCAGATGGTCCAGTTATGCAAAAAAAAACTATGGGTTGGGAAGCTTCATCAGA
AAGAATGTATCCAGAAGATGGTGCTTTAAAAGGTGAAATTAAACAAAGATTGAAATTAAAAGATGGTG
GTCATTATGATGCTGAAGTTAAAACTACTTATAAAGCTAAAAAACCAGTTCAATTACCAGGTGCTTATAA
TGTTAATATTAAATTGGATATTACTTCACATAATGAAGATTATACTATTGTTGAACAATATGAAAGAGCTG
AAGGTAGACATTCAACTGGTGGTATGGATGAATTATATAAAGGTACCGCTCGAGCAGCTGTGATTGAT
*TGA*GTCGACTTGGTTGAACACGTTGCCAAGGCTTAAGTGAATTTACTTTAAATCTTGCATTTAA
ATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAATGACATTTTCGA
TTCGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTA
GTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTAT
CCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA
ATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCT
GTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGC
GCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT
CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAAC
ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTT
TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT
GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC
AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA

GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT
CTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTA
CGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG
ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG
CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG
CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATT
GTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG
CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAAC
GATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTC
CGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA
ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAC
CGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
TCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC
GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATAT
TATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAA
TAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTG
CTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGC
TGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGCT
TCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCT
GCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTT
CTTTTTTGTTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAGATTAC
TTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTT
AAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACT
TCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCG
ATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATT
CTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAA
TGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTA
AAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGG
AGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATAC
TTTTGAGCAATGTTTGTGGAAGCGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCG
TTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTC
CTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCG
CTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCTATATCTG

CGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGT
ACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTC
CATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTC
AATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCA
TTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTC
GGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGG
GGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACTACGTC
GTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACTTTCACCATTATGGGAAATGCTTC
AAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGGTCATTGAGTGTTTTTTATTTGTTGTAT
TTTTTTTTTTTTAGAGAAAATCCTCCAATATCAAATTAGGAATCGTAGTTTCATGATTTTCTGTT
ACACCTAACTTTTTGTGTGGTGCCCTCCTCCTTGTCAATATTAATGTTAAAGTGCAATTCTTTTT
CCTTATCACGTTGAGCCATTAGTATCAATTTGCTTACCTGTATTCCTTTACTATCCTCCTTTTTCT
CCTTCTTGATAAATGTATGTAGATTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTT
GTAATTTCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAA
GAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGG
TACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATC
TTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCAGCAGACAA
GATAGTGGCGATAGGGTCAACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTC
GTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAA
CCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTG
ATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATT
GATGTTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAAT
CTTGAAGAGGCCAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGT
AGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTAT
CCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTC
TCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAAG
AGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTT
AGTAAACCTTGTTCAGGTCTAACACTACCGGTACCCCATTTAGGACCAGCCACAGCACCTAAC
AAAACGGCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGC
ATCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATC
AGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGG
CAAAACGACGATCTTCTTAGGGGCAGACATAGGGGCAGACATTAGAATGGTATATCCTTGAAA
TATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACC
ACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGC
ATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCCTCTCACCTTTCCTTTT
TCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAAT
TTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAA

AAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTT
AACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTCGGCCAAACAACCAATTAC
TTGTTGAGAAATAGAGTATAATTATCCTATAAATATAACGTTTTTGAACACACATGAACAAGGA
AGTACAGGACAATTGATTTTGAAGAGAATGTGGATTTTGATGTAATTGTTGGGATTCCATTTTT
AATAAGGCAATAATATTAGGTATGTGGATATACTAGAAGTTCTCCTCGACCGTCGATATGCGGT
GTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATAT
TTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCG
GCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGA
ACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAA
GCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGA
ACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTG
TAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCG
TCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAG
GGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTAT
AGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAAT
TCCTGCAGCCCGGGGGATCCGTTAGAATCATTTTGAATAAAAAACACGCTTTTTCAGTTCGAG
TTTATCATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAAC
TTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGG
CGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACT
AAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAGAATCCCAGCACCAAA
ATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACA
GGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGAGT
AAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCTTACACCTTCTATTAC
CTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATT
ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATT
TCTTAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACT
TAGTTTCGAATAAACACACATAAACAAACAAAGAATTC (SEQ ID NO: 732)

обозначения:

GFP открытая рамка считывания

ЛИНКЕР, содержащий стоп-кодон +1 рамки считывания, или -1 рамки считывания

mCherry открытая рамка считывания

Остов плазмиды (содержащей промотор GPD, Leu2 маркер и AmpR)

Протоспейсер (подчеркнут)

PAM (выделено жирным)

[1225] Последовательности ДНК плазмид млекопитающих праймирующего редактора и пример плазмиды РЕгРНК

pCMV-PE1:

ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGT
CGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA
CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAG
CATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACC
CGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGC
AAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGA
GTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGG
ACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGAC
AGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCG
GGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTC
ATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGT
GGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCG
CCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC
CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAA
GGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACC
TGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAAC
ACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCA
GGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTT
TCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGA
GTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGA
AGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCA
CCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCT
GAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCC
CTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCAC
CCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGG
ATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTA
CGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAA
AGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAA
CCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGAC
TCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG
CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGA
TATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCT
ATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG
GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCC
TGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC

AGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCA
CGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGA
AGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGA
AATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAA
AACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTA
CGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTC
AGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCG
GGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGG
CAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGA
GAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCG
GCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGA
ATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTC
CGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGC
CTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGT
TCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGA
AATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGA
GATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACC
GGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCC
CCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCC
TGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTA
CGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGG
GCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAG
CAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG
GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAAT
GCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGA
ACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAG
AAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGA
GTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACA
AGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCA
ATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACC
AGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGAC
ACGGATCGACCTGTCTCAGCTGGGAGGTGAC*TCTGGAGGATCTAGCGGAGGATCCTCTGGCAG*
CGAGACACCAGGAACAAGCGAGTCAGCAACACCAGAGAGCAGTGGCGGCAGCAGCGGCGGCAG
*CAGC***ACCCTAAATATAGAAGATGAGTATCGGCTACATGAGACCTCAAAAGAGCCAGATGT**
TTCTCTAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGCGGAAACCGGGGGC
ATGGGACTGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCC
CGTGTCCATAAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACA

TACAGAGACTGTTGGACCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCC
CCTGCTACCCGTTAAGAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAG
AAGTCAACAAGCGGGTGGAAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTT
GAGCGGGCTCCCACCGTCCCACCAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTT
TCTGCCTGAGACTCCACCCCACCAGTCAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCA
GAGATGGGAATCTCAGGACAATTGACCTGGACCAGACTCCCACAGGGTTTCAAAAACA
GTCCCACCCTGTTTGATGAGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCA
CCCAGACTTGATCCTGCTACAGTACGTGGATGACTTACTGCTGGCCGCCACTTCTGAGC
TAGACTGCCAACAAGGTACTCGGGCCCTGTTACAAACCCTAGGGAACCTCGGGTATCGG
GCCTCGGCCAAGAAAGCCCAAATTTGCCAGAAACAGGTCAAGTATCTGGGGTATCTTCT
AAAAGAGGGTCAGAGATGGCTGACTGAGGCCAGAAAAGAGACTGTGATGGGGCAGCCT
ACTCCGAAGACCCCTCGACAACTAAGGGAGTTCCTAGGGACGGCAGGCTTCTGTCGCC
TCTGGATCCCTGGGTTTGCAGAAATGGCAGCCCCCCTGTACCCTCTCACCAAAACGGG
GACTCTGTTTAATTGGGGCCCAGACCAACAAAAGGCCTATCAAGAAATCAAGCAAGCTC
TTCTAACTGCCCCAGCCCTGGGGTTGCCAGATTTGACTAAGCCCTTTGAACTCTTTGTC
GACGAGAAGCAGGGCTACGCCAAAGGTGTCCTAACGCAAAAACTGGGACCTTGGCGTC
GGCCGGTGGCCTACCTGTCCAAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTTG
CCTACGGATGGTAGCAGCCATTGCCGTACTGACAAAGGATGCAGGCAAGCTAACCATGG
GACAGCCACTAGTCATTCTGGCCCCCCATGCAGTAGAGGCACTAGTCAAACAACCCCCC
GACCGCTGGCTTTCCAACGCCCGGATGACTCACTATCAGGCCTTGCTTTTGGACACGGA
CCGGGTCCAGTTCGGACCGGTGGTAGCCCTGAACCCGGCTACGCTGCTCCCACTGCCT
GAGGAAGGGCTGCAACACAACTGCCTTGATATCCTGGCCGAAGCCCACGGAACCCGAC
CCGACCTAACGGACCAGCCGCTCCCAGACGCCGACCACACCTGGTACACGGATGGAAG
CAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCGGTGACCACCGAGACCGAG
GTAATCTGGGCTAAAGCCCTGCCAGCCGGGACATCCGCTCAGCGGGCTGAACTGATAG
CACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTTATACTGATAGC
CGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAAGGCGTGGGTTGCTC
ACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTACTAAAAGCCCT
CTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGGGACACAGCG
CCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCATCACAGA
GACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCACCCTCTGGCGGCTCAAAAA
GAACCGCCGACGGCAGCGAATTCGAGCCCAAGAAGAAGAGGAAAGTCTAACCGGTCATC
ATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCC
ATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTT
TCCTAATAAAATGAGAAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG
GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATACCGTCGACCTCTA
GCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAAT

TCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTAGGGTGCCTAATGAGTGAGCT
AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTT
CCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA
AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA
GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG
CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG
CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC
CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC
ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT
TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG
CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGA
GCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG
GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTT
ACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTAC
TTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA
AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCA
GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCGATCTCCCGAT
CCCCTAGGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCC
CTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAA
GGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAT
GTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG

GTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT
GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTG
ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA
ATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGG
CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAGA
TCCGCGGCCGCTAATACGACTCACTATAGGGAGAGCCGCCACC (SEQ ID NO: 733)

pCMV-PE2:

ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGT
CGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA
CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAG
CATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACC
CGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGC
AAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGA
GTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGG
ACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGAC
AGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCG
GGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTC
ATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGT
GGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCG
CCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC
CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAA
GGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACC
TGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAAC
ACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCA
GGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTT
TCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGA
GTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGA
AGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCA
CCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCT
GAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCC
CTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCAC
CCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGG
ATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTA

CGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAA
AGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAA
CCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGAC
TCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG
CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGA
TATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCT
ATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG
GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCC
TGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC
AGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCA
CGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGA
AGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGA
AATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAA
GCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAA
AACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTA
CGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTC
AGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCG
GGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGG
CAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGA
GAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCG
GCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGA
ATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTC
CGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGC
CTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGT
TCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGA
AATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGA
GATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACC
GGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCC
CCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCC
TGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTA
CGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGG
GCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAG
CAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG
GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAAT
GCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGA
ACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAG
AAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGA
GTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACA

AGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCA
ATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACC
AGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGAC
ACGGATCGACCTGTCTCAGCTGGGAGGTGACTCTGGAGGATCTAGCGGAGGATCCTCTGGCA
GCGAGACACCAGGAACAAGCGAGTCAGCAACACCAGAGAGCAGTGGCGGCAGCAGCGGCG
GCAGCAGCACCCTAAATATAGAAGATGAGTATCGGCTACATGAGACCTCAAAAGAGCCAGAT
GTTTCTCTAGGGTCCACATGGCTGTCTGATTTTCCTCAGGCCTGGGCGGAAACCGGGGGCATG
GGACTGGCAGTTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCTACCCCCGTGTCC
ATAAAACAATACCCCATGTCACAAGAAGCCAGACTGGGGATCAAGCCCCACATACAGAGACT
GTTGGACCAGGGAATACTGGTACCCTGCCAGTCCCCCTGGAACACGCCCCTGCTACCCGTTAA
GAAACCAGGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAGTCAACAAGCGGGTGG
AAGACATCCACCCCACCGTGCCCAACCCTTACAACCTCTTGAGCGGGCTCCCACCGTCCCAC
CAGTGGTACACTGTGCTTGATTTAAAGGATGCCTTTTTCTGCCTGAGACTCCACCCCACCAGT
CAGCCTCTCTTCGCCTTTGAGTGGAGAGATCCAGAGATGGGAATCTCAGGACAATTGACCTG
GACCAGACTCCCACAGGGTTTCAAAAACAGTCCCACCCTGTTTAATGAGGCACTGCACAGAG
ACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTACAGTACGTGGATGACTTAC
TGCTGGCCGCCACTTCTGAGCTAGACTGCCAACAAGGTACTCGGGCCCTGTTACAAACCCTA
GGGAACCTCGGGTATCGGGCCTCGGCCAAGAAAGCCCAAATTTGCCAGAAACAGGTCAAGTA
TCTGGGGTATCTTCTAAAAGAGGGTCAGAGATGGCTGACTGAGGCCAGAAAAGAGACTGTGA
TGGGGCAGCCTACTCCGAAGACCCCTCGACAACTAAGGGAGTTCCTAGGGAAGGCAGGCTTC
TGTCGCCTCTTCATCCCTGGGTTTGCAGAAATGGCAGCCCCCCTGTACCCTCTCACCAAACCG
GGGACTCTGTTTAATTGGGGCCCAGACCAACAAAAGGCCTATCAAGAAATCAAGCAAGCTCT
TCTAACTGCCCCAGCCCTGGGGTTGCCAGATTTGACTAAGCCCTTTGAACTCTTTGTCGACGA
GAAGCAGGGCTACGCCAAAGGTGTCCTAACGCAAAAACTGGGACCTTGGCGTCGGCCGGTG
GCCTACCTGTCCAAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTTGCCTACGGATGGT
AGCAGCCATTGCCGTACTGACAAAGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCA
TTCTGGCCCCCCATGCAGTAGAGGCACTAGTCAAACAACCCCCCGACCGCTGGCTTTCCAAC
GCCCGGATGACTCACTATCAGGCCTTGCTTTTGGACACGGACCGGGTCCAGTTCGGACCGGT
GGTAGCCCTGAACCCGGCTACGCTGCTCCCACTGCCTGAGGAAGGGCTGCAACACAACTGCC
TTGATATCCTGGCCGAAGCCCACGGAACCCGACCCGACCTAACGGACCAGCCGCTCCCAGAC
GCCGACCACACCTGGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCGG
GAGCTGCGGTGACCACCGAGACCGAGGTAATCTGGGCTAAAGCCCTGCCAGCCGGGACATCC
GCTCAGCGGGCTGAACTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCT
AAATGTTTATACTGATAGCCGTTATGCTTTTGCTACTGCCCATATCCATGGAGAAATATACAGAA
GGCGTGGGTGGCTCACATCAGAAGGCAAAGAGATCAAAAATAAAGACGAGATCTTGGCCCTA
CTAAAAGCCCTCTTTCTGCCCAAAAGACTTAGCATAATCCATTGTCCAGGACATCAAAAGGGA
CACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCCGAAAGGCAGCCATCA
CAGAGACTCCAGACACCTCTACCCTCCTCATAGAAAATTCATCACCCTCTGGCGGCTCAAAAA

GAACCGCCGACGGCAGCGAATTCGAGCCCAAGAAGAAGAGGAAAGTCTAACCGGTCATCAT
CACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCAT
CTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTC
CTAATAAAATGAGAAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGG
GTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGG
TGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATACCGTCGACCTCTAGC
TAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC
CACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTAGGGTGCCTAATGAGTGAGCTAA
CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTG
CATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCC
TCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA
GGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACG
CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC
ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT
TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG
CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGA
GCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG
GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTT
ACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTAC
TTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA

AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCA
GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCGATCTCCCGAT
CCCCTAGGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCC
CTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAA
GGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAT
GTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG
GTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCT
GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTG
ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA
ATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGG
CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAGA
TCCGCGGCCGCTAATACGACTCACTATAGGGAGAGCCGCCACC (SEQ ID NO: 734)

N-терминальная NLS + Cas9 H840A

Гибкий линкер

M-MLV обратная транскриптаза + C-концевая NLS

Остов плазмиды (содержащий CMV промотор и AmpR)

pU6-HEK3_PEgPHK_CTTins:

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAG
ATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGT
AATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACC
GTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC*GGCC*
*CAGACTGAGCACGTGA*GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGGACCGAGTCGGTCC**TCTGCCATCAAAGCGTGCTCAGTCTG**TTTTTTT
AAGCTTGGGCCGCTCGAGGTACCTCTCTACATATGACATGTGAGCAAAAGGCCAGCAAAAGG
CCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA
GGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC
AGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA
CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT

TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT
GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA
ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGC
TCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGT
TCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCG
TCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCA
TGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCG
CAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAG
ATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC
GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT
GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTT
TCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGG
AAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTC
ATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC
CCCGAAAAGTGCCACCTGACGTCGCTAGCTGTACAAAAAAGCAGGCTTTAAAGGAACCAATT
CAGTCGACTGGATCCGGTACCAAGGTCGGGCAGGAA (SEQ ID NO: 735)

Последовательность промотора U6

Спейсерная последовательность

каркас сгРНК

3′ удлинение (содержит PBS и матрицу RT)

Остов (содержит AmpR)

pLenti-hSyn-N-PE2-NpuN-P2A-GFP-KASH_U6-DNMT1-PEгРНК:

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGC
AAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGG
CGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTA
ATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGG
TAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCC
CATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCAC

TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGC
CCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC
GGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGT
GGGAGGTCTATATAAGCAGCGCGTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCT
GGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTC
AAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGT
GTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGC
TCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGT
GAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTA
TTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAA
TATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGT
TAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAG
AAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAA
AGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAG
CAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATA
TAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGG
TGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGA
AGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTG
CAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGG
GGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCT
GGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAA
TTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAG
AATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTAT
ATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATA
GTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGA
CCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGAT
TAGTGAACGGATCGGCACTGCGTGCGCCAATTCTGCAGACAAATGGCAGTATTCATCCACAATTTTA
AAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGA
CATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGAC
AGCAGAGATCCAGTTTGGTTAATTAAGGTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGA
TTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAAC
ACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAA
AATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTAT
ATATCTTGTGGAAAGGACGAAACACC**GCGGGCTGGAGCTGTTCGCGCGTTTTAGAGCTAGAAA**
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAA

GATGCAAGCGCGAACAGCTCCAG<u>**TTTTTTTGAATTC**</u>*AGTGCAAGTGGGTTTTAGGACCAGGATGA*
GGCGGGGTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCCAACCCCCATTCCC
CAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGATGCGGCGAGGCGCGTGCGCACTGCC
AGCTTCAGCACCGCGGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGCGCCACCGCCGCCTCAGCACT
GAAGGCGCGCTGACGTCACTCGCCGGTCCCCCGCAAACTCCCCTTCCCGGCCACCTTGGTCGCGTCCG
CGCCGCCGCCGGCCCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGGCACGGGCGCGACCA
TCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCG
*TGCCTGAGAGCGCAG****TCGAGAATCTAGAGCGCTGCCACC***ATGAAACGGACAGCCGACGGAAGCGA
GTTCGAGTCACCAAAGAAGAAGCGGAAAGTCGACAAGAAGTACAGCATCGGCCTGGACATCGGC
ACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGT
GCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCG
GCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAA
CCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCA
CAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCA
ACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTG
GTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTC
CGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCAT
CCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG
CCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAA
CTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACG
ACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAG
AACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCC
CCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCT
CGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACG
CCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAA
AAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGC
GGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGG
CGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTC
CGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAA
GAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAG
AGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCA
CAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGG
AATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGA
CCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAG<u>*TGCCTGTCC*
TACGAGACAGAGATCCTGACAGTGGAGTATGGCCTGCTGCCAATCGGCAAGATCGTGGAGAAGAGGATC
GAGTGTACCGTGTACTCTGTGGATAACAATGGCAACATCTATACACAGCCCGTGGCACAGTGGCACGATA
GGGGAGAGCAGGAGGTGTTCGAGTATTGCCTGGAGGACGGCAGCCTGATCAGGGCAACCAAGGACCAC</u>

AAGTTCATGACAGTGGATGGCCAGATGCTGCCCATCGACGAGATTTTCGAGCGGGAGCTGGACCTGATG
AGAGTGGATAACCTGCCTAATAGCGGAGGCAGTAAAAGAACAGCAGACGGGAGTGAGTTTGAGCCCA
AGAAAAAGAGAAAGGTGGGATCCGGCGCAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATGTCG
AAGAGAATCCTGGACCGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCG
AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC
TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC
GTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC
TTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCA
ACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGG
GCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA
ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACAT
CGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG
CGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTAC
AAGTCCGGACTCAGATCTCGAGAGGAGGAGGAGGAGACAGACAGCAGGATGCCCCACCTCGACAGC
CCCGGCAGCTCCCAGCCGAGACGCTCCTTCCTCTCAAGGGTGATCAGGGCAGCGCTACCGTTGCAG
CTGCTTCTGCTGCTGCTGCTGCTCCTGGCCTGCCTGCTACCTGCCTCTGAAGATGACTACAGCTGCA
CCCAGGCCAACAACTTTGCCCGATCCTTCTACCCCATGCTGCGGTACACCAACGGGCCACCTCCCAC
*CTAA*ACGCGTTAAGTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTT
AACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCC
CGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCC
CGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCAT
TGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTC
ATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTG
TTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGA
CGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGC
TCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCC
CCGCGTCGACTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAA
AGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGT
CTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGC
CTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTA
GAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGGGCCCGTTTAAACCCGCTGAT
CAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGAC
CCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGT
AGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAA
TAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCT
CTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCA
GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGC

CACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCT
TTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
TAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTG
GAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTAT
TGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTT
AGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTC
AGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCA
ATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGC
CCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTG
AGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAG
CTTGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCAT
AGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCAC
CGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCG
TGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGTCCAGGAC
CAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGA
GTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCG
AGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCC
GAGGAGCAGGACTGACACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGC
TTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTC
TTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTT
CACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC
ATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGA
AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGT
GCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACC
TGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT
CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC
ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG
CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA
CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC
TACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA

AAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGT
AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC
CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG
CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT
TGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTA
CAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTC
AGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATG
CGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA
AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA
TCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT
CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATG
TTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG
TGCCACCTGAC (SEQ ID NO: 736)

Промотор U6

РЕгРНК

hSy промотор

N-концевой PE2

N-концевой Npu

P2A-GFP-KASH

pLenti-hSyn-C-PE2-NpuC:

GTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGC
AAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGG
CGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTA
ATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGG
TAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCC
CATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCAC
TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGC
CCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC

GGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGT
GGGAGGTCTATATAAGCAGCGCGTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCT
GGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTC
AAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGT
GTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGC
TCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGT
GAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTA
TTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAA
TATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGT
TAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAG
AAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAA
AGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAG
CAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATA
TAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGG
TGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGA
AGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTG
CAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGG
GGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCT
GGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAA
TTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAG
AATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTAT
ATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATA
GTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGA
CCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGAT
TAGTGAACGGATCGGCACTGCGTGCGCCAATTCTGCAGACAAATGGCAGTATTCATCCACAATTTTA
AAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGA
CATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGAC
AGCAGAGATCCAGTTTGGTTAATTAAGGTACCAGTGCAAGTGGGTTTTAGGACCAGGATGAGGCGG
GGTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCCAACCCCCATTCCCCA
AATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGATGCGGCGAGGCGCGTGCGCACTGC
CAGCTTCAGCACCGCGGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGCGCCACCGCCGCCTCAGC
ACTGAAGGCGCGCTGACGTCACTCGCCGGTCCCCCGCAAACTCCCCTTCCCGGCCACCTTGGTCGC
GTCCGCGCCGCCGCCGGCCCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGGCACGGGC
GCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGA
GTCGTGTCGTGCCTGAGAGCGCAG*TCGAGAATCTAGAGCGCTGCCACCATGAAACGGACAGCCGACG*
*GAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGTC*ATCAAGATTGCTACACGGAAATACCTGGGAAA
GCAGAACGTGTACGACATCGGCGTGGAGCGGGATCACAACTTCGCCCTGAAGAATGGCTTTATCGCCAG

*CAAT*TGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCA
CATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAG
GACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGA
ACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGG
AGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGT
CCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAG
CTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC
AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGG
CATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCC
GAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACA
GCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAA
AGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAG
AATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGT
GGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCA
GAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGA
TGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAAT
CTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGAC
AGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAA
CACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCA
AGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTAC
CACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACC
CTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGAT
CGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCA
TGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATC
GAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGC
GGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGG
CTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAG
GACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGG
TGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGG
GGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAG
GGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCT
GGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACT
GGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGG
GCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGA
CGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGG
ACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAA
TATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACAC
CACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCAC

CAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGTGACTCTG
GCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGCCCAAGAAGAAGAGGAAAGTCTA
AACGCGTTAAGTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAAC
TATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGT
ATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGT
TGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGC
CACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATC
GCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTG
TCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGT
CCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCT
GCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCC
GCGTCGACTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAG
GGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCT
CTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCT
CAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGA
GATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGGGCCCGTTTAAACCCGCTGATCA
GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCC
TGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAG
GTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA
GCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCT
AGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAG
CGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCC
ACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTT
TACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGAT
AGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGG
AACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATT
GGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTA
GGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCA
GCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAA
TTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCC
CATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGA
GCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGC
TTGTATATCCATTTTCGGATCTGATCAGCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATA
GTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACC
GCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTCGT
GGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGTCCAGGACC
AGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAG
TGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGA

GCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCCG
AGGAGCAGGACTGACACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCT
TCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCT
TCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTC
ACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCAT
GTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTG
CCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCT
GTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC
TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC
TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT
TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA
CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
GGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA
ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTC
ATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCC
AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCA
GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTT
GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAG
GCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG
CGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCC
AGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTG
GGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTG
AATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT
ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG

TGCCACCTGAC (SEQ ID NO: 737)

hSyn промотор

C-концевой Npu

C-концевой Cas9 дикого типа

pLenti-U6-DNMT1_вносящая разрыв_crРНК:

TAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAA
GAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAG
ACCCACCTCCCAACCCCGAGGGGACCCAGA*GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATA*
CGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGT
GACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTT
*ACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC**GCCGCGCG***
CGCGAAAAAGCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACCGAGTCGGTGCTTTTTTAAGCTTGGCGTAACTAGATCTTGAGACAAATGGCAGTATT
CATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACA
TAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGG
TTTATTACAGGGACAGCAGAGATCCACTTTGGCGCCGGCTCGAGGGGGCCCGGGTGCAAAGATGG
ATAAAGTTTTAAACAGAGAGGAATCTTTGCAGCTAATGGACCTTCTAGGTCTTGAAAGGAGTGGGA
ATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGG
AGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCG
TGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGA
ACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGC
CTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGAT
TCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCC
CTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGC
ACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCG
ACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTT
TTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCGGCGAGGCGGGGCCTG
CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGG
CCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTTGCGT
GAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCG
GGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGTCGCTTC
ATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTAC
GTCGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGAC
TGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTT
GGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTCAGGTGTCGTGACGTAC
GGCCACCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGCCG
TACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGATCCGGACCGCC

ACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAG
GTGTGGGTCGCGGACGACGGCGCCGCCGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGG
GGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAG
CAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGT
CGGAGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAG
GCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTAC
GAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCAT
GACCCGCAAGCCCGGTGCCTGAACGCGTTAAGTCGACAATCAACCTCTGGATTACAAAATTTGTGA
AAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTT
GTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCT
TTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAAC
CCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCT
ATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGC
ACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCA
CCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTC
CCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGAT
CTCCCTTTGGGCCGCCTCCCCGCGTCGACTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTA
GCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTG
CTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAG
GGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGT
TGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTAC
GTATAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGAAATGAATATCAGAGAGTG
AGAGGAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAA
TAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG
GCTCTAGCTATCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTC
CGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGGACGTACCCAATTCGCCCTATAGTGAGTCGT
ATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC
GCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCG
CGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTT
TCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCT
CCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGT
TCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTA
ATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAA
GGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
TTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATT
TGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAA

TAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGC
ATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTT
GGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCC
CGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTG
ACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCAC
CAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCA
TGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTT
TTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCA
TACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAA
CTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTG
CAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTG
AGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTAT
CTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTC
ACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGT
GAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTT
TTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGG
ATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTG
TTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGC
TCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAG
CTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGC
TTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAC
GAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACT
TGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGG
CCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATT
CTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGC
GCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG
TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCA
ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCG
TATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCC
AAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTGCAAGCTTAATGTAGTCTTATG
CAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAA
AGCACCGTGCATGCCGATTGGTGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGA
CGGGTCTGACATGGATTGGACGAACCACTGAATTGCCGCATTGCAGAGATATTGTATTTAAGTGCCT
AGCTCGATACATAAACGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTA
GGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTG
TTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTG

GCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACGCAGGACTCGG
CTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGAC
TAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGAT
CGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTAT
GGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGT
AGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATA
ATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAG
ACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCA
GACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAA
TTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGC
AGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGT
CAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGC
TGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCT
GGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAG
ATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACAC
TCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAA
ATGGGCAAGTTTGTGGAATTGGTTTAACA (SEQ ID NO: 738)

U6 промотор

сгРНК

[1226] Аминокислотные последовательности вариантов обратной транскриптазы вируса мышиного лейкоза Молони (M-MLV RT), используемых согласно настоящему изобретению.

PE1 M-MLV RT:

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIK
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP
TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF
KNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAK
KAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMA
APLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLP
DADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKL
NVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE
ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 739)

M3 M-MLV RT (D200N, T330P, L603W) (см. Baranauskas et al.[182]):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIK
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP
TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF
KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAK
KAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMA
APLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLP
DADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKL
NVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE
ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 740)

PE2 M-MLV RT (D200N, T306K, W313F, T330P, L603W):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIK
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP
TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF
KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAK
KAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMA
APLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLP
DADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKL
NVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE
ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 741)

M3-deadRT M-MLV RT:

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIK
QYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKLPGTNDYSPVQDLREVNKRVEDIHP
TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF
KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAK
KAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMA
APLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLP
DADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKL
NVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAE
ARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 742)

Ссылочные источники для примера 2

[1227] Каждый из следующих ссылочных источников цитируется в Примере 12, каждая из которых включена в настоящий документ посредством 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).

1. Реализация обратной транскрипции мутагенных цепей ДНК с гидовой РНК в качестве матрицы.

Предпосылки и обоснование

[1228] В рамках предложенной стратегии геномного редактирования с разрывом Cas9 нецелевая 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 пикирования 3 '-конец удаленного от PAM фрагмента высвобождается до диссоциации Cas9. Более того, предполагается, что гРНК могут быть удлинены для функционирования в качестве матриц для синтеза ДНК. Предшествующие исследования показали, что гРНК для SpCas9, SaCas9, и LbCasl2a (ранее Cpfl) переносят удлинение гРНК 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 гРНК с использованием комбинации *in vitro* и клеточных анализов (Фиг.2A-2C).

[1229] Конструкции сконструированных гРНК 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 цепь с разрывом ДНК, и матрицу для синтеза ДНК с помощью обратной транскрипции. 3' удлиненная гРНК содержит участок связывания праймера и матрицу обратной транскрипции. В некоторых случаях 3 ' РНК шпилька ядра гРНК модифицирована для соответствия последовательности-мишени ДНК, поскольку *in vitro* эксперименты показали, что обратная транскрипция приводит к ~3-нуклеотидному удлинению в ядре гРНК для конструктов 3' удлиненной гРНК (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

[1230] Никированная Cas9 ДНК праймирует обратную транскрипцию гРНК матриц. Для оценки доступности нецелевой цепи с разрывом 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 гРНК, 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 гРНК сохраняют аффинность связывания мишени (данные не показаны). Далее была протестирована активность TPRT на подвергнутых предварительному разрыву меченных Cy5 субстратах в виде ДНК дуплексов с использованием dCas9, 5'-удлиненных гРНК, 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 гРНК 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 разрыву субстраты удлинялись до полной 51 п.о. длины ДНК субстрата, что является сильным свидетельством того, что

комплементарная цепь ДНК, а не гРНК, 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:гРНК высвобождает 3 '-конец пикированной не являющейся мишенью цепи, а также тому, что не являющаяся мишенью цепь является доступной для обратной транскрипции.

[1231] Далее оценивали не подвергнутые разрыву 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 гРНК проводили *in vitro* анализы расщепления, как описано ранее[37]. Несмотря на то, что пикирование было нарушено по сравнению со стандартной гРНК, были образованы значительные продукты расщепления (Фиг.4D). Является важным то, что продукты RT также наблюдались при осуществлении TPRT реакций с 5'- удлиненными гРНК и Cas9(H840A), хотя и с более низким выходом, который, вероятно, может быть объяснен сниженной пикирующей активностью (Фиг.4D). Данный результат показывает, что комплексы в виде 5'-удлиненная гРНК:Cas9(H840A) могут пикировать ДНК и служить матрицей для обратной транскрипции.

[1232] Наконец 3' удлинения гРНК оценивали на предмет пикирования Cas9(H840A) и TPRT. В отличие от 5'-удлиненных гРНК, расщепление ДНК за счет 3'-удлиненных гРНК не ухудшалось до какой-либо детектируемой степени по сравнению со стандартной гРНК. Является значимым то, что матрицы в виде 3'-удлиненных гРНК также поддерживают эффективную обратную транскрипцию как с подвергнутыми предварительному пикированию, так и интактными субстратами в виде ДНК дуплексов при поступлении RT M-MLV в транс (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 гРНК. Добавление гомополимерного хвоста продукта с использованием концевой трансферазы с последующим удлинением фрагментом Klenow и секвенированием по Сэнгеру показало, что копировалась полная гРНК матрица для синтеза наряду с 3 концевыми нуклеотидами ядра гРНК. В будущем конец лоскута будет перепрограммирован для модифицирования концевой последовательности гРНК[38,39]. Этот результат показывает, что 3 '-удлиненные гРНК могут служить в качестве эффективных направляющих для нацеливания нуклеаз и могут служить матрицами для обратной транскрипции.

[1233] Cas9-TPRT использует ДНК с разрывом и гРНК в цис. Эксперименты на

основе двойного цвета использовали для определения того, осуществляется ли предпочтительным образом RT реакция с гРНК в цис (связывание с тем же комплексом) (см. Фиг.8). Два отдельных эксперимента проводили для 5'-удлиненных и 3'-удлиненных гРНК. Для указанного эксперименте в отдельных пробирках получали третичные комплексы dCas9, гРНК и ДНК субстрата. В одной пробирке гРНК кодирует длинный RT продукт, а ДНК субстрат помечен Cy3 (красный); в другой пробирке гРНК 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 гРНК матрицы, которая предварительно образовывала комплексы с ДНК субстратом, что свидетельствует о том, что RT реакция вероятно происходит в цис. Эти результаты поддерживают вывод о том. что один комплекс Cas9:гРНК может нацеливаться на участок ДНК и служить матрицей для обратной транскрипции мутагенной цепи ДНК.

Лечение TPRT системы с другими Cas

[1234] Эксперименты, аналогичные описанным в предыдущих разделах, будут проведены с использованием других Cas систем, включая Cas9 из *S. aureus* и Cas 12a 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

[1235] Сначала будет проанализирована группа коммерчески доступных или способных быть подвергнутыми очистке ферментов RT в транс 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 в цис.

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* сконструированных г*РНК в клетках

[1236] Сконструированные гРНК-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 гРНК будут подвергнуты всестороннему сравнению со стандартными гРНК 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 гРНК могут быть оптимизированы как в плане обратной транскрипции на основе матриц, так и эффективного нацеливания Cas9 в клетках.

[1237] Результаты *in vitro* анализов показаны на Фиг.5-7. *In vitro* эксперименты продемонстрировали, что пикированная не являющаяся мишенью цепь ДНК является гибкой и доступной для праймирования синтеза ДНК и что удлинение гРНК может служить матрицей для обратной транскрипции (см. Фиг.5). В рамках данного набора экспериментов использовались 5'-удлиненные гРНК (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 пикированию. 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 в транс. Сначала из очищенных компонентов были образованы комплексы dCas9:гРНК. 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. Следует отметить, что реакции осуществляли при отсутствии полученных с помощью dCas9 продуктов ДНК с длиной в виде 51 нуклеотида

независимо от использованной гРНК. 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 пикированию и используется никаза Cas9 (мутант SpyCas9 H840A). Как показано на геле, никаза эффективно расщепляет цепь ДНК при использовании стандартной гРНК (gRNA_0, дорожка 3). Наблюдается множество продуктов расщепления, что соответствует предыдущим биохимическим исследованиям SpyCas9. 5' удлинение нарушает разрезающую 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 пикированию 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 гРНК. 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

[1238] (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, гРНК 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 гРНК: это скорее всего является проблемой для 5'-удлиненных гРНК. На основе структурных данных[24] могут быть сконструированы и подвергнуты скринингу мутанты Cas9 для идентификации вариантов с большей устойчивостью к удлинению гРНК. Кроме того, библиотеки гРНК могут быть подвергнуты скринингу в клетках на предмет линкеров, которые повышают связанную с нацеливанием активность.

Значимость

[1239] Данные предварительные результаты показывают, что никазы Cas9 и удлиненные гРНК 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 в цис, 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 праймированного редактирования в клетках человека

предпосылки и обоснования

[1240] В рамках предлагаемой стратегии комплекс в виде модифицированной RT-Cas9:гРНК обеспечит внесение мутагенных 3'лоскутов ДНК в геномные целевые участки. Предполагается, что мутагенные 3' лоскуты, содержащие одно несовпадение, будут встроены с помощью механизма репарации ДНК посредством энергетически доступного уравнивания со смежными 5' лоскутами, которые будут предпочтительно удалены (фиг. 1C-1D). Механизмы репликации и репарации ДНК встречают 5'-лоскуты оцДНК при процессировании фрагментов Оказаки[46] и при репарации эксцизии длинных последовательностей оснований (LP-BER)[47]. 5'-Лоскуты являются предпочтительными субстратами для широко экспрессируемой эндонуклеазы FEN 1 лоскутов, 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 Ligl[49]. Действуя в качестве "пояса с инструментами", PCNA ускоряет последовательное расщепление и лигирование лоскутов, 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' лоскутами с последующей скоординированной эксцизией и лигированием 5' лоскутов. 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

[1241] ДНК-лоскуты 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' лоскуты, которые похожи на продукт TPRT. Был создан двойной флуоресцентный репортерный белок, который кодирует стоп-кодон между GFP и mCherry. Мутагенные лоскуты 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 (5. *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 лоскутов предшествует MMR. Данный результат свидетельствует об осуществимости редактирования ДНК с использованием 3' мутагенных цепей.

- ***Системные результаты с модельными лоскутными субстратами***

[1242] На основе описанных выше предварительных результатов будет проанализирован более широкий спектр лоскутных субстратов в клетках HEK для определения принципов эффективного редактирования. 3'-лоскуты оцДНК будут систематически варьироваться для определения влияния образований пар с несовпадениями, расположения мутагенного нуклеотида в лоскутах, 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 гРНК матриц для обратной транскрипции.

[1243] *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 лоскутов-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 гРНК, фермент RT MLV, dNTP], Удлинение гРНК кодирует мутацию в стоп-кодоне. Цепь лоскутов 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) пикированной плазмидой приводит к получению колоний, экспрессирующих только зеленый GFP. Реакция TRT с 5'-удлиненными или 3'-удлиненными гРНК продуцирует смесь зеленых и желтых колоний. Последние экспрессируют как GFP, так и mCherry. Еще больше желтых колоний наблюдается при использовании 3'-удлиненной гРНК. Также показан положительный контроль, который не содержит стоп-кодон.

[1244] Этот результат показывает, что длинные двухцепочечные субстраты могут подвергаться TPRT, и что продукты TPRT индуцируют редактирование в эукариотических клетках.

[1245] Был осуществлен другой эксперимент, аналогичных указанному выше эксперименту на основе праймированного 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 гРНК. 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 гРНК для TRT редактирования. Другие точечные мутации или индели не наблюдались.

Реализация праймированного редактирования в клетках HEK с использованием конструкций RT-Cas9

[1246] 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 гРНК (являющихся

усеченными для предотвращения разрыва). 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 гРНК и продуктов слияния RT-Cas9 (как описано выше). Затем для каждого участка генома будут проанализированы удлиненные гРНК и пары RT-Cas9 на предмет изменения в виде одного нуклеотида. Результаты редактирования будут проанализированы с помощью MiSeq.

[1247] 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 в транс), и модифицированной гРНК с 3' удлинение (см. Фиг. 14). Предварительные исследования показали, что длина участка связывания праймера в удлинении гРНК являлась важной в плане повышения эффективности редактирования в клетках человека (см. Фиг. 15).

Оптимизация параметров праймированного редактирования в клетках HEK

[1248] После идентификации конструкций Cas9-RT, которые могут обеспечивать реализацию праймированного 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 праймированное редактирование в клетках млекопитающих.

[1249] Предварительные результаты. Были сконструированы дополнительные гРНК 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 нуклеотидный спейсер), что, как было показано ранее, предотвращает образование разрезов посредством Cas9, но сохраняет связывание (см. фиг. 16). Для анализа данной стратегии был выбран локус HEK3.

Потенциальные сложности и альтернативы

[1250] 1) Расщепление гРНК в клетках: если концы удлиненной гРНК являются усеченными в клетке, то могут быть внесены стабилизирующие вторичные структуры, или могут быть протестированы синтетические гРНК 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

[1251] Если праймированное 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 праймированное 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.

[1252]

3. Реализация сайт-специфического редактирования патогенных мутаций в культивируемых клетках человека

Предпосылки и обоснование

[1253] 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 праймированном редактировании все транзиции и трансверсии являются теоретически возможными, как и малые вставки и делеции. Более того, в плане PAM окно праймированного 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) мутация Glu6Val в гемоглобине бета при серповидноклеточной анемии (для корректирования требуется трансверсия A•T - T•A), (2) наиболее частый вариант HislO69Gln в ATP7B при болезни Уилсона (для корректирования требуется трансверсия G•C to 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 праймированного 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).***

[1254] 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.***

[1255] 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 ML V обеспечивают улучшение редактирования.***

[1256] Мутантные обратные транскриптазы, описанные в Baranauskas, et al (doi: 10.1093/ protein/gzs034), тестировали в виде С-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' гидовую РНК для праймированного 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.

Разрыв комплементарной цепи с использованием второй гРНК обеспечивает улучшение редактирования.

[1257] 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 гидовые РНК, одна из которых служит матрицей для реакции обратной транскрипции, а другая нацелена на комплементарную цепь ДНК для разрыва. Было протестировано разрезание на различном расстоянии от нуклеотида-мишени (оранжевые треугольники) (см. Фиг.20). Показана эффективность редактирования в паре нуклеотидных оснований-мишеней (синие столбцы) наряду с показателями образования инделей (оранжевые столбцы). Пример «нет» не содержит пикирующей комплементарную цепь гидовой РНК. Показатели редактирования количественно анализировали с помощью высокопроизводительного секвенирования ампликонов геномной ДНК.

[1258] На Фиг. 21 демонстрируются обработанные данные высокопроизводительного секвенирования, показывающие требуемую трансверсивную мутацию в виде конверсии T в A и общее отсутствие других значительных побочных продуктов редактирования генома.

[1259] 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 п.о. против хода транскрипции от 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.

[1260] Новая технология праймированного 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 праймированного редактирования начинаются в участке разрыва ДНК (на 3 основания против хода транскрипции 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, точность и гибкость.

Корректирование патогенных мутаций в полученных у пациента линиях клеток.

[1261] 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 нецелевого редактирования.

[1262] Потенциальное нецелевое редактирование будет подвергнуто скринингу с использованием стандартных способов, таких как GUIDE-seq[55] и CIRCLE-seq[56] с применением гРНК-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.

[1263] (1) Низкая эффективность редактирования: праймирующий редактор (PE) может требовать оптимизации для каждой мишени. В этом случае могут быть протестированы библиотеки гРНК 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.

[1264] Оптимизация конструкций гРНК 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

[1265] Изменения в виде одного нуклеотида в отдельных генах вызывают множество генетических заболеваний. Разработка описанной в настоящем документе 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

[1266] 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 направляемой гомологией 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.

[1267]

ПРИМЕР 3 – ПЕПТИДНОЕ МЕЧЕНИЕ С ПОМОЩЬЮ PE

[1268] Описанные в настоящем документе системы праймированного 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.

[1269] Фиг.25 представляет собой схему, показывающую конструкцию гРНК для генов пептидного мечения в эндогенных геномных локусах и пептидное мечение с помощью геномного редактирования на основе TPRT (то есть праймированного 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), pp 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, pi-кольцо кодирует пептидную последовательность (FCPF) (SEQ ID NO: 622), которая может быть помечена пентафтор-ароматическими субстратами (см.: Nat. Chem. 2016 Feb,8(2):120-8. doi: 10.1038/nchem.2413).

[1270] На Фиг. 26 показано точное внесение His6-метки и FLAG-метки в геномную ДНК. С помощью матрицы обратной транскрипции была сконструирована гидовая РНК, нацеленная на локус HEK3, которая кодирует либо вставку 18-нуклеотидной His-метки, или вставку 24-нуклеотидной FLAG-метки. Эффективность редактирования в трансфицированных HEK клетках анализировали с помощью ампликонного секвенирования. Следует обратить внимание, что полная 24-нуклеотидная последовательность FLAG-метки располагается вне наблюдаемого изображения (секвенирование подтвердило осуществление полной и точной вставки).

ПРИМЕР 4. МЕЧЕНИЕ И МАНИПУЛИРОВАНИЕ РНК С ПОМОЩЬЮ PE

[1271] В настоящем документе 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.

[1272] 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 (и) 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) двухцепочечных 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.

[1273] Праймированное редактирование представляет собой новую технологию геномного редактирования, которая обеспечивает нацеленное редактирование геномных локусов за счет переноса генетической информации из РНК в ДНК. С помощью праймированного редактирования гены РНК могут быть помечены рядом компонентов, таких как РНК аптамеры, рибозимы, или другие мотивы РНК. Праймированное 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. В настоящем документе описан новый способ внесения мотивов РНК в генетические последовательности, которые обеспечивают мечение или иное манипулирование РНК с помощью праймирующих 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 PEгРНК (направляющих РНК для праймированного 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.28А). 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 сгРНК или индуцирования процессирования РНК за счет саморасщепления или РНКаз (см. Фиг.28В). Вследствие гибкости праймированного редактирования является невозможным привести исчерпывающий перечень мотивов РНК, которые могут быть внесены в геном. В настоящем документе 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.

[1274] 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, полиА-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 полиА- хвоста SV40:

SV40 ПОЛИ-А ХВОСТ	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCA ATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATT CTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTA (SEQ ID NO: 331)

[1275] Сигналы посттрансляционной модификации, помимо сигналов полиаденилирования, также могут кодироваться PE. Они включают сигналы, включающие модификации на основе N6-метиладенозина, N1- метиладенозина, 5- метилцитозина, и псевдоуридина[59]. За счет использования PE 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.

[1276] PE 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].

[1277] Примеры:

[1278] Сигнал polyA SV40 приводит к транспорту.

SV40 POLYA TAIL	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTT GTCCAAACTCATCAATGTATCTTA (SEQ ID NO: 331)

[1279] U1 snРНК 3′ приводит к удержанию.

U1 SNPHK 3′ БОКС	TTCATTCAGCAAGTTCAGAGAAATCTGAACTTGCTGGATTTTTGGAGC AGGGAGATGGAATAGGAGCTTGCTCCGTCCACTCCACGCATCGACCTG GTATTGCAGTACCTCCAGGAACGGTGCACCCACTTTCTGGAGTTTCAA AAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTT GGTTTGTGTCTTGGTTGGCGTCTTAA (SEQ ID NO: 625)

[1280] 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.

[1281] Аптамер Broccoli:

BROCCOLI АПТАМЕР	GAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCT C (SEQ ID NO: 357)

[1282] PE 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]).

[1283] PE может обеспечивать вставку последовательностей, которые кодируют сгРНК 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.

[1284] 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[7].

[1285] Рибозим HDV:

HDV РИБОЗИМ	GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATG CTTCGGCATGGCGAATGGGAC (SEQ ID NO: 365)

[1286] Ссылочные источники для примера 4

[1287] Следующие ссылочные источники включены в настоящий документ посредством ссылки во всей своей полноте.

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 5. ВСТАВКА ИММУНОЭПИТОПА С ПОМОЩЬЮ PE

[1288] 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 двухцепочечного 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.

[1289] Недавно разработанные продукты слияния Cas9 с обратными транскриптазами («праймирующие редакторы»), соединенные с PEгРНК, 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.

[1290] Настоящее изобретение относится к способу применения праймированного 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 праймированного редактирования такие вставки могли быть реализованы лишь неэффективным образом с высокими показателями образования инделей из DSB. Праймированное 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 праймированного 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. Праймирующий редактор представляет собой наиболее эффективную и точную технологию реализации таких нацеленных вставок в клетках млекопитающих.

[1291] Ключевая концепция настоящего изобретения состоит в применении праймирующих 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 праймированному редактированию белок (а не только вставленный эпитоп) и потенциально клетки, из которых он был экспрессирован.

[1292] 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.

[1293] 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.

[1294] Отличительные признаки настоящего 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 в настоящем документе 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.

[1295] Праймированное 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.

[1296] 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.

[1297] 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.

[1298] 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 резвившимися генетическими заболеваниями. Ожидается, что текущий стандарт в плане поддерживающих лекарственных средств может быть использован совместно с праймированным 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 праймированного редактирования:

Вакцина	Заболевание	Аминокислотная последовательность эпитопа	Пример последовательности нуклеиновой кислоты (8)
1	Столбнячный токсин	QYIKANSKFIGITEL (SEQ ID NO: 396)	CATGATATAAAAGCAAATTCTAAATTTA TAGGTATAACTGAACTA (SEQ ID NO: 397)
2	Мутант дифтерийного токсина CRM197	GADDVVDSSKSFV MENFSSYHGTKPG YVDSIQKGIQKPKS GTQGNYDDDWKEF YSTDNKYDAAGYS VDNENPLSGKAGG VVKVTYPGLTKVL ALKVDNAETIKKEL GLSLTEPLMEQVGT EEFIKRFGDGASRV VLSLPFAEGSSSVE YINNWEQAKALSV ELEINFETRGKRGQ DAMYEYMAQACA GNRVRRSVGSSLSC INLDWDVIRDKTKT KIESLKEHGPIKNK MSESPNKTVSEEKA KQYLEEFHQTALEH PELSELKTVTGTNP VFAGANYAAWAV NVAQVIDSETADNL EKTTAALSILPGIGS VMGIADGAVHHNT EEIVAQSIALSSLMV AQAIPLVGELVDIG FAAYNFVESIINLFQ VVHNSYNRPAYSP GHKTQPFLHDGYA	GGCGCCGACGACGTGGTGGACAGCAGC AAGAGCTTCGTGATGGAGAACTTCAGC AGCTACCACGGCACCAAGCCCGGCTAC GTGGACAGCATCCAGAAGGGCATCCAG AAGCCCAAGAGCGGCACCCAGGGCAAC TACGACGACGACTGGAAGGAGTTCTAC AGCACCGACAACAAGTACGACGCCGCC GGCTACAGCGTGGACAACGAGAACCCC CTGAGCGGCAAGGCCGGCGGCGTGGTG AAGGTGACCTACCCCGGCCTGACCAAG GTGCTGGCCCTGAAGGTGGACAACGCC GAGACCATCAAGAAGGAGCTGGGCCTG AGCCTGACCGAGCCCCTGATGGAGCAG GTGGGCACCGAGGAGTTCATCAAGAGG TTCGGCGACGGCGCCAGCAGGGTGGTG CTGAGCCTGCCCTTCGCCGAGGGCAGC AGCAGCGTGGAGTACATCAACAACTGG GAGCAGGCCAAGGCCCTGAGCGTGGAG CTGGAGATCAACTTCGAGACCAGGGGC AAGAGGGGCCAGGACGCCATGTACGAG TACATGGCCCAGGCCTGCGCCGGCAAC AGGGTGAGGAGGAGCGTGGGCAGCAGC CTGAGCTGCATCAACCTGGACTGGGAC GTGATCAGGGACAAGACCAAGACCAAG ATCGAGAGCCTGAAGGAGCACGGCCCC ATCAAGAACAAGATGAGCGAGAGCCCC AACAAGACCGTGAGCGAGGAGAAGGCC AAGCAGTACCTGGAGGAGTTCCACCAG ACCGCCCTGGAGCACCCCGAGCTGAGC

		VSWNTVEDSIIRTG FQGESGHDIKITAE NTPLPIAGVLLPTIP GKLDVNKSKTHISV NGRKIRMRCRAIDG DVTFCRPKSPVYVG NGVHANLHVAFHR SSSEKIHSNEISSDSI GVLGYQKTVDHTK VNSKLSLFFEIKS (SEQ ID NO: 630)	GAGCTGAAGACCGTGACCGGCACCAAC CCCGTGTTCGCCGGCGCCAACTACGCCG CCTGGGCCGTGAACGTGGCCCAGGTGA TCGACAGCGAGACCGCCGACAACCTGG AGAAGACCACCGCCGCCCTGAGCATCC TGCCCGGCATCGGCAGCGTGATGGGCA TCGCCGACGGCGCCGTGCACCACAACA CCGAGGAGATCGTGGCCCAGAGCATCG CCCTGAGCAGCCTGATGGTGGCCCAGG CCATCCCCCTGGTGGGCGAGCTGGTGG ACATCGGCTTCGCCGCCTACAACTTCGT GGAGAGCATCATCAACCTGTTCCAGGT GGTGCACAACAGCTACAACAGGCCCGC CTACAGCCCCGGCCACAAGACCCAGCC CTTCCTGCACGACGGCTACGCCGTGAGC TGGAACACCGTGGAGGACAGCATCATC AGGACCGGCTTCCAGGGCGAGAGCGGC CACGACATCAAGATCACCGCCGAGAAC ACCCCCCTGCCCATCGCCGGCGTGCTGC TGCCCACCATCCCCGGCAAGCTGGACGT GAACAAGAGCAAGACCCACATCAGCGT GAACGGCAGGAAGATCAGGATGAGGTG CAGGGCCATCGACGGCGACGTGACCTT CTGCAGGCCCAAGAGCCCCGTGTACGT GGGCAACGGCGTGCACGCCAACCTGCA CGTGGCCTTCCACAGGAGCAGCAGCGA GAAGATCCACAGCAACGAGATCAGCAG CGACAGCATCGGCGTGCTGGGCTACCA GAAGACCGTGGACCACACCAAGGTGAA CAGCAAGCTGAGCCTGTTCTTCGAGATC AAGAGC (SEQ ID NO: 399)
3	паротит	GTYRLIPNARANLT A (SEQ ID NO: 400)	GGCACCTACAGGCTGATCCCCAACGCC AGGGCCAACCTGACCGCC (SEQ ID NO: 401)
4	паротит	PSKFFTISDSATFAP GPVSNA (SEQ ID	Ccgagcaaatttttaccattagcgatagcgcgacctttgcgccg ggcccggtgagcaacgcg (SEQ ID NO: 403)

		NO: 402) PSKLFIMLDNATFAPGPVVNA (SEQ ID NO: 404)	Ccgagcaaactgtttattatgctggataacgcgacctttgcgccgggcccggtggtgaacgcg (SEQ ID NO: 405) ВЫБРАННЫЕ ПРИМЕРЫ ИЗ МНОГООБРАЗИЯ ГЕМАГЛЮТИНИН–НЕЙРАМИНИДАЗЫ (HN) СРЕДИ ВСПЫШЧНЫХ ШТАММОВ (ТАБЛИЦА 1), отклонение МЕЖДУ ВАКЦИННЫМ ШТАММОМ JL5 И ВСПЫШЧНЫМИ ШТАММАМИ (ТАБЛИЦА 2)
5	Вирус краснухи (RV)	TPPPYQVSCGGESDRASARVIDPAAQS (SEQ ID NO: 406)	ACCCCCCCCCCCTACCAGGTGAGCTGCGGCGGCGAGAGCGACAGGGCCAGCGCCAGGGTGATCGACCCCGCCGCCCAGAGC (SEQ ID NO: 407)
6	Гемагглютинин	PEYAYKIVKNKKMEDGFLQGMVDGWYGHHSNEQGSGLMENERTLDKANPNNDLCSWSDHEASSNNTNQEDLLQRESRRKKRIGTSTLNQRGNCNTKCQTEEARLKREEVSLVKSDQCSNGSLQCRANNSTEQVD (SEQ ID NO: 408)	CCCGAGTACGCCTACAAGATCGTGAAGAACAAGAAGATGGAGGACGGCTTCCTGCAGGGCATGGTGGACGGCTGGTACGGCCACCACAGCAACGAGCAGGGCAGCGGCCTGATGGAGAACGAGAGGACCCTGGACAAGGCCAACCCCAACAACGACCTGTGCAGCTGGAGCGACCACGAGGCCAGCAGCAACAACACCAACCAGGAGGACCTGCTGCAGAGGGAGAGCAGGAGGAAGAAGAGGATCGGCACCAGCACCCTGAACCAGAGGGGCAACTGCAACACCAAGTGCCAGACCGAGGAGGCCAGGCTGAAGAGGGAGGAGGTGAGCCTGGTGAAGAGCGACCAGTGCAGCAACGGCAGCCTGCAGTGCAGGGCCAACAACAGCACCGAGCAGGTGGAC (SEQ ID NO: 409)
7	Нейраминидаза	TKSTNSRSGGISGPDNEAPVGEAPSPYGDNPRPNDGNNIRIGSKGYNGIITDTIEESCSCYPDAKVVKSV	ACCAAGAGCACCAACAGCAGGAGCGGCGGCATCAGCGGCCCCGACAACGAGGCCCCCGTGGGCGAGGCCCCCAGCCCCTACGGCGACAACCCCAGGCCCAACGACGGCAACAACATCAGGATCGGCAGCAAGGGC

		ELDSTIWTSGSSPN QKIITIGWDPNGWT GTPMSPNGAYGTD GPSNGQANQHQAE SISAGNSSLCPIRDN WHGSNRSWSWPD GAE (SEQ ID NO: 410)	TACAACGGCATCATCACCGACACCATC GAGGAGAGCTGCAGCTGCTACCCCGAC GCCAAGGTGGTGAAGAGCGTGGAGCTG GACAGCACCATCTGGACCAGCGGCAGC AGCCCCAACCAGAAGATCATCACCATC GGCTGGGACCCCAACGGCTGGACCGGC ACCCCCATGAGCCCCAACGGCGCCTAC GGCACCGACGGCCCCAGCAACGGCCAG GCCAACCAGCACCAGGCCGAGAGCATC AGCGCCGGCAACAGCAGCCTGTGCCCC ATCAGGGACAACTGGCACGGCAGCAAC AGGAGCTGGAGCTGGCCCGACGGCGCC GAG (SEQ ID NO: 411)
8	TAP (презентация транспортного антигена) на гемагглутинине вируса H5N1	EKIVLLLAMMEKIV LLLAKCQTPMGAIK AVDGVTNKCPYLG SPSF (SEQ ID NO: 412)	GAGAAGATCGTGCTGCTGCTGGCCATG ATGGAGAAGATCGTGCTGCTGCTGGCC AAGTGCCAGACCCCCATGGGCGCCATC AAGGCCGTGGACGGCGTGACCAACAAG TGCCCCTACCTGGGCAGCCCCAGCTTC (SEQ ID NO: 413)
9	TAP (презентация транспортного антигена) на нейраминидазе вируса H5N1	IRPCFWVELNPNQK IITIRPCFWVELICYP DAGEIT (SEQ ID NO: 414)	ATCAGGCCCTGCTTCTGGGTGGAGCTGA ACCCCAACCAGAAGATCATCACCATCA GGCCCTGCTTCTGGGTGGAGCTGATCTG CTACCCCGACGCCGGCGAGATCACC (SEQ ID NO: 415)
10	Эпитопы гемагглутинина к HLA класса I	MEKIVLLLAEKIVL LLAMCPYLGSPSFK CQTPMGAIKAVDG VTNK (SEQ ID NO: 416)	ATGGAGAAGATCGTGCTGCTGCTGGCC GAGAAGATCGTGCTGCTGCTGGCCATGT GCCCCTACCTGGGCAGCCCCAGCTTCAA GTGCCAGACCCCCATGGGCGCCATCAA GGCCGTGGACGGCGTGACCAACAAG

			(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	MEKIVLLLAEKIVLLLAMMVSLVKSDQCPYLGSPSFIGTSTLNQRKCQTPMGAIKAVDGVTNK (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)

[1299] Ниже приведены дополнительные примеры эпитопов, которые могут быть встроены в ген-мишень для иммуноэпитопного мечения:

Ссылочные источники для примера 5

[1300] Следующие ссылочные источники включены в настоящий документ

посредством ссылки во всей своей полноте.

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 AVP8* 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 El 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 6. ПРИМЕНЕНИЕ PE ДЛЯ ОБЕСПЕЧЕНИЯ ХИМИЧЕСКИ ИНДУЦИРУЕМОЙ ДИМЕРИЗАЦИИ БЕЛКОВ-МИШЕНЕЙ *IN VIVO*

[1301] Описанные в настоящем документе праймирующие 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 праймированного редактирования, как описано в настоящем документе. С помощью описанных в настоящем документе праймирующих 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 7. ВСТАВКА, УДАЛЕНИЕ И ЗАМЕНА ПОСЛЕДОВАТЕЛЬНОСТИ ДНК С ПОМОЩЬЮ ДВОЙНЫХ ПРАЙМИРУЮЩИХ РЕДАКТОРОВ

[1302] В этом примере описывается использование мультилоскутного праймированного редактирования, в частности, с использованием двойных праймирующих редакторов, для точной вставки новой последовательности ДНК, точного удаления последовательности эндогенной геномной ДНК или замены последовательности эндогенной геномной ДНК новой последовательностью ДНК. См. фиг. 90. Настоящее изобретение направлено на достижение давней цели инженерии генома: программируемая вставка, делеция или замена последовательностей ДНК в геномах живых клеток. Эта технология может применяться для модификации или редактирования генома любой живой клетки в исследовательских, терапевтических или промышленных целях.

[1303] Ранее для вставки, удаления или замены последовательностей ДНК в геномах живых клеток использовалась направляемая гемологией репарация (HDR) индуцированных нуклеазами двухцепочечных разрывов ДНК (DSB). Несмотря на универсальность, низкая эффективность HDR ограничивает его применение во многих типах клеток млекопитающих, особенно в терапевтически значимых типах клеток человека[1-4] . Более того, генерация DSB ведет к преимущественной репарации с помощью подверженных ошибкам механизмов соединения концов, которые производят продукты репарации вставок-делеций (инделей), которые впоследствии не могут быть преобразованы в желаемую отредактированную последовательность ДНК[5-10] . В результате индели обычно представляют собой большинство модифицированных продуктов геномной ДНК, когда нуклеазы используются в клетках млекопитающих. Более того, DSB, генерируемые нуклеазами, могут приводить к нежелательным геномным изменениям, таким как большие делеции и хромосомные транслокации[11] , а также к активации p53[12,13] . Структурные варианты (SV) также могут быть созданы или скорректированы с использованием программируемых нуклеаз, не полагаясь на гомологичную рекомбинацию, хотя стратегии, основанные на нуклеазах, имеют аналогичные недостатки для этих приложений. Например, стратегии двойного разрезания можно использовать для инвертирования целевой последовательности геномной ДНК, но большинство аллелей продукта содержат большие делеции между участками разрывов вместо желаемой инвертированной

последовательности или вставки в отдельных участков разрывов

[1304] Праймированное редактирование представляет собой недавно описанную технологию редактирования генома, которая позволяет вставлять, удалять или заменять последовательности геномной ДНК, не требуя подверженных ошибкам двухцепочечных разрывов ДНК[14] . При праймированном редактировании используется сконструированный слитый белок Cas9-никаза-обратная транскриптаза (PE1 или PE2) в паре со сконструированной гидовой РНК праймированного редактирования (РЕгРНК), которая направляет Cas9 к целевому геномному сайту и кодирует информацию для установки желаемого редактирования. Праймированное редактирование проходит через многоэтапный процесс редактирования: 1) домен Cas9 связывается и надрезает целевой участок геномной ДНК, который определяется спейсерной последовательностью РЕгРНК, 2) домен обратной транскриптазы использует геномную ДНК с разрывом в качестве праймера для инициации синтеза отредактированной цепи ДНК с использованием сконструированного удлинения РЕгРНК в качестве матрицы для обратной транскрипции — это создает одноцепочечный 3'-лоскут, содержащий отредактированную ДНК. последовательность, 3) репарация клеточной ДНК устраняет промежуточный 3'-лоскут за счет вытеснения вида 5'-лоскута, которое происходит за счет инвазии отредактированным 3'-лоскутом, вырезания 5'-лоскута, содержащего исходную последовательность ДНК, и лигирования нового 3'-лоскута для включения отредактированной цепи ДНК, образующий гетеродуплекс из одной отредактированной и одной неотредактированной цепи, и 4) репарация клеточной ДНК заменяет неотредактированную цепь в гетеродуплексе, используя отредактированную цепь в качестве матрицы для репарации, завершая процесс редактирования.

[1305] Эффективное включение желаемого редактирования требует, чтобы вновь синтезированный 3'-лоскут содержал часть последовательности, которая гомологична участку геномной ДНК. Эта гомология позволяет отредактированному 3'-лоскуту конкурировать с эндогенной цепью ДНК (соответствующим 5'-лоскутом) за включение в дуплекс ДНК. Поскольку отредактированный 3'-лоскут будет содержать меньше гомологии последовательностей, чем эндогенный 5'-лоскут, ожидается, что конкуренция будет благоприятствовать цепи 5'-лоскута. Таким образом, 3'-лоскут, который содержит редактирование, может быть не в состоянии эффективно внедриться и сместить цепь 5'-лоскута. Более того, успешная инвазия 3'-лоскута и удаление 5'-лоскута включает редактирование только одной цепи двухцепочечной геномной ДНК. Постоянная установка редактирования требует восстановления клеточной ДНК для замены неотредактированной комплементарной цепи ДНК с использованием отредактированной цепи в качестве

матрицы. В то время как клетка может способствовать замене неотредактированной цепи на отредактированную цепь (стадия 4 выше) путем введения разрыва в неотредактированную цепь, примыкающую к редактированию, с использованием вторичной сгРНК (система PE3), этот процесс все еще зависит от второй стадии репарации ДНК. Эти стадии репарации ДНК могут быть особенно неэффективны для изменений, которые требуют уравновешивания длинных 5'- и 3'-промежуточных лоскутов или содержат длинные негомологичные области, такие как длинные вставки или длинные делеции.

[1306] В этом примере описана система двойного праймированного редактирования (или система двухлоскутного праймированного редактирования), которая решает проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в нередактируемую комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. В системе двухлоскутного праймированного редактирования две PEгРНК используются для нацеливания на противоположные нити геномного сайта и управления синтезом двух комплементарных 3'-лоскутов, содержащих отредактированную последовательность ДНК (**фиг. 1**). В отличие от классического праймированного редактирования, нет необходимости в том, чтобы пара отредактированных цепей ДНК (3'-лоскуты) напрямую конкурировала с 5'-лоскутами в эндогенной геномной ДНК, поскольку вместо этого для гибридизации доступна отредактированная комплементарная цепь. Поскольку обе цепи дуплекса синтезируются как редактируемая ДНК, система двухлоскутного праймированного редактирования устраняет необходимость замены неотредактированной комплементарной цепи ДНК, необходимой для классического праймированного редактирования. Вместо этого механизму репарации клеточной ДНК нужно только вырезать парные 5'-лоскуты (исходная геномная ДНК) и лигировать парные 3'-лоскуты (отредактированная ДНК) в локус. Следовательно, также нет необходимости включать последовательности, гомологичные геномной ДНК, во вновь синтезированные нити ДНК, что позволяет проводить селективную гибридизацию новых цепей и облегчает редактирование, которое содержит минимальную геномную гомологию. Нуклеазоактивные версии праймирующих редакторов, которые разрезают обе нити ДНК, также могут быть использованы для ускорения удаления исходной последовательности ДНК.

[1307] В зависимости от ориентации смещенных фрагментов ДНК, созданных с помощью двойных праймирующих редакторов, система может либо заменить последовательность между двумя фрагментами на новую желаемую последовательность (5'-выступ со смещенной ориентацией), либо вставить новую последовательность ДНК с

сопутствующей дупликацией сайта-мишени (3’ выступ в шахматном порядке). На фиг. 90 показана ориентация (5' смещенных разрывов), где последовательность между разрывами заменена. Следует отметить, что новые 3'-лоскуты не обязательно должны полностью перекрываться по всей длине новой последовательности ДНК, вместо этого они могут быть комплементарными только в части новой последовательности. Таким образом, каждый лоскут может быть короче полной длины новой последовательности ДНК.

Двухлоскутное праймированное редактирование в локусе сайт3 3 *HEK293*

[1308] Двухлоскутное праймированное редактирование в локусе сайта *3 HEK293* в клетках человека (HEK293T) использовали для достижения высокоэффективного введения гетерологичных последовательностей ДНК. Например, замена последовательности из 90 п.о. в локусе новой последовательностью из 22 п.о., кодирующей метку 6xHis, происходила с эффективностью более 80% (таблица аллелей на **фиг. 92).** Точно так же выход >60% был достигнут при замене этой же последовательности из 90 п.о. последовательностью пептида GFP-11 (52 п.о.). Вставку метки GFP-11 также осуществляли на С-конце гена *MYC* с эффективностью >20%, что происходило с сопутствующей запрограммированной делецией части последовательности Мус 3'UT.

РЕгРНК двухлоскутного праймированного редактирования

[1309] Общие конструкции РЕгРНК, использованные для двухлоскутного праймированного редактирования, показаны на фиг. 93. РЕгРНК, применяемые для двухлоскутного праймированного редактирования, имеют конструкции, аналогичный тем, которые используются для классического праймированного редактирования, однако нет необходимости, чтобы матричная область RT кодировала какую-либо гомологию с локусом-мишенью. Вместо этого две РЕгРНК могут содержать матрицы RT, которые кодируют синтез 3'-лоскутов, чьи 3'-концы представляют собой последовательности, обратно комплементарные друг другу. Эта комплементарность между 3'-лоскутами способствует их отжигу и замене эндогенной последовательности ДНК предполагаемой новой последовательностью ДНК. Это требует, чтобы 5'-области RT-матриц в двух РЕгРНК представляли собой последовательности, обратно комплементарные друг другу, и эта степень комплементарности может варьироваться (**фиг. 93**).

Применение двухлоскутного праймированного редактирования для введения сайтов рекомбинации

[1310] Двухлоскутное праймированное редактирование имеет множество потенциальных применений, таких как установка пептидных меток, РНК-меток, иммуноэпитопов, доменов димеризации и целевых сайтов рекомбиназы.

[1311] Одним из таких применений является установка последовательностей

рекомбиназы или интегразы в указанные пользователем места в геноме. На **фиг. 94** проиллюстрирована установка сайтов *attB* (38 п.о.) и *attP* (50 п.о.) рекомбиназы Bxb1 в область-мишень генома человека (*HEK293T сайт 3* или *HEK3*) с одновременной делецией 90 п.о. промежуточной последовательности между двумя участками разрыва. Различные степени комплементарности между 3'-лоскутами обеспечивают успешное редактирование, хотя более длинные последовательности комплементарности дают более благоприятное соотношение желаемых редактирований и инделей. Другие применения двухлоскутного праймированного редактирования включают эндогенную маркировку генов пептидными или белковыми последовательностями или замену экзонов новыми последовательностями ДНК, которые потенциально могут заменить несколько вариантов, для которых мутация попадает в последовательность экзона.

[1312] Двухлоскутное праймированное редактирование можно использовать для введения одного или двух сайтов рекомбиназы в целевые положения генома человека. Если вставляются одиночные рекомбиназные сайты, их можно использовать в качестве посадочных сайтов для опосредованной рекомбиназой реакции между геномным рекомбиназным сайтом и вторым рекомбиназным сайтом в экзогенно поставляемой ДНК, такой как плазмида. Это позволяет целенаправленно интегрировать ДНК. Если два сайта рекомбиназы встроены в соседние области ДНК, в зависимости от ориентации сайтов рекомбиназы, их можно использовать для опосредованного рекомбиназой вырезания или инверсии промежуточной последовательности, или для опосредованного рекомбиназой обмена кассет с экзогенной ДНК для интеграции ДНК. Интеграция совместимых сайтов рекомбиназы на разных хромосомах делает возможной целевую и направленную хромосомную транслокацию. Двухлоскутное праймированное редактирование можно использовать для эффективного введения сайтов рекомбиназы в ряд локусов генома человека (фиг. 94). Таким образом, сочетание интеграции сайта рекомбиназы с помощью двухлоскутного праймированного редактирования с ферментами рекомбиназы ДНК представляет собой мощный подход для достижения многих типов редактирования SV и целевой интеграции ДНК.

ПРИМЕР 8. ВСТАВКА, УДАЛЕНИЕ, ИНВЕРСИЯ, ТРАНСЛОКАЦИЯ И ИНТЕГРАЦИЯ ПОСЛЕДОВАТЕЛЬНОСТИ ДНК С ИСПОЛЬЗОВАНИЕМ ЧЕТЫРЕХЛОСКУТНОГО ПРАЙМИРОВАННОГО РЕДАКТИРОВАНИЯ

[1313] В этом примере описывается использование мультилоскутного праймированного редактирования, в частности четырехлоскутного праймированного редактирования, для целевой вставки, делеции, дупликации и замены последовательностей

эндогенной геномной ДНК, а также для целевой инверсии последовательности геномной ДНК, целевой хромосомной транслокации и целевой интеграции экзогенной ДНК. Эта технология может решить давние задачи инженерии генома: программируемая установка или коррекция структурных вариантов (обычно определяемых как > 50 п.о. изменения геномной последовательности, часто представленные большими делециями, вставкам/дупликациями, инверсиями и транслокациями) и целенаправленная интеграция информации ДНК в определенном месте генома человека. Последнее может позволить использовать терапевтические стратегии увеличения генов и другие приложения биотехнологии. Эта технология потенциально может применяться для модификации или редактирования генома любой живой клетки в исследовательских, терапевтических или промышленных целях.

[1314] Настоящее изобретение относится к системе мультилоскутного праймированного редактирования, которая решает проблемы, связанные с уравновешиванием лоскутов и последующим включением редактирования в нередактируемую комплементарную цепь геномной ДНК путем одновременного редактирования обеих цепей ДНК. Он также расширяет возможности традиционного простого редактирования, позволяя вновь синтезированным 3- лоскутам направлять более крупные перестановки последовательностей.

[1315] Был разработан подход четырехлоскутное праймированное редактирование, который может напрямую выполнять желаемые изменения последовательности без необходимости фермента рекомбиназы (как при двухлоскутном праймированном редактировании), но с аналогичной универсальностью. При четырехлоскутном праймированном редактировании четыре различных последовательности РЕгРНК доставляются в клетки вместе с редактором PE2. Одна пара РЕгРНК является матрицей для синтеза комплементарных лоскутов ДНК, как при двухлоскутном праймированном редактировании, в то время как другая пара РЕгРНК является матрицей для синтеза ортогональной пары комплементарных лоскутов ДНК. Расположение этих лоскутов, их относительная ориентация и комплементарность в паре определяют, какой класс перестройки произойдет. Соединения перестроенной ДНК содержат последовательность, кодируемую 3'-лоскутами матрицы РЕгРНК, которые направляют ориентацию перестройки и могут служить полезной последовательностью ДНК для последующих процессов. Направленную инверсию геномной последовательности выполняли с использованием стратегии четырехлоскутного праймированного редактирования (фиг. 95). Инверсия последовательности размером 2,7 т.о. в локусе AAVS1 в клетках HEK293T была достигнута с использованием четырех РЕгРНК, которые помещают последовательности Bxb1 attB и

attP в инверсионные соединения. Верификацию ожидаемых соединений проводили с помощью секвенирования ампликонов, которое продемонстрировало правильные соединения, содержащие последовательности сайта attB или attP, с ≥90% прочтений секвенирования (фиг. 96).

[1316] Целенаправленная интеграция экзогенных плазмид ДНК также была достигнута с помощью четырехлоскутного праймированного редактирования. В этом подходе две РЕгРНК нацелены на сайт эндогенной геномной ДНК, а две РЕгРНК нацелены на область груза плазмиды. 3'-лоскуты ДНК синтезируют как на геномной ДНК, так и на плазмидной ДНК, так что плазмида и геномная ДНК будут соединены мостиком с помощью гибридизированных 3'-лоскутоы (фиг. 97A). Иссечение 5'-выступов и лигирование надрезов приводит к получению интегрированного плазмидного продукта. Ориентацию интеграции можно контролировать путем спаривания матричных последовательностей РЕгРНК.

[1317] Плазмиду, кодирующую экспрессию слитого белка H2B-GFP, интегрировали в локус AAVS1 в клетках HEK293T с использованием четырехлоскутного праймированного редактирования. Это привело к тому, что 1,6% клеток получили длительную флуоресценцию GFP, локализованную в ядре, по оценке проточной цитометрии. ПЦР предсказанного соединения показала ожидаемый продукт, содержащий последовательность BxB1 attP с матрицей РЕгРНК (фиг. 97B).

[1318] Хромосомные транслокации также выполняли с использованием четырехлоскутного праймированного редактирования (фиг. 98A) для направления хромосомной транслокации между локусом MYC на хромосоме 8 и локусом TIMM44 на хромосоме 19 (фиг. 98B).

[1319] Двухлоскутное праймированное редактирование в локусе IDS выполняли с помощью 12 пар РЕгРНК и PE2. Клетки HEK293T трансфицировали PE2 и различными парами РЕгРНК (*например*, в первой колонке РЕгРНК A1_a и РЕгРНК B2_a с матрицами для установки сайта attP в прямом направлении). Эффективность измеряли с помощью HTS. Эти данные показали, что двухлоскутное праймированное редактирование может успешно вставлять представляющую интерес последовательность в локус IDS с эффективностью до ~80% (фиг. 99).

[1320] Опосредованную двумя лоскутами дупликацию также выполняли в локусе AAVS1 в клетках 293T (фиг. 100A). Двухлоскутную РЕгРНК с PE2 использовали для индукции дупликации генетических последовательностей в AAVS1 (фиг. 100B).

[1321] Проводили транслокацию MYC-CCR5, индуцированную мультилоскутным праймированным редактированием (фиг. 101). Транслокация MYC-CCR5 индуцировалась двухлоскутными РЕгРНК и PE2. События транслокации MYC-CCR5 индуцировались

четырехлоскутными РЕгРНК. В клетках HEK293T тестировали четыре различных набора РЕгРНК. Продукты соединения транслокации между производными chr8 и chr3 амплифицировали с помощью праймеров соединения. Процент прочтений, выровненных с ожидаемыми аллелями соединения, показан на графике. Это показывает, что четырехлоскутное праймированное редактирование может опосредовать транслокацию генов MYC и CCR5 с чистотой продукта приблизительно 100% на соединении 1 и ~ 50% на соединении 2. Репрезентативный график аллелей показывает последовательности, выровненные с ожидаемыми последовательностями аллелей на соединении 1.

[1322] Двухлоскутное и четырехлоскутное праймированное редактирование проводили в других линиях клеток человека (фиг. 102A-102B). Двухлоскутное праймированное редактирование выполняли в четырех различных клеточных линиях человека (фиг. 102A). Клетки HEK293T и HeLa трансфицировали двойными РЕгРНК и PE2 для редактирования трех разных геномных локусов (IDS, MYC и TIMM44). Клетки U2OS и K562 нуклеофицировали теми же компонентами. Двухлоскутное праймированное редактирование показало надежную эффективность редактирования во всех четырех клетках человека в целевых локусах, особенно в клетках HEK293T и K562. Клеточные механизмы, обеспечивающие двухлоскутное праймированное редактирование, сохраняются во многих типах клеток человека. Направленную инверсию последовательности размером 2,7 т.о. с несколькими лоскутами (четырехлоскутная РЕгРНК) также выполняли в локусе AAVS1 в клетках HeLa (фиг. 102B).

[1323] Эффективность инверсии измеряли с помощью HTS в локусе CCR5 (фиг. 103A-103B). Процент ожидаемого инверсионного аллеля редактирования измеряли с помощью HTS. Четыре набора четырехлоскутной РЕгРНК PE2 трансфицировали в клетки HEK293T соответственно. Осуществляли опосредованную двумя лоскутами дупликацию последовательности (~100 нуклеотидов (nt)) в локусе CCR5 в клетках HEK293T (фиг. 103A). Эффективность редактирования достигает ~ 1,5% с помощью анализа секвенирования парных концов HTS 300 циклов. Опосредованную четырьмя лоскутами инверсию последовательности (~95-117 нуклеотидов) также выполняли в локусе CCR5 в клетках HeLa (фиг. 103B). Эффективность редактирования достигает ~ 1,2% с помощью анализа секвенирования парных концов HTS 300 циклов. Этот результат показывает, что мультилоскутное праймированное редактирование может успешно обеспечивать дупликацию и инверсию точно в локусе CCR5. Специфичность редактирования высока, когда целевая последовательность дублируется (процент инделей <2%).

[1324] Затем клетки HEK293T трансфицировали плазмидами, которые экспрессируют Exo1, Fen1, белок красной флуоресценции (контроль), ДНК2, Mlh1 neg и

ингибитор P53 с РЕгРНК и PE2 (фиг. 104A-104D). Эффективность редактирования измеряли с помощью HTS. Эффективность редактирования сравнивали между кандидатом и контролем RFP. Клетки HEK293T также трансфицировали плазмидами siRNA и РЕгРНК и PE2 для каждого локуса-мишени (фиг. 104E). В качестве контроля использовали нецелевую миРНК (siNT). Эффективность редактирования измеряли с помощью HTS. Эффективность редактирования сравнивали между каждым нокдауном siRNA и siNT ctrl в каждом локусе-мишени. Клетки HEK293T трансфицировали двойной РЕгРНК, PE2 и плазмидами, которые экспрессируют Exo1, Fen1, белок красной флуоресценции (ctrl), DNA2, Mlh1 neg и ингибитор P53, соответственно. Эффективность редактирования измеряли с помощью HTS. Эффективность редактирования сравнивали между кандидатом и контролем RFP. Было обнаружено, что сверхэкспрессия FEN1 улучшает эффективность двухлоскутного праймированного редактирования во всех четырех целевых локусах (MYC, TIMM44, IDS, CCR5).

[1325] Опосредованную двумя лоскутами дубликацию последовательности также выполняли в *локусе AAVS1* (фиг. 105A). При использовании двойных РЕгРНК, которые генерируют две уникальные структуры 3'-лоскута, дупликацию последовательности ~300 п.о. индуцировали в локусе AAVS1 в клетках 293T (фиг. 105B). Ожидаемые аллели амплифицировали со специфическими праймерами и подвергали HTS. Приблизительно 94% прочтений соответствовали ожидаемым аллелям с дупликацией. В необработанных образцах продукты дубликации не наблюдались.

[1326] Целенаправленную инверсию геномной последовательности IDS также выполняли с четырехлоскутным праймированным редактированием (фиг. 106). Было показано, что приблизительно у 13% пациентов с синдромом Хантера имеется инверсия последовательностей гена IDS (Bondeson et al., Human Molecular Genetics, 1995). Для индуцирования этой патогенной инверсии геномной последовательности IDS размером приблизительно 40 т.о. в клетках HEK293T применяли четырехлоскутное праймированное редактирование. Шесть наборов четырехлоскутных РЕгРНК тестировали путем трансфекции клеток HEK293T РЕгРНК и PE2. Праймеры использовали для специфической амплификации инвертированных последовательностей в местах соединения «ab» и соединения «cd». Приблизительно 95% ожидаемых последовательностей инвертированных аллелей наблюдались на обоих соединениях с IDS_QF1. Другие наборы РЕгРНК также дают высокий процент ожидаемых аллельных последовательностей на обоих соединениях. В необработанных образцах продукты инвертированного соединения не наблюдаются.

[1327] Было также обнаружено, что модификация 3'-мотива РЕгРНК повышает эффективность двухлоскутного праймированного редактирования в локусе IDS (фиг. 107).

Для дальнейшего повышения эффективности двухлоскутного праймированного редактирования был введен мотив evoPreQ1 для защиты 3'-конца РЕгРНК. При сравнении эффективности редактирования, генерируемой немодифицированными и модифицированными evoPreQ1 двойными РЕгРНК, наблюдалось общее увеличение эффективности редактирования модифицированными РЕгРНК в целевом локусе IDS. Повышение эффективности двухлоскутного праймированного редактирования может достигать в 5,3 раза.

Ссылочные источники (Примеры 7 и 8):

[1328] 1. Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. PNAS 91, 6064–6068 (1994).

[1329] 2. Heyer, W.-D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. Annu. Rev. Genet. 44, 113–139 (2010).

[1330] 3. 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).

[1331] 4. Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. Nature Medicine 21, 121–131 (2015).

[1332] 5. 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).

[1333] 6. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161, 1169–1175 (2002).

[1334] 7. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11, 636–646 (2010).

[1335] 8. Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823–826 (2013).

[1336] 9. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823 (2013).

[1337] 10. Jinek, M. et al. RNA-programmed genome editing in human cells. eLife 2, e00471 (2013).

[1338] 11. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nature Biotechnology 36, 765–771 (2018).

[1339] 12. Ihry, R. J. et al. p53 inhibits CRISPR–Cas9 engineering in human

pluripotent stem cells. Nature Medicine 24, 939–946 (2018).

[1340] 13. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nature Medicine 24, 927–930 (2018).

[1341] 14. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).

ПРИМЕР 9. ПРОГРАММИРУЕМАЯ БОЛЬШАЯ ДЕЛЕТАЦИЯ, ЗАМЕНА, ИНТЕГРАЦИЯ И ИНВЕРСИЯ ДНК С ДВОЙНЫМ ПРАЙМИРОВАННЫМ И САЙТ-СПЕЦИФИЧЕСКИМИ РЕКОМБИНАЗАМИ

[1342] Целенаправленная делеция, замена, интеграция или инверсия последовательностей ДНК в определенных местах генома, в принципе, может использоваться для изучения или лечения многих генетических заболеваний человека. Однако по-прежнему сложно точно и эффективно проводить эти преобразования последовательности ДНК в контексте генома клетки млекопитающих. В настоящем документе представлено двойное праймированное редактирование (twinPE), метод программируемой замены или вырезания последовательности ДНК в эндогенных геномных сайтах человека без необходимости разрыва двухцепочечной ДНК. twinPE использует белок праймирующего редактора (PE) и две годовые РНК праймированного редактирования (PEгРНК), которые являются матрицей для синтеза комплементарных лоскутов ДНК на противоположных цепях геномной ДНК, что приводит к замене последовательности эндогенной ДНК между PE-индуцированными участками разрыва на PEгРНК-кодированные последовательности. Здесь показано, что twinPE может выполнять точные делеции до 780 п.о. и точные замены последовательности геномной ДНК новыми последовательностями до 108 п.о. Комбинируя одиночный или мультиплексный twinPE с сайт-специфичными сериновыми рекомбиназами, показано, что направленную интеграцию ДНК-плазмид размером с ген в безопасные локусы, включая *AAVS1* , *CCR5* и *ALB* , и, кроме того, инверсию 39 т.о. в *IDS* , как также продемонстрировано, можно использовать для коррекции распространенного аллеля синдрома Хантера. TwinPE существенно расширяет возможности редактирования генома, не требуя разрывов двухцепочечной ДНК, и в сочетании с другими инструментами редактирования генома позволяет корректировать или дополнять сложные варианты патогенных аллелей в клетках человека.

[1343] Генетические варианты человека, связанные с заболеванием, возникают в результате различных изменений последовательностей, начиная от замен одной пары оснований и заканчивая делециями и перестройками в масштабе мегануклеотидов.

Подходы к редактированию генома, которые могут установить, исправить или дополнить эти патогенные варианты в клетках человека, могут улучшить понимание генетических заболеваний человека и могут привести к разработке новых терапевтических средств. Недавно для редактирования генома было разработано несколько подходов к редактированию генома, основанных на системах CRISPR-Cas, включая нуклеазы, редакторы оснований и праймирующие редакторы. Нуклеазы CRISPR-Cas при использовании с одной гидовой РНК можно использовать для разрушения генов путем образования двухцепочечных разрывов ДНК (DSB), которые приводят к стохастическим вставкам. Редакторы оснований можно использовать для точной установки изменений пары оснований C•G-в-T•A или A•T-в-G•C или редактирования пары оснований C•G-в-G•C. Основные редакторы могут точно установить любую из двенадцати правок замены пар оснований, а также небольшие и средние вставки и удаления. Однако сами по себе эти стратегии не применимы напрямую для внесения более существенных изменений генома, таких как большие делеции, вставки, замены или инверсии.

[1344] Ранее были разработаны стратегии парных нуклеаз Cas9 для целенаправленной делеции последовательностей геномной ДНК длиной от ~ 50 до > 100 000 пар оснований. Кроме того, путем предоставления линейной последовательности ДНК донора можно осуществлять целевую вставку новых последовательностей ДНК в одиночные участки разрезов или между парными участками разрезов посредством процессов репарации ДНК, направленных на соединение концов или гомологию. Несмотря на универсальность во многих отношениях, подходы с одной нуклеазой и парной нуклеазой сопровождаются рядом недостатков. Во-первых, использование парных нуклеаз для делеций может генерировать множество побочных продуктов на мишени, включая желаемую делецию с инделями, вставки в отдельных сайтах DSB без желаемой делеции, нежелательную инверсию последовательности ДНК между сайтами DSB и непреднамеренную интеграцию экзогенной последовательность ДНК на сайтах DSB. Кроме того, точное расположение делеций ограничено доступностью PAM и соответствующими сайтами разрезов, генерируемыми нуклеазой. Аналогичные ограничения и побочные продукты существуют для нокаута ДНК-донора, который часто происходит без контроля ориентации последовательности, когда используются независимые от гомологии подходы. Наконец, одновременная генерация множественных DSB в целевых и нецелевых участках, что становится более вероятным с дополнительными годовыми РНК, может приводить к нецелевым модификациям генома, хромосомным перестройкам или хромотрипсису. Таким образом, существуют значительные проблемы, связанные с контролем продукта и чистотой при применении стратегий редактирования на

основе нуклеаз для вырезания, замены или вставки более крупных последовательностей ДНК.

[1345] Было показано, что праймированное редактирование способно делать точные вставки до ~ 40 п.о. и делеции до ~ 80 п.о. в клетках человека с высоким соотношением желаемого продукта к побочному продукту. В принципе, праймированное редактирование можно использовать, чтобы обойти многие проблемы, связанные с нуклеазами, избегая создания DSB. Однако еще не было показано, что праймированное редактирование обеспечивает более крупные вставки и делеции размера типичных последовательностей, кодирующих гены, и предполагаемый механизм простых реакций праймированного редактирования делает эти более крупные изменения ДНК затруднительным, требуя длинных матриц обратной транскрипции РЕгРНК. Сайт-специфические ферменты ДНК-рекомбиназы обладают способностью выполнять несколько трансформаций ДНК, включая вырезание, инверсию, интеграцию или обмен большими последовательностями ДНК. Однако проблемы с перепрограммированием сайт-специфических рекомбиназ в настоящее время затрудняют применение для инженерии генома и применений для редактирования.

[1346] В настоящем документе представлена разработка двойного праймированного редактирования (twinPE), которая позволяет с высокой эффективностью удалять, заменять или вставлять последовательность ДНК в эндогенных геномных сайтах в клетках человека. В дополнение к этим изменениям показано, что twinPE можно использовать для интеграции одного или нескольких сайтов распознавания рекомбиназы с высокой эффективностью в целевых сайтах генома человека, которые впоследствии можно использовать в качестве субстратов для сайт-специфических ферментов сериновых рекомбиназ, которые позволяют целевую интеграцию генетического груза или целевую инверсию последовательности геномной ДНК.

Результаты

[1347] Праймированное редактирование использует белок праймирующего редактора, состоящий из слияния каталитически нарушенной никазы Cas9 и фермента обратной транскриптазы MMLV дикого типа (PE1) или сконструированного (PE2), а также гидовой РНК праймированного редактирования (РЕгРНК), которая определяет целевой участок генома и кодирует желаемое редактирование. При распознавании целевого участка комплексы PE•РЕгРНК разрывают цепь ДНК, содержащую PAM, и непосредственно обратно транскрибируют матрицу RT РЕгРНК в геномную ДНК, используя цепь с разрывом в качестве праймера. После обратной транскрипции вновь синтезированный 3'-лоскут вторгается в соседнюю ДНК, которой он в значительной степени гомологичен, чтобы заменить избыточную последовательность 5'-лоскута. Противоположная

неотредактированная цепь затем восстанавливается с использованием отредактированной цепи ДНК в качестве матрицы. Таким образом, этот предложенный путь редактирования предоставляет две возможности, при которых репарация клеточной ДНК может отклонить желаемое редактирование и вернуть последовательность ДНК в ее исходную форму.

[1348] Было предсказано, что обход потенциально неблагоприятных стадий репарации ДНК может позволить праймированному редактированию происходить с большей эффективностью и сделать возможными различные классы редактирования генома. Кроме того, стратегия двойного праймированного редактирования (twinPE), в которой используется пара РЕгРНК, каждая из которых нацелена на одну или другую цепь ДНК, и является матрицей для синтеза последовательности 3'-лоскута ДНК, которая комплементарна последовательности, используеющей в качестве матрицы другую РЕгРНК (**фиг. 108A**). Было высказано предположение, что если бы вновь синтезированные цепи ДНК сильно отличались от эндогенного сайта-мишени, 3'-лоскуты предпочтительно гибридизовались бы с образованием промежуточного продукта, обладающего отожженными 3'-выступами новой последовательности ДНК и отожженными 5'-выступами исходной последовательности ДНК (**фиг. 108A**). Поскольку обе отредактированные цепи синтезируются комплексами праймирующего редактора, нет необходимости во вторжении в целевой участок цепей лоскута, содержащих редактирование, или в том, чтобы редактирование копировалось в комплементарную цепь ДНК. Иссечение исходной последовательности ДНК (отожженные 5'-концы), заполнение пробелов полимеразами и лигирование пары разрывов приведет к замене эндогенной последовательности между сайтами разрывов парными последовательностями 3'-лоскутов (**фиг. 108A**). Из-за гибкости конструкции матрицы редактирование может, в принципе, ввести новую последовательность ДНК, заменить часть последовательности ДНК или удалить часть последовательности ДНК.

[1349] Чтобы оценить стратегию twinPE, локус *сайта 3 HEK* (ранее обозначаемый как *HEK3*) в HEK293T был нацелен на замену 90 п.о. эндогенной последовательности последовательностью прикрепления Bxb1 *attB* из 38 п.о. (**фиг. 108B**). Для каждого протоспейсера конструировали три РЕгРНК с матрицами RT, которые содержали 30, 34 или 38 нуклеотидов последовательности *attB* длиной 38 п.о. (**фиг. 108C**). Предполагается, что парные комбинации этих РЕгРНК генерируют 3'-лоскуты с перекрывающейся комплементарностью в диапазоне от 22 до 38 п.о. (**фиг. 108C**). Интересно, что когда обе РЕгРНК были доставлены в клетки вместе с PE2 при трансфекции плазмидной ДНК, наблюдалась высокая эффективность вставки сайта *attB*, при этом некоторые комбинации РЕгРНК приводили к конверсии >80% в желаемый продукт на основании секвенирования

ампликона (**фиг. 108C**). Аналогичная стратегия вставки последовательности прикрепления attP *Bxb1* из 50 п.о. позволила добиться редактирования с эффективностью, достигающей 58% (**фиг. 108C**). Примечательно, что необязательно, чтобы каждая РЕгРНК кодировала полную последовательность вставки, поскольку частично перекрывающиеся комплементарные лоскуты делают возможной включение полноразмерной последовательности *attB* или *attP*. Однако 3'-лоскуты с большим перекрытием приводили к несколько более высокой эффективности редактирования и меньшему количеству инделей для вставки последовательностей Bxb1 *attB* и *attP* в *сайте 3 HEK*.

[1350] Затем TwinPE исследовали, чтобы определить, может ли оно поддерживать вставку более крупных последовательностей ДНК, чем те, которые были продемонстрированы с использованием систем PE2 или PE3. Сначала исследовали PE3-опосредованные вставки фрагментов кодирующей последовательности FKBP12 размером от 12 до 321 п.о. При нацеливании на локус *сайта 3 HEK* с помощью PE3 умеренная эффективность была достигнута для более коротких вставок из 12 и 36 пар оснований (31,8% и 17,3% соответственно), но очень мало желаемого продукта наблюдалось для вставки из 108 пар оснований (0,8%) и без полноразмерного продукта для вставки 321 п.о. (**фиг. 109A**). Напротив, twinPE обеспечивает эффективность вставки 16% для фрагмента из 108 п.о., что составляет 20-кратное улучшение по сравнению с PE3. Более того, вставки размером 113 п.о. и 103 п.о., содержащие парные сайты рекомбиназы Bxb1, были достигнуты с аналогичной эффективностью (10,7% и 9,7%, соответственно) в локусе CCR5 (данные не показаны). Эти результаты иллюстрируют возможность вставки большей последовательности с использованием twinPE по сравнению с PE3.

[1351] Одним из возможных применений стратегии twinPE является замена экзонной кодирующей последовательности перекодированной последовательностью ДНК, которая поддерживает последовательность белка и может исправить любую мутацию между участками разрыва, нацеленного индуцированными twinPE. Для проверки этого подхода был выбран ген *PAH*, мутации в котором вызывают генетическое нарушение обмена веществ фенилкетонурию (PKU). Рекодирование частей экзона 4 и экзона 7 тестировали в *PAH* клеток *HEK293T* дикого типа, где можно обнаружить корректирующее редактирование. Путем тестирования различных длин перекрытия лоскутов и добавления мотивов evoPreQ1 к 3'-концу РЕгРНК была достигнута желаемая перекодировка последовательности с умеренной эффективностью, достигающей средней эффективности до 9,4% для перекодирования 64 п.о. в экзоне 4, средней эффективности до 22,7% для перекодирования 46 п.о. в экзоне 7 и средней эффективности до 27,4% для перекодирования 64 п.о. в экзоне 7 (**фиг. 109B**). Основываясь на мутациях, аннотированных в базе данных

bioPKU, совместное редактирование этих областей TwinPE позволяет корректировать отдельные мутации, связанные с PKU. Дополнительные исследованные экзоны также могут быть перекодированы, хотя и с меньшей эффективностью (**фиг. 112**). Эти результаты показывают, что twinPE в принципе можно использовать для коррекции множественных мутаций с помощью одной пары РЕгРНК.

[1352] В дополнение к вставке и замене последовательностей ДНК, twinPE также должен иметь возможность точного редактирования удаления. Стратегии делеции парными нуклеазами генерируют делеции, которые охватывают два сайта DSB и, таким образом, ограничиваются доступностью PAM, а желаемые делеции часто сопровождаются нежелательными побочными продуктами инделей. Напротив, twinPE обладает потенциалом для выполнения делеций с большей гибкостью и точностью из-за отсутствия генерации DSB и способности включать дополнительную последовательность ДНК в сайты разрыва, индуцированные twinPE. Чтобы оценить twinPE на предмет точных делеций, сравнивали три стратегии с использованием парных РЕгРНК: стратегию twinPE с «одиночным якорем», стратегию twinPE с «гибридным якорем» и недавно описанную стратегию «PrimeDel» (**фиг. 109C**). Каждая стратегия отличается последовательностью, кодируемой в комплементарных лоскутах, что обеспечивает гибкость в расположении делеции по отношению к участкам разрыва.

[1353] Используя стратегию с одним якорем, комплементарные лоскуты из 13 нуклеотидов использовали для делеции 77 п.о. последовательности, прилегающей к одному из участков разрыва, индуцированных РЕгРНК, с эффективностью 14,8%, а комплементарные лоскуты из 34 нуклеотидов использовали для точного вырезания 56 п.о. последовательности с эффективностью 18,8% (**фиг. 109D**). Используя стратегию гибридного якоря, 64 п.о. были удалены между РЕгРНК-индуцированными сайтами разрывов, так что продукт сохраняет 13 п.о. последовательности 3' каждого разрыва с эффективностью 11,7% (**фиг. 109D**). Наконец, стратегию PrimeDel тестировали на делецию 90 п.о. между участками разрыва, индуцированными РЕгРНК, которая происходила с эффективностью 40,4% (**фиг. 109D**). Следует отметить, что стратегия PrimeDel разрушает последовательности PAM на обеих цепях, что, вероятно, повышает эффективность, а также уменьшает индели. Эффективность редактирования можно повысить в 1,5-2,5 раза путем добавления мотива evoPreQ1 к 3'-концу РЕгРНК, хотя в некоторых случаях это сопровождается увеличением инделей (**фиг. 109D**). Вместе эти данные показывают, что twinPE предлагает стратегию для выполнения целенаправленных делеций с высокой гибкостью, которая не зависит от наличия идеально расположенных сайтов разреза нуклеазами.

[1354] Наконец, терапевтически релевантный локус *DMD* был нацелен на применение twinPE для более крупных делеций. Патогенные аллели *DMD*, ответственные за мышечную дистрофию Дюшенна, обычно содержат большие делеции в областях, содержащих экзоны, что приводит к транскриптам иРНК со сдвигом рамки считывания. Поскольку производство полноразмерного белка дистрофина без замены делетированных экзонов может привести к частичному восстановлению функции белка, в качестве потенциальной терапевтической стратегии было предложено разрушение второго экзона, которое восстанавливает рамку считывания. Три стратегии делеции twinPE были изучены вместе с ранее описанной стратегией делеции нуклеазы Cas9 для вырезания экзона 51 при *DMD*. При использовании подходов делеции twinPE с одним якорем наблюдалась эффективность от 14,4% до 24% для делеции последовательности длиной 780 п.о., содержащей экзон 51, при *DMD* (**фиг. 109E**). В то время как стратегия парной нуклеазы Cas9 достигла более высокой эффективности удаления (в среднем 49,8%) по сравнению со стратегиями twinPE (в среднем от 14% до 24%), опосредованная парной нуклеазой Cas9 делеция также сопровождалась гораздо более высокими уровнями инделей (33% желаемой делеции с инделями, 11% инделей без желаемой делеции) по сравнению с twinPE (от 1% до 12% всех инделей) (**фиг. 109E**). Экзон 51 *DMD* также может быть вырезан с использованием альтернативных пар спейсеров, которые генерируют делецию 627 п.о. при использовании стратегии PrimeDel (средняя эффективность 28,8%) или делецию 590 п.о. при стратегии одинарного twinPE (средняя эффективность 26,4%), или с использованием twinPE для замены последовательности из 589 п.о. последовательностью Bxb1 *attB* из 38 п.о. (средняя эффективность до 40,4%) (**фиг. 109E**). Эти эксперименты показывают, что twinPE и PrimeDel способны генерировать большие делеции в терапевтически значимых локусах в клетках человека с более высокой чистотой продукта по сравнению со стратегиями парных нуклеаз Cas9, хотя и с более низкой средней эффективностью.

Нацеленная интеграция ДНК в локусы safe harbor с помощью twinPE и Bxb1 интегразы

[1355] Несмотря на то, что twinPE способен выполнять более крупные редактирования в виде вставки, чем PE3, верхний предел размера вставки последовательности все еще недостаточен для поддержки интеграции фрагментов ДНК размером с ген. Следовательно после успешной вставки последовательности прикрепления Bxb1 в эндогенные геномные сайты человека с высокой эффективностью с использованием twinPE, twinPE был объединен с сериновыми рекомбиназами для сайт-специфической интеграции груза ДНК. Чтобы определить места для интеграции груза ДНК, TwinPE-опосредованную вставку последовательностей прикрепления Bxb1 *attB* и *attP* в

установленных локусах safe harbor в клетках HEK293T, 32 пары спейсеров, нацеленных на локус *AAVS1,* подвергали скринингу на вставку последовательности attP Bxb1 размером 50 п.о. (данные не показаны). Из оцененных пар спейсеров оптимальные РЕгРНК для 18 из 32 комбинаций спейсеров достигли >50% правильной эффективности вставки при секвенировании ампликона (**фиг. 110A**). Кроме того, 19 пар спейсеров, нацеленных на локус *CCR5*, подвергали скринингу на вставку последовательности attB *Bxb1* размером 38 п.о. (данные не показаны), из которых 6 пар спейсеров достигли >50% желаемой эффективности редактирования с помощью секвенирования ампликона (**фиг. 110B**). Эти результаты демонстрируют, что twinPE можно использовать для встраивания последовательностей прикрепления рекомбиназы в безопасные локусы в клетках человека с высокой эффективностью.

[1356] Затем были исследованы последовательности attB и attP Bxb1 с включенным twinPE, чтобы определить, могут ли они служить в качестве участков для интеграции плазмидной ДНК, содержащей груз, несущий партнерские сайты прикрепления Bxb1. Во-первых, twinPE использовали для создания одноклеточных клонов, несущих гомозиготные вставки сайта attB в локусе CCR5 (данные не показаны). Трансфекция этой клональной клеточной линии плазмидой, экспрессирующей рекомбиназу Bxb1, и плазмидой донорной ДНК, содержащей 5,6 т.о. attP, обеспечила 12-17% нокаута в целевом участке, как было измерено с помощью ddPCR (**фиг. 113**), что согласуется с ранее описанной эффективностью интеграции Bxb1. Воодушевленные этими результатами, были исследованы опосредованная twinPE вставка сайта прикрепления и опосредованная BxB1 интеграция донора ДНК, чтобы определить, могут ли они быть достигнуты за одну стадию трансфекции. Примечательно, что трансфекция клеток HEK293T плазмидами, кодирующими PE2, как РЕгРНК, BxB1, так и донорскую ДНК, приводила к эффективности нокаута 1,4-6,8% при измерении с помощью ddPCR (**фиг. 110C**). Ожидаемые последовательности соединений, содержащие ожидаемые продукты рекомбинации attL и attR, были подтверждены секвенированием ампликонов с чистотой продуктов в диапазоне 71-95% (**фиг. 114**).

[1357] В попытке улучшить эффективность нокаута в «одном сосуде» опосредованное twinPE включение сайтов прикрепления *attB* или *attP*, канонической коровой последовательности (GT) Bxb1, альтернативной коровой последовательности (GA) Bxb1 и лоскутов, различной длины перекрытия, были протестированы. Было обнаружено, что включение сайта *attB* с twinPE работает лучше, чем включение сайта *attP* для однореакторных трансфекций с нокаутом (3,3% против 0,5% нокаута, **фиг. 110D)**. Кроме того, было замечено, что вставки сайтов связывания с канонической коровой

последовательностью GT обычно приводили к более высокой эффективности нокаута, чем вставки сайтов связывания с альтернативной последовательностью GA (3,3% против 0% в отношении *attB*, **фиг. 110D**). Следует отметить, однако, что каноническая и альтернативная коровые последовательности допускают ортогональную рекомбинацию с соответствующими сайтами присоединения *attP*. Уменьшение перекрытия между последовательностями рекомбиназы, кодируемыми РЕгРНК, также улучшило эффективность нокаута с 3,3% при перекрытии 38 пар оснований до 5,5% при перекрытии 20 пар оснований при вставке attB (**фиг. 110D**), возможно, из-за уменьшенной степени последовательность прикрепления в последовательности ДНК плазмиды, экспрессирующей РЕгРНК. Хотя одинаковая эффективность twinPE наблюдалась для длин перекрытия, рекомбинация между донорной ДНК и плазмидой РЕгРНК была снижена (**фиг. 115A-115B**).

[1358] Затем оценивали интеграцию донора в локусе *ALB*, который можно использовать для секреции терапевтического белка из гепатоцитов. Альбумин активно секретируется печенью и составляет приблизительно 60% белков плазмы крови человека. Следовательно, терапевтическая интеграция трансгена в локус *ALB* в относительно небольшом проценте гепатоцитов может обеспечить терапевтические уровни экспрессии белка при многих заболеваниях. Была разработана стратегия, нацеленная на интрон 1 *ALB*, где Bxb1 будет опосредовать интеграцию кольцевой ДНК, содержащей последовательность акцептора сплайсинга, за которой следует кДНК представляющего интерес белка. Интеграция, в принципе, позволила бы осуществить сплайсинг сигнала секреции альбумина, закодированного в экзоне 1, с кодирующей последовательностью терапевтического трансгена. Во-первых, пары РЕгРНК проверяли на вставку последовательности Bxb1 *attB* в интрон 1 *ALB*. После скрининга пар РЕгРНК была идентифицирована комбинация спейсеров, обеспечивающая 43% правильную вставку последовательности *attB* с помощью секвенирования ампликона (**фиг. 110E**). Однократное включение трансфекции в этом локусе было достигнуто с эффективностью 1,3% в HEK293T (**фиг. 110F**). В клетках Huh7 наблюдали 34% правильной вставки последовательности *attB* с эффективностью нокаута одиночной трансфекции 2,6%. Интересно, что нокаут *ALB* в Huh7 был более эффективным, чем нокаут CCR5 (2,6% против 1,1%), несмотря на то, что нокаут CCR5 был более эффективным, чем нокаут *ALB* в клетках HEK293T.

Таблица 14. Последовательности донорской ДНК, используемые в этих примерах

ДНК донор	**Последовательность**	**SEQ**

		ID NO:
attB_Puro_GA донор ДНК:	gatgccagctcattcctcccactcatgatctatagatcccccgggctgcaggaattctacccactctgtcgataccccacc gagaccccattggggccaatacgcccgcgtttcttccttttccccacccccacccccccaagttcgggtgaaggcccagg gctcgcagccaacgtcggggcggcaagcttacatcgagatcccggcttgtcgacgacggcggactccgtcgtcagg atcatccgtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaac ggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgca ccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctac cccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttc aaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctg aagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtct atatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgt gcagctcgccgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcccgacaaccactacctg agcacccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccg ccgggatcactctcggcatggacgagctgtacaagagcggcctgaggagcagagcccaggcgagcaacagcgcc gtggacgccaccatgggcgatcgcccgggaattgactagtgcggccgcctaggatccatgccgatagcgttggttga gtggataaccgtattaccgccaagcttatgcatgtgcccgtcagtgggcagagcgcacatcgcccacagtccccgaga agttgggggggaggggtcggcaattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgt gtactggctccgccttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttttcgcaa cgggtttgccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgc gtgccttgaattacttccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggagagttcg aggccttgcgcttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgcgtgcg aatctggtggcaccttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgcgacgc ttttttttctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggttttttggggccgcgggcggcga cggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcggacgggg ggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcggcaagg ctggcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatgga ggacgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagccgtcg cttcatgtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgcgcttttggagtacgtcgtctttag gttgggggggagggggttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttggcactt gatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaagttttttttctt ccatttcaggtgtcgtgagctagcccaccatgaccgagtacaagcccacggtgcgcctcgccacccgcgacgacgtc ccccgggccgtacgcaccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgacccggaccgc cacatcgagcgggtcaccgagctgcaagaactcttcctcacgcgcgtcgggctcgacatcggcaaggtgtgggtcgc ggacgacggcgccgcggtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccgagatcg gcccgcgcatggccgagttgagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgccgcacc ggcccaaggagcccgcgtggttcctggccaccgtcggcgtctcgcccgaccaccagggcaagggtctgggcagcg ccgtcgtgctccccggagtggaggcggccgagcgcgccggggtgcccgccttcctggagacctccgcgccccgca	4282

	acctccccttctacgagcggctcggcttcaccgtcaccgccgacgtcgaggtgcccgaaggaccgcgcacctggtgc atgacccgcaagcccggtgccggatcgggagagggcagaggaagtctgctaacatgcggtgacgtcgaggagaat cctggcccaccggtcgccaccagcgagctgattaaggagaacatgcacatgaagctgtacatggagggcaccgtgg acaaccatcacttcaagtgcacatccgagggcgaaggcaagccctacgagggcacccagaccatgagaatcaaggt ggtcgagggcggccctctccccttcgccttcgacatcctggctactagcttcctctacggcagcaagaccttcatcaacc acacccagggcatccccgacttcttcaagcagtccttccctgagggcttcacatgggagagagtcaccacatacgaag acgggggcgtgctgaccgctacccaggacaccagcctccaggacggctgcctcatctacaacgtcaagatcagagg ggtgaacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacgctgtacc ccgctgacggcggcctggaaggcagaaacgacatggccctgaagctcgtgggcgggagccatctgatcgcaaacat caagaccacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactatgtggactacagactgga aagaatcaaggaggccaacaacgagacctacgtcgagcagcacgaggtggcagtggccagatactgcgacctccct agcaaactggggcacaagcttaattaagaattctctagaggatccagacatgataagatacattgatgagtttggacaaa ccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaa taaacaagttaacaacaacaattgcattcattttatgtttcaggttcaggggaggtgtgggaggttttttaaagcaagtaa aacctctacaaatgtggtatggctgattatgatcctgcaagcctcgtcgccgcggtttattctgttgacaattaatcatcggc atagtatatcggcatagtataatacgacaaggtgaggaactaaaccatgggatcggccattgaacaagatggattgcac gcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacgatcggctgctctgatgcc gccgtgttccggctgtcagcgcaggggcgcccggttcttttttgtcaagaccgacctgtccggtgccctgaatgaactgc aggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaa gcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaa gtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaagcgaaac atcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagcatcagggg ctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacccatggcg atgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggac cgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgc tttacggtatcgccgcccccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctgagcgggactctggg gttcgaataaagaccgaccaagcgacgtctgagagctccctggcgaattcggtaccaataaaagagctttattttcatga tctgtgtgttggtttttggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgacgct caagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctc ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctatagctcacgct gtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctg cgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaac aggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaga acagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacc accgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatc ttttctagtgtgcg	
attB_Puro донор	gatgccagctcattcctcccactcatgatctatagatcccccgggctgcaggaattctacccactctgtcgatacccccacc gagaccccattggggccaatacgcccgcgtttcttccttttccccaccccaccccccaagttcgggtgaaggcccagg	4283

ДНК:	gctcgcagccaacgtcggggcggcaagcttacatcgagatcccggcttgtcgacgacggcggtctccgtcgtcagga tcatccgtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacg gccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcacc accggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccc cgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaa ggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaa gggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctata tcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgca gctcgccgaccactaccagcagaacaccccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagc acccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccg ggatcactctcggcatggacgagctgtacaagagcggcctgaggagcagagcccaggcgagcaacagcgccgtgg acgccaccatgggcgatcgcccgggaattgactagtgcggccgcctaggatccatgccgatagcgttggttgagtgg ataaccgtattaccgccaagcttatgcatgtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagtt ggggggaggggtcggcaattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtac tggctccgccttttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacgg gtttgccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgcgtgc cttgaattacttccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggagagttcgagg ccttgcgcttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgcgtgcgaatc tggtggcaccttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgcgacgcttttt tctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggttttggggccgcgggcggcgacgg ggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcggacgggggt agtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcggcaaggctg gcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatggagga cgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagccgtcgcttc atgtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgcgcttttggagtacgtcgtctttaggttg gggggagggttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttggcacttgatg taattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaagttttttttcttccat ttcaggtgtcgtgagctagcccaccatgaccgagtacaagcccacggtgcgcctcgccacccgcgacgacgtccccc gggccgtacgcaccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgacccggaccgccaca tcgagcgggtcaccgagctgcaagaactcttcctcacgcgcgtcgggctcgacatcggcaaggtgtgggtcgcggac gacggcgccgcggtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccgagatcggccc gcgcatggccgagttgagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgccgcaccggcc caaggagcccgcgtggttcctggccaccgtcggcgtctcgcccgaccaccagggcaagggtctgggcagcgccgt cgtgctccccggagtggaggcggccgagcgcgccggggtgcccgccttcctggagacctccgcgccccgcaacct ccccttctacgagcggctcggcttcaccgtcaccgccgacgtcgaggtgcccgaaggaccgcgcacctggtgcatga cccgcaagcccggtgccggatcgggagagggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctg gcccaccggtcgccaccagcgagctgattaaggagaacatgcacatgaagctgtacatggagggcaccgtggacaa ccatcacttcaagtgcacatccgagggcgaaggcaagccctacgagggcacccagaccatgagaatcaaggtggtc	

	gagggcggccctctcccctt cgccttcgacatcctggctactagcttcctctacggcagcaagaccttcatcaaccacac ccagggcatccccgacttcttcaagcagtccttccctgagggcttcacatgggagagagtcaccacatacgaagacgg gggcgtgctgaccgctacccaggacaccagcctccaggacggctgcctcatctacaacgtcaagatcagaggggtg aacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacgctgtaccccgct gacggcggcctggaaggcagaaacgacatggccctgaagctcgtgggcgggagccatctgatcgcaaacatcaag accacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactatgtggactacagactggaaaga atcaaggaggccaacaacgagacctacgtcgagcagcacgaggtggcagtggccagatactgcgacctccctagca aactggggcacaagcttaattaagaattctctagaggatccagacatgataagatacattgatgagtttggacaaaccac aactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaa caagttaacaacaacaattgcattcattttatgtttcaggttcaggggggaggtgtgggaggttttttaaagcaagtaaaacc tctacaaatgtggtatggctgattatgatcctgcaagcctcgtcgccgcggtttattctgttgacaattaatcatcggcatag tatatcggcatagtataatacgacaaggtgaggaactaaaccatgggatcggccattgaacaagatggattgcacgca ggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacgatcggctgctctgatgccgcc gtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtccggtgccctgaatgaactgcagg acgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcg ggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtat ccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaagcgaaacatc gcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagcatcaggggctc gcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacccatggcgatg cctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgc tatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgcttta cggtatcgccgcccccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctgagcgggactctggggttc gaataaagaccgaccaagcgacgtctgagagctccctggcgaattcggtaccaataaaagagctttattttcatgatctg tgtgttggtttttggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaa gtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctg ttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgta ggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgc cttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacagg attagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaaca gtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccacc gctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttt ctagtgtgcg	
attP_Puro_GA донор ДНК:	gatgccagctcattcctcccactcatgatctatagatccccgggctgcaggaattctacccactctgtcgataccccacc gagaccccattggggccaatacgcccgcgtttcttccttttccccacccccacccccccaagttcgggtgaaggcccagg gctcgcagccaacgtcggggcggcaagcttacatcgagatcccggtttgtctggtcaaccaccgcggactcagtggtg tacggtacaaaccccgtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcg acgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagtt catctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcag	4284

ccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcacca
tcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatc
gagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccaca
acgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggc
agcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccact
acctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgacc
gccgccgggatcactctcggcatggacgagctgtacaagagcggcctgaggagcagagcccaggcgagcaacag
cgccgtggacgccaccatgggcgatcgcccgggaattgactagtgcggccgcctaggatccatgccgatagcgttgg
ttgagtggataaccgtattaccgccaagcttatgcatgtgcccgtcagtgggcagagcgcacatcgcccacagtcccc
gagaagttgggggggaggggtcggcaattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgat
gtcgtgtactggctccgccttttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttt
cgcaacgggtttgccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggc
ccttgcgtgccttgaattacttccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggag
agttcgaggccttgcgcttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgc
gtgcgaatctggtggcaccttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgc
gacgcttttttttctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggtttttggggccgcgggc
ggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcgg
acgggggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcgg
caaggctggcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaa
atggaggacgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagc
cgtcgcttcatgtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgcgcttttggagtacgtcgt
ctttaggttgggggggagggggttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttg
gcacttgatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaagtt
ttttttcttccatttcaggtgtcgtgagctagcccaccatgaccgagtacaagcccacggtgcgcctcgccacccgcgac
gacgtcccccgggccgtacgcaccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgacccg
gaccgccacatcgagcgggtcaccgagctgcaagaactcttcctcacgcgcgtcgggctcgacatcggcaaggtgtg
ggtcgcggacgacggcgccgcggtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccg
agatcggcccgcgcatggccgagttgagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgc
cgcaccggcccaaggagcccgcgtggttcctggccaccgtcggcgtctcgcccgaccaccagggcaagggtctgg
gcagcgccgtcgtgctccccggagtggaggcggccgagcgcgccggggtgcccgccttcctggagacctccgcgc
cccgcaacctccccttctacgagcggctcggcttcaccgtcaccgccgacgtcgaggtgcccgaaggaccgcgcac
ctggtgcatgacccgcaagcccggtgccggatcgggagagggcagaggaagtctgctaacatgcggtgacgtcga
ggagaatcctggcccaccggtcgccaccagcgagctgattaaggagaacatgcacatgaagctgtacatggagggc
accgtggacaaccatcacttcaagtgcacatccgagggcgaaggcaagccctacgagggcacccagaccatgagaa
tcaaggtggtcgagggcggccctctcccccttcgccttcgacatcctggctactagcttcctctacggcagcaagaccttc
atcaaccacacccagggcatccccgacttcttcaagcagtccttccctgagggcttcacatgggagagagtcaccacat
acgaagacgggggcgtgctgaccgctacccaggacaccagcctccaggacggctgcctcatctacaacgtcaagat
cagaggggtgaacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacg

	ctgtaccccgctgacggcggcctggaaggcagaaacgacatggccctgaagctcgtgggcgggagccatctgatcg caaacatcaagaccacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactatgtggactacag actggaaagaatcaaggaggccaacaacgagacctacgtcgagcagcacgaggtggcagtggccagatactgcga cctccctagcaaactggggcacaagcttaattaagaattctctagaggatccagacatgataagatacattgatgagtttg gacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataa gctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggaggttttttaaag caagtaaaacctctacaaatgtggtatggctgattatgatcctgcaagcctcgtcgccgcggtttattctgttgacaattaat catcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatgggatcggccattgaacaagatg gattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacgatcggctgct ctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtccggtgccctgaa tgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttg tcactgaagcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgc cgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaa gcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagca tcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacc catggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgt ggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctgaccgctt cctcgtgctttacggtatcgccgcccccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctgagcggg actctggggttcgaataaagaccgaccaagcgacgtctgagagctccctggcgaattcggtaccaataaaagagcttta ttttcatgatctgtgtgttggtttttggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaa tcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgt gcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatag ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccg accgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccac tggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacact agaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa acaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagat cctttgatcttttctagtgtgcg	
attP_Puro донор ДНК:	gatgccagctcattcctcccactcatgatctatagatcccccgggctgcaggaattctacccactctgtcgataccccacc gagaccccattggggccaatacgcccgcgtttcttccttttccccacccccacccccccaagttcgggtgaaggcccagg gctcgcagccaacgtcggggcggcaagcttacatcgagatcccggtttgtctggtcaaccaccgcggtctcagtggtg tacggtacaaaccccgtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcg acgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagtt catctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcag ccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcacca tcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatc gagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccaca acgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggc	4285

agcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccact
acctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgacc
gccgccgggatcactctcggcatggacgagctgtacaagagcggcctgaggagcagagcccaggcgagcaacag
cgccgtggacgccaccatgggcgatcgcccgggaattgactagtgcggccgcctaggatccatgccgatagcgttgg
ttgagtggataaccgtattaccgccaagcttatgcatgtgcccgtcagtgggcagagcgcacatcgcccacagtcccc
gagaagttgggggggaggggtcggcaattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgat
gtcgtgtactggctccgcctttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttt
cgcaacgggtttgccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggc
ccttgcgtgccttgaattacttccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggag
agttcgaggccttgcgcttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgc
gtgcgaatctggtggcaccttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgc
gacgctttttttctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggttttttggggccgcgggc
ggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcgg
acgggggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcgg
caaggctggcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaa
atggaggacgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagc
cgtcgcttcatgtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgcgcttttggagtacgtcgt
ctttaggttgggggggagggtttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttg
gcacttgatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaagtt
ttttttcttccatttcaggtgtcgtgagctagcccaccatgaccgagtacaagcccacggtgcgcctcgccacccgcgac
gacgtcccccgggccgtacgcaccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgacccg
gaccgccacatcgagcgggtcaccgagctgcaagaactcttcctcacgcgcgtcgggctcgacatcggcaaggtgtg
ggtcgcggacgacggcgccgcggtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccg
agatcggcccgcgcatggccgagttgagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgc
cgcaccggcccaaggagcccgcgtggttcctggccaccgtcggcgtctcgcccgaccaccagggcaagggtctgg
gcagcgccgtcgtgctccccggagtggaggcggccgagcgcgccggggtgcccgccttcctggagacctccgcgc
cccgcaacctccccttctacgagcggctcggcttcaccgtcaccgccgacgtcgaggtgcccgaaggaccgcgcac
ctggtgcatgacccgcaagcccggtgccggatcgggagagggcagaggaagtctgctaacatgcggtgacgtcga
ggagaatcctggcccaccggtcgccaccagcgagctgattaaggagaacatgcacatgaagctgtacatggagggc
accgtggacaaccatcacttcaagtgcacatccgagggcgaaggcaagccctacgagggcacccagaccatgagaa
tcaaggtggtcgagggcggccctctcccccttcgccttcgacatcctggctactagcttcctctacggcagcaagaccttc
atcaaccacacccagggcatccccgacttcttcaagcagtccttccctgagggcttcacatgggagagagtcaccacat
acgaagacgggggcgtgctgaccgctacccaggacaccagcctccaggacggctgcctcatctacaacgtcaagat
cagaggggtgaacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacg
ctgtaccccgctgacggcggcctggaaggcagaaacgacatggccctgaagctcgtgggcgggagccatctgatcg
caaacatcaagaccacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactatgtggactacag
actggaaagaatcaaggaggccaacaacgagacctacgtcgagcagcacgaggtggcagtggccagatactgcga
cctccctagcaaactggggcacaagcttaattaagaattctctagaggatccagacatgataagatacattgatgagtttg

	gacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataa gctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcaggggggaggtgtgggaggttttttaaag caagtaaaacctctacaaatgtggtatggctgattatgatcctgcaagcctcgtcgccgcggtttattctgttgacaattaat catcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatgggatcggccattgaacaagatg gattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacgatcggctgct ctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtccggtgccctgaa tgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttg tcactgaagcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgc cgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaa gcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagca tcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacc catggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgt ggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctgaccgctt cctcgtgctttacggtatcgccgcccccgattcgcagcgcatcgccttctatcgccttcttgacgagttcttctgagcggg actctggggttcgaataaagaccgaccaagcgacgtctgagagctccctggcgaattcggtaccaataaaagagcttta ttttcatgatctgtgtgttggtttttggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaa tcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgt gcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatag ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccg accgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccac tggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacact agaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa acaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagat cctttgatcttttctagtgtgcg	

Мультиплексное редактирование и инверсия, опосредованная рекомбиназой

[1359] Большие структурные варианты обнаруживаются во многих патогенных аллелях человека, например, большая инверсия размером 0,6 Мб в локусе *F8,* которая вызывает приблизительно 50% случаев тяжелой гемофилии А. В результате высокой эффективности вставок сайтов присоединения рекомбиназы с использованием twinPE было высказано предположение, что мультиплексирование вставок сайтов присоединения как *attB, так* и *attP* Bxb1 можно использовать для коррекции более сложных генетических вариантов путем однонаправленной делеции или инверсии промежуточной последовательности ДНК. Чтобы проверить этот подход на терапевтически релевантном локусе, была предпринята попытка инверсии 39 т.о. между *IDS* и его псевдогеном *IDS2*. Инверсии между этими сайтами наблюдались приблизительно в 13% случаев синдрома Хантера, а характеристика точек разрыва в патогенных аллелях показала, что инверсия

часто происходит в горячей точке рекомбинации интрона 7, которая присутствует как в *IDS* , так и в *IDS2*. Следовательно, области, фланкирующие эти горячие точки рекомбинации, были нацелены на вставку последовательностей прикрепления Bxb1 *attP* и *attB*, которые при ориентации в противоположных направлениях можно было бы использовать в качестве субстрата для однонаправленной инверсии с помощью Bxb1 для коррекции последовательности патогенного аллеля в клетках пациентов (**фиг. 111A**). После скрининга 12 пар спейсеров РЕгРНК и оптимизации характеристик последовательности РЕгРНК было обнаружено, что пары РЕгРНК способны вставлять сайты присоединения *attP* или *attB* в левый и правый целевые участки, с эффективностью 73,7% и 69,8%, соответственно (фиг. **111B**).

[1360] Затем исследовали опосредованную мультиплексным twinPE вставку сайтов прикрепления как *attP, так* и *attB* с опосредованной рекомбиназой Bxb1 инверсией последовательности 39 т.о. в *IDS* и *IDS2*. Первоначально были исследованы последовательные трансфекции ДНК с использованием компонентов twinPE с последующей рекомбиназой Bxb1. Был протестирован первый набор РЕгРНК (набор 1), который устанавливает прямо ориентированную последовательность *attP* в интроне 3 *IDS2* и обратно ориентированную последовательность *attB* в интроне 7 *IDS*. Кроме того, второй набор РЕгРНК (набор 2) использовали для установки обратно ориентированной последовательности *attP* в *IDS2* и направленной вперед последовательности *attB* в *IDS* при первой трансфекции. 52,3% или 54,5% вставки последовательности *attP* и 27,6% или 30,9% вставки последовательности *attB* наблюдали для мультиплексного редактирования с набором 1 и набором 2 РЕгРНК? соответственно (**фиг. 111C**). Когда отредактированные клетки впоследствии трансфицировали рекомбиназой Bxb1, наблюдали значительно сниженный процент амплифицированных аллелей, содержащих последовательности attP и attB, по сравнению с имитацией контроля трансфекции (значение $p < 0,05$), что позволяет предположить, что некоторые отредактированные аллели использовались в качестве субстратов для Bxb1- опосредованной рекомбинация (**фиг. 111C**). Амплификация ожидаемых инверсионных соединений с последующим секвенированием ампликона показала сигнатуры последовательностей attL и attR опосредованной Bxb1 рекомбинации с чистотой продукта на уровне 80% или выше на обоих соединениях (**фиг. 111D**).

[1361] Для того, чтобы осуществить инверсию, опосредованную TwinPE и Bxb1 в одном сосуде, и избежать нежелательной рекомбинации между плазмидной ДНК РЕгРНК, все компоненты РНК, содержащие иРНК PE2, синтетические РЕгРНК из набора 1 и иРНК Bxb1, подвергали нуклеофекции. С помощью секвенирования ампликонов были захвачены ожидаемые соединения инвертированных аллелей, содержащие последовательности *attR* и

attL. Кроме того, для количественной оценки эффективности инверсии был разработан обратный праймер, который может связываться с идентичной последовательностью как в неинвертированном, так и в инвертированном аллелях и, следовательно, амплифицировать обе модификации с использованием одной и той же пары праймеров (**фиг. 116**). Наблюдалась эффективность инверсии ~4,6-7,2% и ~1,5-2,3% для «последовательной» и «однореакторной» стратегии, соответственно (**фиг. 111E**). В совокупности эти данные подтверждают, что комбинирование twinPE с сайт-специфическими сериновыми рекомбиназами может быть использовано для коррекции распространенной инверсии 39 т.о., обнаруженной в аллелях при синдроме Хантера, и может служить многообещающей терапевтической платформой для нацеливания на другие сложные патогенные варианты.

Обсуждение

[1362] Как представлено в настоящем документе, был разработан подход двойного праймированного редактирования, который можно использовать для удаления, замены или вставки последовательностей ДНК в целевом участке в геномах клеток человека. В сочетании с сайт-специфичными ферментами сериновой рекомбиназы, twinPE может также поддерживать интеграцию груза ДНК или инверсию последовательности ДНК.

Способы

Общие способы

[1363] Амплификацию ДНК проводили с помощью ПЦР с использованием Phusion U Green Multiplex PCR Master Mix (ThermoFisher Scientific) или Q5 Hot Start High-Fidelity 2x Master Mix (New England BioLabs), если не указано иное. Олигонуклеотиды ДНК были получены от Integrated DNA Technologies. Плазмиды, экспрессирующие сгРНК, были сконструированы путем лигирования отожженных олигонуклеотидов в акцепторный вектор, расщепленный *Bsm* BI. Плазмиды, экспрессирующие РЕгРНК, конструировали с помощью сборки Гибсона или сборки Golden Gate, как описано ранее. Последовательности конструкций сгРНК и РЕгРНК, использованные в этом примере, перечислены в таблицах 11-12.

Таблица 11. Последовательности сгРНК, использованные в этих примерах

сгРНК	**Спейсерная последовательность**	**SEQ ID NO:**
HEK3_3b_+90_разрыв	GTCAACCAGTATCCCGGTGC	616
pU6-Sp-sgRNA-DMDExo51-A1	GATTGGCTTTGATTTCCCTA	3971
pU6-Sp-sgRNA-DMDExo51-A3	GTATATGATTGTTACTGAGA	3972
pU6-Sp-sgRNA-DMDExo51-B1	Gcagttgcctaagaactggt	3973

Таблица 12. Последовательности РЕгРНК, использованные в этих примерах

РЕгРНК	**Спейсерная последовательность**	**Спейсер SEQ ID NO:**	**3' удлинение**	**SEQ ID NO удлинения:**
HEK3_attB_A_38	ggcccagactgagcacgtga	383	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCGTGCTCAGTCTG	4029
HEK3_attB_A_34	ggcccagactgagcacgtga	383	TCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCGTGCTCAGTCTG	4030
HEK3_attB_A_30	ggcccagactgagcacgtga	383	GACGACGGAGACCGCCGTCGTCGACAAGCCCGTGCTCAGTCTG	4031
HEK3_attB_B_38	gtcaaccagtatcccggtgc	616	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGATACTGG	4032
HEK3_attB_B_34	gtcaaccagtatcccggtgc	616	TGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGATACTGG	4033
HEK3_attB_B_30	gtcaaccagtatcccggtgc	616	GACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGATACTGG	4034
HEK3_attP_A_43	ggcccagactgagcacgtga	383	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGTGCTCAGTCTG	4035
HEK3_attP_A_39	ggcccagactgagcacgtga	383	GTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGTGCTCAGTCTG	4036
HEK3_attP_A_35	ggcccagactgagcacgtga	383	ACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGTGCTCAGTCTG	4037
HEK3_attP_B_44	gtcaaccagtatcccggtgc	616	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCGGGATACTGG	4038
HEK3_attP_B_40	gtcaaccagtatcccggtgc	616	GGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCGGGATACTGG	4039

HEK3_attP_B_36	gtcaaccagtatccggtgc	616	AACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCGGGATACTGG	4040
PE3_HEK3_FKBP_ins12bp	ggcccagactgagcacgtga	383	tggaggaagcagggcttcctttcctctgccatcacaccctgcactcccgtgctcagtctg	4041
PE3_HEK3_FKBP_ins36bp	ggcccagactgagcacgtga	383	tggaggaagcagggcttcctttcctctgccatcacccgtctcctggggagatggtttccacctgcactcccgtgctcagtctg	4042
PE3_HEK3_FKBP_ins108bp	ggcccagactgagcacgtga	383	tggaggaagcagggcttcctttcctctgccatcaaaatttctttccatcttcaagcatcccggtgtagtgcaccacgcaggtctggccgcgcttggggaaggtgcgcccgtctcctggggagatggtttccacctgcactcccgtgctcagtctg	4043
PE3_HEK3_FKBP_ins321bp	ggcccagactgagcacgtga	383	tggaggaagcagggcttcctttcctctgccatcattccagttttagaagctccacatcgaagacgagagtggcatgtggtgggatgatgcctgggtgcccagtggcaccataggcataatctggagatatagtcagtttggctctctgacccacactcatctgggcaaccccttcttcccagcctcggatcacctcctgcttgcctagcataaacttaaagggcttgtttctgtcccgggaggaatcaaatttctttccatcttcaagcatcccggtgtagtgcaccacgcaggtctggccgcgcttggggaaggtgcgcccgtctcctggggagatggtttccacctgcactcccgtgctcagtctg	4044
twinPE_HEK3_FKBP_ins108bp_32bp_overhang_A	ggcccagactgagcacgtga	383	accacgcaggtctggccgcgcttggggaaggtgcgcccgtctcctggggagatggtttccacctgcactcccgtgctcagtctg	4045
twinPE_HEK3_FKBP_ins108bp_32bp_overhang_B	gtcaaccagtatccggtgc	616	accttccccaagcgcggccagacctgcgtggtgcactacaccgggatgcttgaagatggaaagaaatttccgggatactgg	4046
pU6-Sp-gRNA-HEK3_del_DF_A_13ntTEMP_Basic	ggcccagactgagcacgtga	383	tcctctgccatcacgtgctcagtctg	4047
pU6-Sp-gRNA-HEK3_del_DF_B_13ntTEMP_Basic	gtcaaccagtatccggtgc	616	tgatggcagaggaccgggatactgg	4048
pU6-Sp-gRNA-HEK3_del_DF_A_34	ggcccagactgagcacgtga	383	tggaggaagcagggcttcctttcctctgccatcacgtgctcagtctg	4049

ntTEMP_Basic				
pU6-Sp-gRNA-HEK3_del_DF_B_34 ntTEMP_Basic	gtcaaccagtatccc ggtgc	616	tgatggcagaggaaaggaagccctgcttcctccacc gggatactgg	4050
pU6-Sp-gRNA-HEK3_del_DF_A_13-13TEMP_Hybrid	ggcccagactgag cacgtga	383	tgcaggagctgcatcctctgccatcacgtgctcagtct g	4051
pU6-Sp-gRNA-HEK3_del_DF_B_13-13TEMP_Hybrid	gtcaaccagtatccc ggtgc	616	tgatggcagaggatgcagctcctgcaccgggatact gg	4052
pU6-Sp-gRNA-HEK3_del_DF_A_31 ntTEMP_Shendure	ggcccagactgag cacgtga	383	gcccagccaaacttgtcaaccagtatcccggcgtgct cagtctg	4053
pU6-Sp-gRNA-HEK3_del_DF_B_31 ntTEMP_Shendure	gtcaaccagtatccc ggtgc	616	gggtcaatccttggggcccagactgagcacgccgg gatactgg	4054
pU6-Sp-gRNA-HEK3_del_DF_A_13 ntTEMP_Basic_EvoPreQ1	ggcccagactgag cacgtga	383	tcctctgccatcacgtgctcagtctgTTAAATA ACGCGGTTCTATCTAGTTACGCG TTAAACCAACTAGAA	4055
pU6-Sp-gRNA-HEK3_del_DF_B_13 ntTEMP_Basic_EvoPreQ1	gtcaaccagtatccc ggtgc	616	tgatggcagaggaccgggatactggAAAATA TACGCGGTTCTATCTAGTTACGC GTTAAACCAACTAGAA	4056
pU6-Sp-gRNA-HEK3_del_DF_A_34 ntTEMP_Basic_EvoPreQ1	ggcccagactgag cacgtga	383	tggaggaagcagggcttcctttcctctgccatcacgt gctcagtctgTTAATAATCGCGGTTCT ATCTAGTTACGCGTTAAACCAAC TAGAA	4057
pU6-Sp-gRNA-HEK3_del_DF_B_34 ntTEMP_Basic_EvoPreQ1	gtcaaccagtatccc ggtgc	616	tgatggcagaggaaaggaagccctgcttcctccacc gggatactggAAAAAAAACGCGGTTC TATCTAGTTACGCGTTAAACCAA CTAGAA	4058
pU6-Sp-gRNA-HEK3_del_DF_A_13-13TEMP_Hybrid_EvoPreQ1	ggcccagactgag cacgtga	383	tgcaggagctgcatcctctgccatcacgtgctcagtct gATAAATAACGCGGTTCTATCTA GTTACGCGTTAAACCAACTAGA A	4059
pU6-Sp-gRNA-	gtcaaccagtatcc	616	tgatggcagaggatgcagctcctgcaccgggatact	4060

HEK3_del_DF_B_13-13TEMP_Hybrid_EvoPreQ1	cggtgc		ggAAAAAAGGCGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA	
pU6-Sp-gRNA-HEK3_del_DF_A_31ntTEMP_Shendure_EvoPreQ1	ggcccagactgagcacgtga	383	gcccagccaaacttgtcaaccagtatcccggcgtgctcagtctgTTAAATACCGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA	4061
pU6-Sp-gRNA-HEK3_del_DF_B_31ntTEMP_Shendure_EvoPreQ1	gtcaaccagtatccgcggtgc	616	gggtcaatccttggggcccagactgagcacgccgggatactggAAAATAATCGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA	4062
pU6-Sp-PErPHK-DMDExo51-A1_a_Basic	GATTGGCTTTGATTTCCCTA	3971	ttgaataggaagtaaattaatttgaagctggaccctagGGAAATCAA	4063
pU6-Sp-PErPHK-DMDExo51-A1_b_Basic	GATTGGCTTTGATTTCCCTA	3971	ttgaataggaagtaaattaatttgaagctggaccctagGGAAATCAAAG	4064
pU6-Sp-PErPHK-DMDExo51-A1_c_Basic	GATTGGCTTTGATTTCCCTA	3971	ttgaataggaagtaaattaatttgaagctggaccctagGGAAATCAAAGCC	4065
pU6-Sp-PErPHK-DMDExo51-B1_A1_b_Basic	Gcagttgcctaagaactggt	3973	tagggtccagcttcaaattaatttacttcctattcaaagttcttaggc	4066
pU6-Sp-PErPHK-DMDExo51-B1_A1_c_Basic	Gcagttgcctaagaactggt	3973	tagggtccagcttcaaattaatttacttcctattcaaagttcttaggcaa	4067
pU6-Sp-PErPHK-DMDExo51-A3_c_Basic	GTATATGATTGTTACTGAGA	3972	atgcctggacaagtaacttaagttaaataagccttctCAGTAACAATCAT	4068
pU6-Sp-PErPHK-DMDExo51-B1_A3_a_Basic	Gcagttgcctaagaactggt	3973	agaaggcttatttaacttaagttacttgtccaggcatagttcttag	4069
pu6-sp-PErPHK-dmdexo51-a3-b1_a_PrimeDel	GTATATGATTGTTACTGAGA	3972	tgctgagagagaaacagttgcctaagaactCAGTAACAAT	4070
pu6-sp-PErPHK-dmdexo51-b1-a3_a-	Gcagttgcctaagaactggt	3973	accacttccacaatgtatatgattgttactgagttcttagg	4071

PrimeDel				
pu6-sp-PErPHK-dmdexo51-b1-a3_b-PrimeDel	Gcagttgcctaagaactggt	3973	accacttccacaatgtatatgattgttactgagttcttaggcaa	4072
pU6-Sp-PErPHK-DMDExo51-A3_b_attB_fwd	GTATATGATTGTTACTGAGA	3972	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCAGTAACAATC	4073
pU6-Sp-PErPHK-DMDExo51-A3_c_attB_fwd	GTATATGATTGTTACTGAGA	3972	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCAGTAACAATCATA	4074
pU6-Sp-PErPHK-DMDExo51-B1_b_attB_fwd	Gcagttgcctaagaactggt	3973	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATagttcttaggc	4075
pU6-Sp-PErPHK-DMDExo51-B1_c_attB_fwd	Gcagttgcctaagaactggt	3973	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATagttcttaggcaac	4076
pU6-Sp-PErPHK-DMDExo51-B2_b_attB_fwd	Gaggagagtaaagtgattgg	3974	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATatcactttactc	4077
pU6-Sp-PErPHK-DMDExo51-B2_c_attB_fwd	Gaggagagtaaagtgattgg	3974	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATatcactttactctcc	4078
A_1077a	GCAGAGCCAGGAACCCCTGT	3975	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGGTTCCT	4079
A_1077b	GCAGAGCCAGGAACCCCTGT	3975	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGGTTCCTGG	4080
A_1077c	GCAGAGCCAGGAACCCCTGT	3975	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGGTTCCTGGCT	4081
A_1098a	GGGAAGGGGCAGGAGAGCCA	3976	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCTCCTGC	4082
A_1098b	GGGAAGGGGCAGGAGAGCCA	3976	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCTCCTGCCC	4083

A_1098c	GGGAAGGGGCAGGAGAGCCA	3976	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCTCCTGCCCCT	4084
A_1246a	GAATATGTCCCAGATAGCAC	3977	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTATCTGGG	4085
A_1246b	GAATATGTCCCAGATAGCAC	3977	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTATCTGGGAC	4086
A_1246c	GAATATGTCCCAGATAGCAC	3977	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTATCTGGGACAT	4087
A_1267a	GGGGACTCTTTAAGGAAAGA	3978	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTTTCCTTAAA	4088
A_1267b	GGGGACTCTTTAAGGAAAGA	3978	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTTTCCTTAAAGA	4089
A_1267c	GGGGACTCTTTAAGGAAAGA	3978	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTTTCCTTAAAGAGT	4090
A_1293a	GAGAAAGAGAAAGGGAGTAG	3979	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCCCTTTC	4091
A_1293b	GAGAAAGAGAAAGGGAGTAG	3979	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCCCTTTCTC	4092
A_1293c	GAGAAAGAGAAAGGGAGTAG	3979	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTCCCTTTCTCTT	4093
A_1307a	GAGTAGAGGCGGCCACGACC	3980	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGTGGCCGC	4094
A_1307b	GAGTAGAGGCGGCCACGACC	3980	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCGTGGCCGCCT	4095
A_1307c	GAGTAGAG	3980	TACCGTACACCACTGAGACCGC	4096

	GCGGCCAC GACC		GGTGGTTGACCAGACAAACCTC GTGGCCGCCTCT	
A_1582a	GATCAGTG AAACGCAC CAGA	3981	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTG GTGCGTTT	4097
A_1582b	GATCAGTG AAACGCAC CAGA	3981	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTG GTGCGTTTCA	4098
A_1582c	GATCAGTG AAACGCAC CAGA	3981	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTG GTGCGTTTCACT	4099
A_1615a	GCAGCTCA GGTTCTGGG AGA	3982	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC CCAGAACC	4100
A_1615b	GCAGCTCA GGTTCTGGG AGA	3982	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC CCAGAACCTG	4101
A_1615c	GCAGCTCA GGTTCTGGG AGA	3982	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC CCAGAACCTGAG	4102
A_1647a	GTGGCCACT GAGAACCG GGC	3983	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC GGTTCTCA	4103
A_1647b	GTGGCCACT GAGAACCG GGC	3983	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC GGTTCTCAGT	4104
A_1647c	GTGGCCACT GAGAACCG GGC	3983	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTC GGTTCTCAGTGG	4105
A_1810a	GAATCTGCC TAACAGGA GGT	3984	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTT CCTGTTAG	4106
A_1810b	GAATCTGCC TAACAGGA GGT	3984	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTT CCTGTTAGGC	4107
A_1810c	GAATCTGCC TAACAGGA	3984	TACCGTACACCACTGAGACCGC GGTGGTTGACCAGACAAACCTT	4108

	GGT		CCTGTTAGGCAG	
A_1890a	GTCACCAATCCTGTCCCTAG	3985	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGACAGGA	4109
A_1890b	GTCACCAATCCTGTCCCTAG	3985	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGACAGGATT	4110
A_1890c	GTCACCAATCCTGTCCCTAG	3985	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGACAGGATTGG	4111
A_3786a	GCTGGCCCCCCACCGCCCCA	3986	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCGGTGGG	4112
A_3786b	GCTGGCCCCCCACCGCCCCA	3986	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCGGTGGGGG	4113
A_3786c	GCTGGCCCCCCACCGCCCCA	3986	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCGGTGGGGGGC	4114
A_3835a	GACGTCACGGCGCTGCCCCA	3987	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCAGCGCC	4115
A_3835b	GACGTCACGGCGCTGCCCCA	3987	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCAGCGCCGT	4116
A_3835c	GACGTCACGGCGCTGCCCCA	3987	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCAGCGCCGTGA	4117
A_3856a	GGTGTGCTGGGCAGGTCGCG	3988	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGACCTGCCC	4118
A_3856b	GGTGTGCTGGGCAGGTCGCG	3988	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGACCTGCCCAG	4119
A_3856c	GGTGTGCTGGGCAGGTCGCG	3988	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGACCTGCCCAGCA	4120

B_1154a	GTCCTTGGCAAGCCCAGGAG	3989	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCTGGGCTTG	4121
B_1154b	GTCCTTGGCAAGCCCAGGAG	3989	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCTGGGCTTGCC	4122
B_1154c	GTCCTTGGCAAGCCCAGGAG	3989	cTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCTGGGCTTGCCAA	4123
B_1314a	GTGCGTCCTAGGTGTTCACC	3990	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAACACCTA	4124
B_1314b	GTGCGTCCTAGGTGTTCACC	3990	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAACACCTAGG	4125
B_1314c	GTGCGTCCTAGGTGTTCACC	3990	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAACACCTAGGAC	4126
B_1376a	GTCCTGGCAGGGCTGTGGTG	3991	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCACAGCCCT	4127
B_1376b	GTCCTGGCAGGGCTGTGGTG	3991	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCACAGCCCTGC	4128
B_1376c	GTCCTGGCAGGGCTGTGGTG	3991	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCACAGCCCTGCCA	4129
B_1640a	GTGACCTGCCCGGTTCTCAG	3992	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTAGAACCGGG	4130
B_1640b	GTGACCTGCCCGGTTCTCAG	3992	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTAGAACCGGGCA	4131
B_1640c	GTGACCTGCCCGGTTCTCAG	3992	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTAGAACCGGGCAGG	4132
B_1676a	GAGCTTGG	3993	GTCTGGTCAACCACCGCGGTCTC	4133

	CAGGGGGT GGGA		AGTGGTGTACGGTACAAACCTC ACCCCCTG	
B_1676b	GAGCTTGG CAGGGGGT GGGA	3993	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTC ACCCCCTGCC	4134
B_1676c	GAGCTTGG CAGGGGGT GGGA	3993	TCTGGTCAACCACCGCGGTCTCA GTGGTGTACGGTACAAACCTCA CCCCCTGCCAA	4135
B_1701a	GAGCCAGA GAGGATCC TGGG	3994	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTA GGATCCTC	4136
B_1701b	GAGCCAGA GAGGATCC TGGG	3994	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTA GGATCCTCTC	4137
B_1701c	GAGCCAGA GAGGATCC TGGG	3994	TCTGGTCAACCACCGCGGTCTCA GTGGTGTACGGTACAAACCTAG GATCCTCTCTG	4138
B_1705a	GATGGAGC CAGAGAGG ATCC	3995	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTTC CTCTCTG	4139
B_1705b	GATGGAGC CAGAGAGG ATCC	3995	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTTC CTCTCTGGC	4140
B_1705c	GATGGAGC CAGAGAGG ATCC	3995	TCTGGTCAACCACCGCGGTCTCA GTGGTGTACGGTACAAACCTTCC TCTCTGGCTC	4141
B_1883a	GGGGCCAC TAGGGACA GGAT	3996	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTCT GTCCCTA	4142
B_1883b	GGGGCCAC TAGGGACA GGAT	3996	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTCT GTCCCTAGT	4143
B_1883c	GGGGCCAC TAGGGACA GGAT	3996	TCTGGTCAACCACCGCGGTCTCA GTGGTGTACGGTACAAACCTCTG TCCCTAGTGG	4144
B_1902a	GTCCCCTCC ACCCCACA	3997	GTCTGGTCAACCACCGCGGTCTC AGTGGTGTACGGTACAAACCTT	4145

	GTG		GTGGGGTG	
B_1902b	GTCCCCTCCACCCCACAGTG	3997	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGTGGGGTGGA	4146
B_1902c	GTCCCCTCCACCCCACAGTG	3997	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGTGGGGTGGAGG	4147
B_1962a	GGGACCACCTTATATTCCCA	3998	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAATATAAG	4148
B_1962b	GGGACCACCTTATATTCCCA	3998	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAATATAAGGT	4149
B_1962c	GGGACCACCTTATATTCCCA	3998	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGAATATAAGGTGG	4150
B_3839a	GACCTGCCCAGCACACCCTG	3999	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGGTGTGCTG	4151
B_3839b	GACCTGCCCAGCACACCCTG	3999	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGGTGTGCTGGG	4152
B_3839c	GACCTGCCCAGCACACCCTG	3999	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGGTGTGCTGGGCA	4153
B_3903a	GCGACTCCTGGAAGTGGCCA	4000	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCACTTCCA	4154
B_3903b	GCGACTCCTGGAAGTGGCCA	4000	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCACTTCCAGG	4155
B_3903c	GCGACTCCTGGAAGTGGCCA	4000	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCACTTCCAGGAG	4156
B_3930a	GGACTTCCCAGTGTGCATCG	4001	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGCACACTG	4157

B_3930b	GGACTTCCCAGTGTGCATCG	4001	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGCACACTGGG	4158
B_3930c	GGACTTCCCAGTGTGCATCG	4001	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGCACACTGGGAA	4159
A_1_223_a	TTTGCAGTTTATCAGGATGA	4002	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGCAGTTTATC	4160
A_1_223_b	TTTGCAGTTTATCAGGATGA	4002	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGCAGTTTATCAG	4161
A_1_223_c	TTTGCAGTTTATCAGGATGA	4002	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGCAGTTTATCAGG	4162
A_2_260_a	GGTTGAGCAGGTAGATGTCA	4003	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCTGAGCAGGT	4163
A_2_260_b	GGTTGAGCAGGTAGATGTCA	4003	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCTGAGCAGGTAG	4164
A_2_260_c	GGTTGAGCAGGTAGATGTCA	4003	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCTGAGCAGGTAGAT	4165
A_3_325_a	CTGGGCGGCAGCATAGTGAG	4004	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCGGCAGC	4166
A_3_325_b	CTGGGCGGCAGCATAGTGAG	4004	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCGGCAGCAT	4167
A_3_325_c	CTGGGCGGCAGCATAGTGAG	4004	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCGGCAGCATAG	4168
A_4_360_a	GTCAAGAGTTGACACATTGT	4005	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCAAGAGTTGAC	4169
A_4_360_b	GTCAAGAG	4005	ATGATCCTGACGACGGAGACCG	4170

	TTGACACAT TGT		CCGTCGTCGACAAGCCAAGAGT TGACAC	
A_4_360_c	GTCAAGAG TTGACACAT TGT	4005	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAAGAGT TGACACA	4171
A_5_507_a	CCCAAGTG ATCACACTT GTC	4006	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAAGTGA TCAC	4172
A_5_507_b	CCCAAGTG ATCACACTT GTC	4006	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAAGTGA TCACAC	4173
A_5_507_c	CCCAAGTG ATCACACTT GTC	4006	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAAGTGA TCACACTT	4174
A_6_510_a	CCACCCAA GTGATCAC ACTT	4007	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCCCCAAG TGA	4175
A_6_510_b	CCACCCAA GTGATCAC ACTT	4007	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCCCCAAG TGATC	4176
A_6_510_c	CCACCCAA GTGATCAC ACTT	4007	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCCCCAAG TGATCAC	4177
A_7_532_a	GGGAGAGA CGCAAACA CAGC	4008	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAGAGAC GCA	4178
A_7_532_b	GGGAGAGA CGCAAACA CAGC	4008	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAGAGAC GCAAA	4179
A_7_532_c	GGGAGAGA CGCAAACA CAGC	4008	ATGATCCTGACGACGGAGACCG CCGTCGTCGACAAGCCAGAGAC GCAAACA	4180
B_1_272_a	TCTCTGACC TGTTTTTCC TT	4009	GGCTTGTCGACGACGGCGGTCTC CGTCGTCAGGATCATCTGACCTG	4181
B_1_272_b	TCTCTGACC TGTTTTTCC	4009	GGCTTGTCGACGACGGCGGTCTC CGTCGTCAGGATCATCTGACCTG	4182

	TT		T	
B_1_272_c	TCTCTGACCTGTTTTTCCTT	4009	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCTGACCTGTTT	4183
B_2_291_a	TCTTACTGTCCCCTTCTGGG	4010	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATTACTGTCCC	4184
B_2_291_b	TCTTACTGTCCCCTTCTGGG	4010	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATTACTGTCCCCT	4185
B_2_291_c	TCTTACTGTCCCCTTCTGGG	4010	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATTACTGTCCCCTTC	4186
B_3_305_a	TCTGGGCTCACTATGCTGCC	4011	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGGCTCACT	4187
B_3_305_b	TCTGGGCTCACTATGCTGCC	4011	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGGCTCACTAT	4188
B_3_305_c	TCTGGGCTCACTATGCTGCC	4011	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGGCTCACTATGC	4189
B_4_326_a	CCCAGTGGGACTTTGGAAAT	4012	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATAGTGGGACT	4190
B_4_326_b	CCCAGTGGGACTTTGGAAAT	4012	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATAGTGGGACTTTG	4191
B_4_326_c	CCCAGTGGGACTTTGGAAAT	4012	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATAGTGGGACTTTGG	4192
B_5_330_a	GTGGGACTTTGGAAATACAA	4013	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGACTTTGG	4193
B_5_330_b	GTGGGACTTTGGAAATACAA	4013	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGACTTTGGAA	4194

B_5_330_c	GTGGGACTTTGGAAATACAA	4013	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGGACTTTGGAAAT	4195
B_6_414_a	CCTGACAATCGATAGGTACC	4014	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGACAATCGA	4196
B_6_414_b	CCTGACAATCGATAGGTACC	4014	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGACAATCGATA	4197
B_6_414_c	CCTGACAATCGATAGGTACC	4014	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGACAATCGATAGG	4198
B_7_538_a	GGAATCATCTTTACCAGATC	4015	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATATCATCTTTA	4199
B_7_538_b	GGAATCATCTTTACCAGATC	4015	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATATCATCTTTACC	4200
B_7_538_c	GGAATCATCTTTACCAGATC	4015	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATATCATCTTTACCAG	4201
B_8_584_a	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATT	4202
B_8_584_b	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4203
B_8_584_c	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTCCA	4204
B_9_601_a	TACAGTCAGTATCAATTCTG	4017	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATAGTCAGTAT	4205
B_9_601_b	TACAGTCAGTATCAATTCTG	4017	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATAGTCAGTATCA	4206
B_9_601_c	TACAGTCA	4017	GGCTTGTCGACGACGGCGGTCTC	4207

	GTATCAATTCTG		CGTCGTCAGGATCATAGTCAGTATCAAT	
A_3_325_a	CTGGGCGGCAGCATAGTGAG	4004	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCGGCAGC	4166
B_6_414_b	CCTGACAATCGATAGGTACC	4014	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGACAATCGATA	4197
A_5_507_c	CCCAAGTGATCACACTTGTC	4006	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCAAGTGATCACACTT	4174
A_6_510_b	CCACCCAAGTGATCACACTT	4007	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCCCAAGTGATC	4176
A_7_532_b	GGGAGAGACGCAAACACAGC	4008	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCAGAGACGCAAA	4179
B_8_584_b	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4203
A_1077c	GCAGAGCCAGGAACCCCTGT	3975	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGGGTTCCTGGCT	4081
B_1154c	GTCCTTGGCAAGCCCAGGAG	3989	cTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCTGGGCTTGCCAA	4123
A_3786c	GCTGGCCCCCCACCGCCCCA	3986	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGGCGGTGGGGGGC	4114
B_3903c	GCGACTCCTGGAAGTGGCCA	4000	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCACTTCCAGGAG	4156
B_3930c	GGACTTCCCAGTGTGCATCG	4001	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTTGCACACTGGGAA	4159
A_7_532_b	GGGAGAGACGCAAACA	4008	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCAGAGAC	4179

	CAGC		GCAAA	
B_8_584_b	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4203
A7_attB_30	GGGAGAGACGCAAACACAGC	4008	TCCTGACGACGGAGACCGCCGTCGTCGACAAGCCAGAGACGCAAA	4208
B8_attB_30	GCAGCTCTCATTTTCCATAC	4016	TGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4209
A7_attB_20	GGGAGAGACGCAAACACAGC	4008	ACGACGGAGACCGCCGTCGTCGACAAGCCAGAGACGCAAA	4210
B8_attB_20	GCAGCTCTCATTTTCCATAC	4016	ACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4211
A7_attB_GA_38	GGGAGAGACGCAAACACAGC	4008	ATGATCCTGACGACGGAGTCCGCCGTCGTCGACAAGCCAGAGACGCAAA	4212
B8_attB_GA_38	GCAGCTCTCATTTTCCATAC	4016	GGCTTGTCGACGACGGCGGACTCCGTCGTCAGGATCATGCTCTCATTTTC	4213
A7_attB_GA_30	GGGAGAGACGCAAACACAGC	4008	TCCTGACGACGGAGTCCGCCGTCGTCGACAAGCCAGAGACGCAAA	4214
B8_attB_GA_30	GCAGCTCTCATTTTCCATAC	4016	TGTCGACGACGGCGGACTCCGTCGTCAGGATCATGCTCTCATTTTC	4215
A7_attB_GA_20	GGGAGAGACGCAAACACAGC	4008	ACGACGGAGTCCGCCGTCGTCGACAAGCCAGAGACGCAAA	4216
B8_attB_GA_20	GCAGCTCTCATTTTCCATAC	4016	ACGACGGCGGACTCCGTCGTCAGGATCATGCTCTCATTTTC	4217
A7_attP_50	GGGAGAGACGCAAACACAGC	4008	GGGTTTGTACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4218

B8_attP_50	GCAGCTCTCATTTTCCATAC	4016	TGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4219
A7_attP_40	GGGAGAGACGCAAACACAGC	4008	TGTACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4220
B8_attP_40	GCAGCTCTCATTTTCCATAC	4016	TGTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4221
A7_attP_30	GGGAGAGACGCAAACACAGC	4008	CGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4222
B8_attP_30	GCAGCTCTCATTTTCCATAC	4016	GGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4223
A7_attP_GA_50	GGGAGAGACGCAAACACAGC	4008	GGGTTTGTACCGTACACCACTGAGTCCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4224
B8_attP_GA_50	GCAGCTCTCATTTTCCATAC	4016	TGGTTTGTCTGGTCAACCACCGCGGACTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4225
A7_attP_GA_40	GGGAGAGACGCAAACACAGC	4008	TGTACCGTACACCACTGAGTCCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4226
B8_attP_GA_40	GCAGCTCTCATTTTCCATAC	4016	TGTCTGGTCAACCACCGCGGACTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4227
A7_attP_GA_30	GGGAGAGACGCAAACACAGC	4008	CGTACACCACTGAGTCCGCGGTGGTTGACCAGACAAACCAAGAGACGCAAA	4228
B8_attP_GA_30	GCAGCTCTCATTTTCCATAC	4016	GGTCAACCACCGCGGACTCAGTGGTGTACGGTACAAACCCGCTCTCATTTTC	4229
A277_b	GACTGAAACTTCACAGAATA	4018	AACTTCACAGAGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCAT	4230
A277_c	GACTGAAA	4018	GAAACTTCACAGAGGCTTGTCG	4231

	CTTCACAGAATA		ACGACGGCGGTCTCCGTCGTCAGGATCAT	
B358_b	GATTTATGAGATCAACAGCAC	4019	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCTGTTGATCTC	4232
B358_c	GATTTATGAGATCAACAGCAC	4019	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCTGTTGATCTCAT	4233
A7_attB_20	GGGAGAGACGCAAACACAGC	4008	ACGACGGAGACCGCCGTCGTCGACAAGCCAGAGACGCAAA	4210
B8_attB_20	GCAGCTCTCATTTTCCATAC	4016	ACGACGGCGGTCTCCGTCGTCAGGATCATGCTCTCATTTTC	4211
A277_c	GACTGAAACTTCACAGAATA	4018	GAAACTTCACAGAGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCAT	4231
B358_c	GATTTATGAGATCAACAGCAC	4019	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCTGTTGATCTCAT	4233
pU6-Sp-gRNA-IDS_DF_A1_a_attP_fwd.dna	GACACCAAAAACTGCCACAGG	4020	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGTGGCAGTTA	4234
pU6-Sp-gRNA-IDS_DF_A1_a_attP_rev.dna	GACACCAAAAACTGCCACAGG	4020	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGTGGCAGTTA	4235
pU6-Sp-gRNA-IDS_DF_A1_c_attP_fwd.dna	GACACCAAAAACTGCCACAGG	4020	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTGTGGCAGTTTTA	4236
pU6-Sp-gRNA-IDS_DF_A1_c_attP_rev.dna	GACACCAAAAACTGCCACAGG	4020	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTGTGGCAGTTTTA	4237
pU6-Sp-gRNA-IDS_DF_A4_b_attP_fwd.dna	GCACTCATTTCCTCCAAGCTC	4021	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCTTGGAGGAAA	4238
pU6-Sp-gRNA-IDS_DF_A4_b_attP_r	GCACTCATTTCCTCCAAG	4021	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCT	4239

ev.dna	CTC		TGGAGGAAA	
pU6-Sp-gRNA-IDS_DF_B2_a_attP_fwd.dna	GTAGGTACAGGACAGGGCAG	4022	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCCTGTCC	4240
pU6-Sp-gRNA-IDS_DF_B2_a_attP_rev.dna	GTAGGTACAGGACAGGGCAG	4022	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCCCTGTCC	4241
pU6-Sp-gRNA-IDS_DF_B3_a_attP_fwd.dna	GAGATAGGTAGGTACAGGACA	4023	TCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCTGTACCT	4242
pU6-Sp-gRNA-IDS_DF_B3_a_attP_rev.dna	GAGATAGGTAGGTACAGGACA	4023	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCCTGTACCT	4243
pU6-Sp-gRNA-IDS_DF_B4_b_attP_fwd.dna	GTGAAAAGATAGGTAGGTAC	4024	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTCCTACCTATCTA	4244
pU6-Sp-gRNA-IDS_DF_B4_b_attP_rev.dna	GTGAAAAGATAGGTAGGTAC	4024	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTCCTACCTATCTA	4245
pU6-Sp-gRNA-IDS_DF_B7_b_attP_fwd.dna	GTTATGGTTTACTCCATCTA	4025	GTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCTATGGAGTAAACCT	4246
pU6-Sp-gRNA-IDS_DF_B7_b_attP_rev.dna	GTTATGGTTTACTCCATCTA	4025	TACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCTATGGAGTAAACC	4247
pU6-Sp-gRNA-IDS_DF_C2_c_attB_fwd.dna	GTTTTGGTTTACCCTATCTA	4026	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCATAGGGTAAACCA	4248
pU6-Sp-gRNA-IDS_DF_C2_c_attB_rev.dna	GTTTTGGTTTACCCTATCTA	4026	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATATAGGGTAAACCA	4249
pU6-Sp-gRNA-IDS_DF_D1_b_attB_fwd.dna	GCTGTGGAACTGCAACACACT	4027	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGTGTTGCAGT	4250
pU6-Sp-gRNA-IDS_DF_D1_b_attB_rev.dna	GCTGTGGAACTGCAACACACT	4027	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGTGTTGCAGT	4251

pU6-Sp-gRNA-IDS_DF_D2_c_attB_fwd.dna	GTGCCACCTAACAGTGAGCTG	4028	GGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCTCACTGTTAGGT	4252
pU6-Sp-gRNA-IDS_DF_D2_c_attB_rev.dna	GTGCCACCTAACAGTGAGCTG	4028	ATGATCCTGACGACGGAGACCGCCGTCGTCGACAAGCCCTCACTGTTAGGT	4253

[1364] Все векторы для экспериментов с клетками млекопитающих очищали с использованием наборов Plasmid Plus Midiprep (Qiagen) или наборов плазмидных минипрепаратов PureYield (Promega) или наборов QIAprep Spin Miniprep.

Общие условия культивирования клеток млекопитающих

[1365] Клетки HEK293T (ATCC CRL-3216), U2OS (ATTC HTB-96), K562 (CCL-243) и HeLa (CCL-2) были приобретены у ATCC, культивированы и пассированы в модифицированной Дульбекко среде Игла (DMEM) плюс GlutaMAX (ThermoFisher Scientific), McCoy's 5A среде (Gibco), RPMI среде 1640 плюс GlutaMAX (Gibco) или минимальной питательной среде Игла (EMEM, ATCC) соответственно, каждая из которых содержит 10% (об./об.) эмбриональной телячьей сыворотки (Gibco, квалифицированная) и 1x пенициллин-стрептомицин (Corning). Все типы клеток инкубировали, поддерживали и культивировали при 37°C с 5% CO_2. Линии клеток были аутентифицированы их соответствующими поставщиками и дали отрицательный результат на микоплазму.

Протокол трансфекции культур тканей HEK293T, HeLa и Huh7 и подготовка геномной ДНК

[1366] Выращенные клетки HEK293T высевали на 48-луночные планшеты, покрытые поли-D-лизином (Corning). Через 16–24 ч после посева клетки трансфицировали приблизительно при 60% конфлюэнтности 1 мкл липофектамина 2000 (Thermo Fisher Scientific) в соответствии с протоколами производителя и либо: 750 нг плазмиды PE2, 125 нг PEгРНК 1 и 125 нг PEгРНК 2 (для трансфекции twinPE), 750 нг плазмиды PE2, 250 нг плазмиды PEгРНК и 83 нг плазмиды сгРНК (для трансфекции PE3), или 750 нг Cas9 и 125 нг сгРНК 1 и 125 нг сгРНК 2 (для трансфекции парными нуклеазами Cas9). Если не указано иное, клетки культивировали через 3 дня после трансфекции, после чего среду удаляли, клетки промывали 1x раствором PBS (Thermo Fisher Scientific) и экстрагировали геномную ДНК добавлением 150 мкл свежеприготовленного буфера для лизиса (10 мМ Tris-HCl, pH 7,5, 0,05% SDS, 25 мкг/мл протеиназы K (ThermoFisher Scientific)) непосредственно в каждую лунку планшета для тканевых культур. Смесь геномной ДНК инкубировали при

37°C в течение 1-2 часов, после чего следовала стадия инактивации фермента при 80°C в течение 30 мин. Праймеры, используемые для амплификации геномной ДНК клеток млекопитающих, перечислены в таблице 13.

Таблица 13. Последовательности праймеров, используемых в Примерах

Праймеры	**Последовательность**	**SEQ ID NO:**
HEK3_fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATGTGGGCTGCCTAGAAAGG	3811
HEK3_rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCAGCCAAACTTGTCAACC	3812
HEK3_fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATGTGGGCTGCCTAGAAAGG	3811
HEK3_rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCAGCCAAACTTGTCAACC	3812
DMD_UMI_fwd1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNNtgctggccagtttactaacaat	4254
DMD_UMI_fwd2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNNcagaaagaagatcttatcccatcttg	4255
DMD_rev0	TGGAGTTCAGACGTGTGCTCTTCCGATCTggctacttttgttatttgcatt	4256
AAVS1_AVA1713	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCAGAGCAGGGTCCCGCTTC	4257
AAVS1_AVA1717	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACGGGGCTCAGTCTGAAGAG	4258
AAVS1_AVA1651	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCAAGGACTCAAACCCAGA	4259
AAVS1_AVA1652	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCGTGCGTCAGTTTTACCT	4260
AAVS1_AVA1653	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAACTGCTTCTCCTCTTGGGAA	4261
AAVS1_AVA1715	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCTTGCCAGAACCTCTAAGGT	4262
AAVS1_AVA1655	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATCCTCTCTGGCTCCATCGTA	4263
AAVS1_AVA1656	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCACTTCAGGACAGCATGTTT	4264
AAVS1_AVA1707	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCGCCGGGAACTGCCGCTGGC	4265

AAVS1_AVA1710	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAGGAGGCCCTCATCTGGCG	4266
CCR5_AVA1678	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAATCAATGTGAAGCAAATCGCAGC	4267
CCR5_AVA1679	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCGATTGTCAGGAGGATGATGAA	4268
CCR5_AVA1680	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCTTCTTACTGTCCCCTTCTGGGC	4269
CCR5_AVA1681	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCAAACACAGCATGGACGAC	4270
CCR5_AVA1682	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACAATCGATAGGTACCTGGCTGTC	4271
CCR5_AVA1683	TGGAGTTCAGACGTGTGCTCTTCCGATCTACCAGCCCCAAGATGACTAT	4272
IDS_AVA1763	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTGAAAACCTGAGCTTGGAGG	4273
IDS_AVA1764	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTACTCCAGCTTAATGGAAGTGG	4274
IDS_AVA1765	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAGAAGATGTGGAAATGCCTCAC	4275
IDS_AVA1766	TGGAGTTCAGACGTGTGCTCTTCCGATCTAATCAACATGAAGGGTTGTGTTGT	4276
IDS_AVA1769	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTTCCCACACATGCGTTCCTC	4277
IDS_AVA1770	TGGAGTTCAGACGTGTGCTCTTCCGATCTGGCATGAAGGGTTGTTTTTAATTGA	4278
IDS_UMI_junc1_fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNGCCCTGTCCTGTACCTACCT	4279
IDS_junc2_fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTCAGTTTGGGGTATGTGCCA	4280
IDS_universal_rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCAAATTTACCCGTGGCAGC	4281

[1367] Клетки HeLa выращивали и высевали с плотностью, аналогичной плотности клеток HEK293T, и в каждую лунку с клетками трансфицировали 190 нг PE2_P2A_Blast (ген устойчивости к бластицидину), 31,5 нг PEгРНК1 и 31,5 нг PEгРНК 2 (для трансфекции twinPE) с 0,75 мкл реагента HeLa monster (Mirus). Затем клетки обрабатывали 10 мкг/мл для

отбора через 24 часа после трансфекции. Остальные стадии сбора геномной ДНК идентичны клеткам HEK293T. Клетки Huh7 высевали по 150 тысяч/лунка на 24-луночные планшеты, покрытые поли-D-лизином (Corning). Через 16-24 ч после посева клетки трансфицировали при высокой конфлюентности 2 мкл липофектамина 2000 (Thermo Fisher Scientific) в соответствии с протоколами производителя и до 800 нг плазмидной ДНК. Сбор геномной ДНК проводили, как указано выше для HEK293T.

Высокопроизводительное секвенирование ДНК образцов геномной ДНК

[1368] Представляющие интерес геномные участки амплифицировали из образцов геномной ДНК и секвенировали на Illumina MiSeq, как описано ранее, со следующими модификациями. Кратко, праймеры амплификации, содержащие прямые и обратные адаптеры Illumina (Таблица 13), использовали для первой стадии 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 частности, 25 мкл данной реакции ПЦР-2 содержали по 0,5 мкМ каждой уникальной пары праймеров Illumina для прямого и обратного кодирования, 1 мкл неочищенной реакционной смеси для ПЦР-1 и 12,5 мкл Phusion U Green Multiplex PCR 2x Master Mix. Реакции ПЦР-2 проводили следующим образом: 98 °C в течение 2 мин, затем 12 циклов [98 °C в течение 10 с, 61 °C в течение 20 с и 72 °C в течение 30 с], после чего конечное повышение до 72 ° C в течение 2 мин. Продукты ПЦР оценивали аналитически с помощью электрофореза в 1,5% агарозном геле. Продукты ПЦР 2 (объединенные общими ампликонами) очищали электрофорезом в 1,5% агарозном геле с использованием набора QIAquick Gel Extraction Kit (Qiagen), элюируя 40 мкл воды. Концентрацию ДНК измеряли с помощью флуорометрического количественного определения (Qubit, ThermoFisher Scientific) или количественной ПЦР (набор для количественного анализа библиотеки KAPA-Illumina, KAPA Biosystems) и секвенировали на устройстве Illumina MiSeq в соответствии с протоколами производителя.

[1369] Прочтения секвенирования демультиплексировали с использованием MiSeq Reporter (Illumina). Выравнивание последовательностей ампликонов с эталонной последовательностью выполняли с помощью CRISPResso2. Для количественной оценки выхода всего праймированного редактирования эффективность основного редактирования вычисляли как: % от [количества прочтений с желаемым редактированием,

не содержащих инделей] ÷ [количество прочтений]. Для количественной оценки редактирования точечных мутаций CRISPResso2 запускали в стандартном режиме с включенным «discard_indel_reads». Затем праймированное редактирование для установки точечных мутаций было точно вычисляли как: [частота указанной точечной мутации в неотброшенных прочтениях] x [количество неотброшенных прочтений] ÷ [общее количество чтений]. Для количественной оценки редактирования CRISPResso2 запускали в режиме HDR, используя желаемый аллель в качестве ожидаемого аллеля (e flag) и с включенным параметром «discard_indel_reads». Эффективность редактирования вычисляли как: [количество неотброшенных прочтений, выровненных по HDR] ÷ [общее число прочтений]. Выходы инделей рассчитывались как: [количество отброшенных прочтений, содержащих индель] ÷ [общее количество прочтений]. Уникальные молекулярные идентификаторы (UMI) применяли для количественной оценки эффективности делеции и оценки систематической ошибки ПЦР в этом трехэтапном протоколе ПЦР. Вкратце, на первой стадии линейной амплификации 1 мкл экстракта геномной ДНК линейно амплифицировали с 0,1 мкМ одинарного прямого праймера UMI, который содержит уникальные молекулярные идентификаторы из 15 или 16 п.о., и 12,5 мкл Phusion U Green Multiplex PCR Master Mix в реакция 25 мкл. Затем продукты ПЦР очищали с помощью гранул AmPure 1,6X (Beckman Coulter) и элюировали 20 мкл элюирующего буфера QIAGEN. На второй стадии 1 или 2 мкл линейно амплифицированных ПЦР-элюентов затем амплифицировали в течение 30 циклов с использованием по 0,5 мкМ каждого прямого и обратного праймера, смеси Phusion U Green Multiplex для ПЦР в 25-мкл реакционной смеси. В этом случае прямой праймер отжигается с адаптерной последовательностью P5 Illumina, которая находится на 5'-конце праймера UMI и против хода транскрипции от UMI. Продукты ПЦР очищали с помощью ампул 1X и элюировали 25 мкл элюирующего буфера. На третьей стадии 1 мкл элюента амплифицировали в течение 12 циклов, как указано выше, для добавления уникального признака Illumina. Для того, чтобы оценить большую делецию в локусе DMD, верхнюю полосу (неотредактированный большой ампликон) и нижнюю полосу (отредактированные ампликоны с делециями) вырезали отдельно из 1,5% агарозы, и конечную библиотеку (60 пМ верхних полос против 20 пМ нижних полос) загружали в два отдельных набора Miseq, чтобы избежать слабой кластеризации больших ампликонов (~ 1 т.о.) по сравнению с короткими (~ 300 п.о.).

[1370] Необработанные прочтения секвенирования дедуплицировали UMI с использованием AmpUMI. Для прочтения с парными концами SeqKit использовали для конкатенации (слияния без перекрытия) R1 с обратным комплементом R2. Объединенные R1+R2 дедуплицировали UMI с использованием UMI на 5'-конце R1. Дедуплицированные

UMI R1 или объединенные R1+R2 анализировали с помощью CRISPResso2. Для анализа конкатенированных R1+R2 была предоставлена соответствующая объединенная эталонная последовательность ампликона, чтобы свести к минимуму выравнивания последовательности в результате объединения.

Нуклеофекция U2OS и K562

[1371] Нуклеофекцию осуществляли в рамках всех экспериментов с использованием клеток K562 и U2OS. На одну нуклеофекцию использовали 200 000 клеток. Подсчитанные клетки центрифугировали и промывали PBS, а затем ресуспендировали в растворе для нуклеофекции в соответствии с инструкциями для набора Lonza SE Cell Line 4D-Nucleofector. После нуклеофекции клеток в устройстве клеткам давали возможность инкубироваться в кювете при комнатной температуре в течение 10 минут. Содержимое кюветы переносили в 48-луночный планшет, заполненный предварительно подогретой средой. Геномную ДНК экстрагировали и готовили для устройства Illumina Miseq, как указано выше.

Однореакторный twinPE и BxbI-опосредованный нокин информации и большая инверсия

[1372] Для эксперимента с большой инверсией, в одном сосуде, проводили последовательную плазмидную трансфекцию и нуклеофекцию иРНК в одном сосуде. В эксперименте по последовательной плазмидной трансфекции клетки HEK293T трансфицировали 750 нг PE2, 62,5 нг каждой PEгРНК (всего четыре), используя липофектамин 2000, как указано выше. Через три дня клетки обрабатывали трипсином, высевали на 24-луночный планшет и пассировали в течение приблизительно семи дней. Затем 20 тысяч клеток снова высевали и трансфицировали 500 нг плазмиды BxBI. Геномную ДНК экстрагировали и готовили для устройства Illumina Miseq, как указано выше. В эксперименте по нуклеофекции иРНК в одном сосуде 200000 клеток подвергали нуклеофекции 1 мкг иРНК PE2, 30 пмоль каждой синтетической PEгРНК (всего 4 PEгРНК) и 750 нг иРНК BxBI. Синтетические PEгРНК получали у IDT, а транскрипты иРНК транскрибировали *in vitro в* соответствии с опубликованным протоколом. 20 мкл смеси буфера Lonza и клеток подвергали электропорации с помощью программы CM-130 и извлекали с помощью 80 мкл подогретой среды в течение пяти минут. Затем в каждую лунку 48-луночного планшета добавляли 25 мкл образцов из кюветы и инкубировали при 37°C в течение 72 часов.

Другие ссылочные источники, указанные в настоящем документе

Следующие ссылочные источники включены в настоящий документ посредством ссылки во всей своей полноте.

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).

Эквиваленты и объем

[1373] В настоящем документе единственное число может означать один или несколько, если не указано иное или иным образом не следует из контекста. Варианты осуществления или описания, которые включают «или» между одним или несколькими членами группы, считаются удовлетворительными, если один, несколько или все члены группы присутствуют, используются или иным образом имеют отношение к данному продукту или способу, если не указано иное или иным образом очевидно из контекста. Настоящее изобретение включает варианты осуществления, в которых ровно один член группы присутствует, используется или иным образом имеет отношение к данному продукту или способу. Настоящее изобретение включает варианты осуществления, в которых более одного или все члены группы присутствуют, используются или иным образом относятся к данному продукту или способу.

[1374] Кроме того, настоящее изобретение охватывает все варианты, комбинации и перестановки, в которых одно или несколько ограничений, элементов, признаков и описательных терминов из одного или нескольких перечисленных пунктов формулы изобретения вводятся в другой пункт формулы изобретения. Например, любой пункт формулы изобретения, который зависит от другого пункта формулы изобретения, может быть изменен с включением одного или нескольких ограничений, обнаруженных в любых других пунктах формулы изобретения, которые зависят от того же основного пункта. Когда элементы представлены в виде перечня, *например*, в формате группы Маркуша, каждая подгруппа элементов также раскрывается, и любой элемент (элементы) может быть удален из группы. Следует понимать, что в целом, когда настоящее изобретение или аспекты изобретения упоминаются как содержащие определенные элементы и/или признаки, определенные варианты осуществления раскрытия или аспекты раскрытия состоят или состоят по существу из таких элементов и/или признаков. В целях простоты эти варианты осуществления не были специально изложены в данном документе *in haec verba*. Также отмечается, что термины «содержит» и «содержащий» предназначены быть открытыми и

допускают включение дополнительных элементов или стадий. Там, где указаны диапазоны, включены конечные точки. Кроме того, если не указано иное или иное не очевидно из контекста и понимания специалиста в данной области техники, значения, которые выражены в виде диапазонов, могут принимать любое конкретное значение или поддиапазон в пределах указанных диапазонов в различных вариантах осуществления настоящего изобретения, вплоть до десятой части единицы нижнего предела диапазона, если из контекста явно не следует иное.

[1375] Настоящая заявка относится к различным выданным патентам, опубликованным патентным заявкам, статьям и другим публикациям, все из которых включены в настоящий документ посредством ссылки. В случае противоречия между любой из включенных ссылок и настоящим описанием, описание имеет преимущественную силу. Кроме того, любой конкретный вариант осуществления настоящего изобретения, относящийся к известному уровню техники, может быть явно исключен из любого одного или нескольких вариантов осуществления. Поскольку считается, что такие варианты осуществления известны специалистам в данной области техники, они могут быть исключены, даже если это исключение прямо не указано в настоящем документе. Любой конкретный вариант осуществления изобретения может быть исключен из любого варианта осуществления по любой причине, независимо от того, связана она или нет с известным уровнем техники.

[1376] Специалисты в данной области распознают или смогут установить, используя не более чем обычные эксперименты, многие эквиваленты конкретных вариантов осуществления, описанных в настоящем документе. Объем настоящих вариантов осуществления, описанных в настоящем документе, не предназначен для ограничения вышеприведенным описанием, а скорее изложен в прилагаемых вариантах осуществления. Специалистам в данной области техники понятно, что в это описание могут быть внесены различные изменения и модификации без отклонения от сущности или объема настоящего изобретения, как определено в следующих вариантах осуществления.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, где система содержит первый комплекс праймированного редактирования и второй комплекс праймированного редактирования, где каждый из первого и второго комплексов праймированного редактирования содержит (1) праймирующий редактор, содержащий (i) программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и (ii) полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и (2) РЕгРНК, содержащую спейсерную последовательность, ядро гРНК, матрицу для синтеза ДНК и участок связывания праймера, где матрица для синтеза ДНК РЕгРНК первого комплекса праймированного редактирования кодирует первую последовательность одноцепочечной ДНК, и матрица для синтеза ДНК РЕгРНК второго комплекса праймированного редактирования кодирует вторую последовательность одноцепочечной ДНК, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть, по сравнению с последовательностью ДНК в целевом участке, подлежащем редактированию, который вставляется в целевой участок, подлежащий редактированию.

2. Система по п. 1, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

3. Система по п. 1 или 2, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

4. Система по п. 1 или 2, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат разные праймирующие редакторы.

5. Система по любому из пп. 1-4, где napDNAbp представляет собой домен Cas9 или его вариант.

6. Система по любому из пп. 1-5, где napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью,

никазный домен Cas9 или их вариант.

7. Система по любому из пп. 1-4, где napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

8. Система по любому из пп. 1-7, где napDNAbp обладает никазной активностью.

9. Система по любому из пп. 1-4, где napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

10. Система по любому из пп. 1-9, где полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

11. Система по любому из пп. 1-10, где полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

12. Система по любому из пп. 1-11, где праймирующий редактор дополнительно содержит линкер, который соединяет napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

13. Система по п. 12, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

14. Система по п. 12 или 13, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

15. Система по любому из пп. 1-14, где РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244 или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

16. Система по любому из пп. 1-15, где спейсерная последовательность РЕгРНК первого комплекса праймированного редактирования связывается с первым участком

связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

17. Система по любому из пп. 1-16, где спейсерная последовательность РЕгРНК второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

18. Система по п. 17, где связывание спейсерной последовательности РЕгРНК первого и второго комплексов праймированного редактирования в присутствии последовательности РАМ приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям РАМ на каждой цепи, посредством napDNAbp каждого из первого и второго комплексов праймированного редактирования.

19. Система по п. 18, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

20. Система по п. 19, где непрерывная область между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований

или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

21. Система по п. 19, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

22. Система по любому из пп. 1-21, где матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по

меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

23. Система по любому из пп. 1-22, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

24. Система по любому из пп. 1-22, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

25. Система по любому из пп. 1-24, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

26. Система по п. 25, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине желаемой последовательности ДНК и ее комплемента.

27. Система по п. 25, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть.

28. Система по любому из пп. 1-27, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3'-концы.

29. Система по п. 21, где замещение непрерывной области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

30. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, причем указанная система содержит:

a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

b) второй комплекс праймированного редактирования, содержащий:

iii. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

iv. вторую гидовую РНК праймированного редактирования (вторую PEгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

где первая PEгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, и вторая PEгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, и где первая цепь и вторая цепь последовательности двухцепочечной ДНК комплементарны друг другу.

31. Система по п. 30, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

32. Система по п. 30, где первая и вторая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

33. Система по п. 30, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, содержащий отредактированную часть, которая вставляется в целевой участок, подлежащий редактированию.

34. Система по п. 30, где первая последовательность одноцепочечной ДНК и/или вторая последовательность одноцепочечной ДНК не имеют гомологии последовательности по сравнению с последовательностью ДНК в целевом участке.

35. Система по п. 30, где первая РЕгРНК дополнительно содержит первую спейсерную последовательность, первое ядро гРНК и первый участок связывания праймера, и вторая РЕгРНК дополнительно содержит вторую спейсерную последовательность, второе ядро гРНК и второй участок связывания праймера.

36. Система по любому из пп. 30-35, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

37. Система по любому из пп. 30-36, где первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, являются одинаковыми.

38. Система по любому из пп. 30-37, где первый и/или второй napDNAbp представляет собой домен Cas9 или его вариант.

39. Система по любому из пп. 30-38, где первый и/или второй napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, никазный домен Cas9 или их вариант.

40. Система по любому из пп. 30-39, где первый и/или второй napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

41. Система по любому из пп. 30-40, где napDNAbp обладает никазной активностью.

42. Система по любому из пп. 30-41, где первый и/или второй napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

43. Система по любому из пп. 30-42, где первый и/или второй полипептид,

обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

44. Система по любому из пп. 30-43, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

45. Система по любому из пп. 30-44, где первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

46. Система по п. 45, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

47. Система по п. 45, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

48. Система по любому из пп. 30-47, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

49. Система по любому из пп. 30-48, где первая спейсерная последовательность связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

50. Система по п. 35, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй

цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

51. Система по п. 35, где связывание спейсерной последовательности первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM, на каждой цепи посредством первого и/или второго napDNAbp каждого из первого и второго комплексов праймированного редактирования.

52. Система по п. 30, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

53. Система по п. 52, где непрерывная область между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

54. Система по п. 52, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

55. Система по любому из пп. 30-54, где первая и/или вторая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по

меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

56. Система по любому из пп. 30-55, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

57. Система по любому из пп. 30-56, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

58. Система по любому из пп. 30-57, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

59. Система по любому из пп. 30-58, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

60. Система по любому из пп. 30-59, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс.

61. Система по любому из пп. 30-60, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

62. Система по п. 54, где замещение непрерывной области дуплексом, содержащим

отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

63. Полинуклеотид, кодирующий систему по любому из пп. 1-62.

64. Полинуклеотид, кодирующий первый и второй праймирующие редакторы по любому из пп. 1-62.

65. Полинуклеотид, кодирующий первую и/или вторую РЕгРНК по любому из пп. 1-62.

66. Полинуклеотид, кодирующий первую и/или вторую РЕгРНК по любому из пп. 30-62.

67. Полинуклеотид, кодирующий первую и вторую РЕгРНК по любому из пп. 1-29.

68. Полинуклеотид, кодирующий первую и вторую РЕгРНК по любому из пп. 30-62.

69. Вектор, содержащий полинуклеотид по любому из пп. 63-68.

70. Клетка, содержащая полинуклеотид по любому из пп. 63-68 или вектор по п. 69.

71. Фармацевтическая композиция, содержащая полинуклеотид по любому из пп. 63-68, вектор по п. 69 или клетку по п. 70 и фармацевтическое вспомогательное вещество.

72. Набор, содержащий полинуклеотид по любому из пп. 63-68 или вектор по п. 69 и инструкции по одновременному редактированию обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию.

73. Способ одновременного редактирования первой и второй комплементарных цепей последовательности двухцепочечной ДНК в целевом участке, причем указанный способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем указанная пара содержит:

a. первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности геномной ДНК против хода транскрипции от целевого участка,

b. второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок, (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности

геномной ДНК по ходу транскрипции от целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв в последовательности, комплементарной первому участку связывания, и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв в последовательности, комплементарной второму участку связывания, и последующую полимеризацию второй последовательности одноцепочечной ДНК, имеющей 3'-конец, от доступного 5'-конца, образованного вторым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, содержащий редактирование, и

где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

74. Способ по п. 73, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

75. Способ по п. 73, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

76. Способ по п. 73, где первый комплекс праймированного редактирования и второй комплекс праймированного редактирования содержат разные праймирующие редакторы.

77. Способ по любому из пп. 73-76, где первый и/или второй napDNAbp представляет собой домен Cas9 или его вариант.

78. Способ по любому из пп. 73-76, где первый и/или второй napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, никазный домен Cas9 или их вариант.

79. Способ по любому из пп. 73-78, где первый и/или второй napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, CasX, CasY и Argonaute.

80. Способ по любому из пп. 73-79, где первый и/или второй napDNAbp обладает никазной активностью.

81. Способ по любому из пп. 73-80, где первый и/или второй napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 18-88 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 18-88.

82. Способ по любому из пп. 73-81, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

83. Способ по любому из пп. 73-82, где первый и/или второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159 и 700-766 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 89-100, 105-122, 128-129 и 132.

84. Способ по любому из пп. 73-83, где первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

85. Способ по п. 84, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 166-177 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 166-177.

86. Способ по п. 84 или 85, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

87. Способ по любому из пп. 74-86, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244 или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

88. Способ по любому из пп. 73-87, где первая и/или вторая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244 или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 325-330, 499-505, 101-104, 181-183 и 223-244.

89. Способ по любому из пп. 73-88, где первая спейсерная последовательность первого комплекса праймированного редактирования связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

90. Способ по любому из пп. 73-89, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

91. Способ по п. 90, где связывание первой и второй спейсерных последовательностей первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого первого и второго комплекса праймированного редактирования.

92. Способ по п. 91, где первый и второй участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

93. Способ по п. 92, где непрерывная область между первым и вторым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований,

по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104

нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

94. Способ по п. 92 или 93, где непрерывная область замещается дуплексом, содержащим отредактированную часть.

95. Способ по любому из пп. 73-94, где первая и/или вторая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по

меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

96. Способ по любому из пп. 73-95, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют одинаковые длины.

97. Способ по любому из пп. 73-96, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК имеют разные длины.

98. Способ по любому из пп. 73-97, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК.

99. Способ по любому из пп. 73-98, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100

нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине желаемой последовательности ДНК и ее комплемента.

100. Способ по любому из пп. 73-99, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК образуют дуплекс.

101. Способ по любому из пп. 73-100, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3'-концы.

102. Способ по п. 94, где замещение непрерывной области дуплексом, содержащим отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

103. Способ по любому из пп. 73-102, где дуплекс содержит редактирование, содержащее один или несколько сайтов рекомбиназы.

104. Способ по п. 103, дополнительно предусматривающий предоставление рекомбиназы для стимуляции опосредованного рекомбиназой редактирования последовательности ДНК в целевом участке, подлежащем редактированию.

105. Пара РЕгРНК для применения в двойном праймированном редактировании, где пара содержит:

a. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции целевого участка, подлежащего редактированию, где первая РЕгРНК содержит первую спейсерную последовательность, первое ядро гРНК, первую матрицу для синтеза ДНК и первый участок связывания праймера, где первая матрица для синтеза ДНК кодирует первую последовательность одноцепочечной ДНК, содержащую отредактированную часть,

b. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где вторая РЕгРНК содержит вторую спейсерную последовательность, второе ядро гРНК, вторую матрицу для синтеза ДНК и второй участок связывания праймера, где вторая матрица для синтеза ДНК кодирует вторую последовательность

одноцепочечной ДНК, содержащую отредактированную часть, где вторая последовательность одноцепочечной ДНК содержит область комплементарности с первой последовательностью одноцепочечной ДНК.

106. Пара РЕгРНК по п. 105, где первая РЕгРНК содержит структуру:

5′-[первая спейсерная последовательность гРНК]-[первое ядро гРНК]-[первая матрица для синтеза ДНК]-[первый участок связывания праймера]-3′, и

где вторая РЕгРНК содержит структуру:

5′-[вторая спейсерная последовательность гРНК]-[второе ядро гРНК]-[вторая матрица для синтеза ДНК]-[второй участок связывания праймера]-3′.

107. Пара РЕгРНК по п. 105 или 106, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК образуют дуплекс, который интегрируется в последовательность двухцепочечной ДНК в целевом участке, подлежащем редактированию.

108. Пара РЕгРНК по п. 107, где интеграция дуплекса приводит к вставке, делеции или замене ДНК в целевом участке.

109. Пара РЕгРНК по п. 108, где вставка составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

110. Пара РЕгРНК по п. 109, где делеция составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по

меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

111. Пара РЕгРНК по п. 1110, где замена составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77

нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

112. Полинуклеотид, кодирующий пару РЕгРНК по любому из пп. 105-111.

113. Вектор, содержащий полинуклеотид по п. 112 для экспрессии пары РЕгРНК в клетке.

114. Клетка, содержащая вектор по п. 113.

115. Фармацевтическая композиция, содержащая пару РЕгРНК по любому из пп. 109-116, вектор по п. 113 или клетку по п. 114 и фармацевтическое вспомогательное вещество.

116. Пара комплексов праймированного редактирования, причем указанная пара содержит:

a. первый комплекс праймированного редактирования, содержащая:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первой целевой последовательностью на цепи последовательности геномной ДНК против хода транскрипции от целевого участка,

b. второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со второй целевой последовательностью на второй цепи последовательности геномной ДНК по ходу транскрипции от целевого участка.

117. Пара комплексов праймированного редактирования по п. 116, где праймирующий редактор первого комплекса праймированного редактирования, второго комплекса праймированного редактирования, или как первого комплекса праймированного редактирования, так и второго комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

118. Полинуклеотид, кодирующий пару комплексов праймированного редактирования по п. 116 или 117.

119. Вектор, содержащий полинуклеотид по п. 118 для экспрессии пары комплексов праймированного редактирования в клетке.

120. Клетка, содержащая вектор по п. 119.

121. Фармацевтическая композиция, содержащая пару комплексов праймированного редактирования по п. 116 или 117, вектор по п. 119 или клетку по п. 120 и фармацевтическое вспомогательное вещество.

122. Система по любому из пп. 1-62 или способ по любому из пп. 73-104, где интеграция дуплекса, содержащего отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

123. Система или способ по п. 122, где вставка представляет собой пептидную метку.

124. Система или способ по п. 123, 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.

125. Система или способ по п. 123, где пептидная метка имеет аминокислотную последовательность, выбранную из группы, состоящей из последовательности SEQ ID NO: 1-6, 245-249, 252-262, 264-273, 275-276, 281, 278-288 и 622.

126. Система или способ по п. 122, где вставка представляет собой метку иммуноэпитопа.

127. Система или способ по п. 126, где метка иммуноэпитопа выбрана из группы, состоящей из столбнячного токсина (SEQ ID NO: 396), мутанта дифтерийного токсина CRM197 (SEQ ID NO: 630), иммуноэпитопа паротита 1 (SEQ ID NO: 400), иммуноэпитопа

паротита 2 (SEQ ID NO: 402), иммуноэпитоп паротита 3 (SEQ ID NO: 404), вируса краснухи (SEQ ID NO: 406), гемагглютинина (SEQ ID NO: 408), нейраминидазы (SEQ ID NO: 410), TAP1 (SEQ ID NO: 412), TAP2 (SEQ ID NO: 414), эпитопов гемагглютинина к HLA класса I (SEQ ID NO: 416), эпитопов нейраминидазы к HLA класса I (SEQ ID NO: 418), эпитопов гемагглютинина к HLA класса II (SEQ ID NO: 420), эпитопов нейраминидазы к HLA класса II (SEQ ID NO: 422), эпитопов гемагглютинина к HLA класса I и класса II, H5N1-связанного, (SEQ ID NO: 424) и эпитопов нейраминидазы к HLA класса I и класса II, H5N1-связанного (SEQ ID NO: 426).

128. Система или способ по п. 122, где вставка представляет собой домен димеризации.

129. Система или способ по п. 128, где домен димеризации представляет собой домен связывания малой молекулы FKBP12 последовательности SEQ ID NO: 488, FKBP12-F36V последовательности SEQ ID NO: 489 или циклофилин последовательности SEQ ID NO: 490 или 493-494.

130. Система одновременного редактирования обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию, причем система содержит:

a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую PEгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую PEгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию,

c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, и вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, третья РЕгРНК содержит третью матрицу для синтеза ДНК, кодирующую третью последовательность одноцепочечной ДНК, и четвертая РЕгРНК содержит четвертую матрицу для синтеза ДНК, кодирующую четвертую последовательность одноцепочечной ДНК,

где первая и третья последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где вторая и четвертая последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу.

131. Система для редактирования одной или нескольких последовательностей двухцепочечной ДНК, причем система содержит:

a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем

редактированию,

b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи первой последовательности двухцепочечной ДНК на первом целевом участке, подлежащем редактированию,

c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с первым участком связывания на первой цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий четвертый программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается со вторым участком связывания на второй цепи второй последовательности двухцепочечной ДНК на втором целевом участке, подлежащем редактированию,

где первая РЕгРНК содержит первую матрицу для синтеза ДНК, кодирующую первую последовательность одноцепочечной ДНК, вторая РЕгРНК содержит вторую матрицу для синтеза ДНК, кодирующую вторую последовательность одноцепочечной ДНК, третья РЕгРНК содержит третью матрицу для синтеза ДНК, кодирующую третью последовательность одноцепочечной ДНК, и четвертая РЕгРНК содержит четвертую матрицу для синтеза ДНК, кодирующую четвертую последовательность одноцепочечной ДНК,

где первая и третья последовательность одноцепочечной ДНК каждая содержит область комплементарности друг другу, и где вторая и четвертая последовательность

одноцепочечной ДНК каждая содержит область комплементарности друг другу.

132. Система по п. 131, где праймирующий редактор первого комплекса праймированного редактирования, второй комплекс праймированного редактирования, третьего комплекса праймированного редактирования и/или четвертого комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

133. Система по любому из пп. 130-132, где

(i) первая РЕгРНК дополнительно содержит первую спейсерную последовательность, первое ядро гРНК и первый участок связывания праймера,

(ii) вторая РЕгРНК дополнительно содержит вторую спейсерную последовательность, второе ядро гРНК и второй участок связывания праймера,

(iii) третья РЕгРНК дополнительно содержит третью спейсерную последовательность, третье ядро гРНК и третий участок связывания праймера, и

(iv) четвертая РЕгРНК содержит четвертую спейсерную последовательность, четвертое ядро гРНК, четвертую матрицу для синтеза ДНК и четвертый участок связывания праймера.

134. Система по п. 131, где область комплементарности первой и третьей последовательности одноцепочечной ДНК и/или область комплементарности второй и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

135. Система по п. 134, где первая спейсерная последовательность связывается с первым участком связывания первой последовательности двухцепочечной ДНК против хода транскрипции от первого целевого участка, подлежащего редактированию, и вторая спейсерная последовательность связывается со вторым участком связывания на первой последовательности двухцепочечной ДНК по ходу транскрипции от первого целевого участка, подлежащего редактированию.

136. Система по п. 134 или 135, где третья спейсерная последовательность связывается с первым участком связывания второй последовательности двухцепочечной ДНК против хода транскрипции от второго целевого участка, подлежащего редактированию, и четвертая спейсерная последовательность связывается со вторым участком связывания на второй последовательности двухцепочечной ДНК по ходу транскрипции от второго целевого участка, подлежащего редактированию.

137. Система по любому из пп. 134-136, где первый и второй участки связывания определяют оба конца первой непрерывной области первой двухцепочечной ДНК.

138. Система по любому из пп. 134-137, где третий и четвертый участки связывания

определяют оба конца второй непрерывной области второй двухцепочечной ДНК.

139. Система по п. 134, где непрерывная область между первым и вторым участками связывания или третьим и четвертым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72

нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

140. Система по любому из пп. 134-139, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК и/или область комплементарности второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

141. Система по п. 140, где дуплекс встраивается в первую двухцепочечную ДНК и/или вторую двухцепочечную ДНК.

142. Система по любому из пп. 134-141, где система приводит к замещению первой непрерывной области второй непрерывной областью в первой двухцепочечной ДНК.

143. Система по любому из пп. 134-142, где система приводит к замене первой непрерывной области и второй непрерывной области между первой двухцепочечной ДНК и второй двухцепочечной ДНК.

144. Система по любому из пп. 134-143, где первый комплекс праймированного редактирования, второй комплекс праймированного редактирования, третий комплекс праймированного редактирования и четвертый комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

145. Система по любому из пп. 134-144, где первый, второй, третий и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, являются одинаковыми.

146. Система по любому из пп. 134-145, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 или его вариант.

147. Система по любому из пп. 134-146, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью, или никазный домен Cas9, или их вариант.

148. Система по любому из пп. 134-147, где первый, второй, третий и/или четвертый napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c и Argonaute.

149. Система по любому из пп. 134-148, где первый, второй, третий и/или четвертый napDNAbp имеет никазную активность.

150. Система по любому из пп. 134-149, где первый, второй, третий и/или четвертый napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 2-65, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 2-65.

151. Система по любому из пп. 134-150, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

152. Система по любому из пп. 134-151, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110.

153. Система по любому из пп. 134-152, где

(i) первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(ii) второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(iii) третий праймирующий редактор дополнительно содержит линкер, который соединяет третий napDNAbp и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или

(iv) четвертый праймирующий редактор дополнительно содержит линкер, который соединяет четвертый napDNAbp и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

154. Система по п. 153, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 119-128, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 119-128.

155. Система по п. 153 или 154, где линкер составляет 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, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

156. Система по любому из пп. 134-155, где первая, вторая, третья и/или четвертая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 192-203, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 192-203.

157. Система по любому из пп. 134-156, где первая спейсерная последовательность связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

158. Система по любому из пп. 134-157, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

159. Система по любому из пп. 134-158, где третья спейсерная последовательность связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

160. Система по любому из пп. 134-159, где четвертая спейсерная последовательность четвертого комплекса праймированного редактирования связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

161. Система по п. 160, где связывание спейсерной последовательности первого и второго комплексов праймированного редактирования в присутствии последовательности РАМ приводит к разрыву первой и второй цепей, соответственно, в участке разрыва,

проксимальном по отношению к последовательностям PAM, на каждой цепи посредством первого и/или второго napDNAbp каждого из первого и второго комплексов праймированного редактирования.

162. Система по п. 161, где связывание спейсерной последовательности третьего и четвертого комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательностям PAM, на каждой цепи посредством третьего и/или четвертого napDNAbp каждого третьего и четвертого комплексов праймированного редактирования.

163. Система по п. 133, где первый и третий участки связывания, или второй и четвертый участки связывания определяют два конца непрерывной области двухцепочечной ДНК.

164. Система по п. 163, где непрерывная область между первым и третьим участками связывания или вторым и четвертым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

165. Система по любому из пп. 122-1163, где область комплементарности первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК и/или область комплементарности второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

166. Система по п. 165, где непрерывная область замещается дуплексом.

167. Система по п. 166, где дуплекс встраивается в двухцепочечную ДНК в целевом участке.

168. Система по любому из пп. 122-167, где первая, вторая, третья и/или четвертая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по

меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

169. Система по любому из пп. 122-168, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК имеют одинаковые длины.

170. Система по любому из пп. 122-169, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК, и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК имеют разные длины.

171. Система по любому из пп. 122-170, где область комплементарности первой, второй, третьей и/или четвертой последовательности одноцепочечной ДНК составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК или второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК.

172. Система по любому из пп. 122-171, где область комплементарности первой, второй, третьей и/или четвертой последовательности одноцепочечной ДНК составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95

нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину, и либо меньше, либо равна длине первой последовательности одноцепочечной ДНК и третьей последовательности одноцепочечной ДНК, и второй последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК.

173. Система по любому из пп. 130-172, где первая, вторая, третья и четвертая последовательность одноцепочечной ДНК представляет собой каждая одноцепочечные цепи ДНК, имеющие 3'-концы.

174. Система по п. 173, где включение дуплекса, содержащего отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

175. Полинуклеотид, кодирующий первый, второй, третий и четвертый праймирующие редакторы по любому из пп. 130-175.

176. Полинуклеотид, кодирующий первую, вторую, третью и четвертую РЕгРНК по любому из пп. 130-176.

177. Вектор, содержащий полинуклеотид по п. 175 или 1176.

178. Клетка, содержащая полинуклеотид по п. 175 или 176 или вектор по п. 177.

179. Фармацевтическая композиция, содержащая полинуклеотид по п. 175 или 176, вектор по п. 177 или клетку по п. 178 и фармацевтическое вспомогательное вещество.

180. Набор, содержащий полинуклеотид по п. 175 или 176 или вектор по п. 177 и инструкции по одновременному редактированию обеих цепей последовательности двухцепочечной ДНК в целевом участке, подлежащем редактированию.

181. Способ одновременного редактирования первой и второй комплементарных цепей последовательности двухцепочечной ДНК в целевом участке, причем способ предусматривает контакт последовательности двухцепочечной ДНК с парой комплексов праймированного редактирования, причем пара содержит:

(a) первый комплекс праймированного редактирования, содержащий:

i. первый праймирующий редактор, содержащий первый программируемый нуклеиновой кислотой ДНК-связывающий белок (первый napDNAbp) и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первой целевой последовательностью на первой цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(b) второй комплекс праймированного редактирования, содержащий:

i. второй праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (второй napDNAbp) и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со второй целевой последовательностью на второй цепи последовательности геномной ДНК против хода транскрипции целевого участка,

(c) третий комплекс праймированного редактирования, содержащий:

i. третий праймирующий редактор, содержащий третий программируемый нуклеиновой кислотой ДНК-связывающий белок (третий napDNAbp) и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с третьей целевой последовательностью на первой цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

(d) четвертый комплекс праймированного редактирования, содержащий:

i. четвертый праймирующий редактор, содержащий второй программируемый нуклеиновой кислотой ДНК-связывающий белок (четвертый napDNAbp) и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и

ii. четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертой целевой последовательностью на второй цепи последовательности геномной ДНК по ходу транскрипции целевого участка,

где первый комплекс праймированного редактирования вызывает первый разрыв на первой целевой последовательности и последующую полимеризацию первой последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного первым разрывом,

где второй комплекс праймированного редактирования вызывает второй разрыв на второй целевой последовательности и последующую полимеризацию второй последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного вторым разрывом,

где третий комплекс праймированного редактирования вызывает третий разрыв на третьей целевой последовательности и последующую полимеризацию третьей

последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного третьим разрывом,

где четвертый комплекс праймированного редактирования вызывает четвертый разрыв на четвертой целевой последовательности и последующую полимеризацию четвертой последовательности одноцепочечной ДНК, имеющей 3’-конец, от доступного 5’-конца, образованного четвертым разрывом,

где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК, и

где третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК обратно комплементарны по меньшей мере в области комплементарности и образуют дуплекс, где дуплекс замещает разорванные первую и вторую комплементарные цепи последовательности двухцепочечной ДНК.

182. Способ по п. 181, где праймирующий редактор первого комплекса праймированного редактирования, второй комплекс праймированного редактирования, третьего комплекса праймированного редактирования и/или четвертого комплекса праймированного редактирования представляет собой слитый белок, содержащий napDNAbp и полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

183. Способ по п. 181 или 182, где первый комплекс праймированного редактирования, второй комплекс праймированного редактирования, третий комплекс праймированного редактирования и четвертый комплекс праймированного редактирования содержат один и тот же праймирующий редактор.

184. Способ по любому из пп. 181-183, где первый, второй, третий и/или четвертый комплекс праймированного редактирования содержат разные праймирующие редакторы.

185. Способ по любому из пп. 181-184, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 или его вариант.

186. Способ по любому из пп. 181-185, где первый, второй, третий и/или четвертый napDNAbp представляет собой домен Cas9 с нуклеазной активностью, домен Cas9 с инактивированной нуклеазной активностью или никазный домен Cas9, или их вариант.

187. Способ по любому из пп. 181-184, где первый, второй, третий и/или четвертый napDNAbp выбран из группы, состоящей из Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c.

188. Способ по любому из пп. 181-187, где первый, второй, третий и/или четвертый

napDNAbp обладает никазной активностью.

189. Способ по любому из пп. 181-188, где первый, второй, третий и/или четвертый napDNAbp содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 2-65 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 2-65.

190. Способ по любому из пп. 181-189, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, представляет собой обратную транскриптазу.

191. Способ по любому из пп. 181-190, где первый, второй, третий и/или четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110 или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 37, 68-79, 82-98, 81, 98 и 110.

192. Способ по любому из пп. 181-191, где

(i) первый праймирующий редактор дополнительно содержит линкер, который соединяет первый napDNAbp и первый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(ii) второй праймирующий редактор дополнительно содержит линкер, который соединяет второй napDNAbp и второй полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы,

(iii) третий праймирующий редактор дополнительно содержит линкер, который соединяет третий napDNAbp и третий полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы, и/или

(iv) четвертый праймирующий редактор дополнительно содержит линкер, который соединяет четвертый napDNAbp и четвертый полипептид, обладающий активностью РНК-зависимой ДНК-полимеразы.

193. Способ по п. 192, где линкер содержит аминокислотную последовательность согласно любой из последовательностей SEQ ID NO: 119-128, или аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 119-128.

194. Способ по п. 192 или 193, где линкер составляет 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, 38, 38,

39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 или 50 аминокислот в длину.

195. Способ по любому из пп. 192-194, где первая, вторая, третья и/или четвертая РЕгРНК содержит нуклеотидную последовательность согласно любой из последовательностей SEQ ID NO: 192-203, или нуклеотидную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 95% или 99% идентичности последовательности с любой из последовательностей SEQ ID NO: 192-203.

196. Способ по любому из пп. 192-195, где первая спейсерная последовательность первого комплекса праймированного редактирования связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

197. Способ по любому из пп. 192-196, где вторая спейсерная последовательность второго комплекса праймированного редактирования связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

198. Способ по любому из пп. 192-197, где третья спейсерная последовательность третьего комплекса праймированного редактирования связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию.

199. Способ по любому из пп. 192-198, где четвертая спейсерная последовательность четвертого комплекса праймированного редактирования связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию.

200. Способ по п. 199, где связывание первой и второй спейсерных последовательностей первого и второго комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого первого и второго комплекса праймированного редактирования.

201. Способ по п. 200, где связывание третьей и четвертой спейсерных последовательностей третьего и четвертого комплексов праймированного редактирования в присутствии последовательности PAM приводит к разрыву первой и второй цепей, соответственно, в участке разрыва, проксимальном по отношению к последовательности PAM, на каждой цепи посредством napDNAbp каждого третьего и четвертого комплекса праймированного редактирования.

202. Способ по п. 200 или 201, где первый и второй участки связывания и/или третий и четвертый участки связывания определяют два конца непрерывной области двухцепочечной ДНК, содержащей целевой участок, подлежащий редактированию.

203. Способ по п. 202, где непрерывная область между первым и вторым участками связывания и/или непрерывная область между третьим и четвертым участками связывания составляет по меньшей мере 10 пар нуклеотидных оснований, по меньшей мере 11 пар нуклеотидных оснований, по меньшей мере 12 пар нуклеотидных оснований, по меньшей мере 13 пар нуклеотидных оснований, по меньшей мере 14 пар нуклеотидных оснований, по меньшей мере 15 пар нуклеотидных оснований, по меньшей мере 16 пар нуклеотидных оснований, по меньшей мере 17 пар нуклеотидных оснований, по меньшей мере 18 пар нуклеотидных оснований, по меньшей мере 19 пар нуклеотидных оснований, по меньшей мере 20 пар нуклеотидных оснований, по меньшей мере 21 пар нуклеотидных оснований, по меньшей мере 22 пар нуклеотидных оснований, по меньшей мере 23 пар нуклеотидных оснований, по меньшей мере 24 пар нуклеотидных оснований, по меньшей мере 25 пар нуклеотидных оснований, по меньшей мере 26 пар нуклеотидных оснований, по меньшей мере 27 пар нуклеотидных оснований, по меньшей мере 28 пар нуклеотидных оснований, по меньшей мере 29 пар нуклеотидных оснований, по меньшей мере 30 пар нуклеотидных оснований, по меньшей мере 31 пар нуклеотидных оснований, по меньшей мере 32 пар нуклеотидных оснований, по меньшей мере 33 пар нуклеотидных оснований, по меньшей мере 34 пар нуклеотидных оснований, по меньшей мере 35 пар нуклеотидных оснований, по меньшей мере 36 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 37 пар нуклеотидных оснований, по меньшей мере 38 пар нуклеотидных оснований, по меньшей мере 39 пар нуклеотидных оснований или по меньшей мере 40 пар нуклеотидных оснований, 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по

меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

204. Способ по п. 202 или 203, где непрерывная область замещается дуплексом.

205. Способ по любому из пп. 185-204, где первая, вторая, третья и/или четвертая матрица для синтеза ДНК составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

206. Способ по любому из пп. 185-205, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК и/или третья последовательность одноцепочечной ДНК и четвертая последовательность

одноцепочечной ДНК имеют одинаковые длины.

207. Способ по любому из пп. 185-206, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК и/или третья последовательность одноцепочечной и четвертая последовательность одноцепочечной ДНК имеют разные длины.

208. Способ по любому из пп. 185-207, где область комплементарности составляет по меньшей мере 99%, по меньшей мере 95%, по меньшей мере 90%, по меньшей мере 85%, по меньшей мере 80%, по меньшей мере 75%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 60%, по меньшей мере 55%, по меньшей мере 50%, по меньшей мере 45%, по меньшей мере 40%, по меньшей мере 35%, по меньшей мере 30%, по меньшей мере 25%, по меньшей мере 20%, по меньшей мере 15% или по меньшей мере 10% от общей длины первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК или третьей последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК.

209. Способ по любому из пп. 185-208, где область комплементарности составляет 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по

меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

210. Способ по любому из пп. 188-209, где область комплементарности первой последовательности одноцепочечной ДНК и второй последовательности одноцепочечной ДНК и/или третьей последовательности одноцепочечной ДНК и четвертой последовательности одноцепочечной ДНК образуют дуплекс.

211. Способ по любому из пп. 188-210, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК, представляет собой каждая одноцепочечные цепи ДНК, имеющие 3’-концы.

212. Способ по п. 204, где замещение непрерывной области дуплексом, содержащим

отредактированную часть, приводит к вставке, делеции или замене ДНК в целевом участке.

213. Способ по любому из пп. 181-212, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК находятся на концах целевой последовательности ДНК, и где третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на противоположных концах той же целевой последовательности ДНК.

214. Способ по п. 213, где целевая последовательность ДНК является инвертированной.

215. Способ по любому из пп. 185-214 дополнительно предусматривает предоставление донора кольцевой ДНК.

216. Способ по п. 215, где первая последовательность одноцепочечной ДНК и третья последовательность одноцепочечной ДНК находятся на противоположных концах целевой последовательности ДНК, и где вторая последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на доноре кольцевой ДНК.

217. Способ по п. 216, где часть донора кольцевой ДНК между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК замещает целевую последовательность между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК.

218. Способ по любому из пп. 185-217, где первая последовательность одноцепочечной ДНК и третья последовательность одноцепочечной ДНК находятся на первой молекуле нуклеиновой кислоты, и где вторая последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК находятся на второй молекуле нуклеиновой кислоты.

219. Способ по п. 218, где часть первой молекулы нуклеиновой кислоты между первой последовательностью одноцепочечной ДНК и третьей последовательностью одноцепочечной ДНК встраивается во вторую молекулу нуклеиновой кислоты, и где часть второй молекулы нуклеиновой кислоты между второй последовательностию одноцепочечной ДНК и четвертой последовательностью одноцепочечной ДНК встраивается в первую молекулу нуклеиновой кислоты.

220. Способ по п. 218 или 219, где первая молекула нуклеиновой кислоты представляет собой первую хромосому, и вторая молекула нуклеиновой кислоты представляет собой вторую хромосому.

221. Набор РЕгРНК для применения в мультилоскутном праймированном

редактировании, где РЕгРНК содержат:

(a) первую гидовую РНК праймированного редактирования (первую РЕгРНК), которая связывается с первым участком связывания на первой цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию, где первая РЕгРНК содержит первую спейсерную последовательность, первое ядро гРНК, первую матрицу для синтеза ДНК и первый участок связывания праймера, где первая матрица для синтеза ДНК кодирует первую последовательность одноцепочечной ДНК,

(b) вторую гидовую РНК праймированного редактирования (вторую РЕгРНК), которая связывается со вторым участком связывания на второй цепи последовательности двухцепочечной ДНК против хода транскрипции от целевого участка, подлежащего редактированию, где вторая РЕгРНК содержит вторую спейсерную последовательность, второе ядро гРНК, вторую матрица для синтеза ДНК и второй участок связывания праймера, где вторая матрица для синтеза ДНК кодирует вторую последовательность одноцепочечной ДНК, содержащую отредактированную часть, где вторая последовательность одноцепочечной ДНК комплементарна первой последовательности одноцепочечной ДНК.

(c) третью гидовую РНК праймированного редактирования (третью РЕгРНК), которая связывается с третьим участком связывания на первой цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где третья РЕгРНК содержит третью спейсерную последовательность, третье ядро гРНК, третью матрицу для синтеза ДНК и третий участок связывания праймера, где третья матрица для синтеза ДНК кодирует третью последовательность одноцепочечной ДНК,

(d) четвертую гидовую РНК праймированного редактирования (четвертую РЕгРНК), которая связывается с четвертым участком связывания на второй цепи последовательности двухцепочечной ДНК по ходу транскрипции от целевого участка, подлежащего редактированию, где четвертая РЕгРНК содержит четвертую спейсерную последовательность, четвертое ядро гРНК, четвертую матрица для синтеза ДНК и четвертый участок связывания праймера, где четвертая матрица для синтеза ДНК кодирует четвертую последовательность одноцепочечной ДНК, где четвертая последовательность одноцепочечной ДНК содержит область комплементарности с третьей последовательностью одноцепочечной ДНК.

222. Набор РЕгРНК по п. 221, где

первая РЕгРНК содержит структуру:

5′-[первая спейсерная последовательность гРНК]-[первое ядро гРНК]-[первая матрица для синтеза ДНК]-[первый участок связывания праймера]-3′,

вторая РЕгРНК содержит структуру:

5′-[вторая спейсерная последовательность гРНК]-[второе ядро гРНК]-[вторая матрица для синтеза ДНК]-[второй участок связывания праймера]-3′,

третья РЕгРНК содержит структуру:

5′-[третья спейсерная последовательность гРНК]-[третье ядро гРНК]-[третья матрица для синтеза ДНК]-[третий участок связывания праймера]-3′, и

четвертая РЕгРНК содержит структуру:

5′-[четвертая спейсерная последовательность гРНК]-[четвертое ядро гРНК]-[четвертая матрица для синтеза ДНК]-[четвертый участок связывания праймера]-3′.

223. Набор РЕгРНК по п. 221 или 222, где первая последовательность одноцепочечной ДНК и вторая последовательность одноцепочечной ДНК и/или третья последовательность одноцепочечной ДНК и четвертая последовательность одноцепочечной ДНК образуют дуплекс, который интегрируется в последовательность двухцепочечной ДНК в целевом участке, подлежащем редактированию.

224. Набор РЕгРНК по любому из пп. 221-223, где интеграция дуплекса приводит к вставке, делеции или замене ДНК в целевом участке.

225. Набор РЕгРНК по п. 224, где вставка составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов в длину.

226. Набор РЕгРНК по п. 224, где делеция составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 на 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67 нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по

меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

227. Набор РЕгРНК по п. 224, где замена составляет по меньшей мере 1 нуклеотид, по меньшей мере 2 нуклеотида, 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 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, по меньшей мере 50 нуклеотидов, по меньшей мере 51 нуклеотидов, по меньшей мере 52 нуклеотидов, по меньшей мере 53 нуклеотидов, по меньшей мере 54 нуклеотидов, по меньшей мере 55 нуклеотидов, по меньшей мере 56 нуклеотидов, по меньшей мере 57 нуклеотидов, по меньшей мере 58 нуклеотидов, по меньшей мере 59 нуклеотидов, по меньшей мере 60 нуклеотидов, по меньшей мере 61 нуклеотидов, по меньшей мере 62 нуклеотидов, по меньшей мере 63 нуклеотидов, по меньшей мере 64 нуклеотидов, по меньшей мере 65 нуклеотидов, по меньшей мере 66 нуклеотидов, по меньшей мере 67

нуклеотидов, по меньшей мере 68 нуклеотидов, по меньшей мере 69 нуклеотидов, по меньшей мере 70 нуклеотидов, по меньшей мере 71 нуклеотидов, по меньшей мере 72 нуклеотидов, по меньшей мере 73 нуклеотидов, по меньшей мере 74 нуклеотидов, по меньшей мере 75 нуклеотидов, по меньшей мере 76 нуклеотидов, по меньшей мере 77 нуклеотидов, по меньшей мере 78 нуклеотидов, по меньшей мере 79 нуклеотидов, по меньшей мере 80 нуклеотидов, по меньшей мере 81 нуклеотидов, по меньшей мере 82 нуклеотидов, по меньшей мере 83 нуклеотидов, по меньшей мере 84 нуклеотидов, по меньшей мере 85 нуклеотидов, по меньшей мере 86 нуклеотидов, по меньшей мере 87 нуклеотидов, по меньшей мере 88 нуклеотидов, по меньшей мере 89 нуклеотидов, по меньшей мере 90 нуклеотидов, по меньшей мере 91 нуклеотидов, по меньшей мере 92 нуклеотидов, по меньшей мере 93 нуклеотидов, по меньшей мере 94 нуклеотидов, по меньшей мере 95 нуклеотидов, по меньшей мере 96 нуклеотидов, по меньшей мере 97 нуклеотидов, по меньшей мере 98 нуклеотидов, по меньшей мере 99 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 101 нуклеотидов, по меньшей мере 102 нуклеотидов, по меньшей мере 103 нуклеотидов, по меньшей мере 104 нуклеотидов, по меньшей мере 105 нуклеотидов, по меньшей мере 106 нуклеотидов, по меньшей мере 107 нуклеотидов, по меньшей мере 108 нуклеотидов, по меньшей мере 109 нуклеотидов или по меньшей мере 110 нуклеотидов, или более в длину.

228. Полинуклеотид, кодирующий набор РЕгРНК по любому из пп. 221-227.

229. Вектор, содержащий полинуклеотид по п. 228 для экспрессии множества РЕгРНК в клетке.

230. Клетка, содержащая вектор по п. 229.

231. Фармацевтическая композиция, содержащая множество РЕгРНК по любому из пп. 221-227, вектор по п. 229 или клетку по п. 230 и фармацевтическое вспомогательное вещество.

232. Способ модификации генома в клетке, предусматривающий:

(i) приведение генома в контакт с системой двухлоскутного праймированного редактирования, причем указанная система содержит первый комплекс праймирующего редактора/РЕгРНК, способный устанавливать первый 3’-лоскут нуклеиновой кислоты в первом участке разрыва на первой цепи целевого участка, и второй комплекс праймирующего редактора/РЕгРНК, способный устанавливать второй 3’-лоскут нуклеиновой кислоты во втором участке разрыва на второй цепи целевого участка,

где первый 3’-лоскут нуклеиновой кислоты и второй 3’-лоскут нуклеиновой кислоты представляют собой обратно комплементарные последовательности, которые образуют

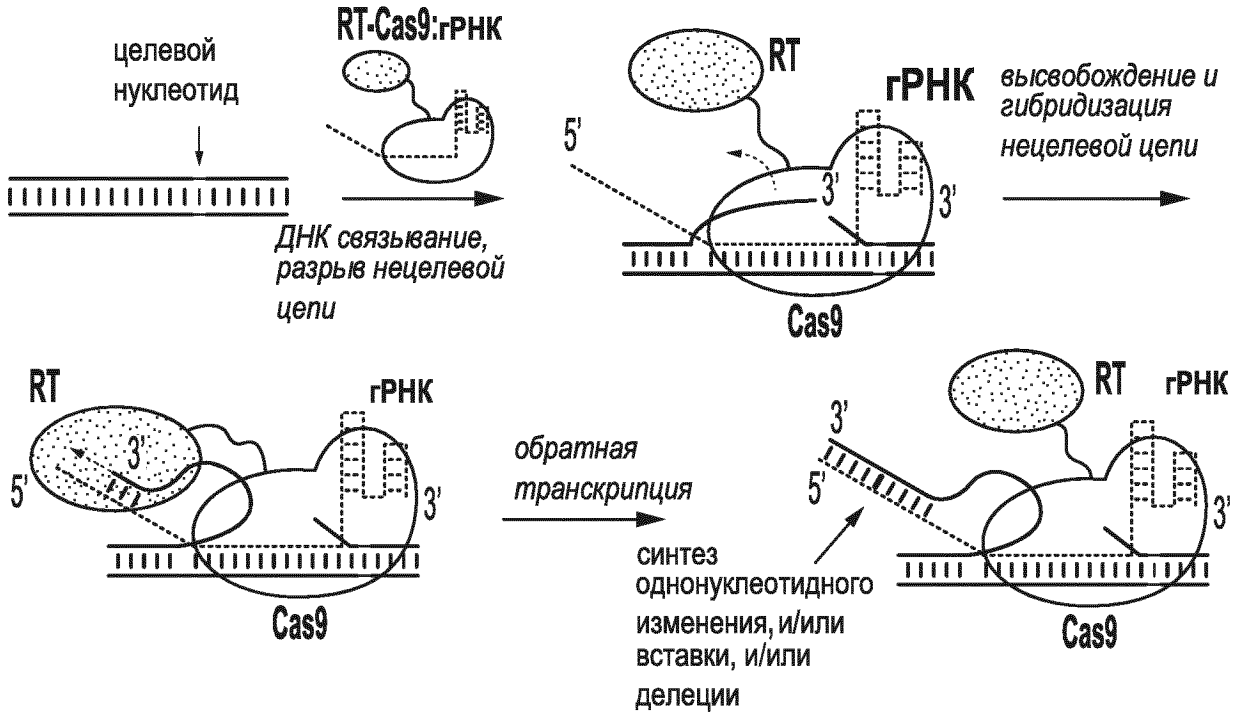
дуплекс, и

где дуплекс содержит первый сайт рекомбиназы, и

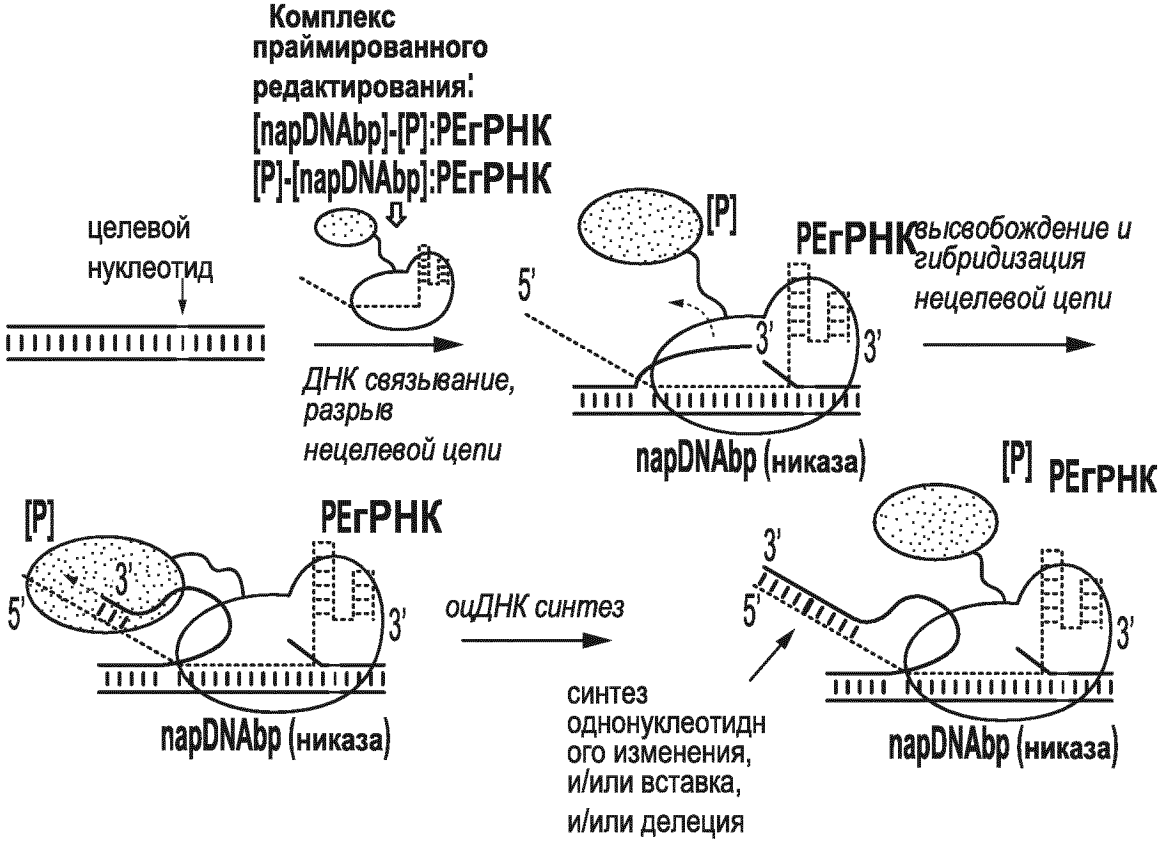
где клетка устанавливает дуплекс, содержащий первый сайт рекомбиназы, в целевом участке вместо эндогенного дуплекса, расположенного между первым и вторым участками разрыва,

(ii) контакт первого сайта рекомбиназы с донорной нуклеиновой кислотой, содержащей второй сайт рекомбиназы, и

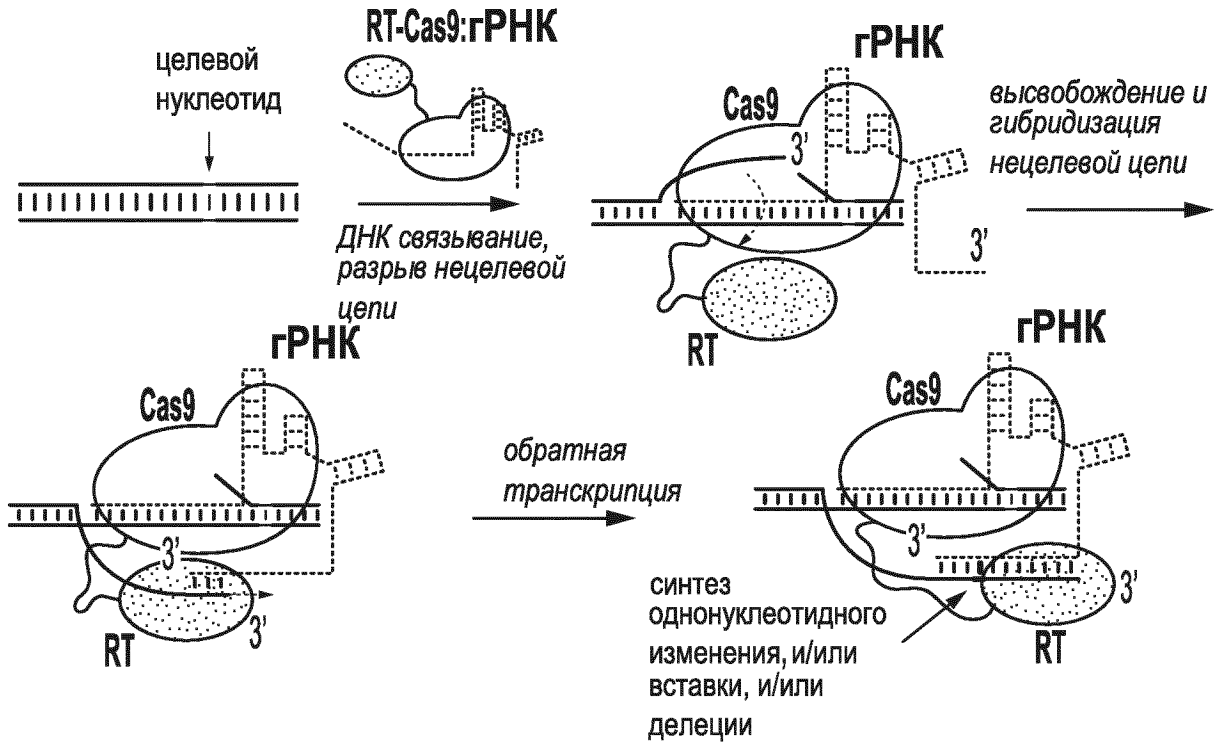
(iii) контакт генома с рекомбиназой, способной к рекомбинации между первым и вторым сайтами рекомбиназы, тем самым интегрируя донорную нуклеиновую кислоту в целевой участок в геноме.


целевой
нуклеотид
RT-Cas9:гРНК
ДНК связывание,
разрыв нецелевой
цепи
RT
гРНК
5'
3'
3'
Cas9
высвобождение и
гибридизация
нецелевой цепи
RT
гРНК
5'
3'
3'
Cas9
обратная
транскрипция
RT
гРНК
3'
5'
3'
Cas9
синтез
однонуклеотидного
изменения, и/или
вставки, и/или
делеции

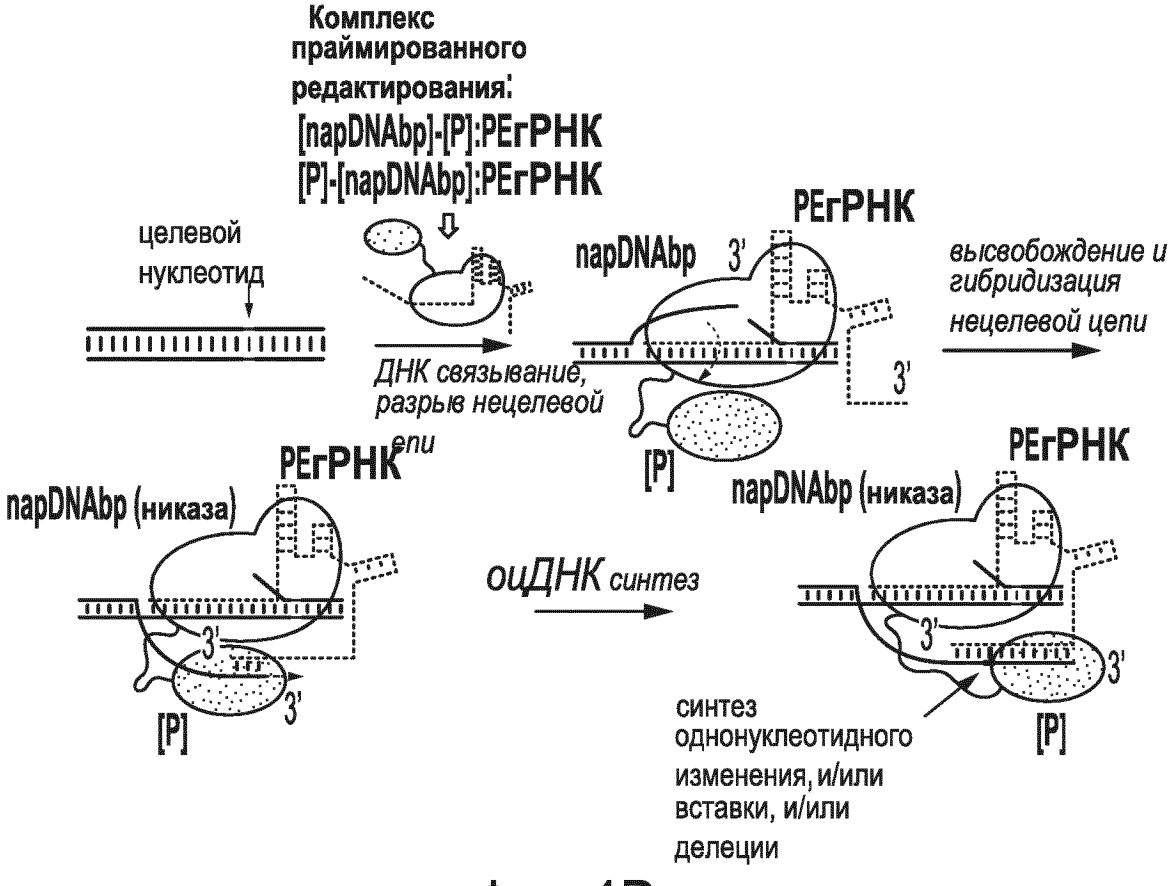
Фиг.1A


Комплекс
праймированного
редактирования:
[napDNAbp]-[P]:PEгРНК
[P]-[napDNAbp]:PEгРНК
целевой
нуклеотид
ДНК связывание,
разрыв
нецелевой цепи
[P]
PEгРНК
5'
3'
3'
napDNAbp (никаза)
высвобождение и
гибридизация
нецелевой цепи
[P]
PEгРНК
5'
3'
3'
napDNAbp (никаза)
оцДНК синтез
[P]
PEгРНК
3'
5'
3'
napDNAbp (никаза)
синтез
однонуклеотидн
ого изменения,
и/или вставка,
и/или делеция

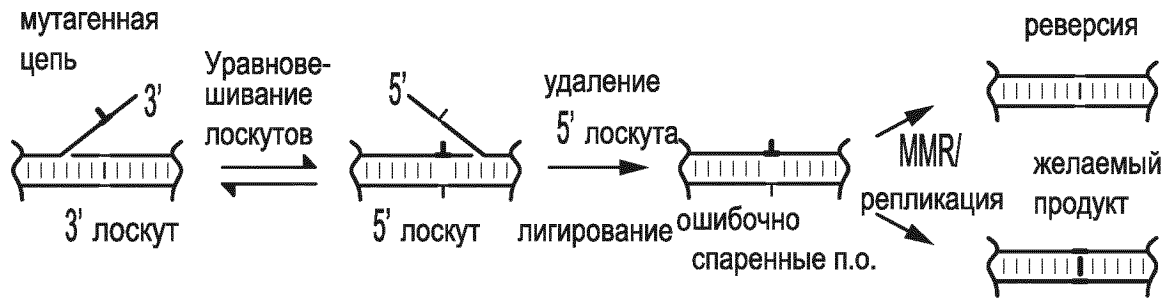
Фиг. 1B


целевой
нуклеотид
RT-Cas9:гРНК
ДНК связывание,
разрыв нецелевой
цепи
гРНК
Cas9
3'
RT
высвобождение и
гибридизация
нецелевой цепи
обратная
транскрипция
синтез
однонуклеотидного
изменения, и/или
вставки, и/или
делеции

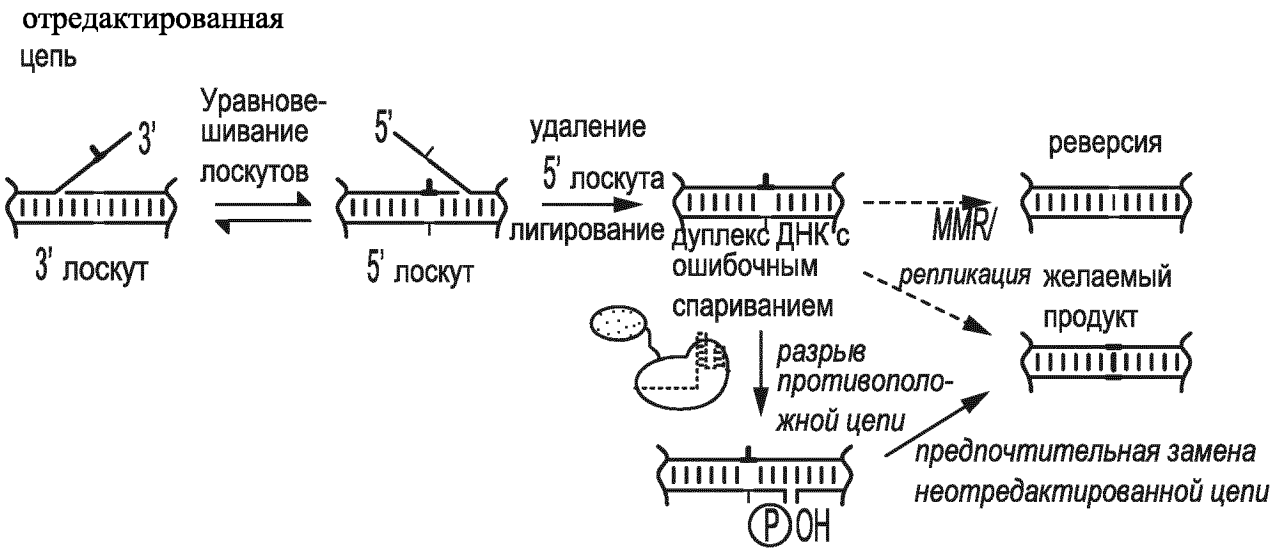
Фиг. 1C


Комплекс
праймированного
редактирования:
[napDNAbp]-[P]:PEгРНК
[P]-[napDNAbp]:PEгРНК
целевой
нуклеотид
ДНК связывание,
разрыв нецелевой
цепи
PEгРНК
napDNAbp
3'
[P]
высвобождение и
гибридизация
нецелевой цепи
napDNAbp (никаза)
оцДНК синтез
синтез
однонуклеотидного
изменения, и/или
вставки, и/или
делеции

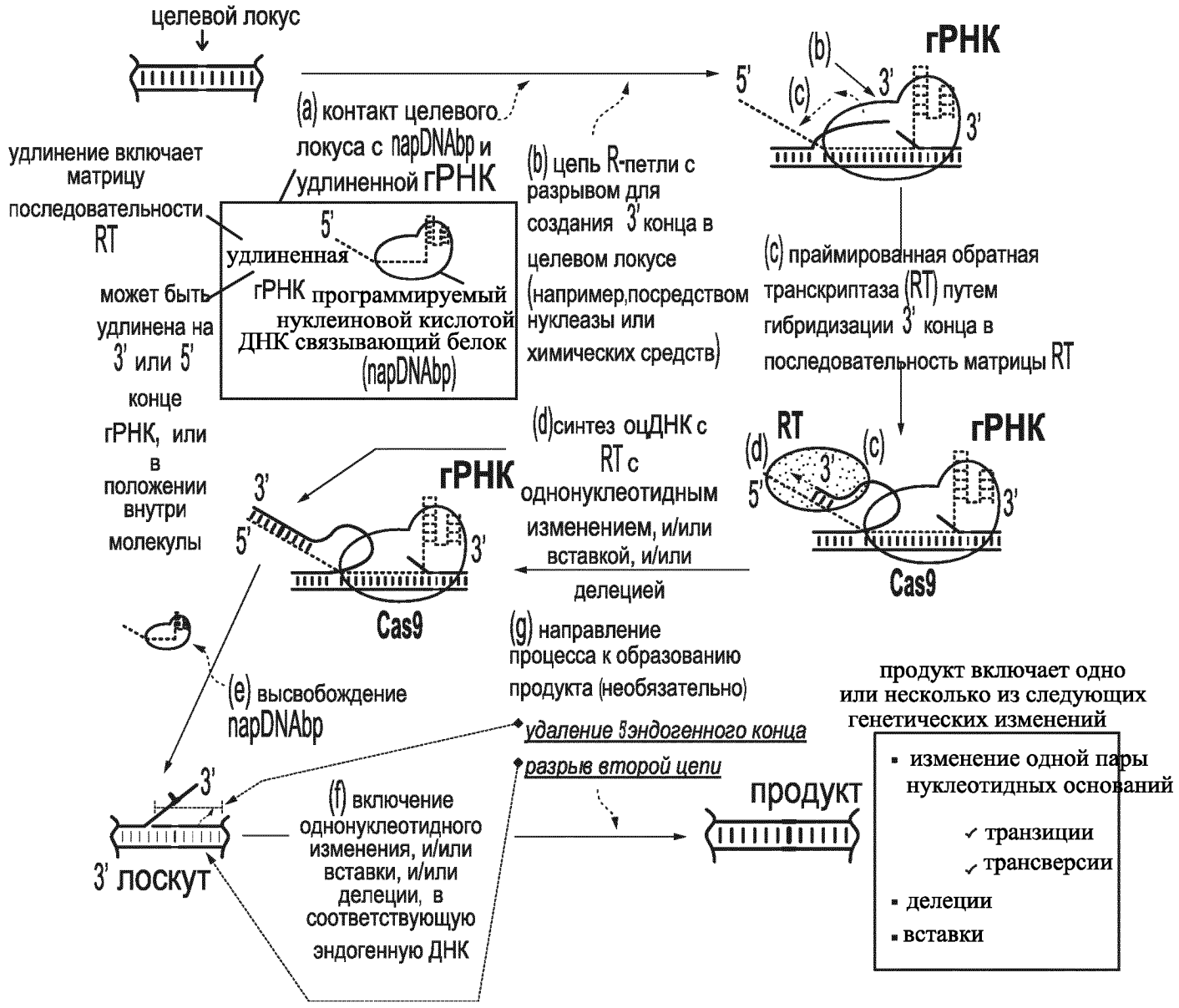
Фиг. 1D


мутагенная
цепь
3'
3' лоскут
Уравнове-
шивание
лоскутов
5'
5' лоскут
удаление
5' лоскута
лигирование
ошибочно
спаренные п.о.
MMR/
репликация
реверсия
желаемый
продукт

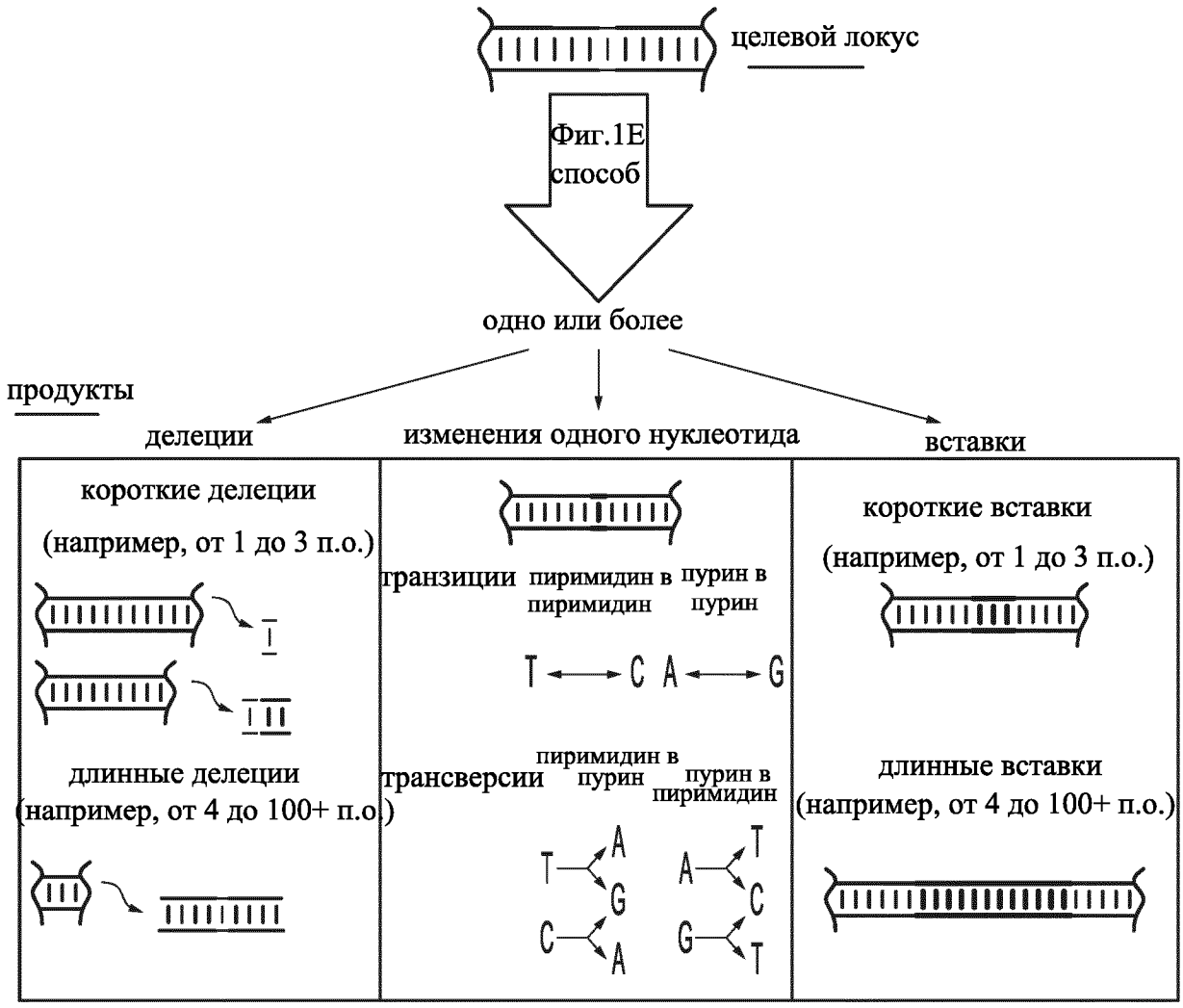
Фиг. 1E


отредактированная
цепь
3'
3' лоскут
Уравнове-
шивание
лоскутов
5'
5' лоскут
удаление
5' лоскута
лигирование
дуплекс ДНК с
ошибочным
спариванием
MMR/
репликация
реверсия
желаемый
продукт
разрыв
противополо-
жной цепи
предпочтительная замена
неотредактированной цепи
POH

Фиг. 1F


целевой локус
(a) контакт целевого локуса с napDNAbp и удлиненной гРНК
удлинение включает матрицу последовательности RT
может быть удлинена на 3' или 5' конце гРНК, или в положении внутри молекулы
5'
удлиненная гРНК программируемый нуклеиновой кислотой ДНК связывающий белок (napDNAbp)
(b) цепь R-петли с разрывом для создания 3' конца в целевом локусе (например,посредством нуклеазы или химических средств)
гРНК
(b)
(c)
5'
3'
3'
(c) праймированная обратная транскриптаза (RT) путем гибридизации 3' конца в последовательность матрицы RT
RT
гРНК
(d)
(c)
5'
3'
3'
Cas9
(d)синтез оцДНК с RT с однонуклеотидным изменением, и/или вставкой, и/или делецией
гРНК
3'
5'
3'
Cas9
(e) высвобождение napDNAbp
(g) направление процесса к образованию продукта (необязательно)
удаление 3эндогенного конца
разрыв второй цепи
продукт
продукт включает одно или несколько из следующих генетических изменений
▪ изменение одной пары нуклеотидных оснований
✓ транзиции
✓ трансверсии
▪ делеции
▪ вставки
3'
3' лоскут
(f) включение однонуклеотидного изменения, и/или вставки, и/или делеции, в соответствующую эндогенную ДНК

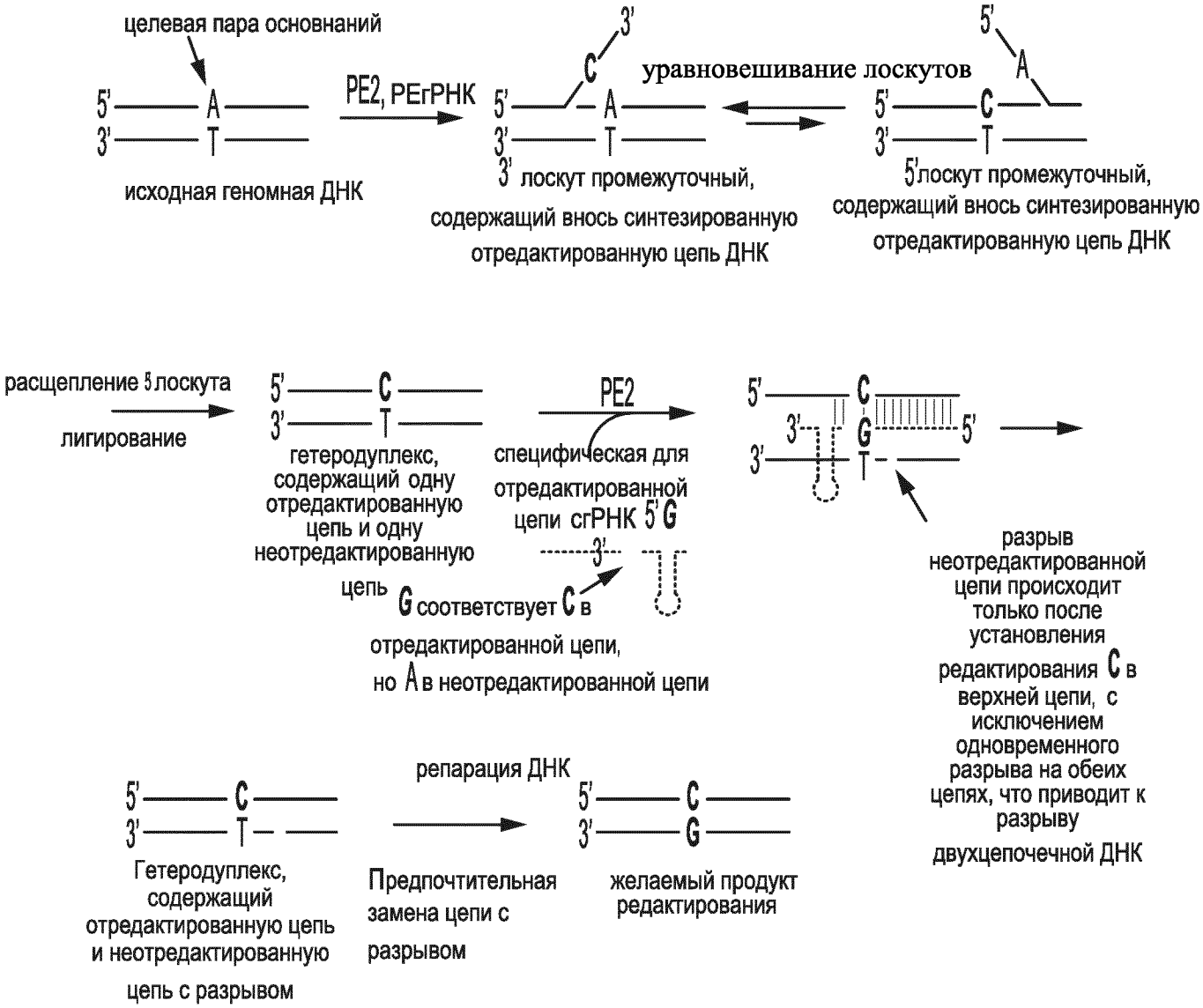
Фиг. 1G


целевой локус
Фиг.1E
способ
одно или более
продукты
делеции
изменения одного нуклеотида
вставки
короткие делеции
(например, от 1 до 3 п.о.)
транзиции
пиримидин в пиримидин
пурин в пурин
T C A G
короткие вставки
(например, от 1 до 3 п.о.)
длинные делеции
(например, от 4 до 100+ п.о.)
трансверсии
пиримидин в пурин
пурин в пиримидин
T A G C
A T C G
длинные вставки
(например, от 4 до 100+ п.о.)

Фиг.1H

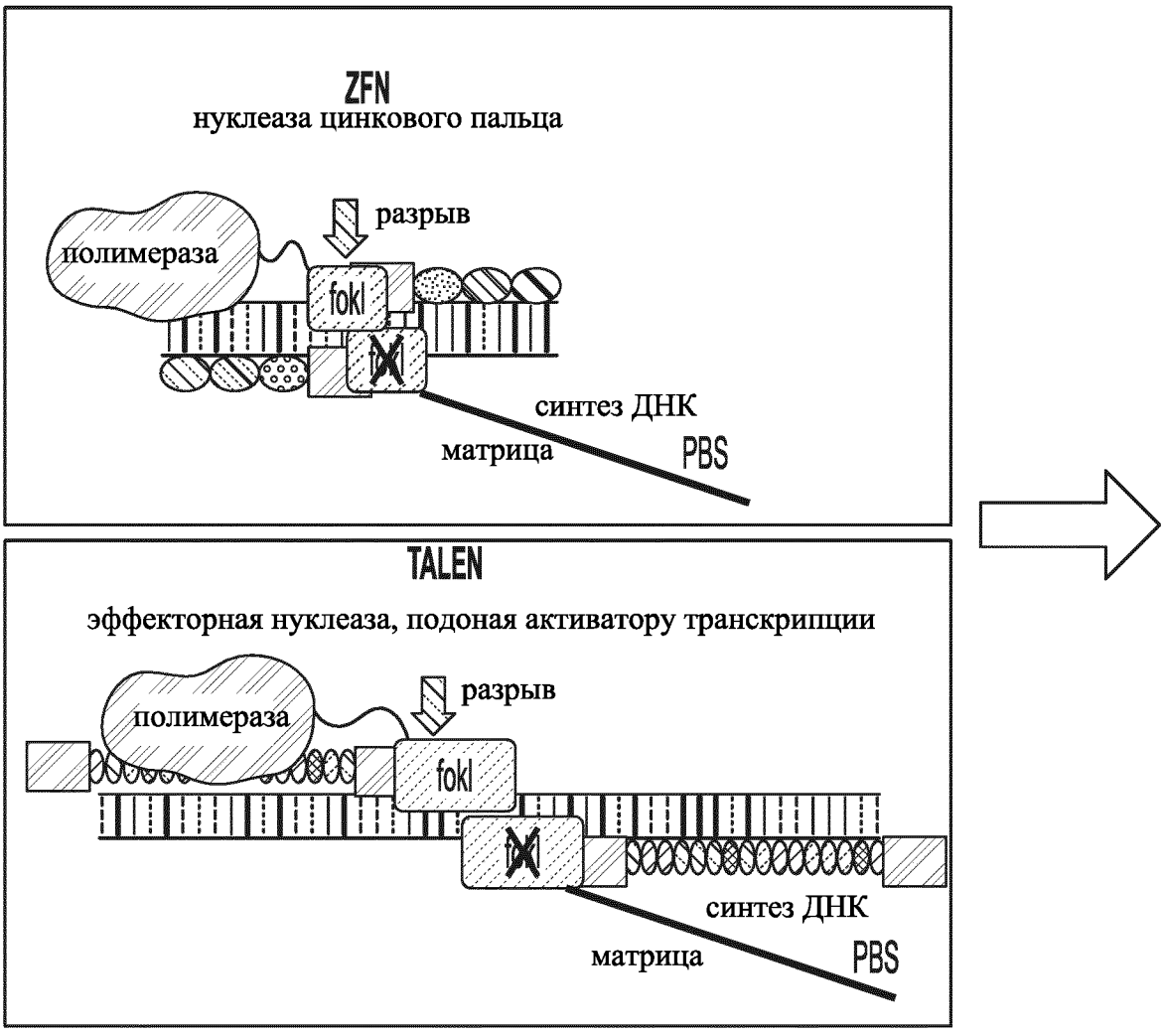
PE3b временной разрыв.

PE3b система редактирования: специфический для редактирования разрыв комплементарной (нередактируемой)цепи ДНК



Фиг. 1I

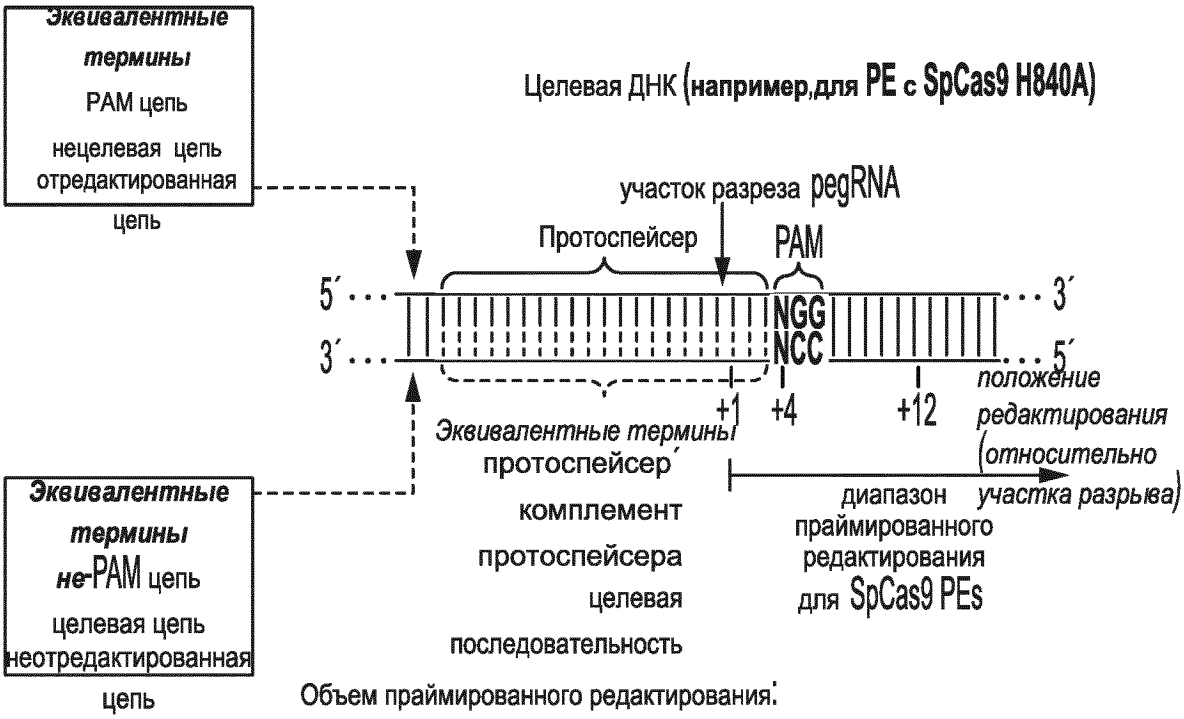
Праймированное редактированное с любой программируемой нуклеазой



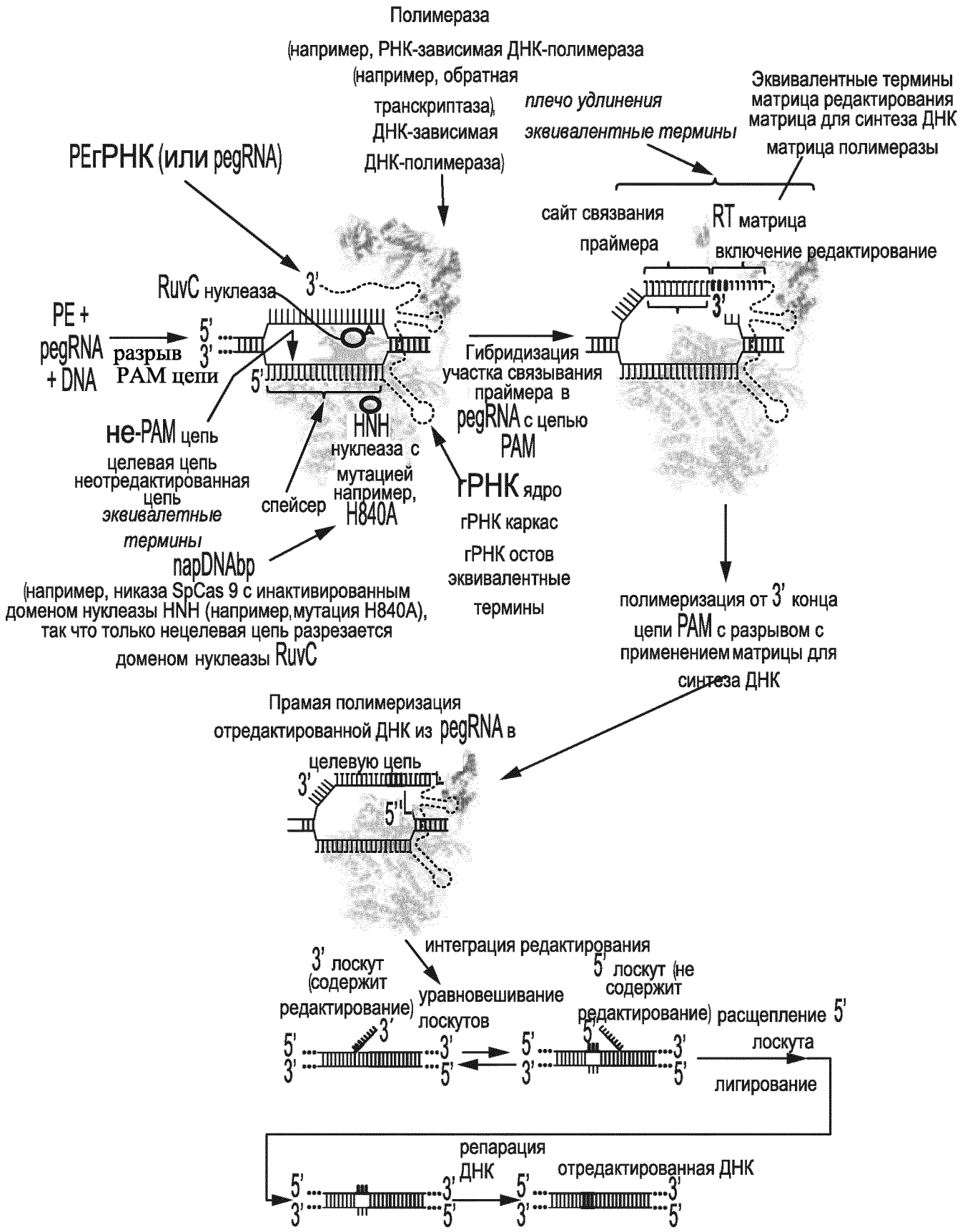
Фиг.1J

Полимераза-зависимый синтез замещающей цепи, кодируемой полимеразой относительно матрицы для синтеза ДНК, и которая содержит желаемое редактирование. Механизм подобен праймированному редактированию. Дополнительные факторы или эффекторы могут быть добавлены в транс или как слияния для облегчения реакции в целом (например, (a) хеликаза для раскручивания ДНК в сайте разреза с получением разрезанной цепи с 3' концом, доступным в качестве праймера, (b) FEN1 для помощи в удалении эндогенной цепи на разрезанной цепи для направления реакции в сторону замены эндогенной цепи синтезированной цепью, или (c) комплекс nCas9:гРНК для создания второго участка разреза на противоположной цепи, что может помочь направить интеграцию синтезированной репарации посредством благоприятной клеточной репарации неотредактированной цепи)

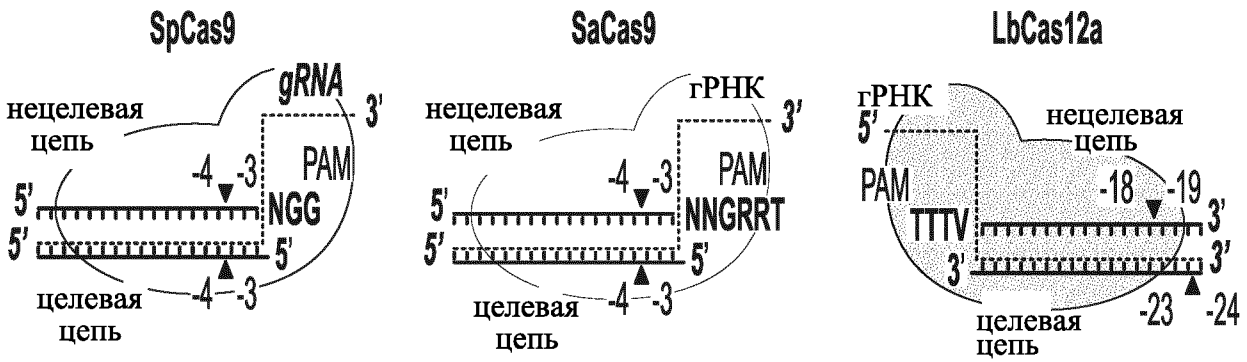
Фиг. 1K


Эквивалентные термины
PAM цепь
нецелевая цепь
отредактированная цепь
Целевая ДНК (например,для PE с SpCas9 H840A)
участок разреза pegRNA
Протоспейсер
PAM
5´
3´
NGG
NCC
3´
5´
+1
+4
+12
положение редактирования (относительно участка разрыва)
Эквивалентные термины
протоспейсер
комплемент протоспейсера
целевая последовательность
диапазон праймированного редактирования для SpCas9 PEs
Эквивалентные термины не-PAM цепь
целевая цепь
неотредактированная цепь
Объем праймированного редактирования:
все 4 точечные мутации транзиции
все 8 точечных мутаций трансверсии
вставки (1 п.о. до ≥ 44 п.о.)
делеции (1 п.о. до ≥ 80 п.о.)
комбинации указанного

Фиг. 1L


Полимераза
(например, РНК-зависимая ДНК-полимераза
(например, обратная
транскриптаза),
ДНК-зависимая
ДНК-полимераза)
плечо удлинения
эквивалентные термины
Эквивалентные термины
матрица редактирования
матрица для синтеза ДНК
матрица полимеразы
PEгРНК (или pegRNA)
сайт связывания
праймера
RT матрица
включение редактирование
RuvC нуклеаза
3'
PE +
pegRNA
+ DNA
разрыв
PAM цепи
5'
3'
5'
Гибридизация
участка связывания
праймера в
pegRNA с цепью
PAM
не-PAM цепь
целевая цепь
неотредактированная
цепь
эквивалетные
термины
HNH
нуклеаза с
мутацией
например,
H840A
спейсер
гРНК ядро
гРНК каркас
гРНК остов
эквивалентные
термины
napDNAbp
(например, никаза SpCas 9 с инактивированным
доменом нуклеазы HNH (например, мутация H840A),
так что только нецелевая цепь разрезается
доменом нуклеазы RuvC
полимеризация от 3' конца
цепи PAM с разрывом с
применением матрицы для
синтеза ДНК
Прамая полимеризация
отредактированной ДНК из pegRNA в
целевую цепь
3'
5'
интеграция редактирования
3' лоскут
(содержит
редактирование)
уравновешивание
лоскутов
5' лоскут (не
содержит
редактирование)
расщепление 5'
лоскута
лигирование
репарация
ДНК
отредактированная ДНК

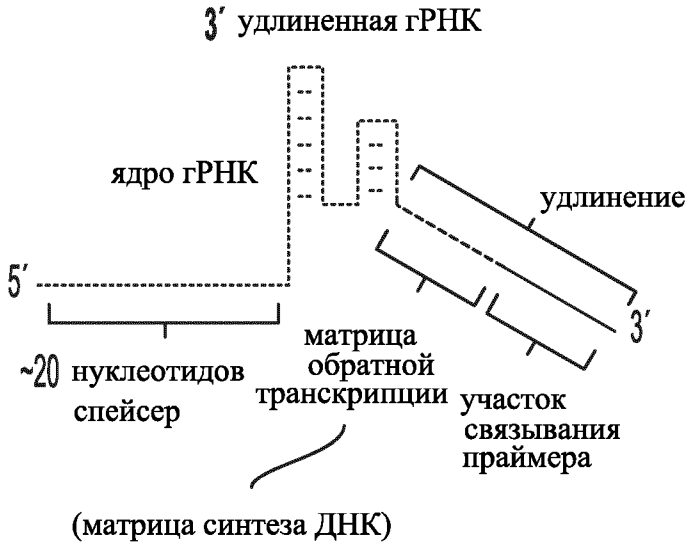
Фиг. 1M


SpCas9
нецелевая цепь
gRNA
3'
PAM
-4
-3
5'
NGG
5'
5'
целевая цепь
-4
-3
SaCas9
нецелевая цепь
гРНК
3'
PAM
-4
-3
5'
NNGRRT
5'
целевая цепь
-4
-3
LbCas12a
гРНК
5'
нецелевая цепь
PAM
TTTV
-18
-19
3'
3'
3'
целевая цепь
-23
-24

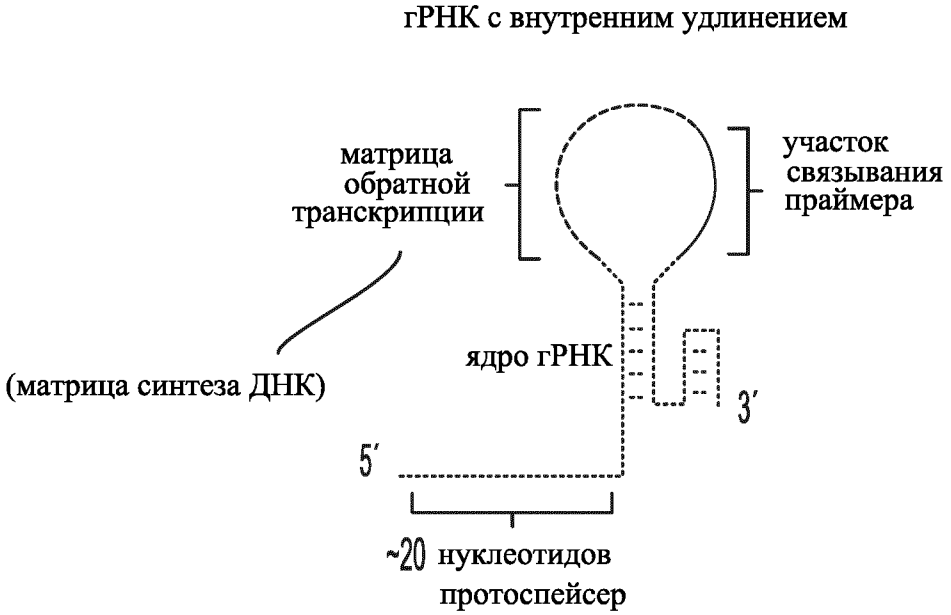
Фиг. 2


5' удлиненная гРНК
5'
удлинение
ядро гРНК
матрица обратной транскрипции
3'
сайт связывания праймера
5-20 нуклеотидов
линкер
~20 нуклеотидов
спейсер
(матрица синтеза ДНК)

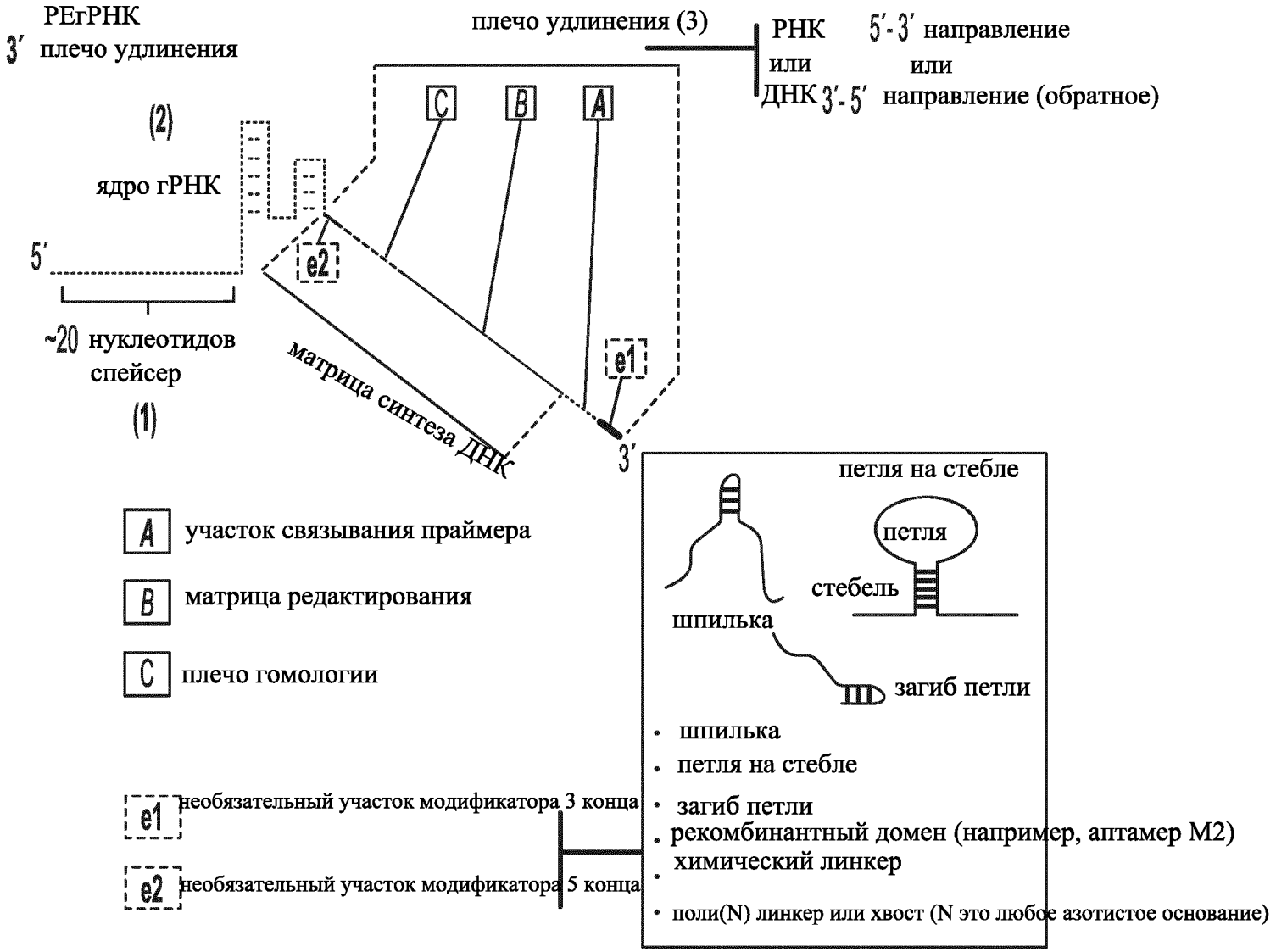
Фиг. 3А


3′ удлиненная гРНК
ядро гРНК
удлинение
5′
3′
~20 нуклеотидов
спейсер
матрица
обратной
транскрипции
участок
связывания
праймера
(матрица синтеза ДНК)

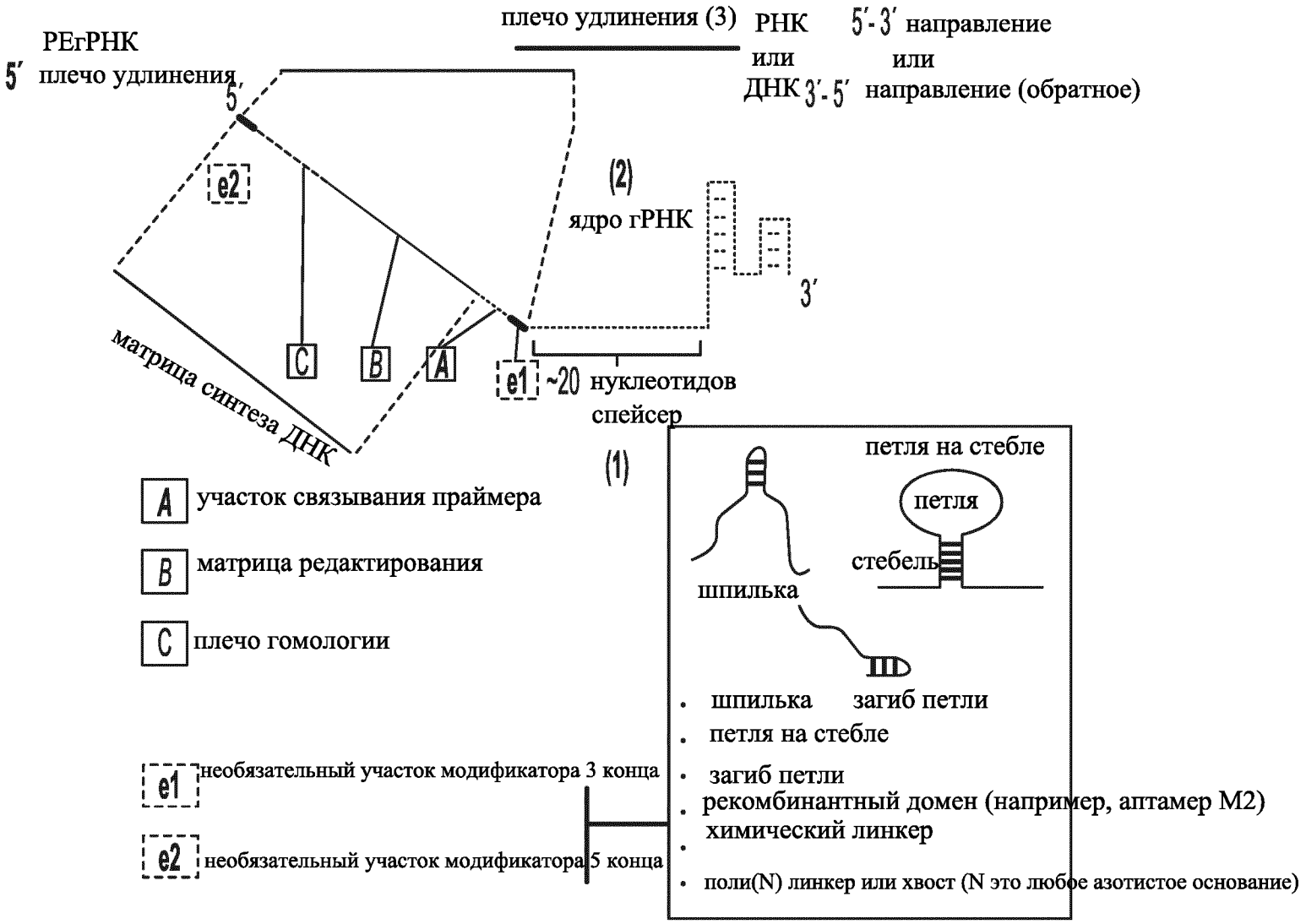
Фиг. 3B


гРНК с внутренним удлинением
матрица
обратной
транскрипции
участок
связывания
праймера
(матрица синтеза ДНК)
ядро гРНК
3′
5′
~20 нуклеотидов
протоспейсер

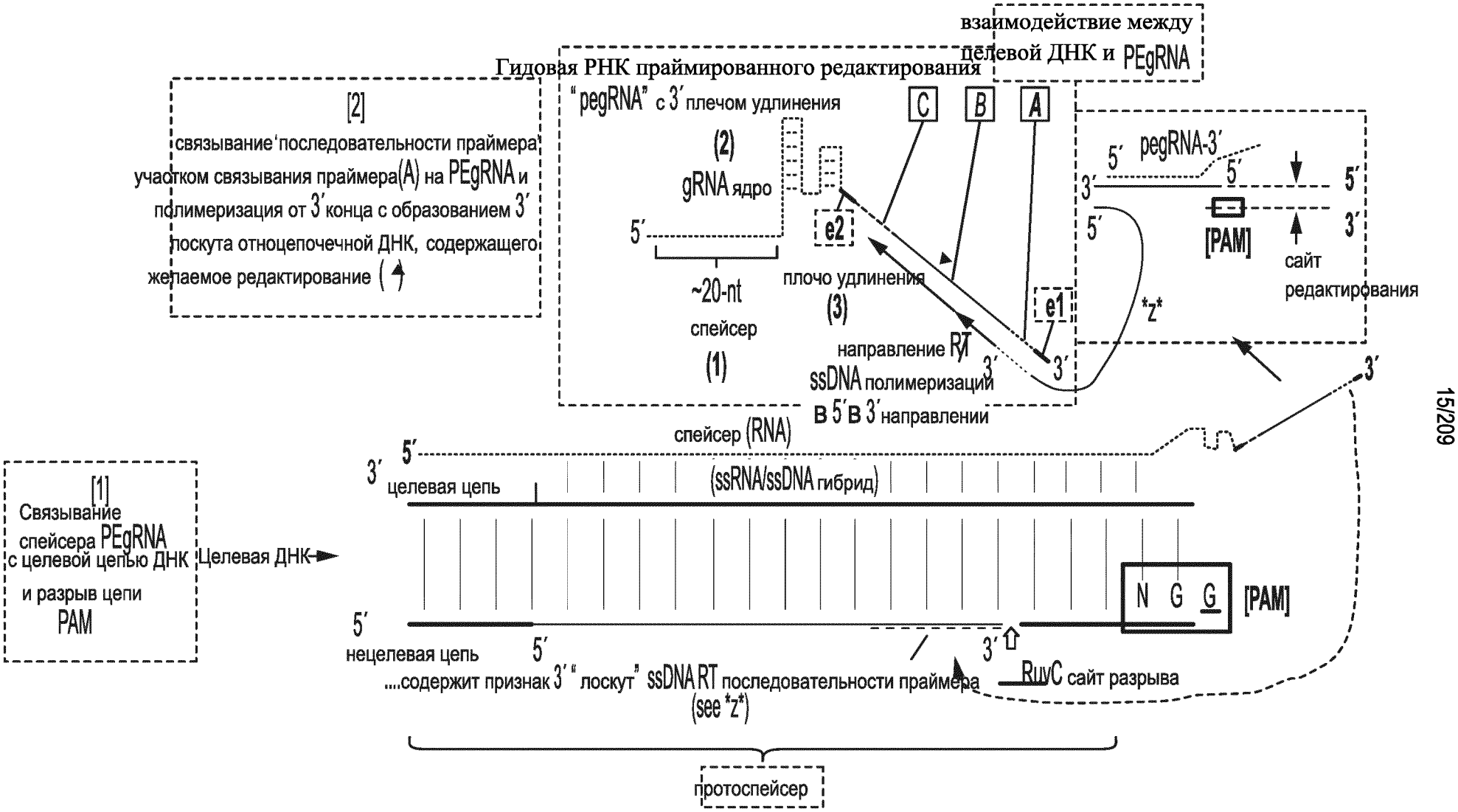
Фиг. 3C


РЕгРНК
3′ плечо удлинения
(2)
ядро гРНК
5′
~20 нуклеотидов
спейсер
(1)
плечо удлинения (3)
РНК или ДНК
5′-3′ направление или
3′-5′ направление (обратное)
C
B
A
e2
e1
матрица синтеза ДНК
3′
A участок связывания праймера
B матрица редактирования
C плечо гомологии
e1 необязательный участок модификатора 3 конца
e2 необязательный участок модификатора 5 конца
петля на стебле
петля
стебель
шпилька
загиб петли
• шпилька
• петля на стебле
• загиб петли
• рекомбинантный домен (например, аптамер M2)
• химический линкер
• поли(N) линкер или хвост (N это любое азотистое основание)

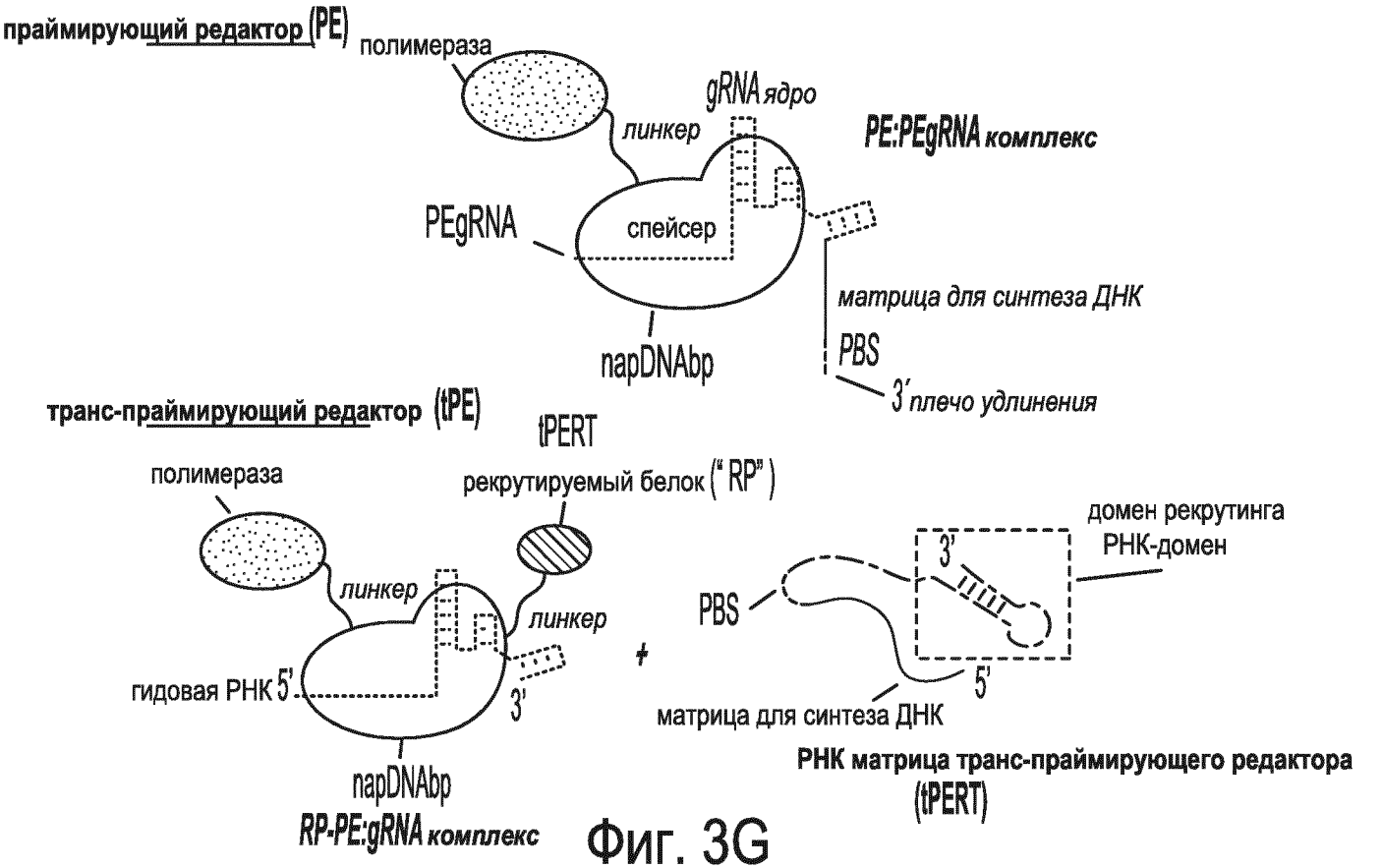
Фиг. 3D


плечо удлинения (3)
РНК или ДНК
5′-3′ направление или
3′-5′ направление (обратное)
РЕгРНК
5′ плечо удлинения
5′
e2
(2)
ядро гРНК
3′
C
B
A
e1
~20 нуклеотидов
спейсер
(1)
матрица синтеза ДНК
A участок связывания праймера
B матрица редактирования
C плечо гомологии
e1 необязательный участок модификатора 3 конца
e2 необязательный участок модификатора 5 конца
петля на стебле
петля
стебель
шпилька
шпилька
загиб петли
петля на стебле
загиб петли
рекомбинантный домен (например, аптамер M2)
химический линкер
поли(N) линкер или хвост (N это любое азотистое основание)

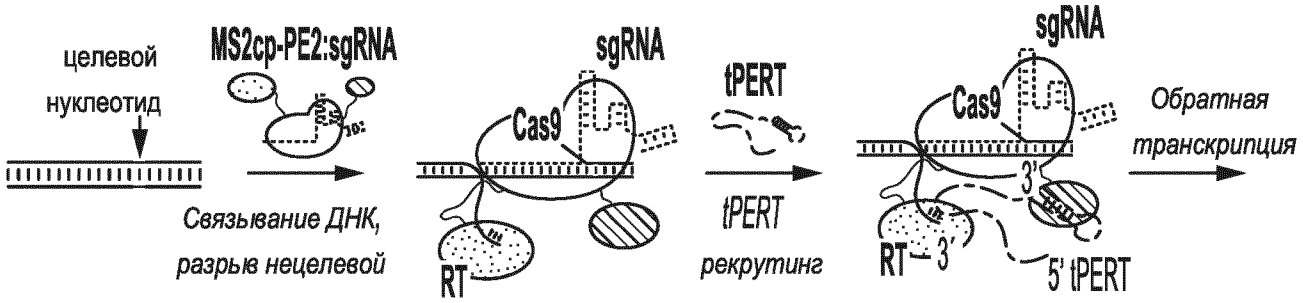
Фиг. 3E


взаимодействие между
целевой ДНК и PEgRNA
Гидовая РНК праймированного редактирования
"pegRNA" с 3′ плечом удлинения
[2]
связывание последовательности праймера участком связывания праймера (A) на PEgRNA и полимеризация от 3′ конца с образованием 3′ лоскута отноцепочечной ДНК, содержащего желаемое редактирование
(2)
gRNA ядро
C
B
A
5′
e2
~20-nt
спейсер
(1)
плочо удлинения
(3)
направление RT
ssDNA полимеризаций
в 5′ в 3′ направлении
e1
3′
5′
pegRNA-3′
5′
3′
5′
3′
[PAM]
сайт
редактирования
z
3′
спейсер (RNA)
[1]
Связывание
спейсера PEgRNA
с целевой цепью ДНК
и разрыв цепи
PAM
Целевая ДНК
3′
5′
целевая цепь
(ssRNA/ssDNA гибрид)
N G G
[PAM]
5′
нецелевая цепь
5′
3′
RuvC сайт разрыва
....содержит признак 3′ "лоскут" ssDNA RT последовательности праймера
(see *z*)
протоспейсер

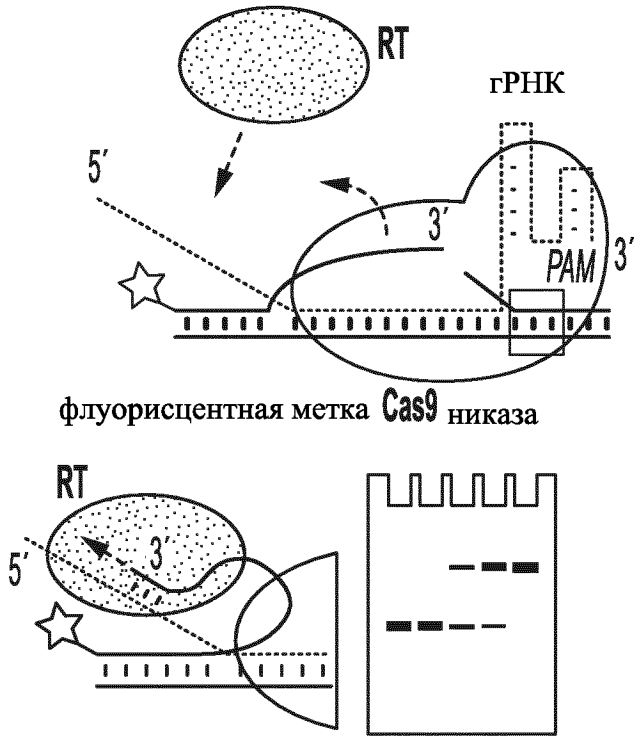
Фиг. 3F


праймирующий редактор (PE)
полимераза
gRNA ядро
линкер
PE:PEgRNA комплекс
PEgRNA
спейсер
матрица для синтеза ДНК
napDNAbp
PBS
3′ плечо удлинения
транс-праймирующий редактор (tPE)
tPERT
полимераза
рекрутируемый белок (" RP")
домен рекрутинга
РНК-домен
линкер
линкер
PBS
3'
гидовая РНК 5'
3'
5'
матрица для синтеза ДНК
napDNAbp
РНК матрица транс-праймирующего редактора
(tPERT)
RP-PE:gRNA комплекс

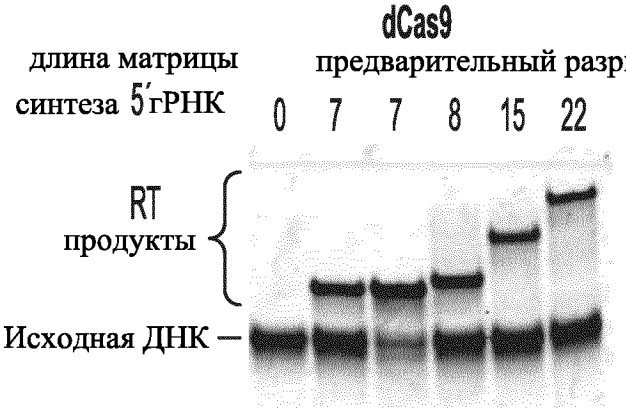
Фиг. 3G


целевой
нуклеотид
MS2cp-PE2:sgRNA
sgRNA
Cas9
tPERT
sgRNA
Cas9
Обратная
транскрипция
Связывание ДНК,
разрыв нецелевой
цепи
RT
tPERT
рекрутинг
RT
3'
5' tPERT

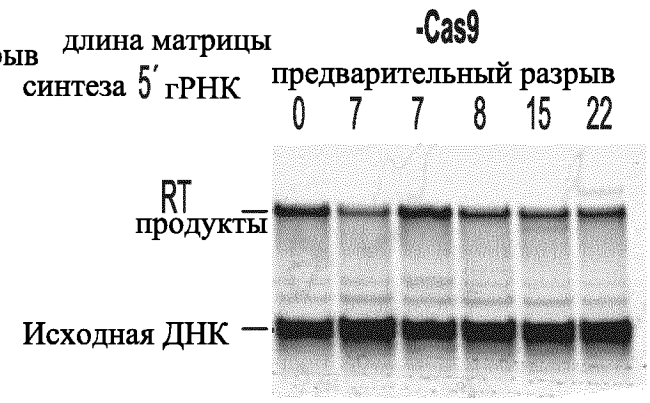
Фиг. 3H



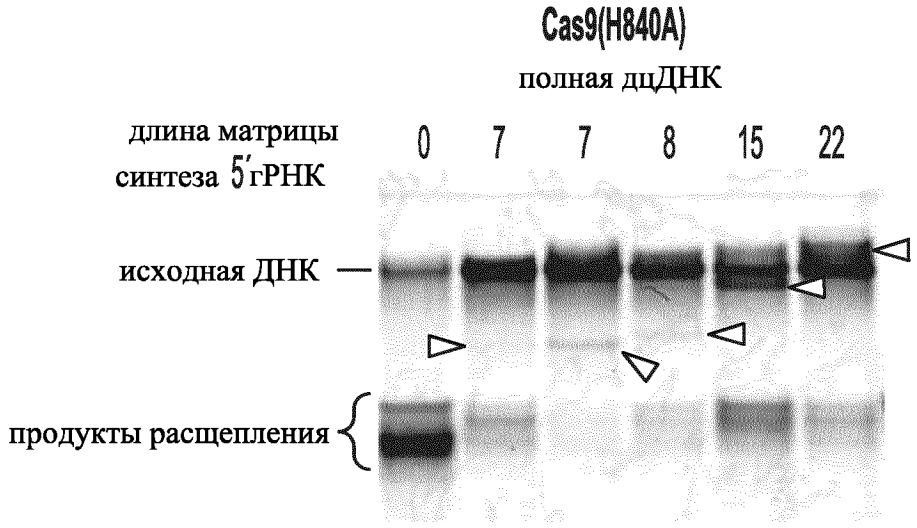
Фиг. 4A



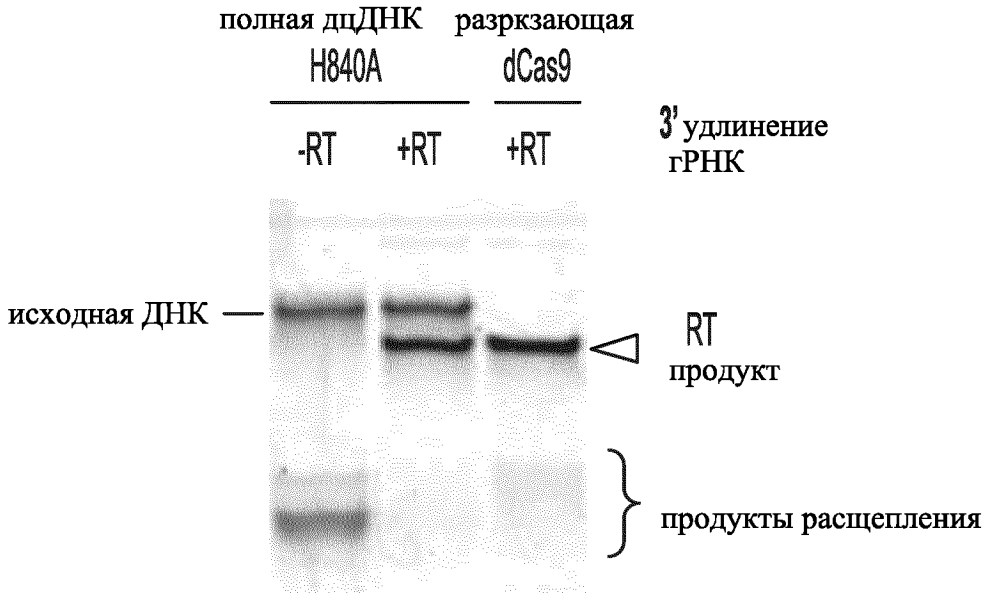
Фиг. 4B



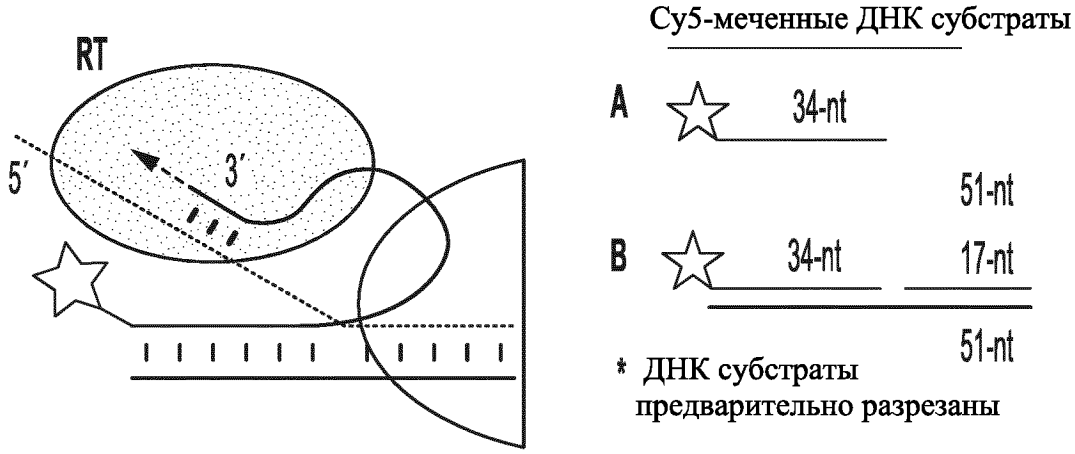
Фиг. 4C

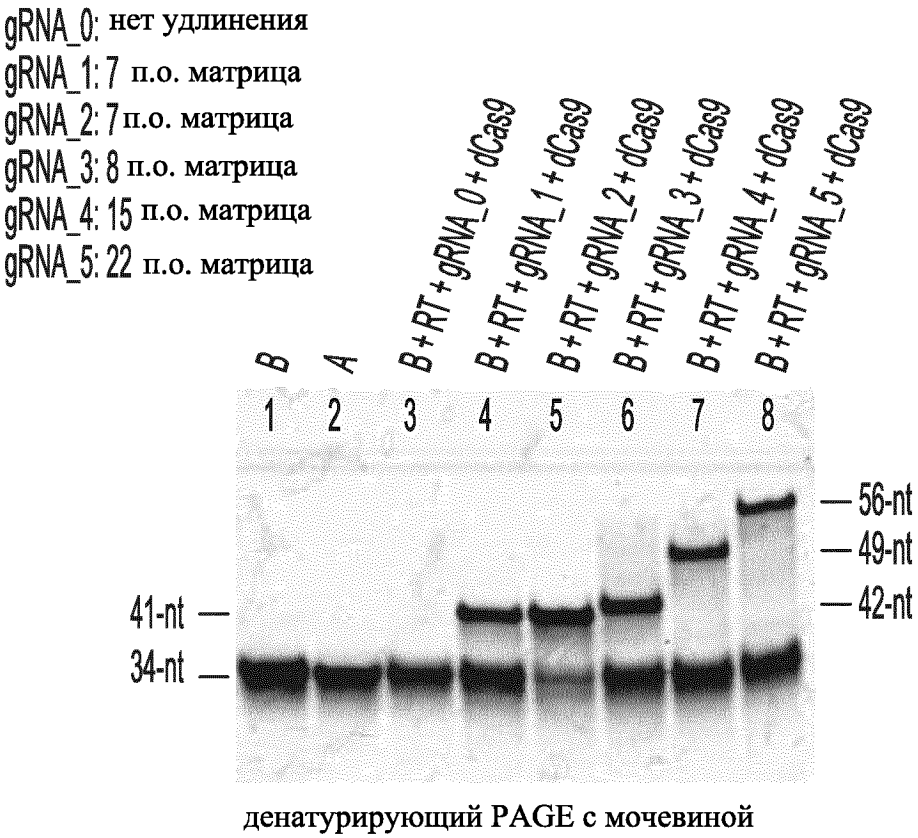

Cas9(H840A)
полная дцДНК
длина матрицы синтеза 5′гРНК
0
7
7
8
15
22
исходная ДНК
продукты расщепления

Фиг. 4D

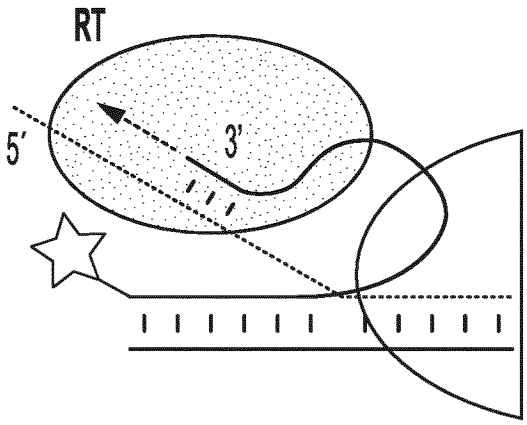

полная дцДНК
разрезающая
H840A
dCas9
-RT
+RT
+RT
3' удлинение гРНК
исходная ДНК
RT продукт
продукты расщепления

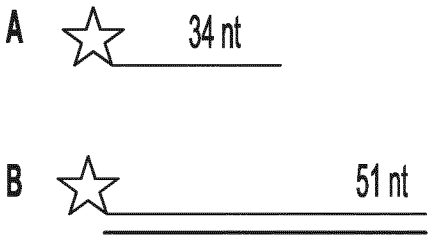
Фиг. 4E


RT
5′
3′
Cy5-меченные ДНК субстраты
A
34-nt
51-nt
B
34-nt
17-nt
51-nt
* ДНК субстраты
предварительно разрезаны


gRNA_0: нет удлинения
gRNA_1: 7 п.о. матрица
gRNA_2: 7 п.о. матрица
gRNA_3: 8 п.о. матрица
gRNA_4: 15 п.о. матрица
gRNA_5: 22 п.о. 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 с мочевиной

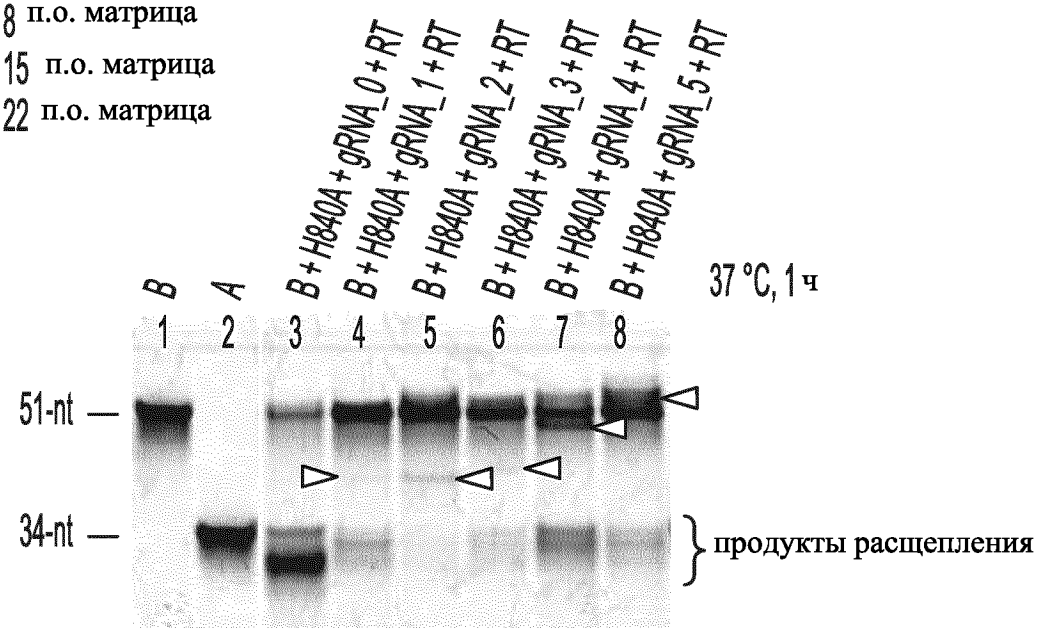
Фиг.5





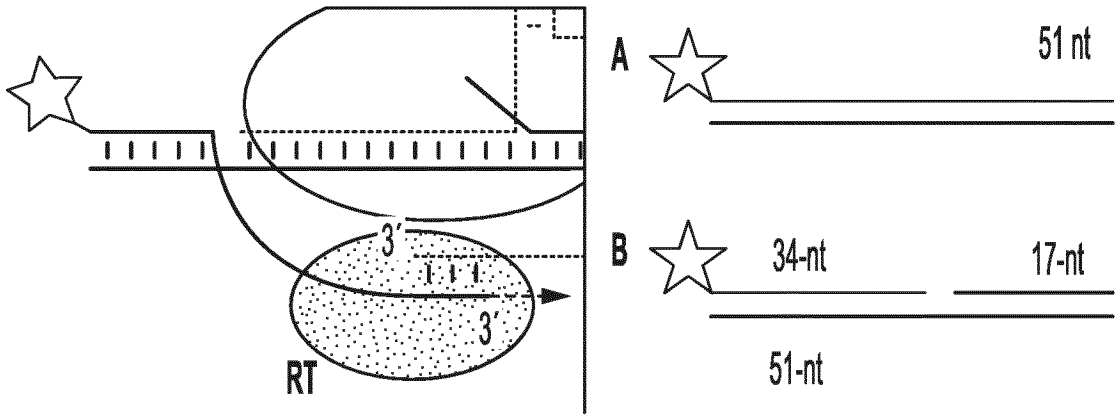
* ДНК субстраты
предварительно не разрезаны

gRNA_0: нет удлинения
gRNA_1: 7 п.о. матрица
gRNA_2: 7 п.о. матрица
gRNA_3: 8 п.о. матрица
gRNA_4: 15 п.о. матрица
gRNA_5: 22 п.о. матрица

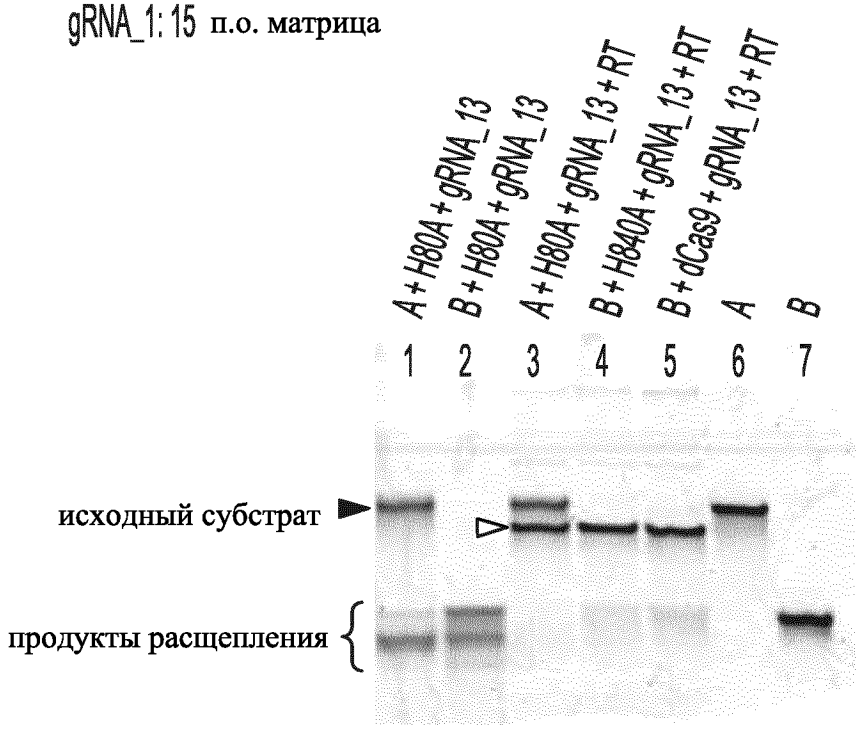


денатурирующий PAGE с 15 % 4мочевиной

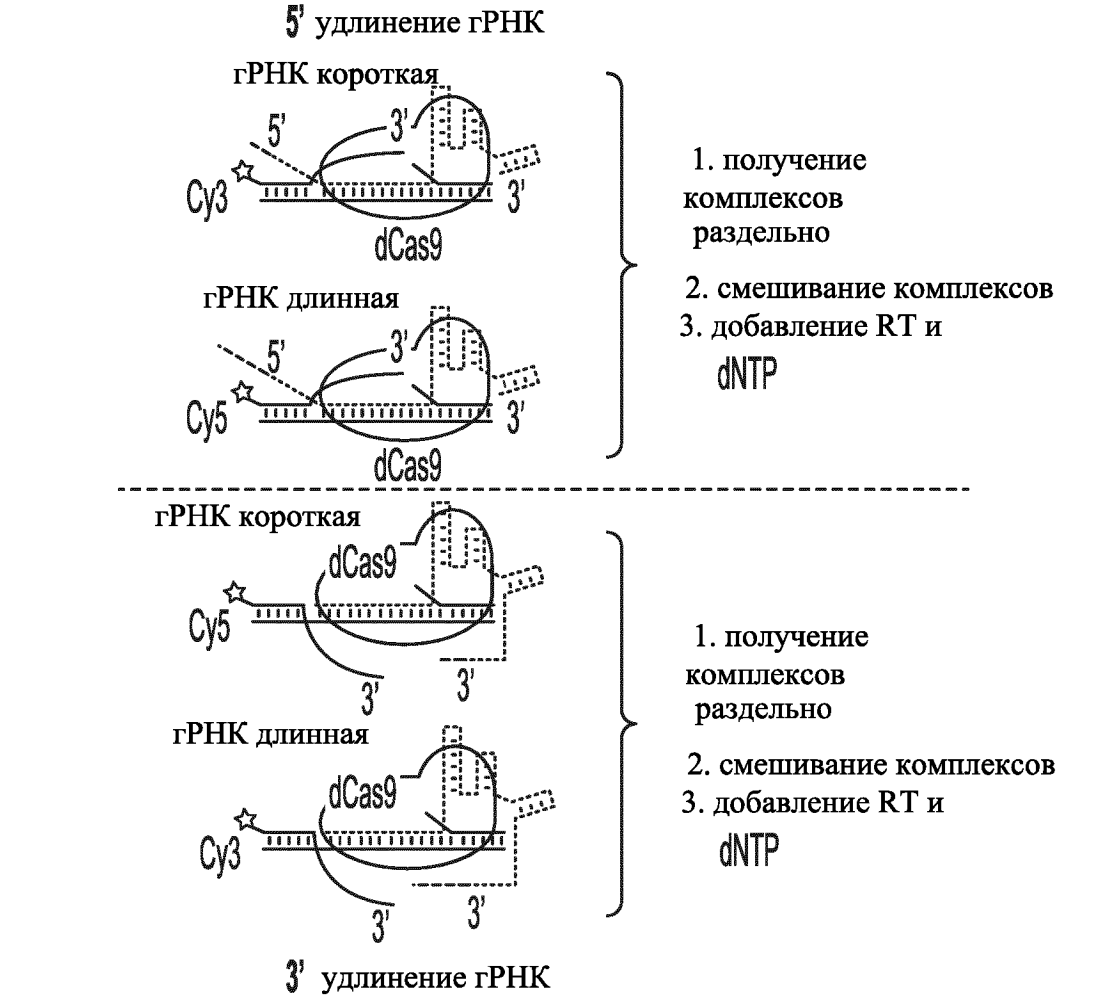
Фиг.6

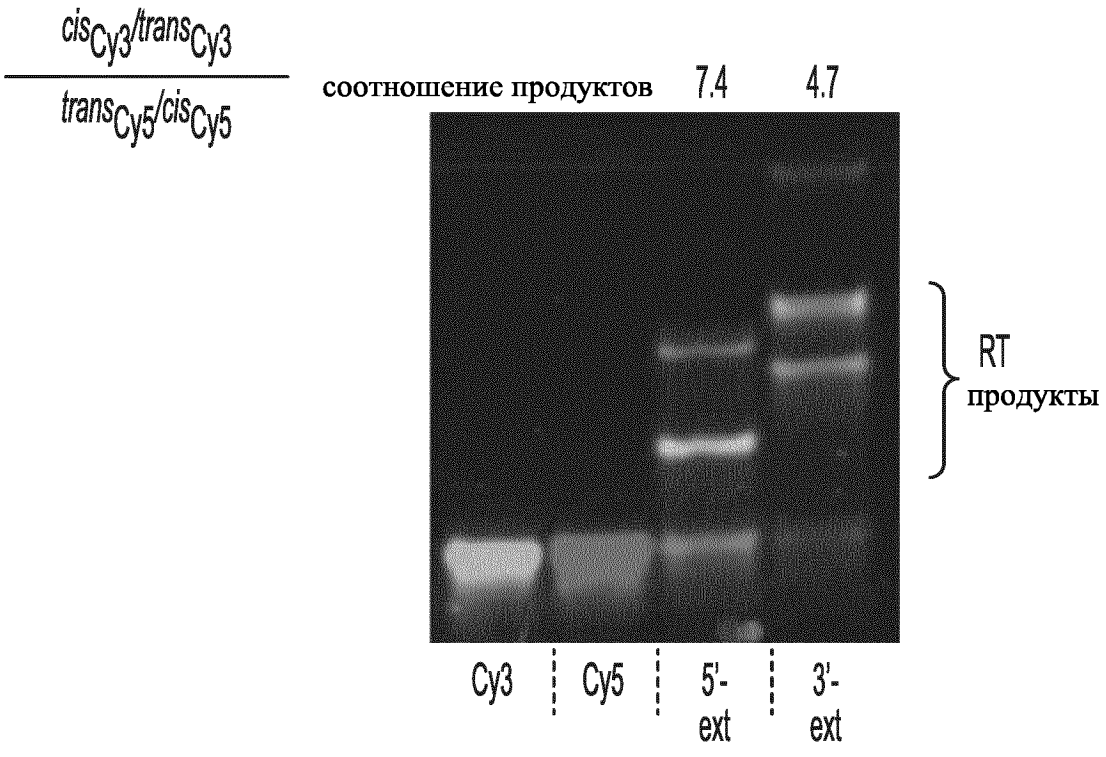


gRNA_0: нет удлинения
gRNA_1: 15 п.о. матрица

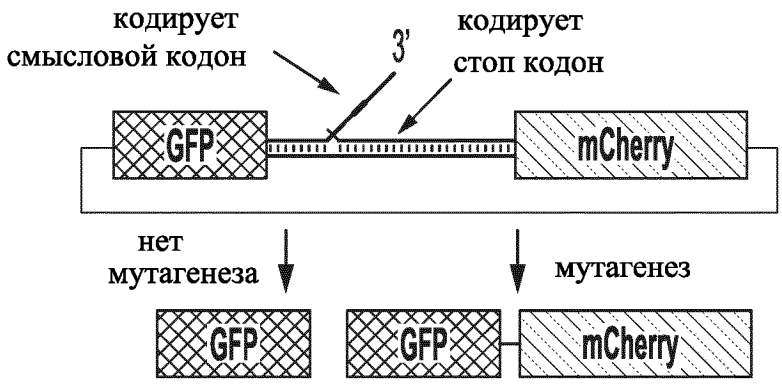


Фиг.7

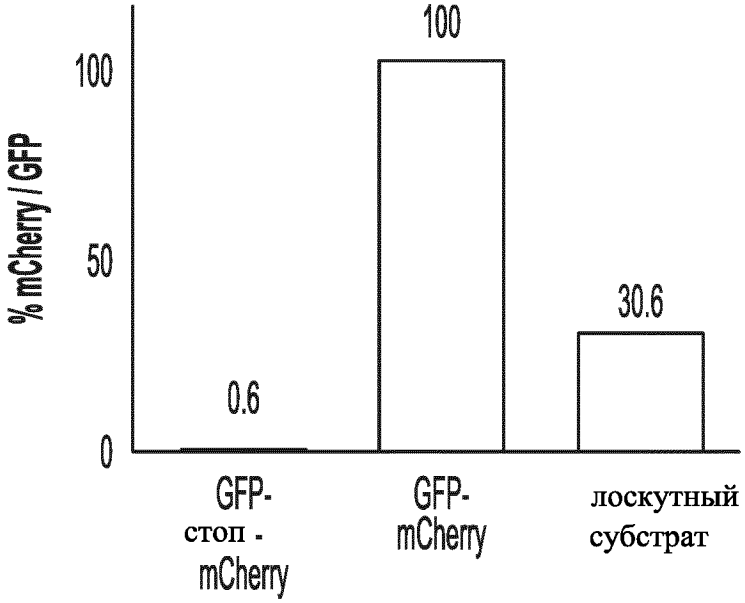

5' удлинение гРНК
гРНК короткая
5'
3'
Cy3
3'
dCas9
гРНК длинная
5'
3'
Cy5
3'
dCas9
1. получение комплексов раздельно
2. смешивание комплексов
3. добавление RT и dNTP
гРНК короткая
dCas9
Cy5
3'
3'
гРНК длинная
dCas9
Cy3
3'
3'
1. получение комплексов раздельно
2. смешивание комплексов
3. добавление RT и dNTP
3' удлинение гРНК


cisCy3/transCy3
transCy5/cisCy5
соотношение продуктов
7.4
4.7
RT продукты
Cy3
Cy5
5'-ext
3'-ext

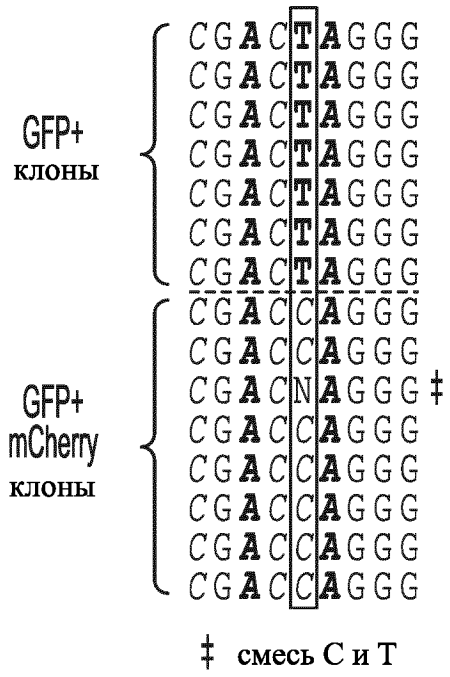
Фиг.8


кодирует
смысловой кодон
3'
кодирует
стоп кодон
GFP
mCherry
нет
мутагенеза
мутагенез
GFP
GFP
mCherry

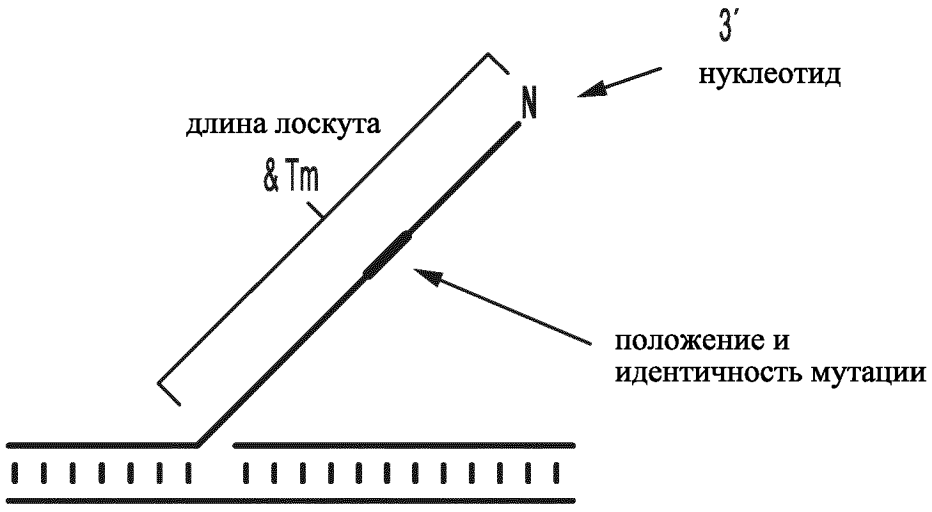
Фиг. 9A


% mCherry / GFP
100
50
0
0.6
100
30.6
GFP-
стоп -
mCherry
GFP-
mCherry
лоскутный
субстрат

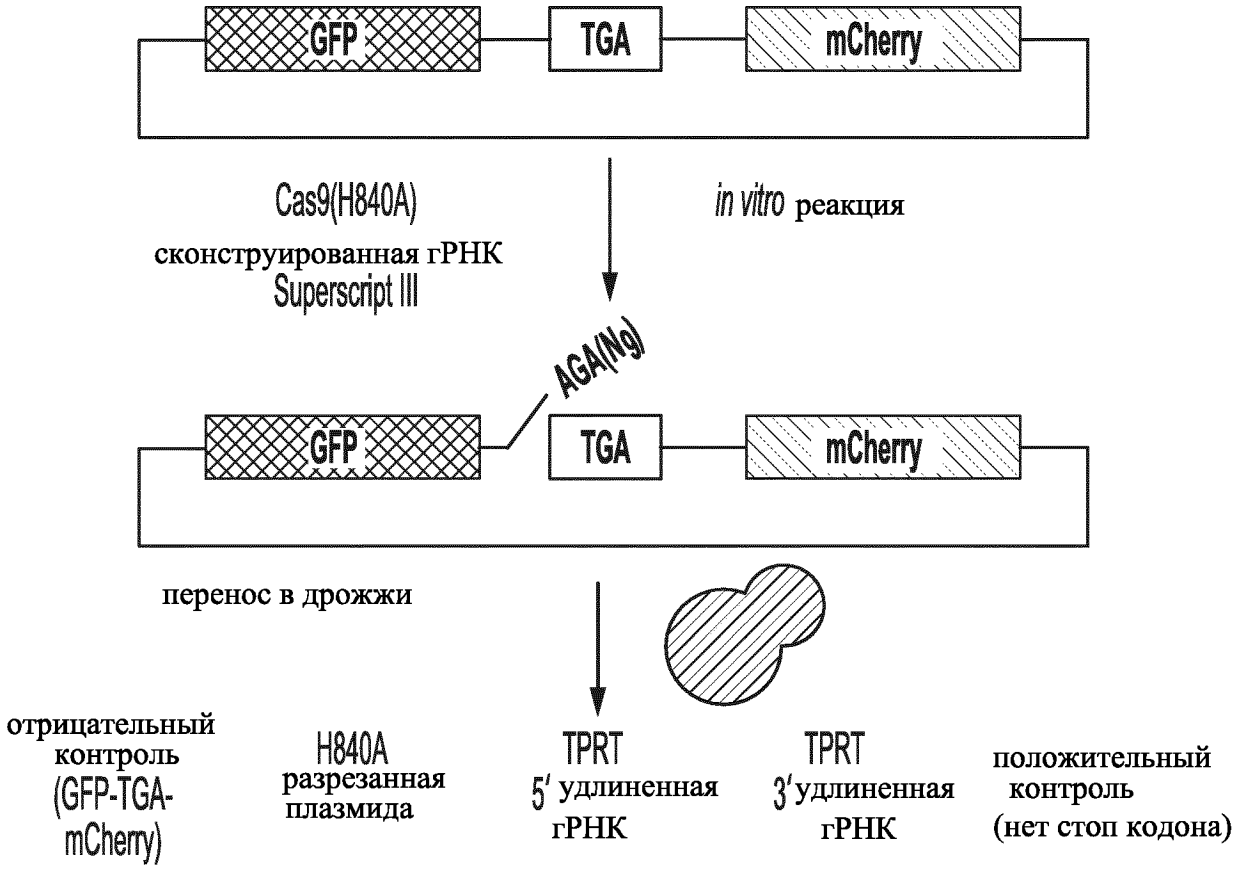
Фиг. 9B


GFP+
клоны
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
GFP+
mCherry
клоны
CGACCAGGG
CGACCAGGG
CGACNAGGG ‡
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
‡ смесь C и T

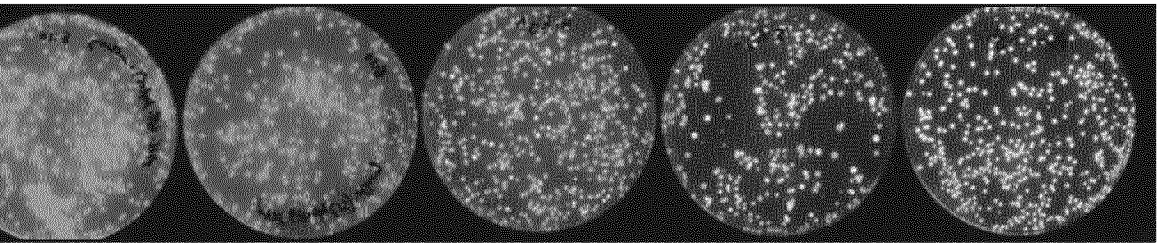
Фиг. 9C


3′
нуклеотид
N
длина лоскута
& Tm
положение и
идентичность мутации

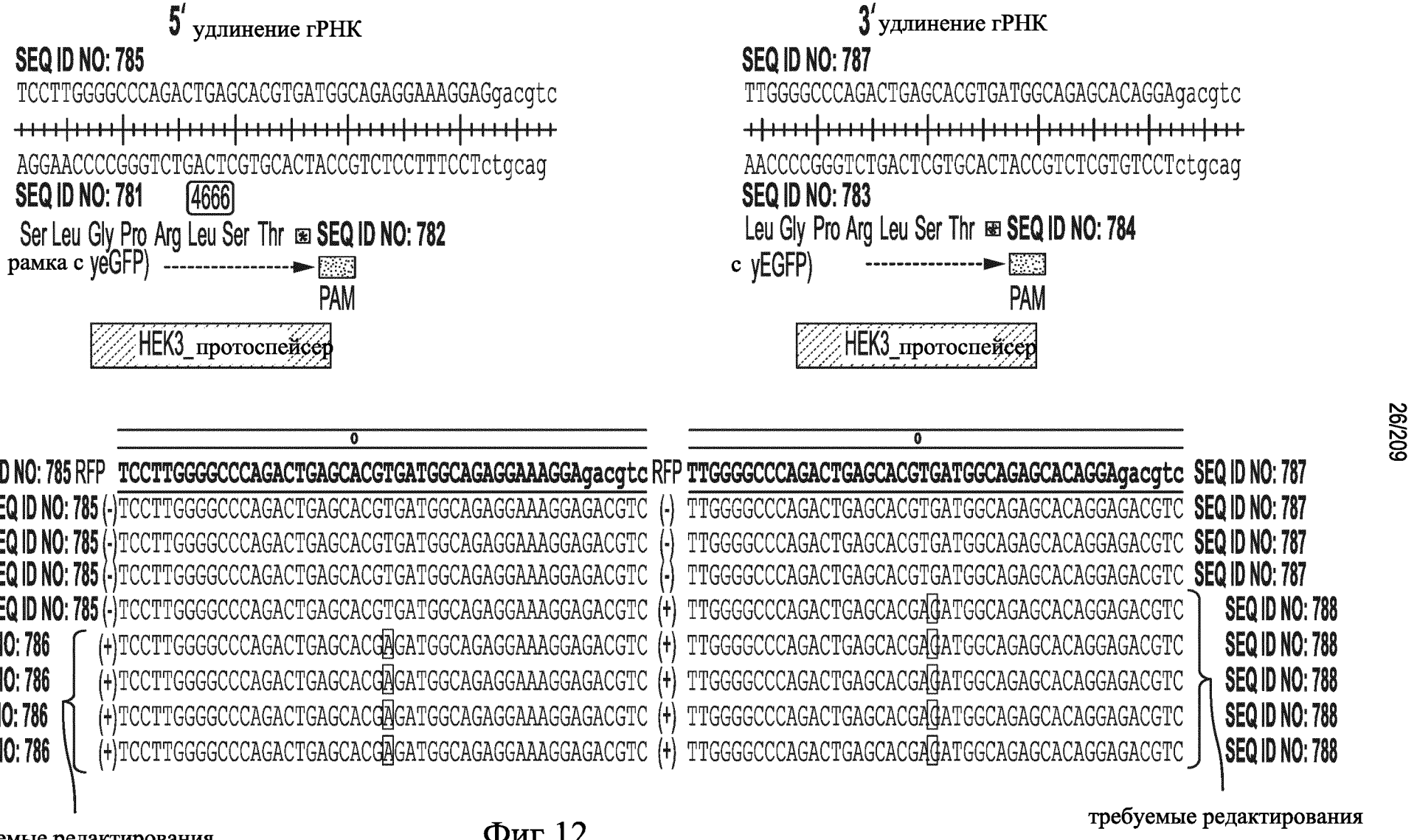
Фиг. 9D


GFP
TGA
mCherry
Cas9(H840A)
сконструированная гРНК
Superscript III
in vitro реакция
AGA(Ng)
GFP
TGA
mCherry
перенос в дрожжи
отрицательный контроль (GFP-TGA-mCherry)
H840A разрезанная плазмида
TPRT 5′ удлиненная гРНК
TPRT 3′удлиненная гРНК
положительный контроль (нет стоп кодона)

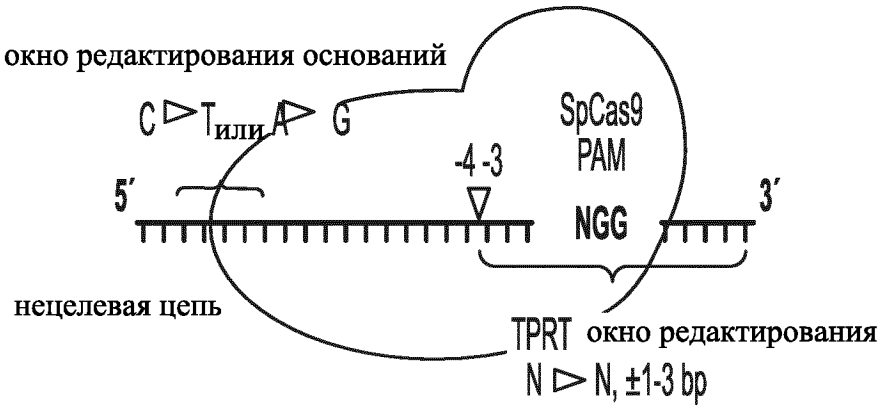
Фиг.10


Коррекция вставки (TPRT делеция 1 нуклеотида)
Коррекция делеции (TPRT вставка 1 нуклеотида)

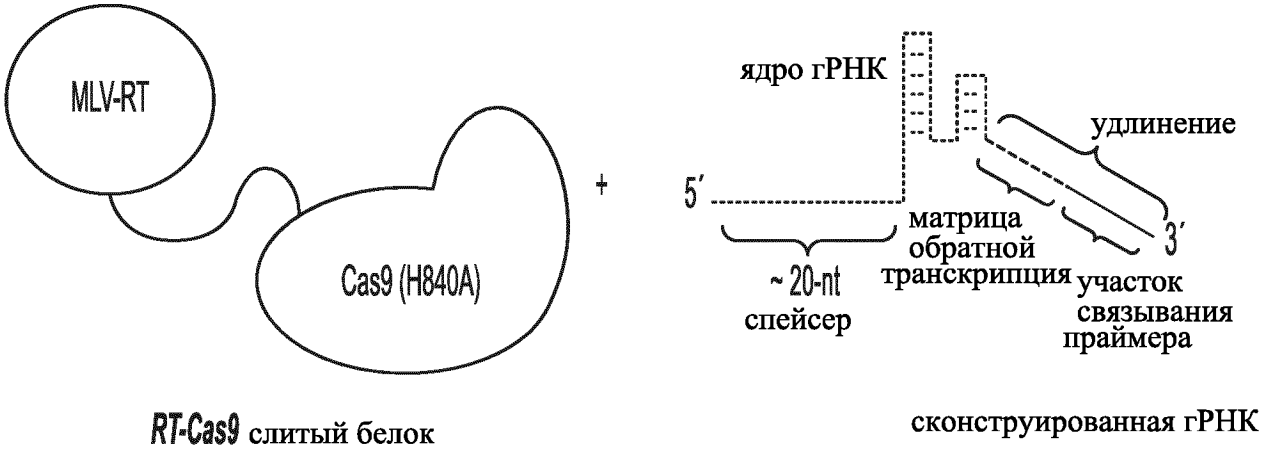
Фиг.11


5′ удлинение гРНК
SEQ ID NO: 785
TCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAGgacgtc
AGGAACCCCGGGTCTGACTCGTGCACTACCGTCTCCTTTCCTctgcag
SEQ ID NO: 781
4666
Ser Leu Gly Pro Arg Leu Ser Thr
SEQ ID NO: 782
рамка с yeGFP)
PAM
HEK3_протоспейсер
3′ удлинение гРНК
SEQ ID NO: 787
TTGGGGCCCAGACTGAGCACGTGATGGCAGAGCACAGGAgacgtc
AACCCCGGGTCTGACTCGTGCACTACCGTCTCGTGTCCTctgcag
SEQ ID NO: 783
Leu Gly Pro Arg Leu Ser Thr
SEQ ID NO: 784
с yEGFP)
PAM
HEK3_протоспейсер
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 редактирования
требуемые редактирования

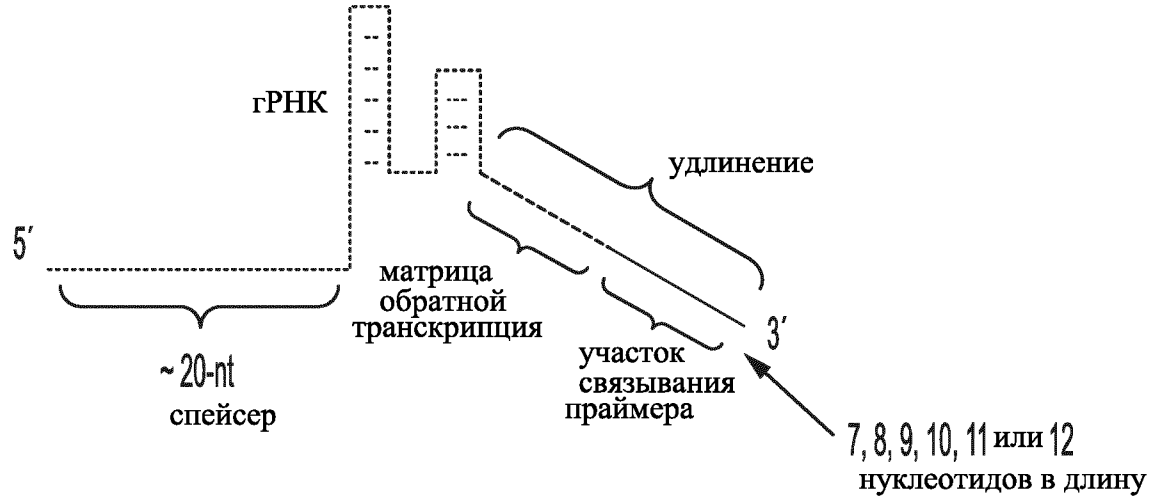
Фиг.12


окно редактирования оснований
C ▷ T или A ▷ G
SpCas9
PAM
-4 -3
5′
3′
NGG
нецелевая цепь
TPRT
окно редактирования
N ▷ N, ±1-3 bp

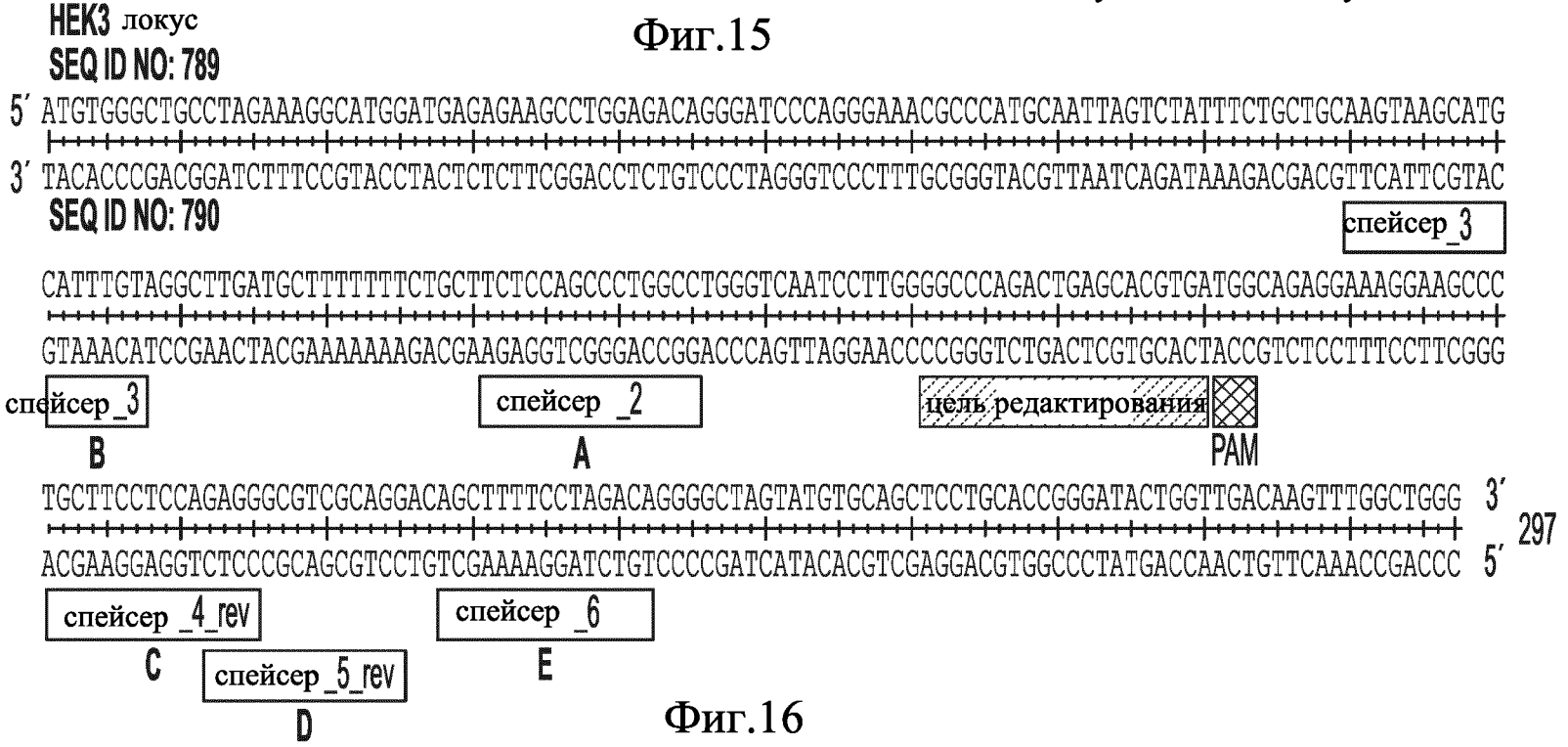
Фиг.13


MLV-RT
Cas9 (H840A)
+
ядро гРНК
удлинение
5′
3′
матрица
обратной
транскрипция
~ 20-nt
спейсер
участок
связывания
праймера
RT-Cas9 слитый белок
сконструированная гРНК

Фиг.14


гРНК
удлинение
5′
матрица
обратной
транскрипция
3′
участок
связывания
праймера
~ 20-nt
спейсер
7, 8, 9, 10, 11 или 12
нуклеотидов в длину

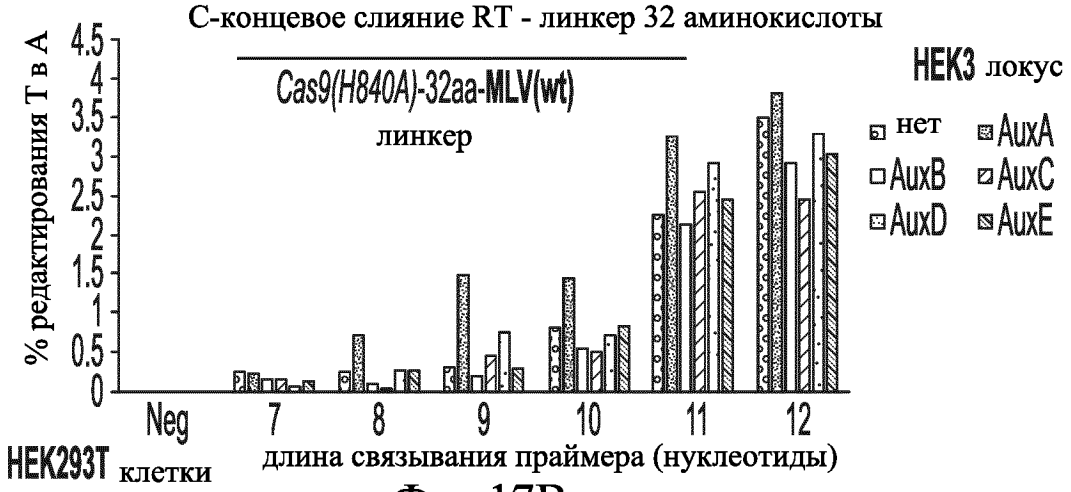
Фиг.15


HEK3 локус
SEQ ID NO: 789
5′ ATGTGGGCTGCCTAGAAAGGCATGGATGAGAGAAGCCTGGAGACAGGGATCCCAGGGAAACGCCCATGCAATTAGTCTATTTCTGCTGCAAGTAAGCATG
3′ TACACCCGACGGATCTTTCCGTACCTACTCTCTTCGGACCTCTGTCCCTAGGGTCCCTTTGCGGGTACGTTAATCAGATAAAGACGACGTTCATTCGTAC
SEQ ID NO: 790
спейсер_3
CATTTGTAGGCTTGATGCTTTTTTTCTGCTTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAGCCC
GTAAACATCCGAACTACGAAAAAAAGACGAAGAGGTCGGGACCGGACCCAGTTAGGAACCCCGGGTCTGACTCGTGCACTACCGTCTCCTTTCCTTCGGG
спейсер_3
B
спейсер _2
A
цель редактирования
PAM
TGCTTCCTCCAGAGGGCGTCGCAGGACAGCTTTTCCTAGACAGGGGCTAGTATGTGCAGCTCCTGCACCGGGATACTGGTTGACAAGTTTGGCTGGG 3′
297
ACGAAGGAGGTCTCCCGCAGCGTCCTGTCGAAAAGGATCTGTCCCCGATCATACACGTCGAGGACGTGGCCCTATGACCAACTGTTCAAACCGACCC 5′
спейсер _4_rev
C
спейсер _5_rev
D
спейсер _6
E

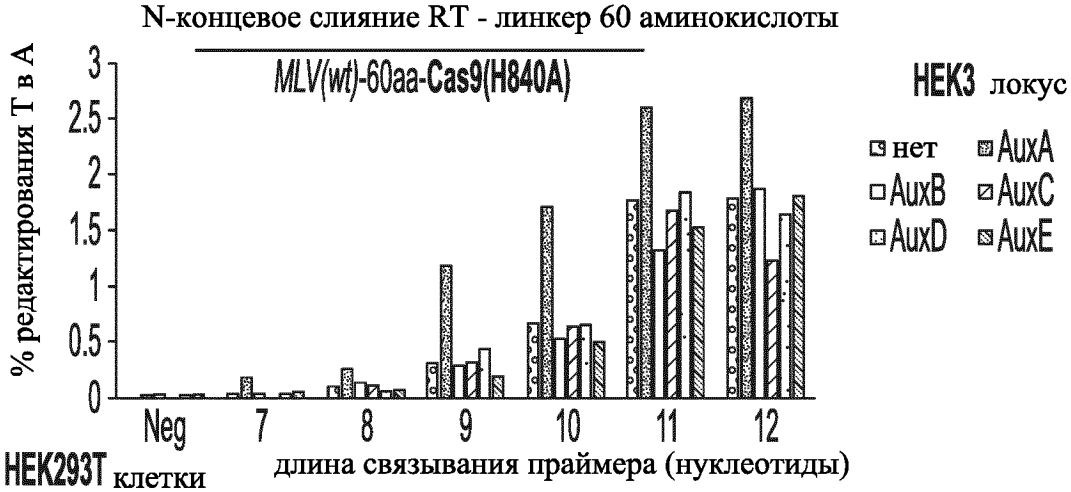
Фиг.16



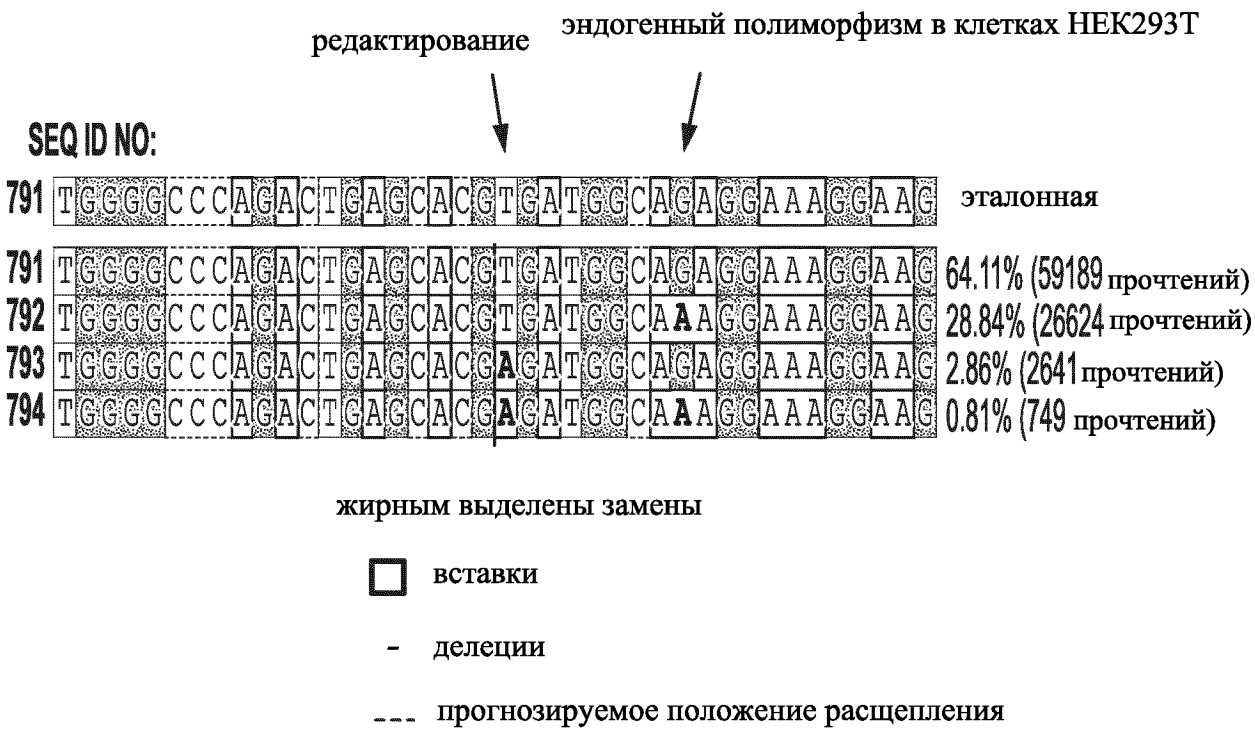
Фиг.17A



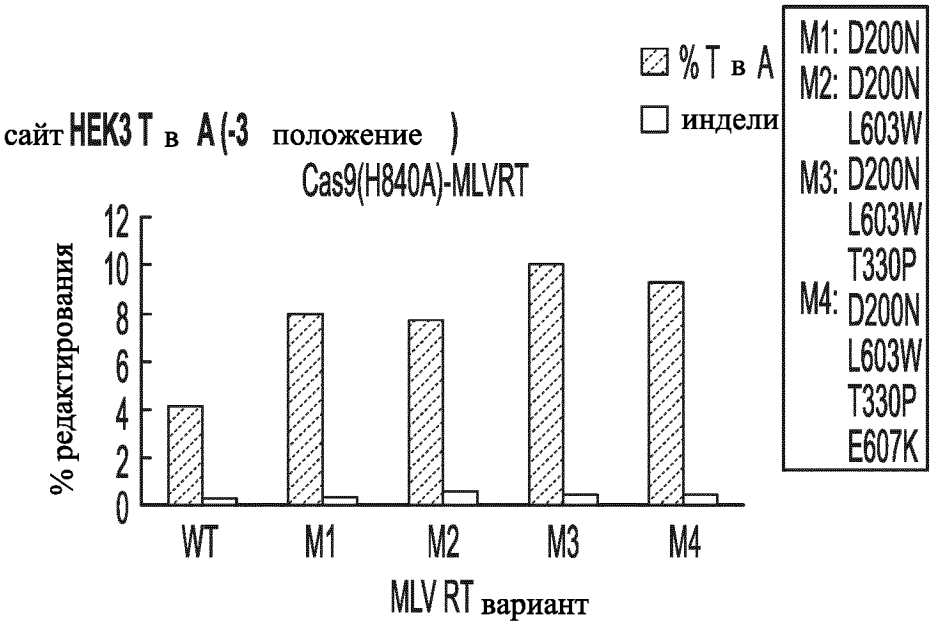
Фиг.17B



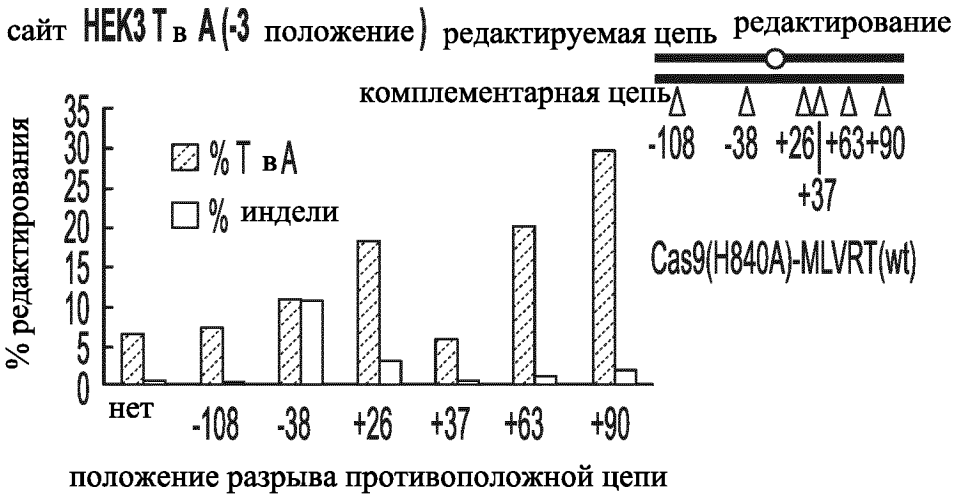
Фиг.17C


редактирование
эндогенный полиморфизм в клетках HEK293T
SEQ ID NO:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG эталонная
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 64.11% (59189 прочтений)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 28.84% (26624 прочтений)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 2.86% (2641 прочтений)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 0.81% (749 прочтений)
жирным выделены замены
вставки
- делеции
--- прогнозируемое положение расщепления

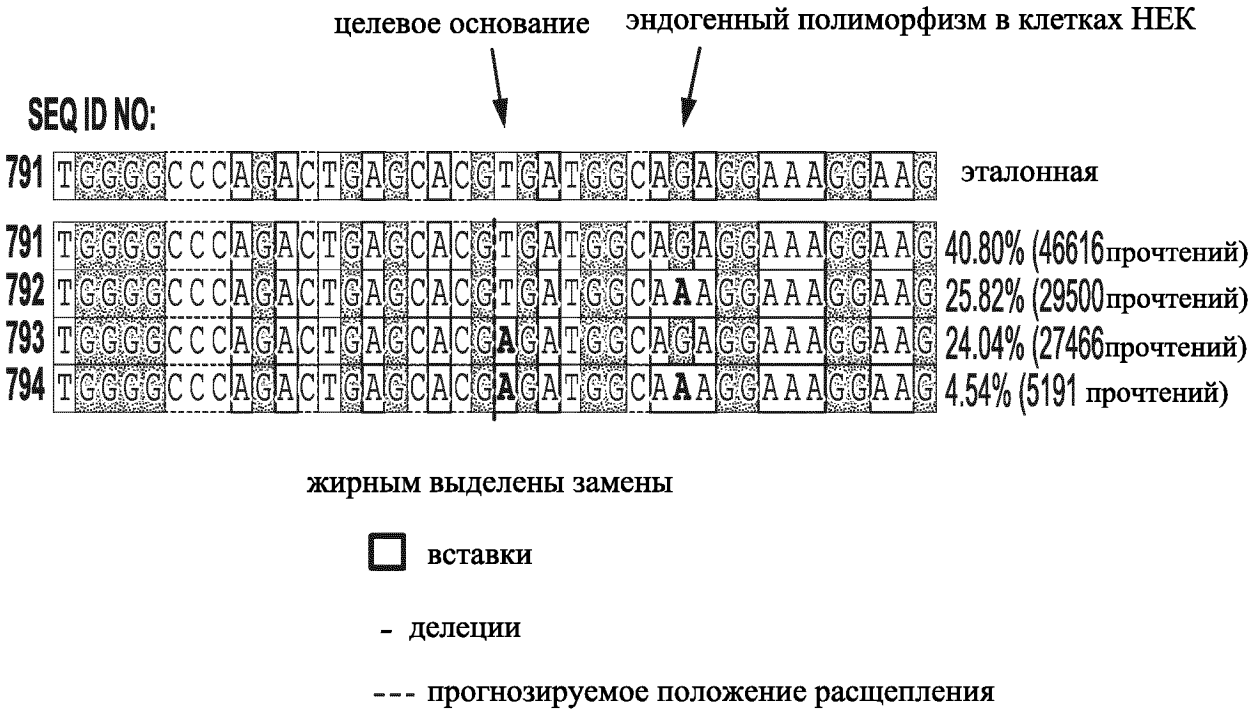
Фиг.18


сайт HEK3 T в A (-3 положение)
Cas9(H840A)-MLVRT
% редактирования
12
10
8
6
4
2
0
WT
M1
M2
M3
M4
MLV RT вариант
% T в A
индели
M1: D200N
M2: D200N
L603W
M3: D200N
L603W
T330P
M4: D200N
L603W
T330P
E607K

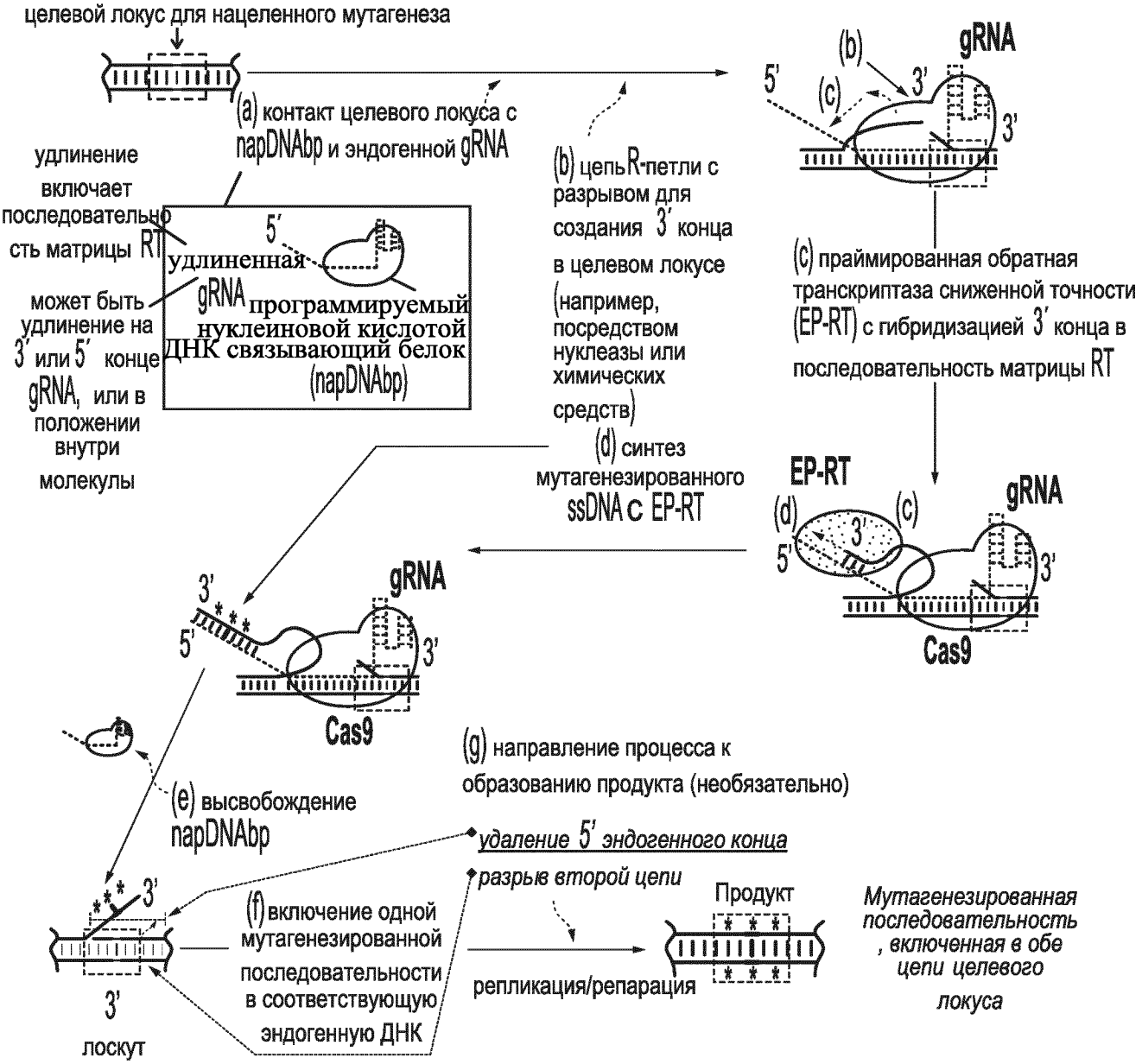
Фиг.19


сайт HEK3 T в A (-3 положение) редактируемая цепь
редактирование
комплементарная цепь
-108
-38
+26
+37
+63
+90
Cas9(H840A)-MLVRT(wt)
% редактирования
35
30
25
20
15
10
5
0
% T в A
% индели
нет
-108
-38
+26
+37
+63
+90
положение разрыва противоположной цепи

Фиг.20

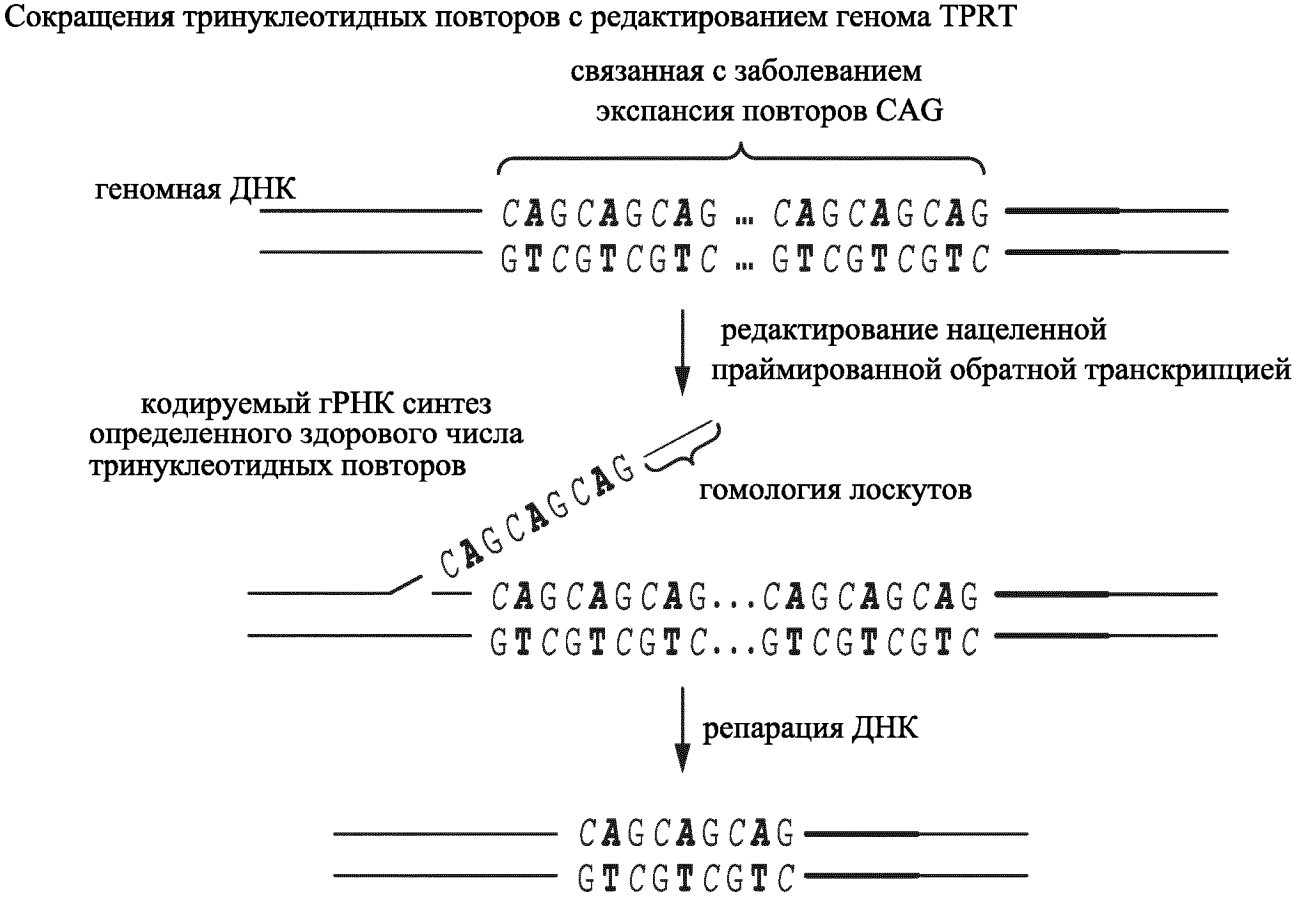

целевое основание
эндогенный полиморфизм в клетках HEK
SEQ ID NO:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG эталонная
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 40.80% (46616прочтений)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 25.82% (29500прочтений)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 24.04% (27466прочтений)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 4.54% (5191 прочтений)
жирным выделены замены
вставки
- делеции
--- прогнозируемое положение расщепления

Фиг.21


целевой локус для нацеленного мутагенеза
(a) контакт целевого локуса с napDNAbp и эндогенной gRNA
удлинение включает последовательность матрицы RT
может быть удлинение на 3′ или 5′ конце gRNA, или в положении внутри молекулы
5′
удлиненная gRNA
программируемый нуклеиновой кислотой ДНК связывающий белок (napDNAbp)
(b) цепь R-петли с разрывом для создания 3′ конца в целевом локусе (например, посредством нуклеазы или химических средств)
(b)
gRNA
5′
(c)
3′
3′
(c) праймированная обратная транскриптаза сниженной точности (EP-RT) с гибридизацией 3′ конца в последовательность матрицы RT
EP-RT
(d)
(c)
gRNA
5′
3′
3′
Cas9
(d) синтез мутагенезированного ssDNA с EP-RT
3′
5′
gRNA
3′
Cas9
(e) высвобождение napDNAbp
(g) направление процесса к образованию продукта (необязательно)
удаление 5′ эндогенного конца
разрыв второй цепи
Продукт
3′
3′
лоскут
(f) включение одной мутагенезированной последовательности в соответствующую эндогенную ДНК
репликация/репарация
Мутагенезированная последовательность, включенная в обе цепи целевого локуса

Фиг. 22


Конструирование гРНК для сокращения последовательностей тринуклеотидных повторов
5′
протоспейсер
каркас гРНК
гомология
повтор
отжиг
3′
нацеливание на ген
кодирует желаемую гомологию
и длину повторов
связыват праймер с
инициацией обратной транскрипции

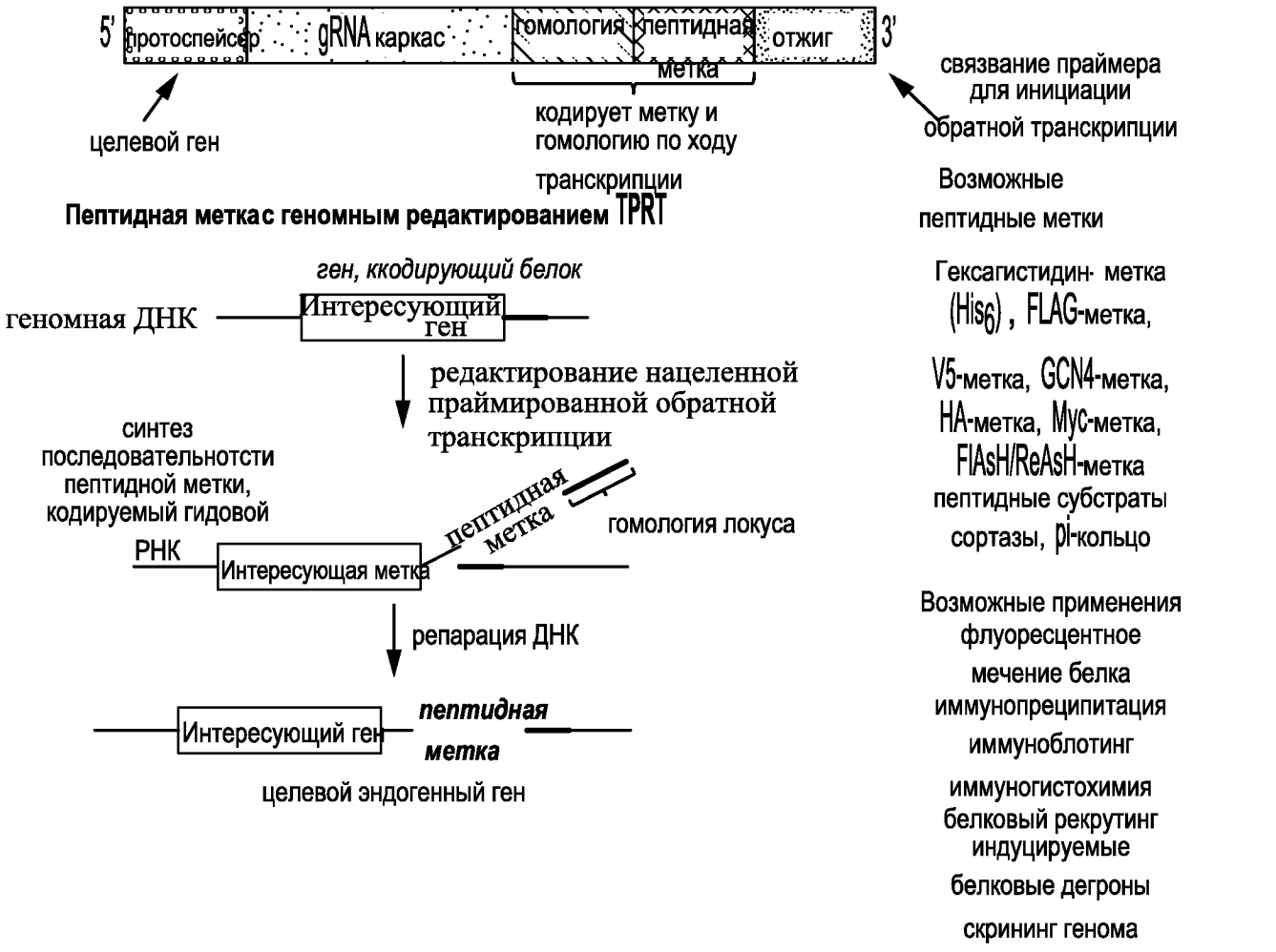

Сокращения тринуклеотидных повторов с редактированием генома TPRT
связанная с заболеванием
экспансия повторов CAG
геномная ДНК
CAGCAGCAG … CAGCAGCAG
GTCGTCGTC … GTCGTCGTC
редактирование нацеленной
праймированной обратной транскрипцией
кодируемый гРНК синтез
определенного здорового числа
тринуклеотидных повторов
CAGCAGCAG
гомология лоскутов
CAGCAGCAG...CAGCAGCAG
GTCGTCGTC...GTCGTCGTC
репарация ДНК
CAGCAGCAG
GTCGTCGTC

Фиг.23

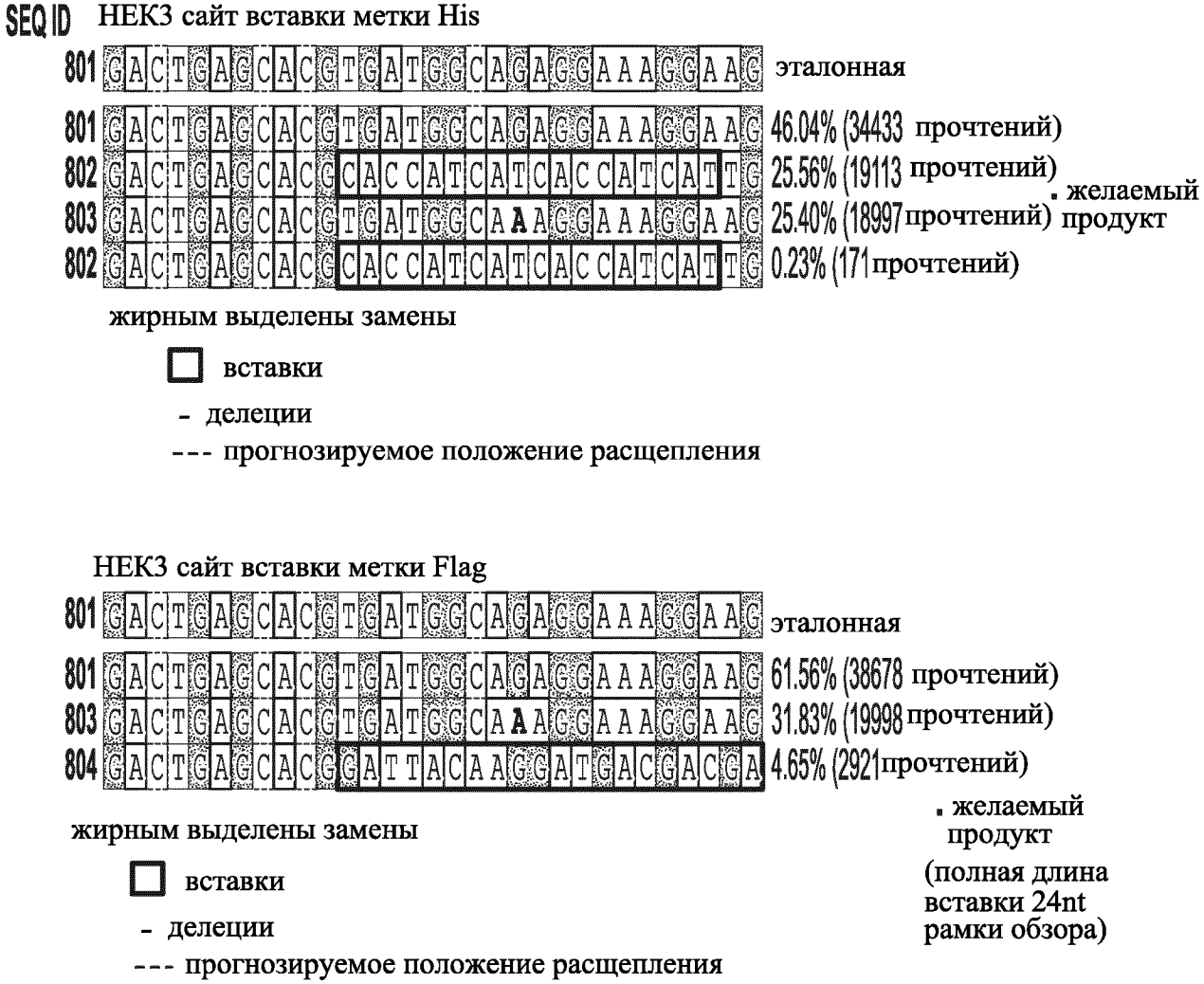


Фиг. 24

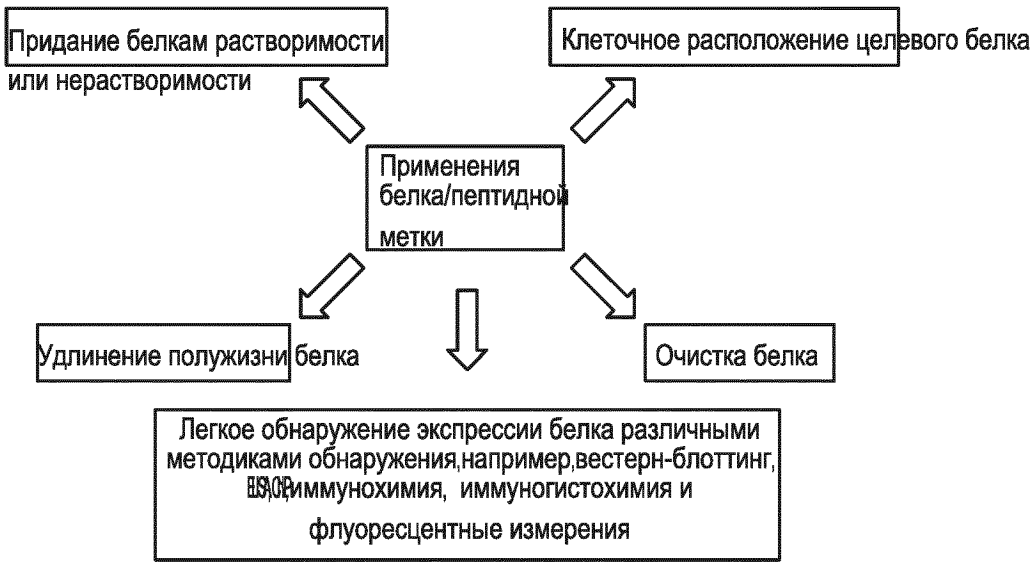
Конструирование гРНК для пептидного нацеливания на гены в эндогенном геномном локусе



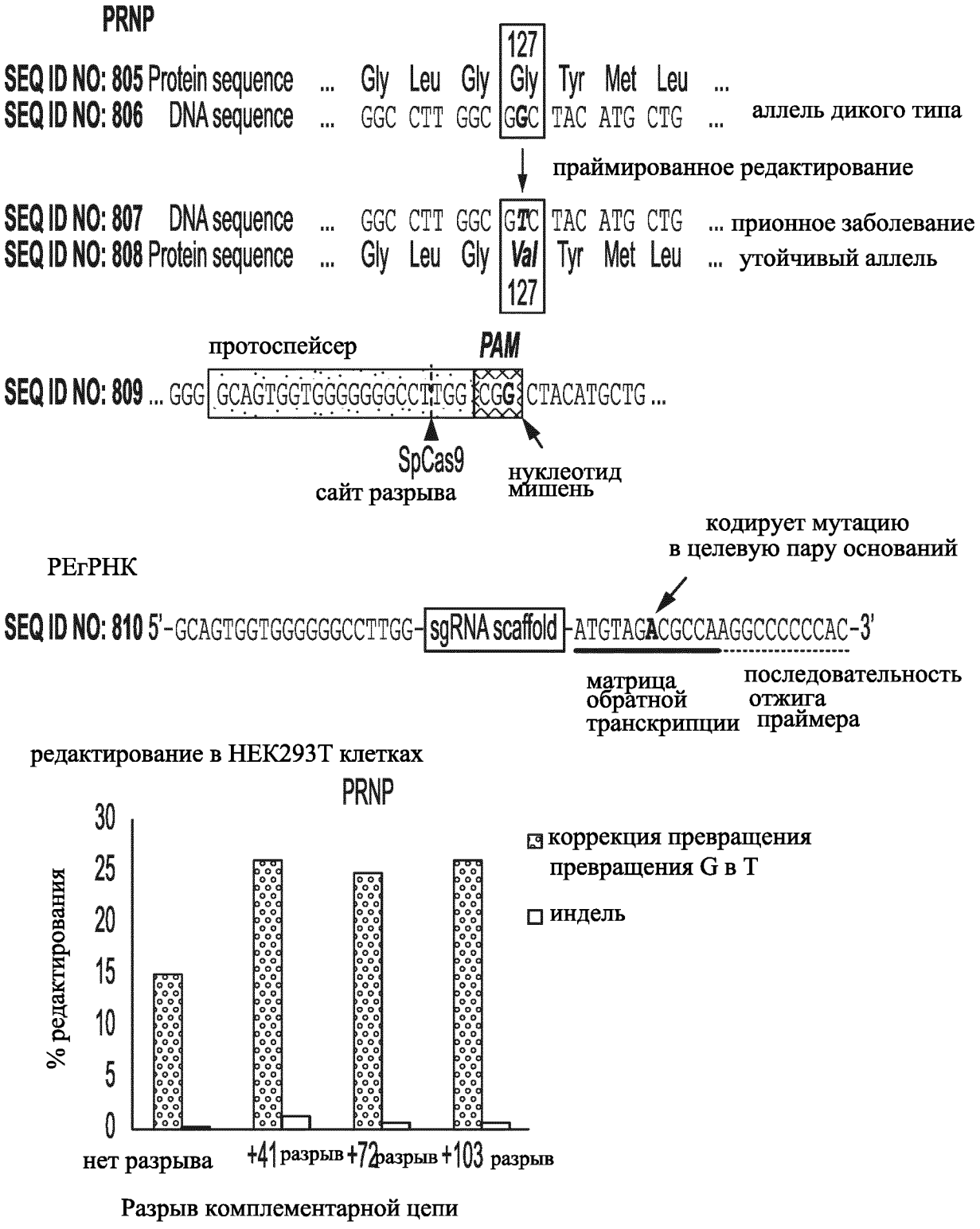
Фиг. 25



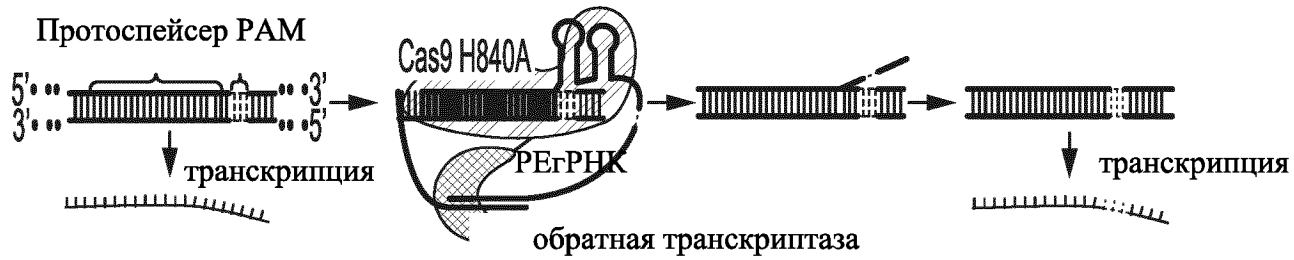
Фиг. 26A



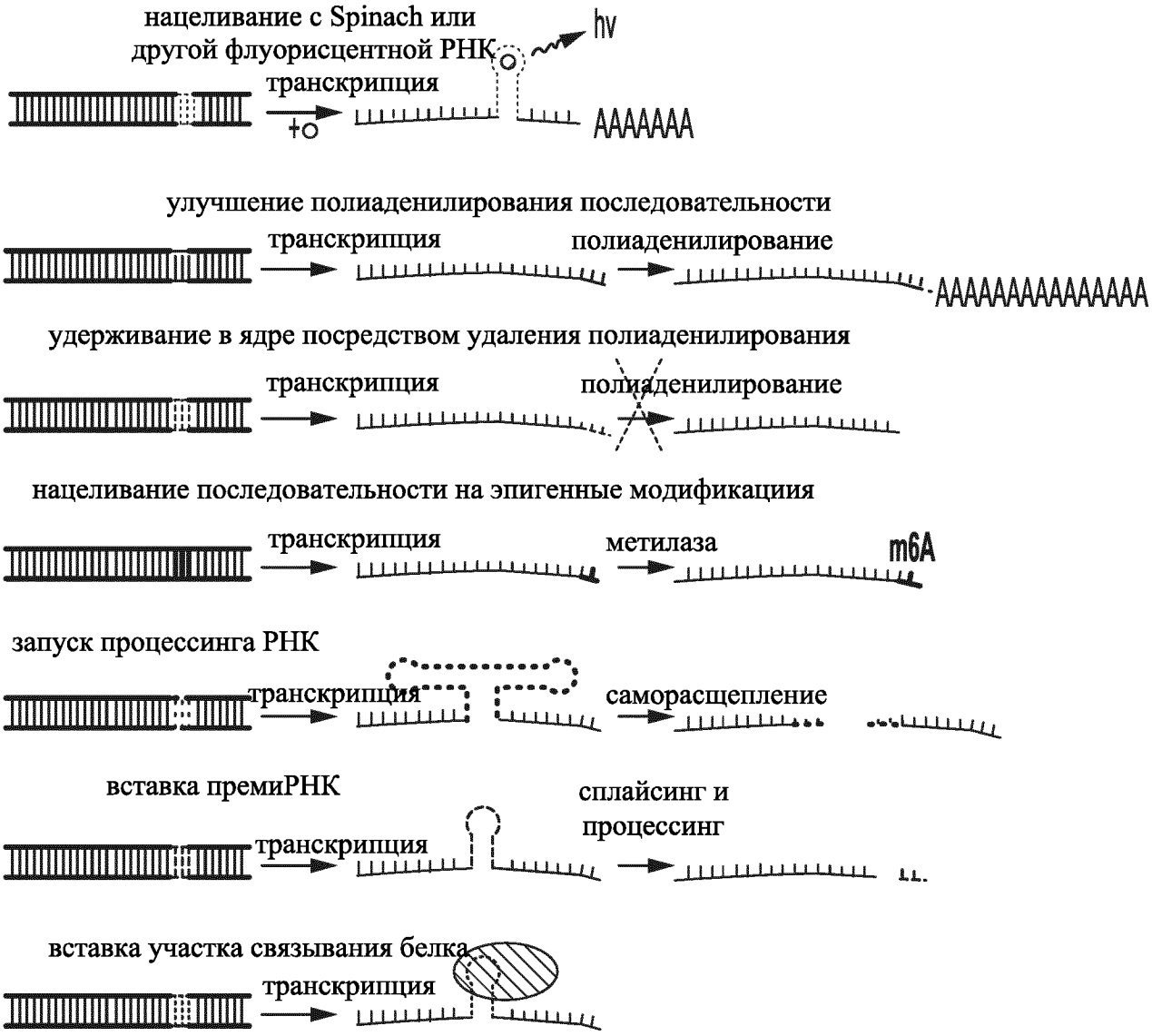
Фиг. 26B


PRNP
SEQ ID NO: 805 Protein sequence ... Gly Leu Gly 127 Gly Tyr Met Leu ...
SEQ ID NO: 806 DNA sequence ... GGC CTT GGC GGC TAC ATG CTG ... аллель дикого типа
праймированное редактирование
SEQ ID NO: 807 DNA sequence ... GGC CTT GGC GTC TAC ATG CTG ... прионное заболевание
SEQ ID NO: 808 Protein sequence ... Gly Leu Gly Val Tyr Met Leu ... утойчивый аллель
127
протоспейсер
PAM
SEQ ID NO: 809 ... GGG GCAGTGGTGGGGGGGCCTTGG CGG CTACATGCTG ...
SpCas9
сайт разрыва
нуклеотид мишень
кодирует мутацию в целевую пару оснований
РЕгРНК
SEQ ID NO: 810 5'-GCAGTGGTGGGGGGCCTTGG-sgRNA scaffold-ATGTAGACGCCAAGGCCCCCCAC-3'
матрица обратной транскрипции
последовательность отжига праймера
редактирование в HEK293T клетках
PRNP
% редактирования
30
25
20
15
10
5
0
коррекция превращения превращения G в T
индель
нет разрыва
+41 разрыв
+72 разрыв
+103 разрыв
Разрыв комплементарной цепи

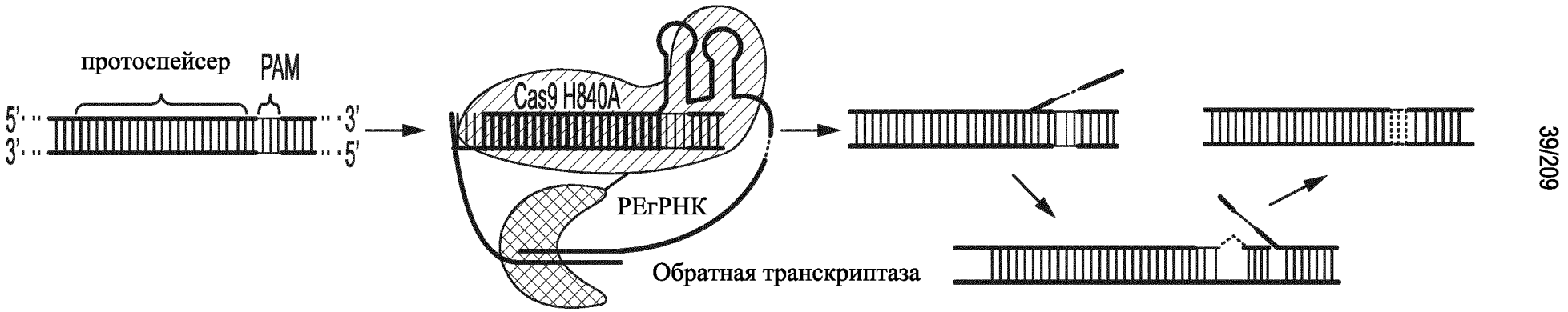
Фиг. 27



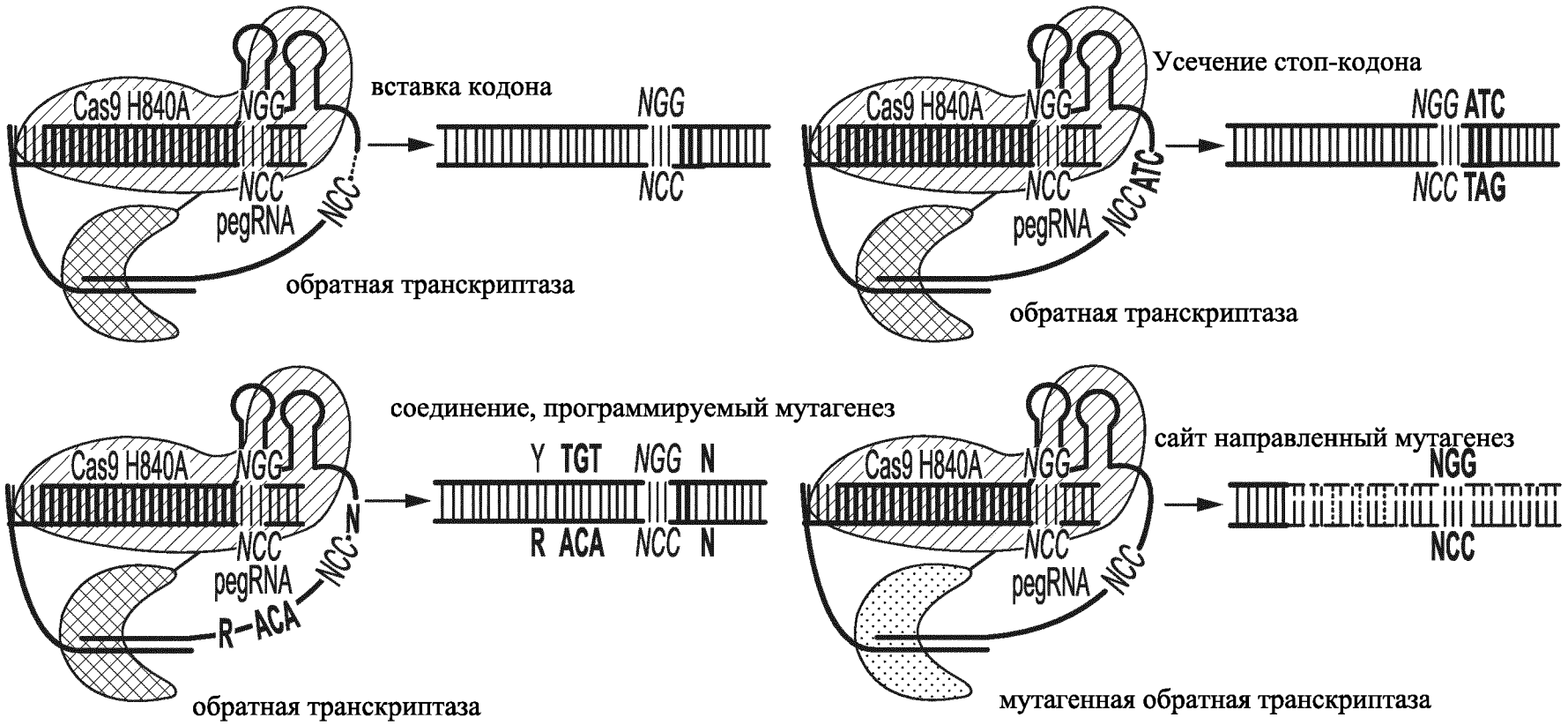
Фиг.28A



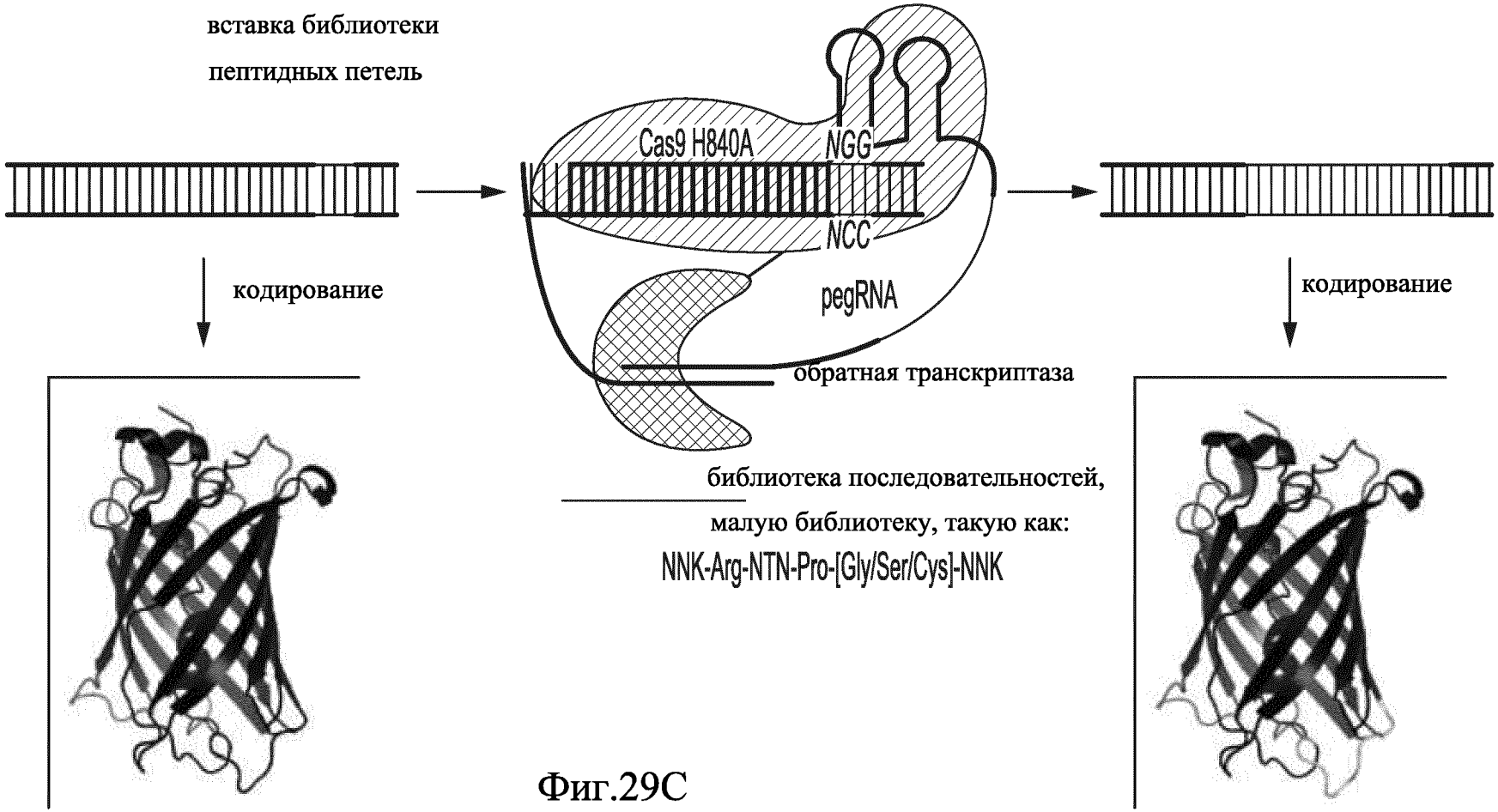
Фиг.28B


протоспейсер
PAM
5'
3'
3'
5'
Cas9 H840A
РЕгРНК
Обратная транскриптаза

Фиг.29А


Cas9 H840A
NGG
NCC
pegRNA
NCC
обратная транскриптаза
вставка кодона
NGG
NCC
Усечение стоп-кодона
NCC ATC
NGG ATC
NCC TAG
соединение, программируемый мутагенез
Y TGT NGG N
R ACA NCC N
R–ACA
N
сайт направленный мутагенез
NGG
NCC
мутагенная обратная транскриптаза

Фиг.29B

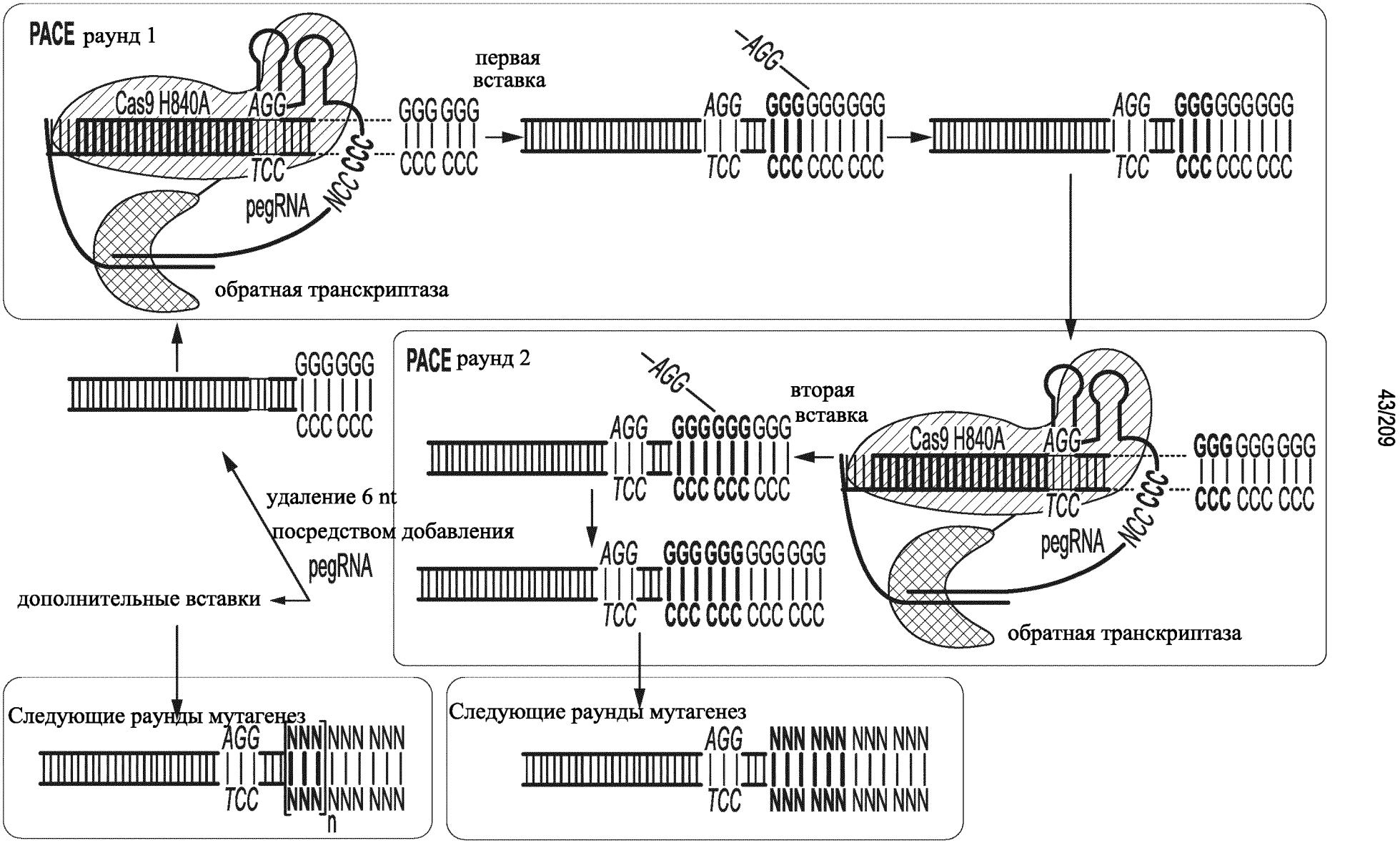

вставка библиотеки
пептидных петель
Cas9 H840A
NGG
NCC
pegRNA
обратная транскриптаза
кодирование
кодирование
библиотека последовательностей,
малую библиотеку, такую как:
NNK-Arg-NTN-Pro-[Gly/Ser/Cys]-NNK

Фиг.29C

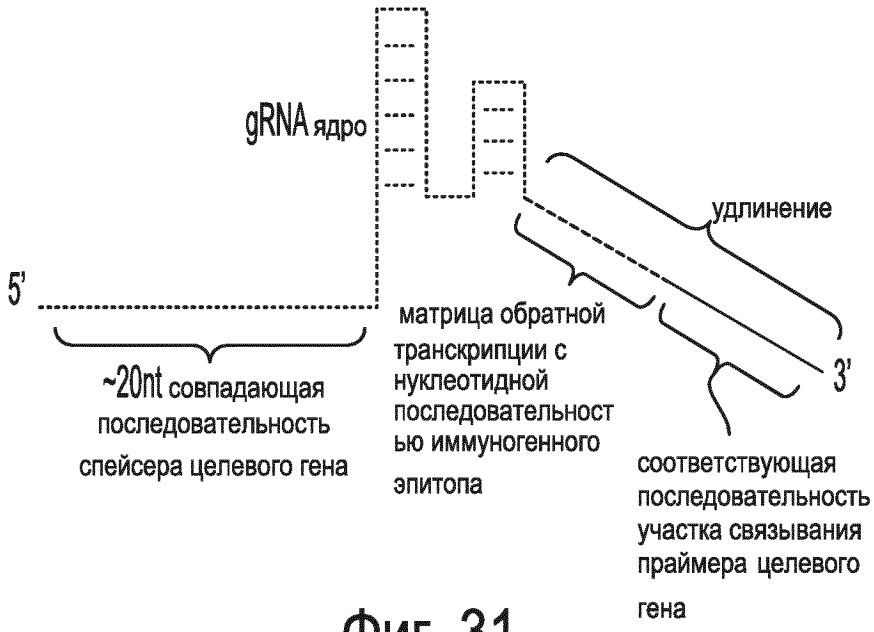
программируемое удаление кодона и усечения

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----------

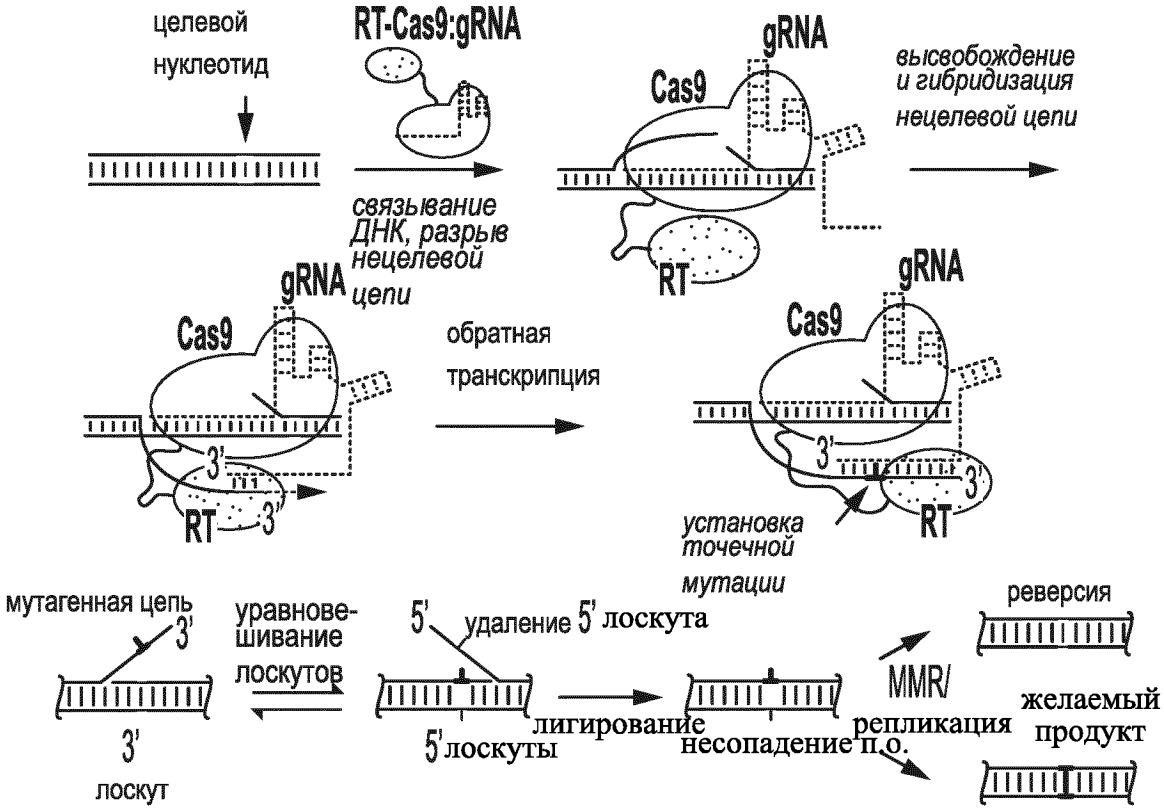
Фиг.29D


PACE раунд 1
Cas9 H840A
AGG
TCC
pegRNA
NCC
CCC
GGG GGG
CCC CCC
обратная транскриптаза
первая
вставка
AGG
GGG GGG GGG
TCC
CCC CCC CCC
PACE раунд 2
вторая
вставка
удаление 6 nt
посредством добавления
pegRNA
дополнительные вставки
Следующие раунды мутагенез
NNN NNN NNN
n

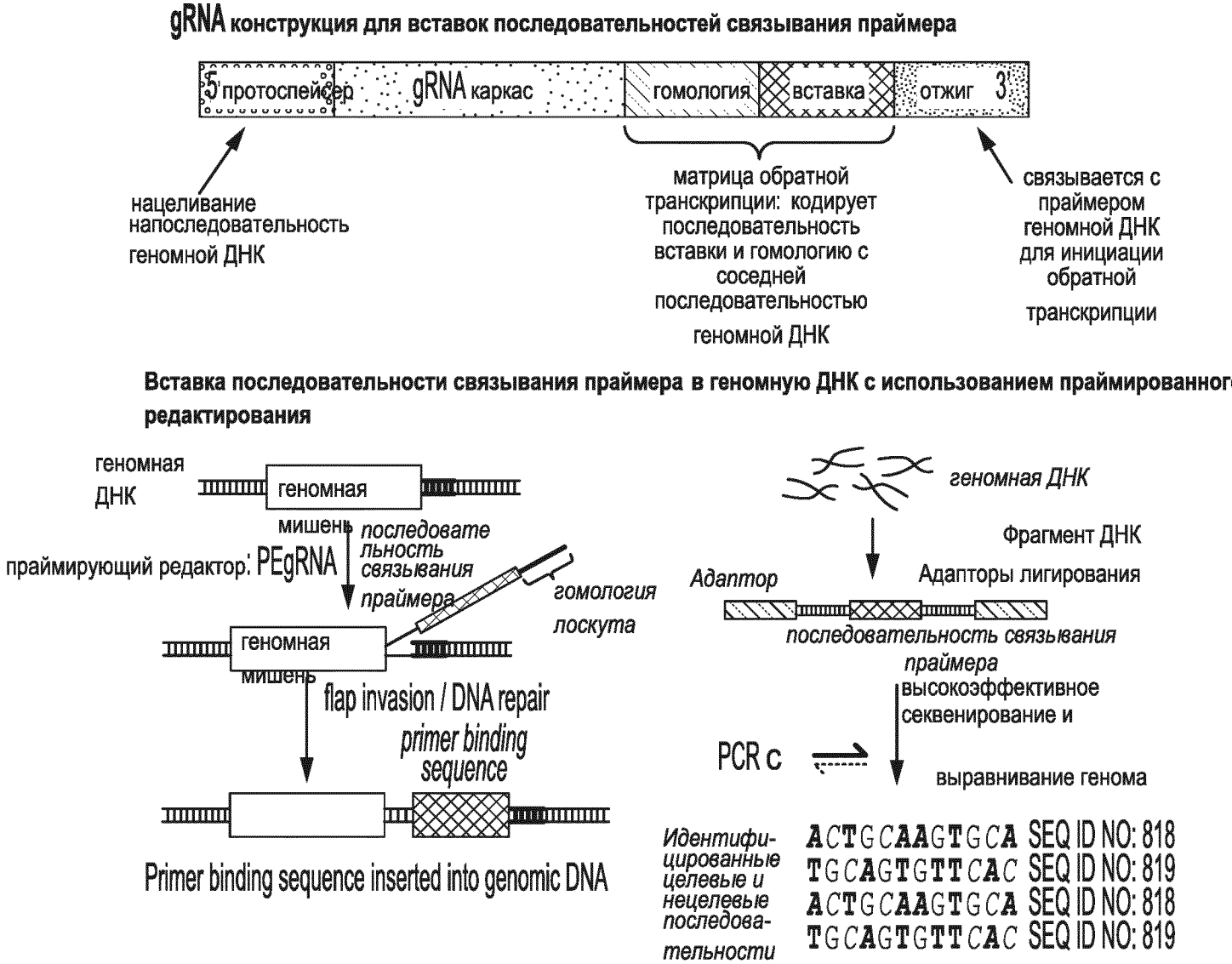
Фиг.30


gRNA ядро
удлинение
5'
3'
~20nt совпадающая последовательность спейсера целевого гена
матрица обратной транскрипции с нуклеотидной последовательностью иммуногенного эпитопа
соответствующая последовательность участка связывания праймера целевого гена

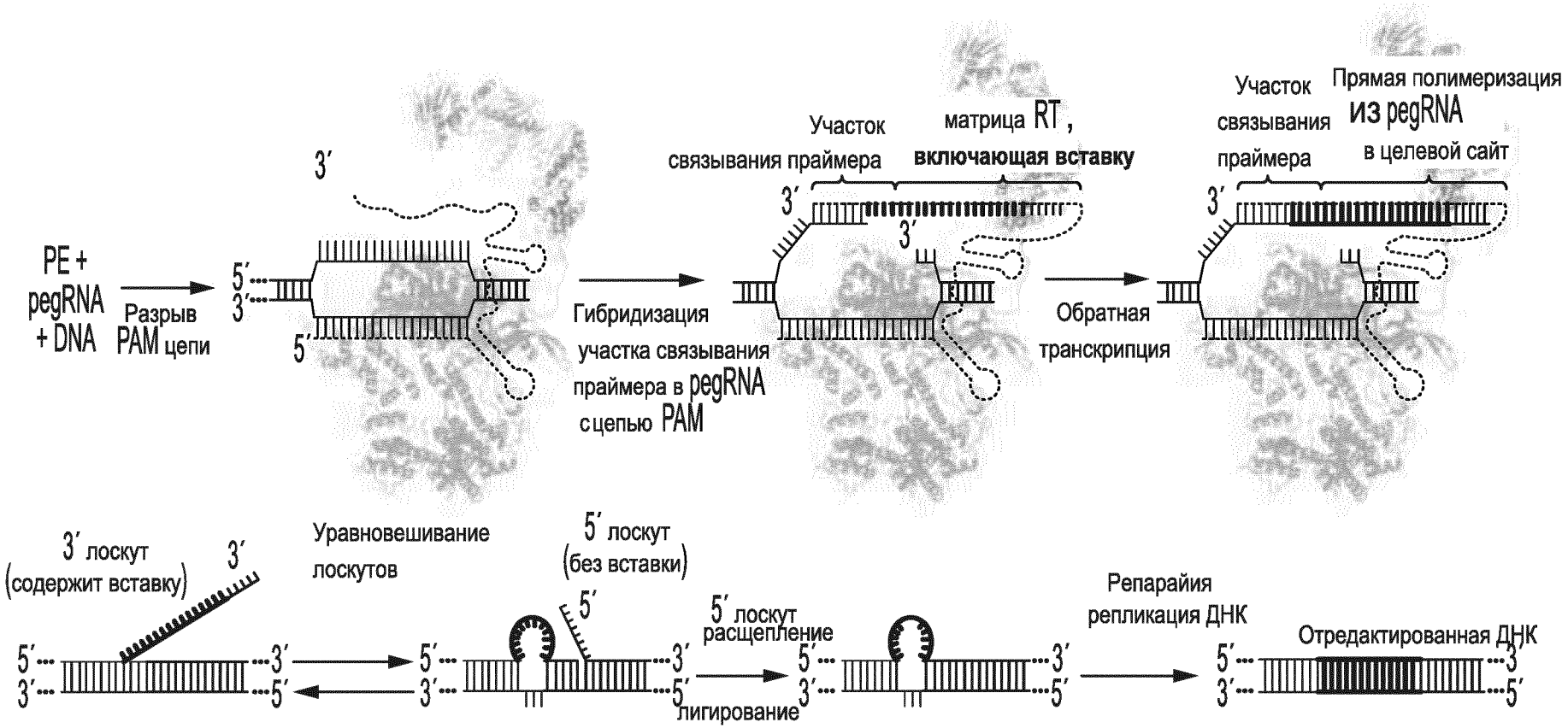
Фиг. 31


целевой нуклеотид
RT-Cas9:gRNA
связывание ДНК, разрыв нецелевой цепи
gRNA
Cas9
RT
высвобождение и гибридизация нецелевой цепи
обратная транскрипция
3'
3''
установка точечной мутации
мутагенная цепь
уравновешивание лоскутов
5'
удаление 5' лоскута
5' лоскуты
лигирование
несопадение п.о.
MMR/ репликация
реверсия
желаемый продукт
лоскут

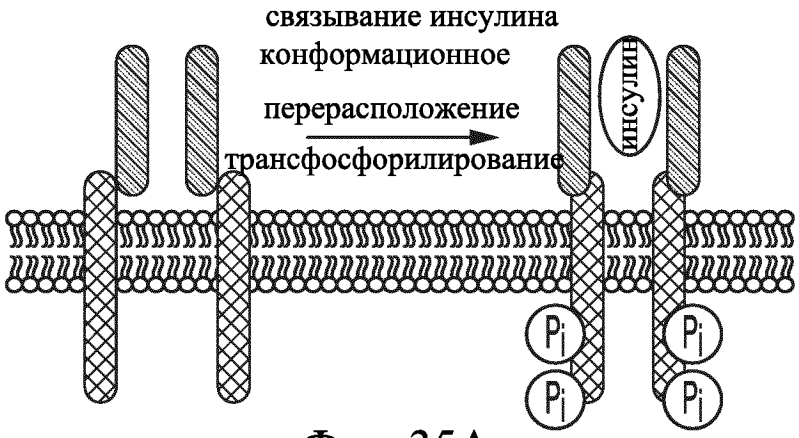
Фиг. 32


gRNA конструкция для вставок последовательностей связывания праймера
5' протоспейсер
gRNA каркас
гомология
вставка
отжиг 3'
нацеливание напоследовательность геномной ДНК
матрица обратной транскрипции: кодирует последовательность вставки и гомологию с соседней последовательностью геномной ДНК
связывается с праймером геномной ДНК для инициации обратной транскрипции
Вставка последовательности связывания праймера в геномную ДНК с использованием праймированного редактирования
геномная ДНК
геномная
мишень
праймирующий редактор: PEgRNA
последовательность связывания праймера
гомология
лоскута
геномная
мишень
flap invasion / DNA repair
primer binding sequence
Primer binding sequence inserted into genomic DNA
геномная ДНК
Фрагмент ДНК
Адаптор
Адаптеры лигирования
последовательность связывания праймера
высокоэффективное секвенирование и
PCR с
выравнивание генома
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

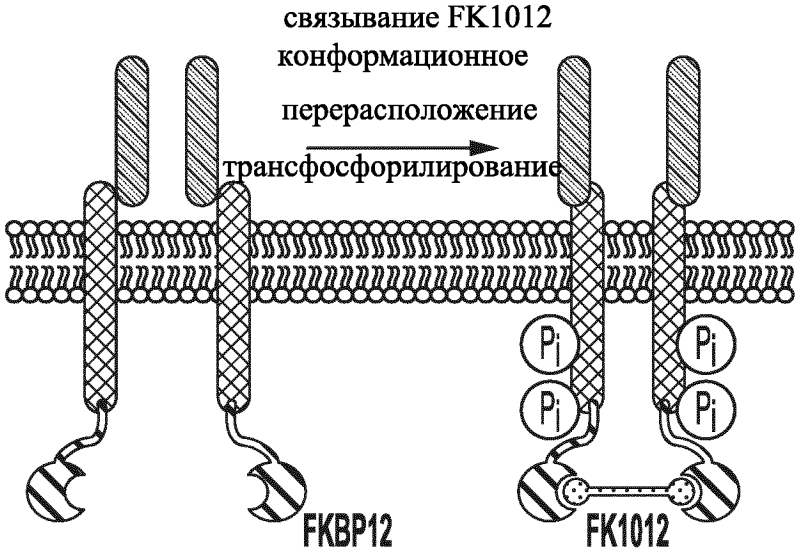
Фиг. 33


PE +
pegRNA
+ DNA
Разрыв
PAM цепи
3′
5′
3′
5′
Гибридизация
участка связывания
праймера в pegRNA
с цепью PAM
Участок
связывания праймера
матрица RT,
включающая вставку
3′
3′
Обратная
транскрипция
Участок
связывания
праймера
Прямая полимеризация
из pegRNA
в целевой сайт
3′
3′
3′ лоскут
(содержит вставку)
3′
Уравновешивание
лоскутов
5′ лоскут
(без вставки)
5′
5′ лоскут
расщепление
Лигирование
Репарайия
репликация ДНК
Отредактированная ДНК
5′
3′
3′
5′

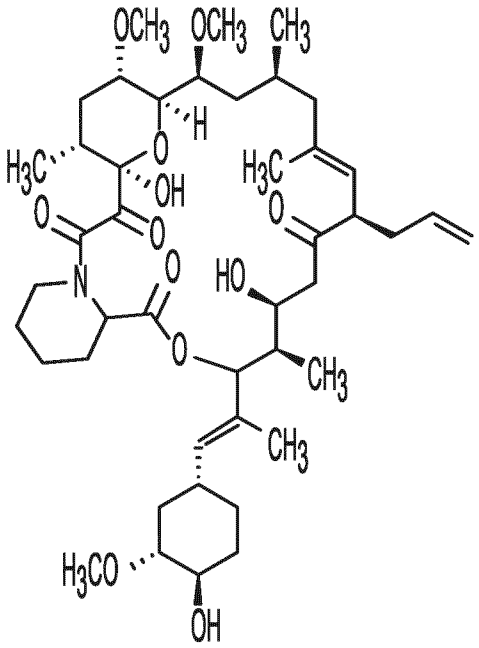
Фиг. 34



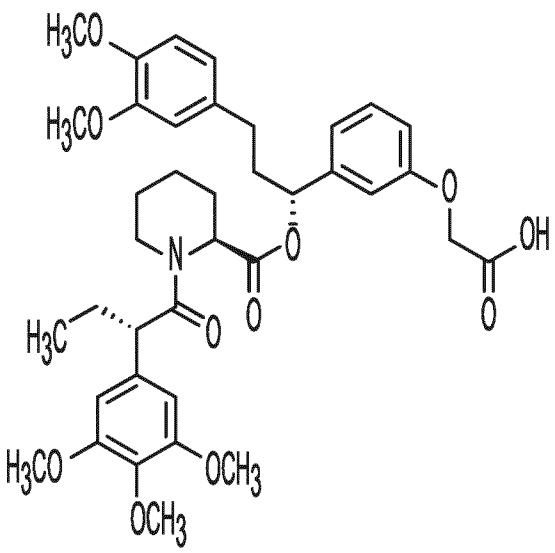
Фиг. 35A



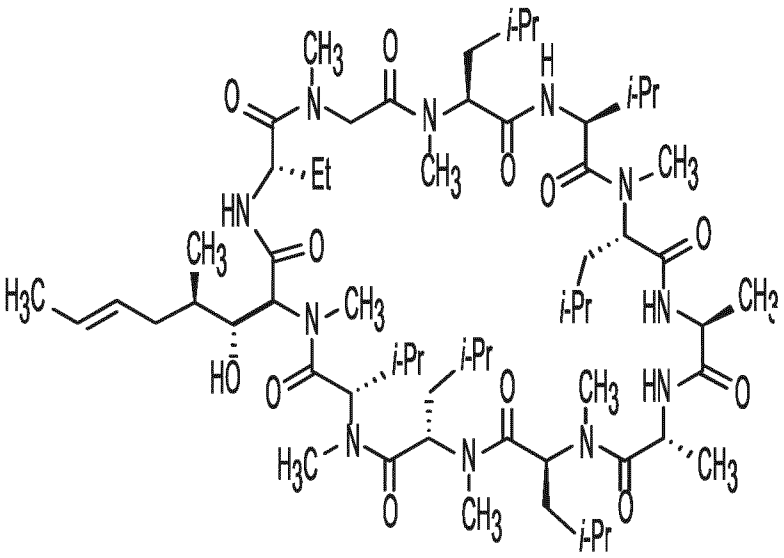
Фиг. 35B



1
FK506
мишень FKBP12

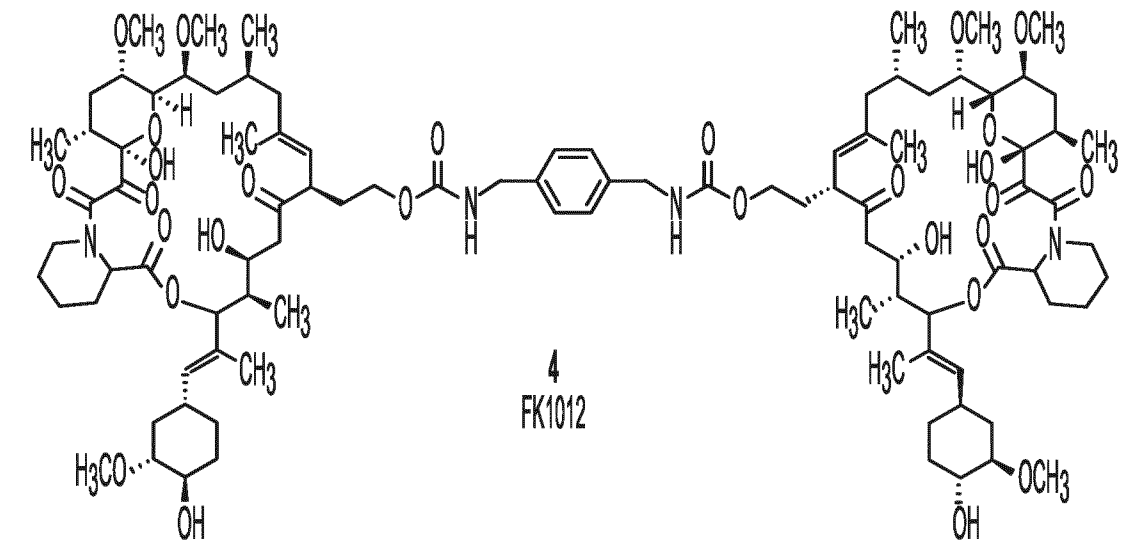


2
bumped FK506 миметик
мишень : FKBP12 F36V

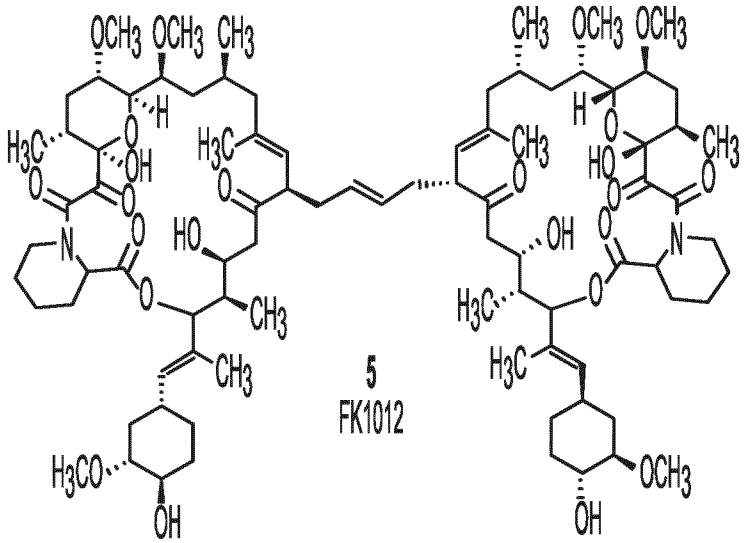


3
циклоспорин А
мишень: циклофилин

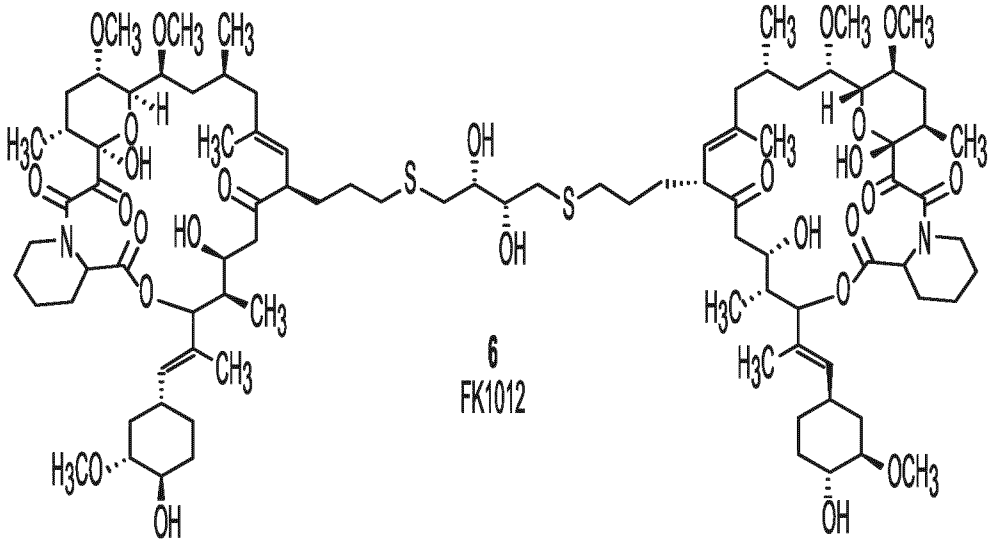
Фиг.36



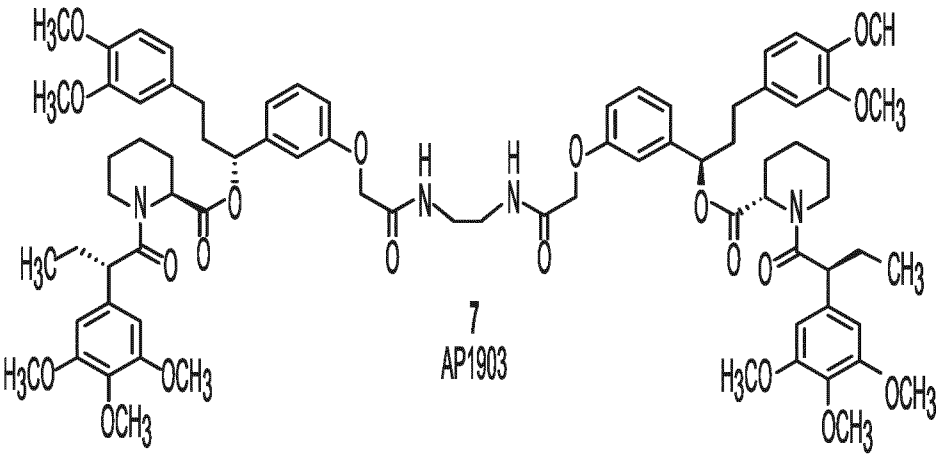
4
FK1012



5
FK1012

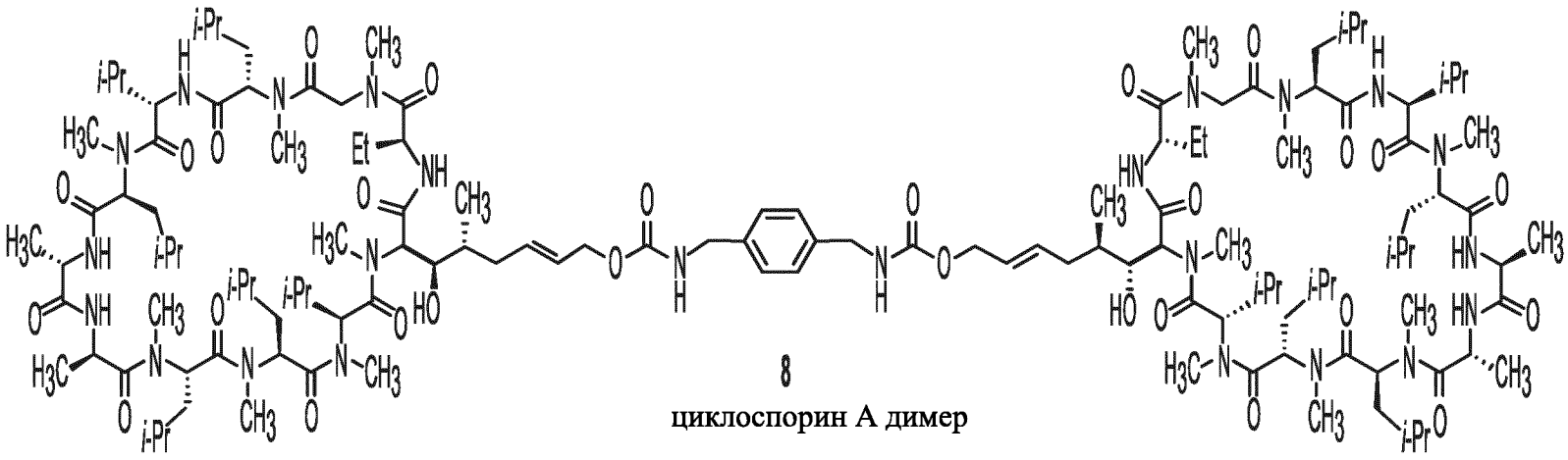


6
FK1012

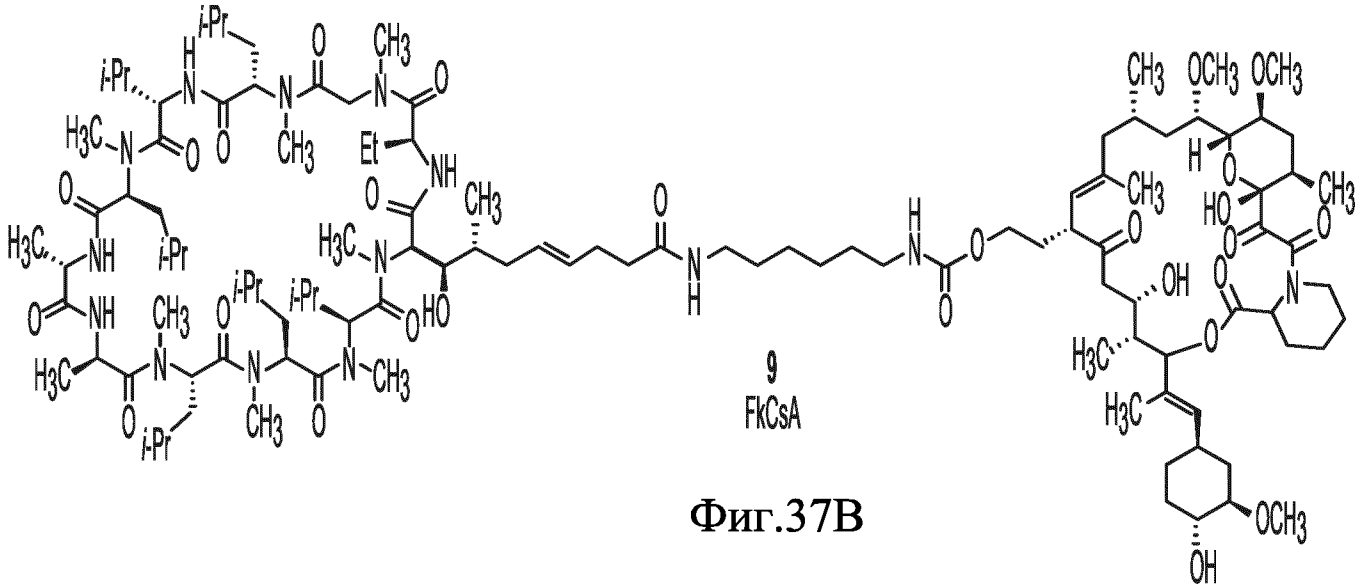


7
AP1903

Фиг.37A

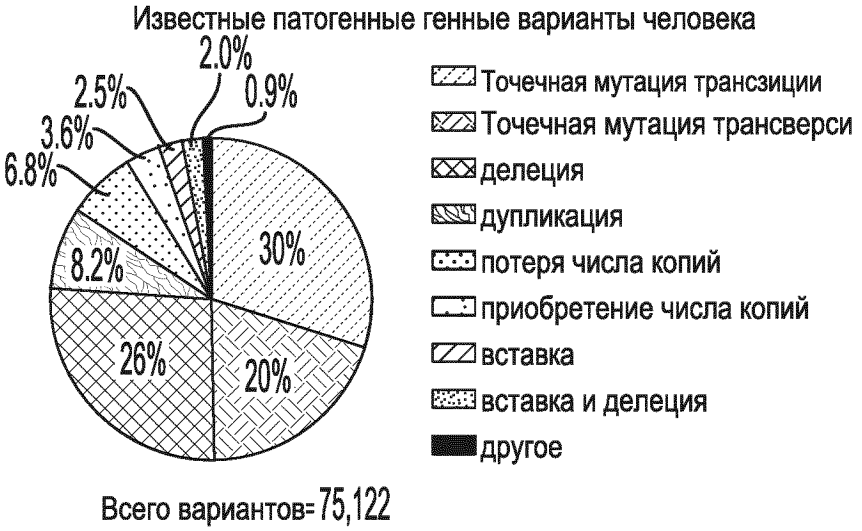


8
циклоспорин А димер

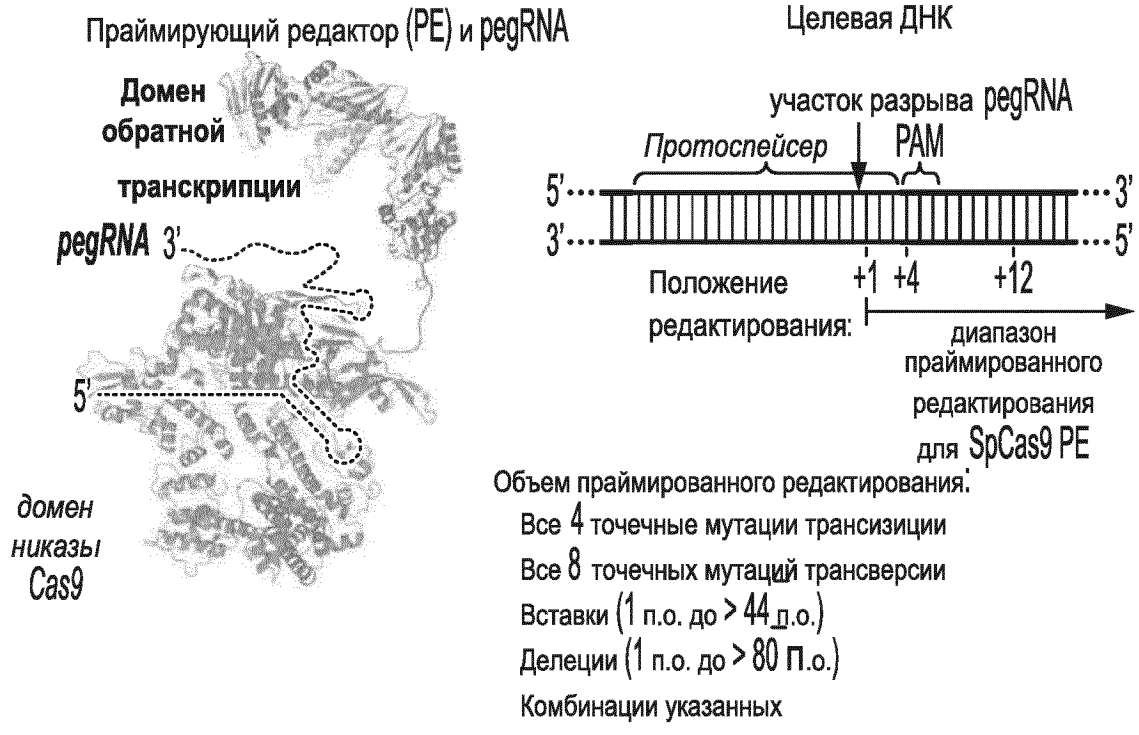


9
FkCsA

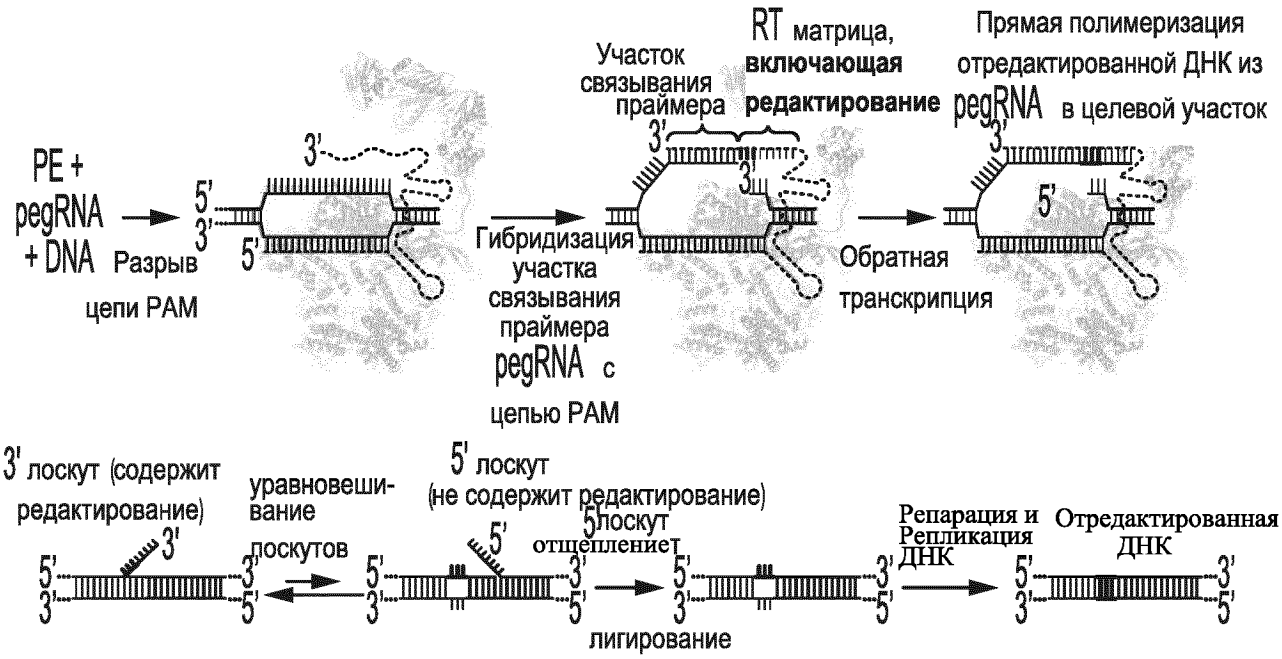
Фиг.37В



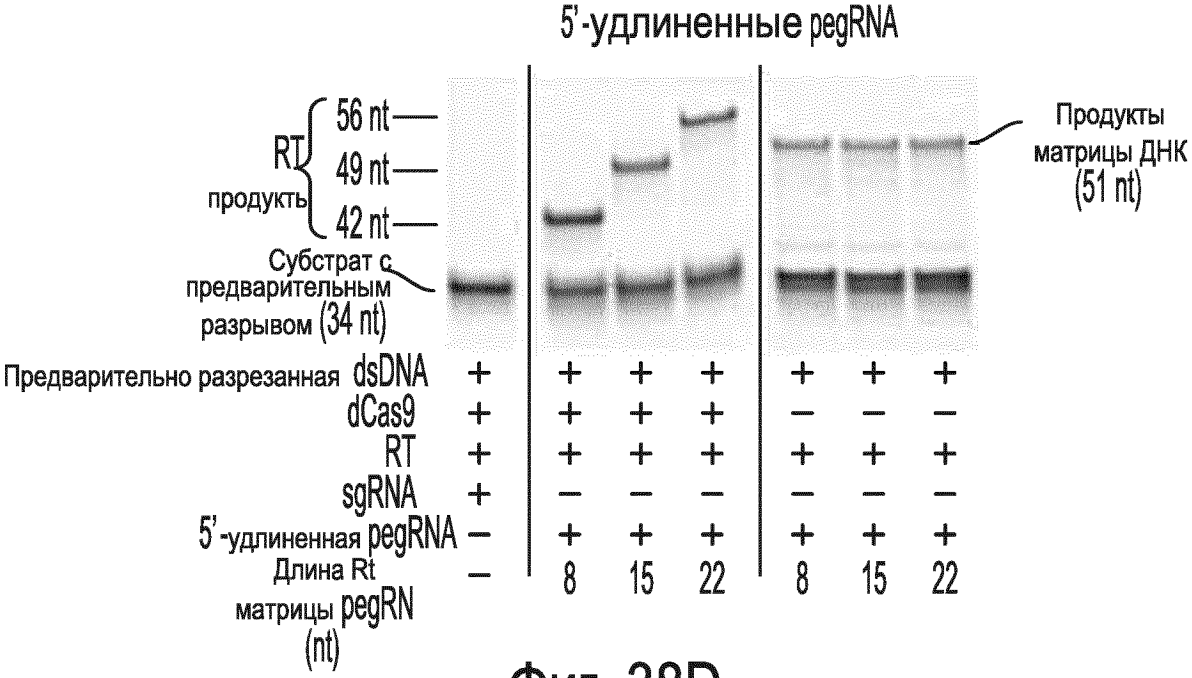
Фиг. 38A



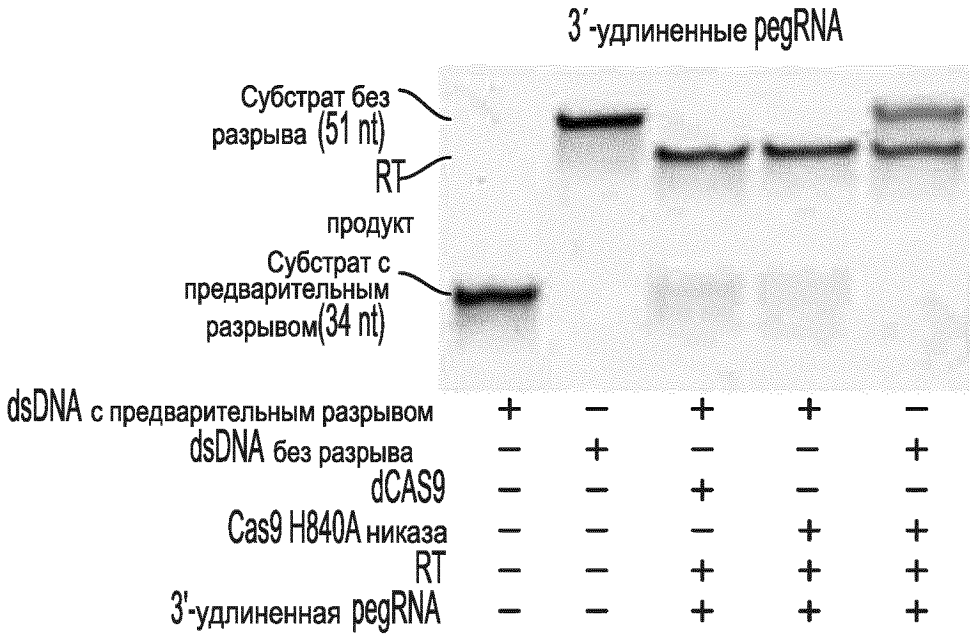
Фиг. 38B



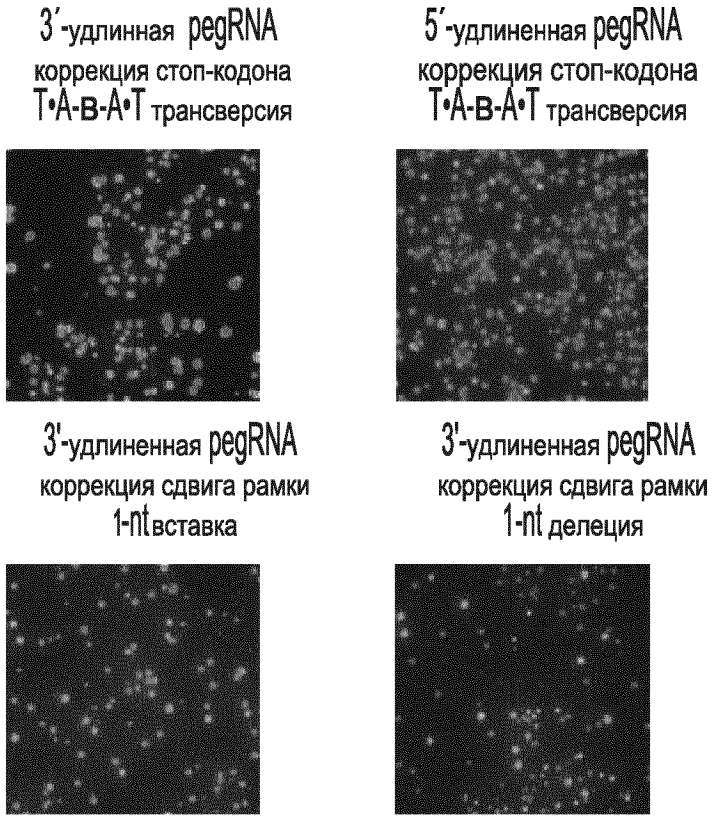
Фиг. 38C



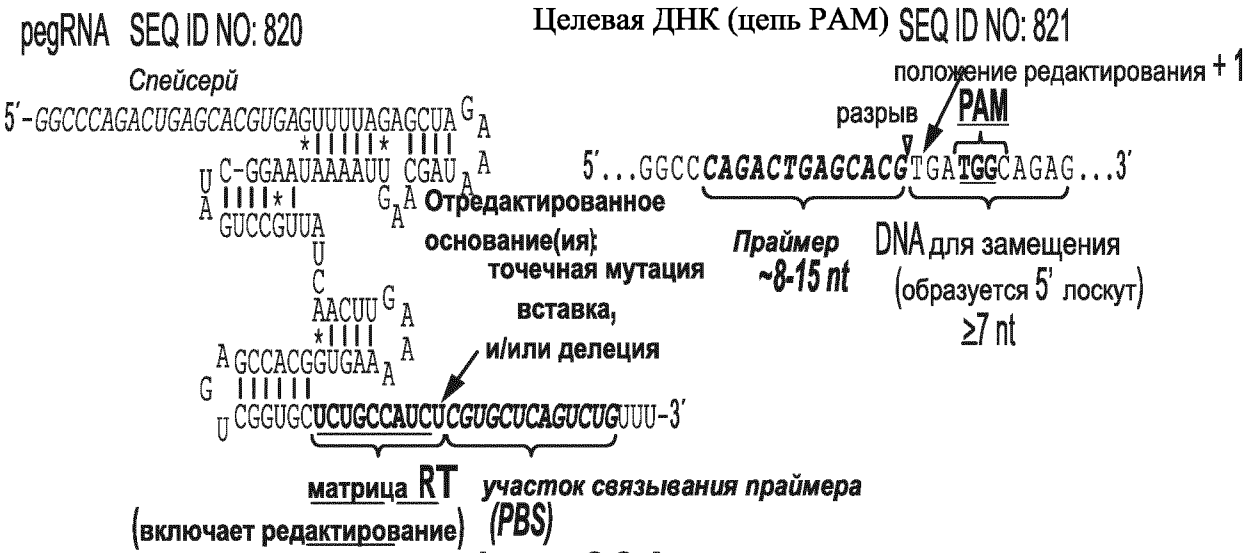
Фиг. 38D



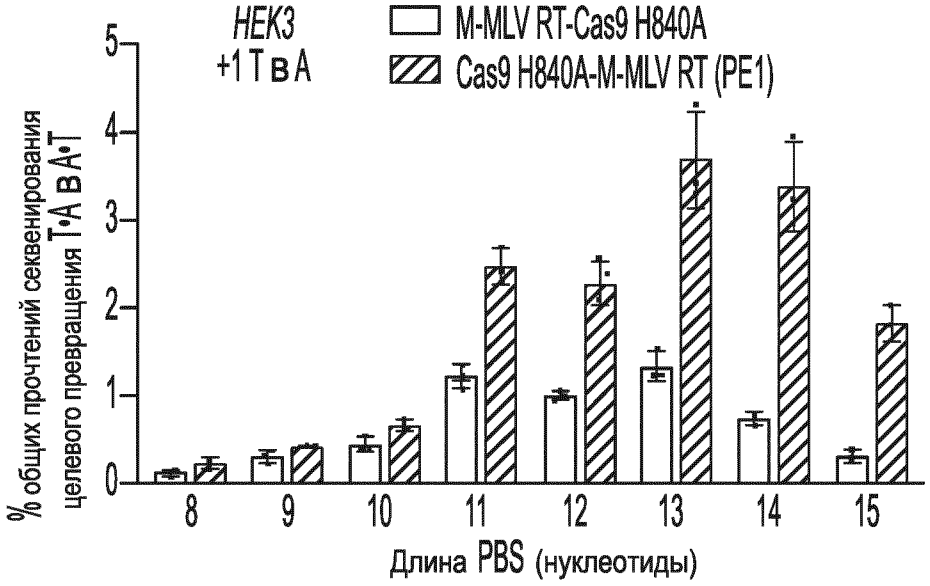
Фиг. 38E



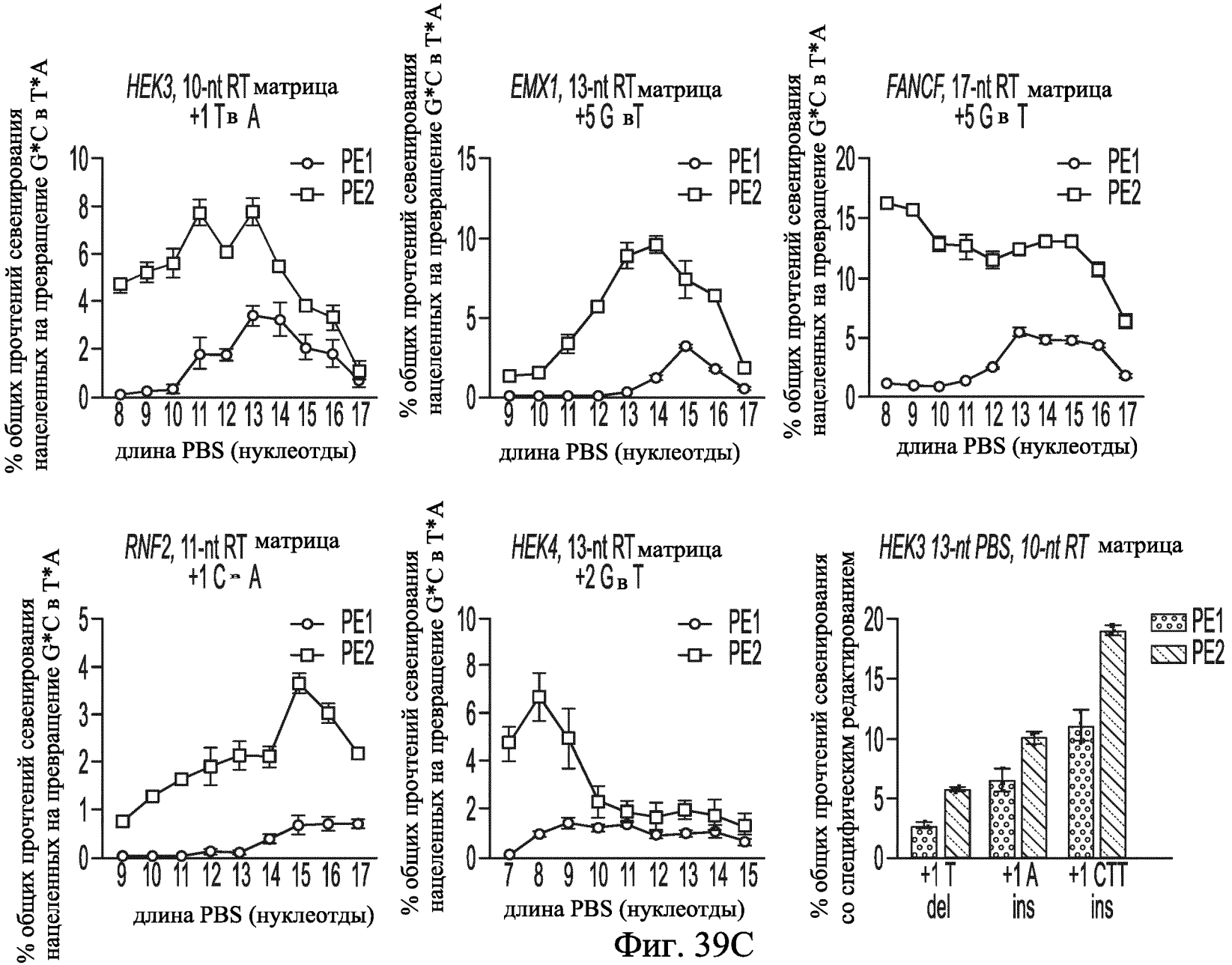
Фиг. 38F



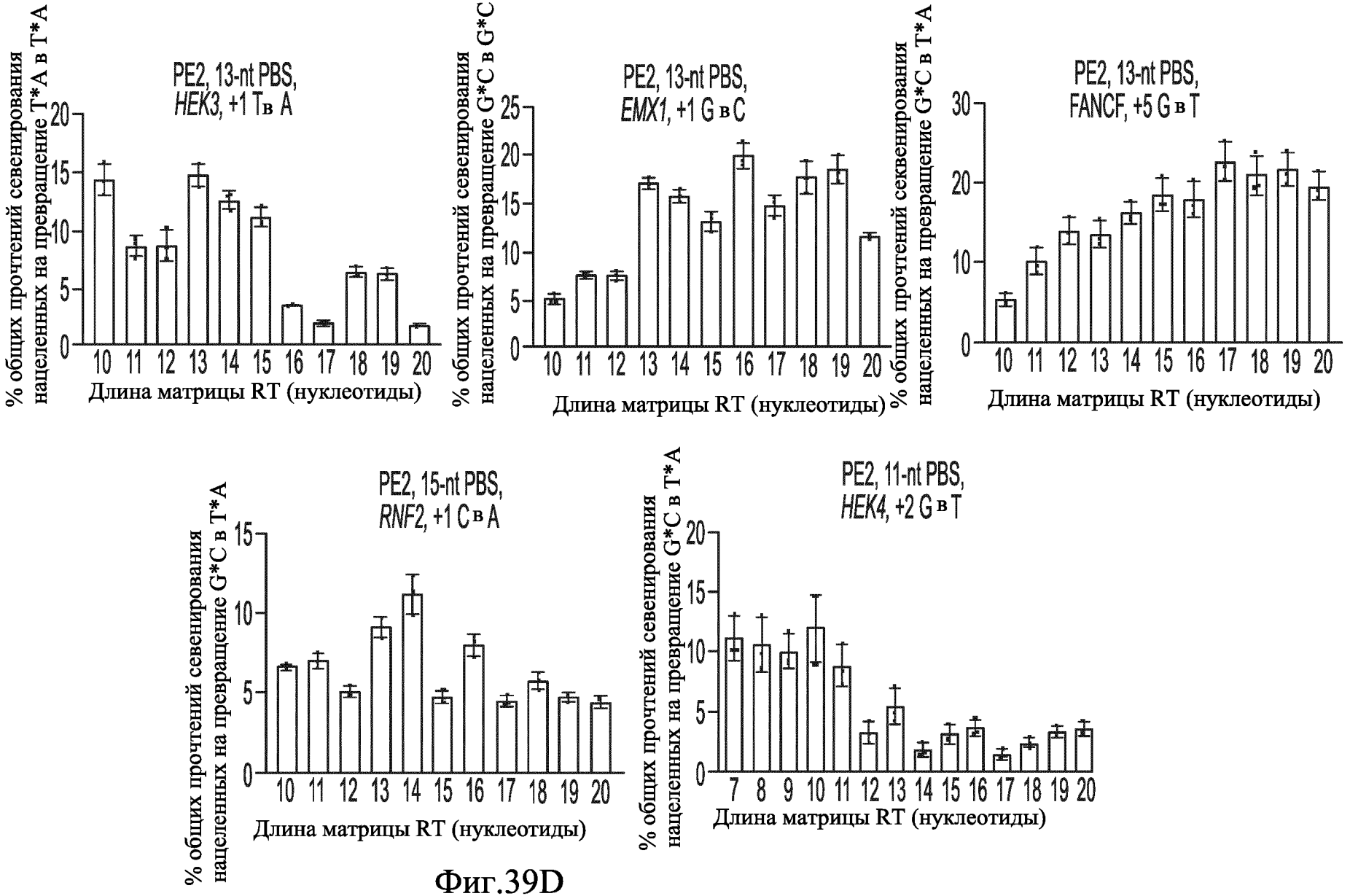
Фиг. 39A



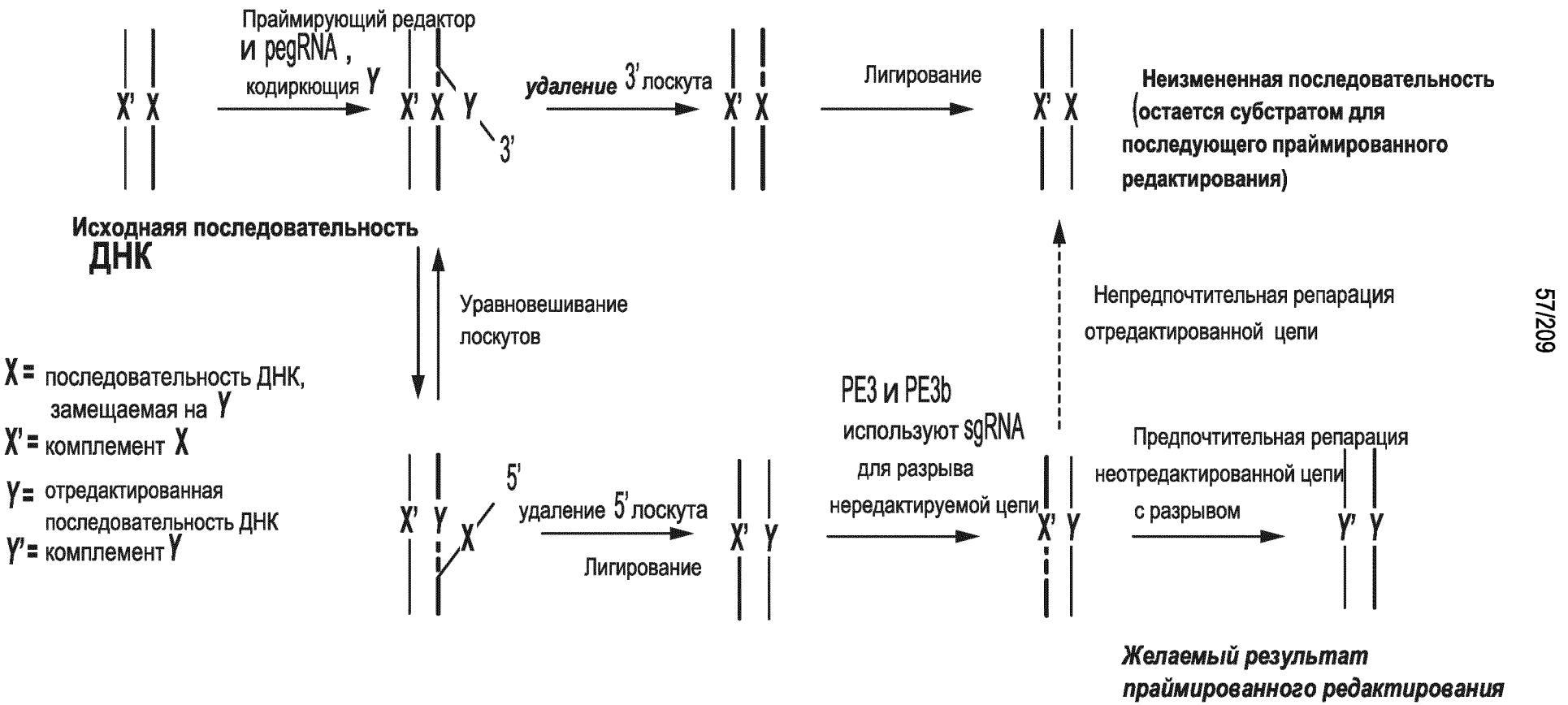
Фиг. 39B


HEK3, 10-nt RT матрица
+1 T в A
EMX1, 13-nt RT матрица
+5 G в T
FANCF, 17-nt RT матрица
+5 G в T
RNF2, 11-nt RT матрица
+1 C в A
HEK4, 13-nt RT матрица
+2 G в T
HEK3 13-nt PBS, 10-nt RT матрица
PE1
PE2
% общих прочтений севенирования нацеленных на превращение G*C в T*A
% общих прочтений севенирования со специфическим редактированием
длина PBS (нуклеотды)
+1 T del
+1 A ins
+1 CTT ins

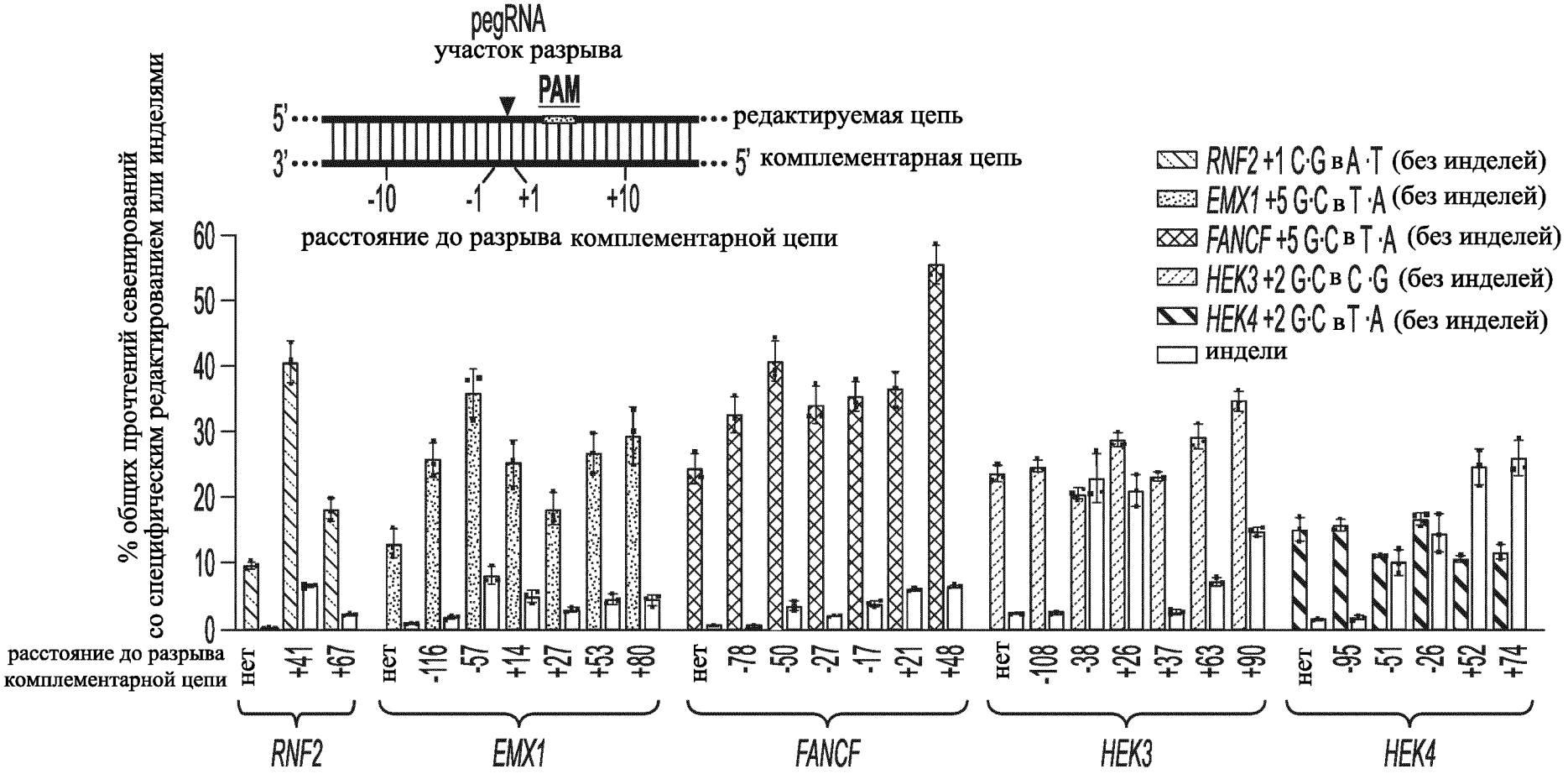
Фиг. 39C


PE2, 13-nt PBS,
HEK3, +1 T в A
% общих прочтений севенирования нацеленных на превращение T*A в T*A
Длина матрицы RT (нуклеотиды)
PE2, 13-nt PBS,
EMX1, +1 G в C
% общих прочтений севенирования нацеленных на превращение G*C в G*C
Длина матрицы RT (нуклеотиды)
PE2, 13-nt PBS,
FANCF, +5 G в T
% общих прочтений секвенирования нацеленных на превращение G*C в T*A
Длина матрицы RT (нуклеотиды)
PE2, 15-nt PBS,
RNF2, +1 C в A
% общих прочтений севенирования нацеленных на превращение G*C в T*A
Длина матрицы RT (нуклеотиды)
PE2, 11-nt PBS,
HEK4, +2 G в T
% общих прочтений севенирования нацеленных на превращение G*C в T*A
Длина матрицы RT (нуклеотиды)

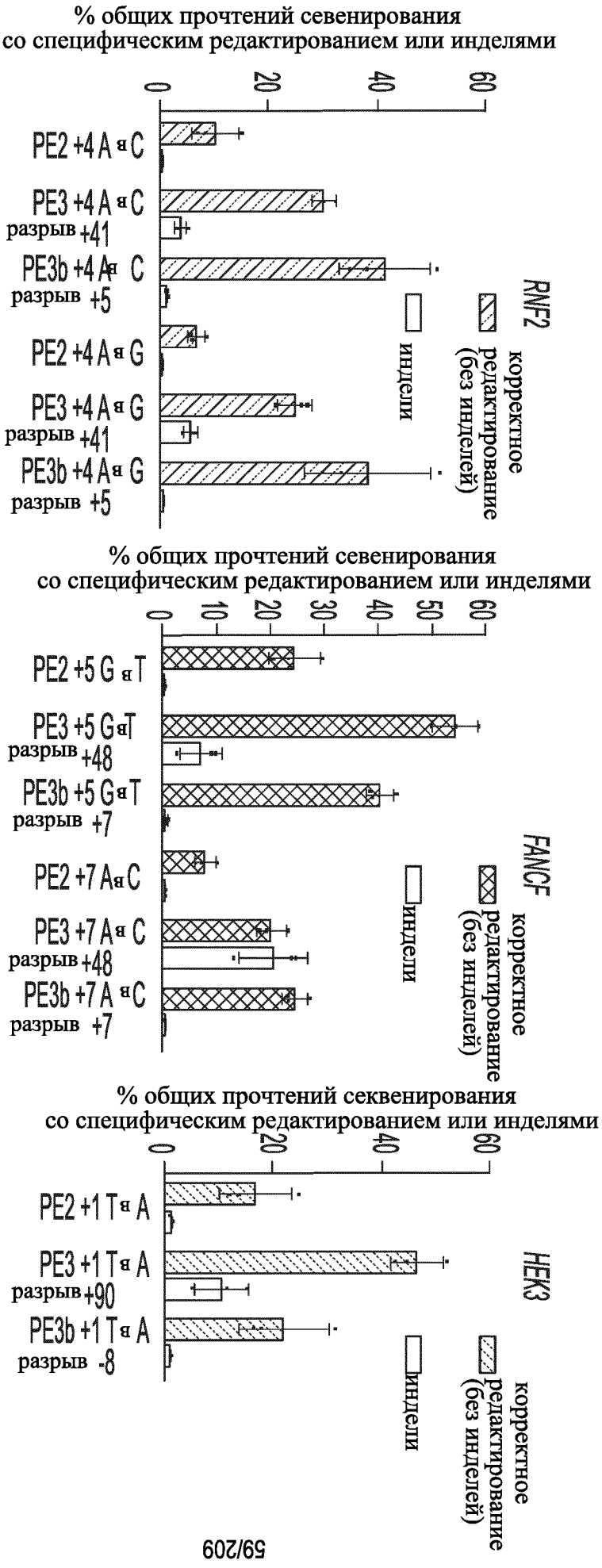
Фиг.39D


Праймирующий редактор
и pegRNA ,
кодирующая Y
X' X
X' X Y
3'
удаление 3' лоскута
X' X
Лигирование
X' X
Неизмененная последовательность
(остается субстратом для
последующего праймированного
редактирования)
Исходная последовательность
ДНК
Уравновешивание
лоскутов
Непредпочтительная репарация
отредактированной цепи
X= последовательность ДНК,
замещаемая на Y
X'= комплемент X
Y= отредактированная
последовательность ДНК
Y'= комплемент Y
PE3 и PE3b
используют sgRNA
для разрыва
нередактируемой цепи
X' Y
Предпочтительная репарация
неотредактированной цепи
с разрывом
Y' Y
5'
X' Y X
удаление 5' лоскута
X' Y
Лигирование
Желаемый результат
праймированного редактирования

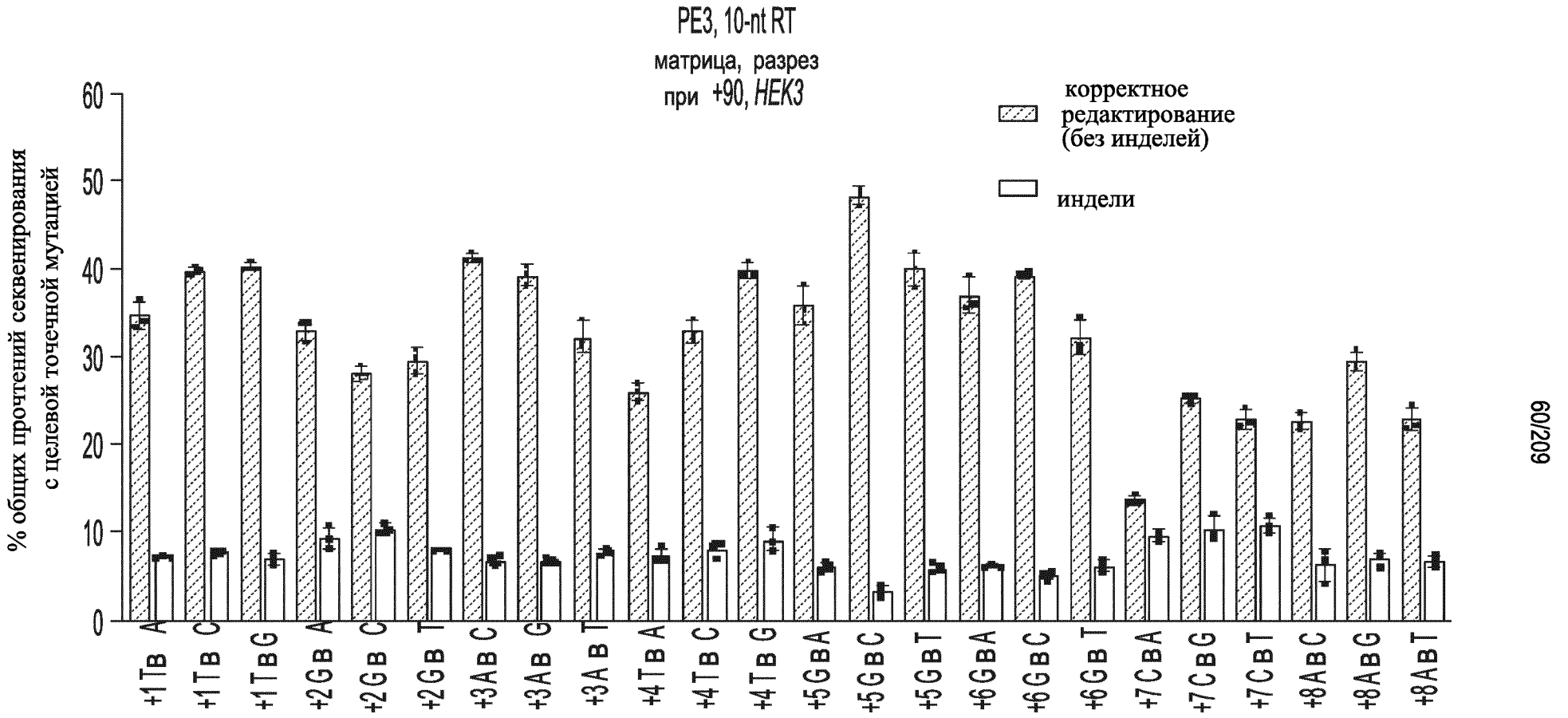
Фиг. 40A


pegRNA
участок разрыва
PAM
5'
3'
редактируемая цепь
5' комплементарная цепь
-10
-1
+1
+10
расстояние до разрыва комплементарной цепи
% общих прочтений севенирований со специфическим редактированием или инделями
60
50
40
30
20
10
0
RNF2 +1 C·G в A·T (без инделей)
EMX1 +5 G·C в T·A (без инделей)
FANCF +5 G·C в T·A (без инделей)
HEK3 +2 G·C в C·G (без инделей)
HEK4 +2 G·C в T·A (без инделей)
индели
расстояние до разрыва комплементарной цепи
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

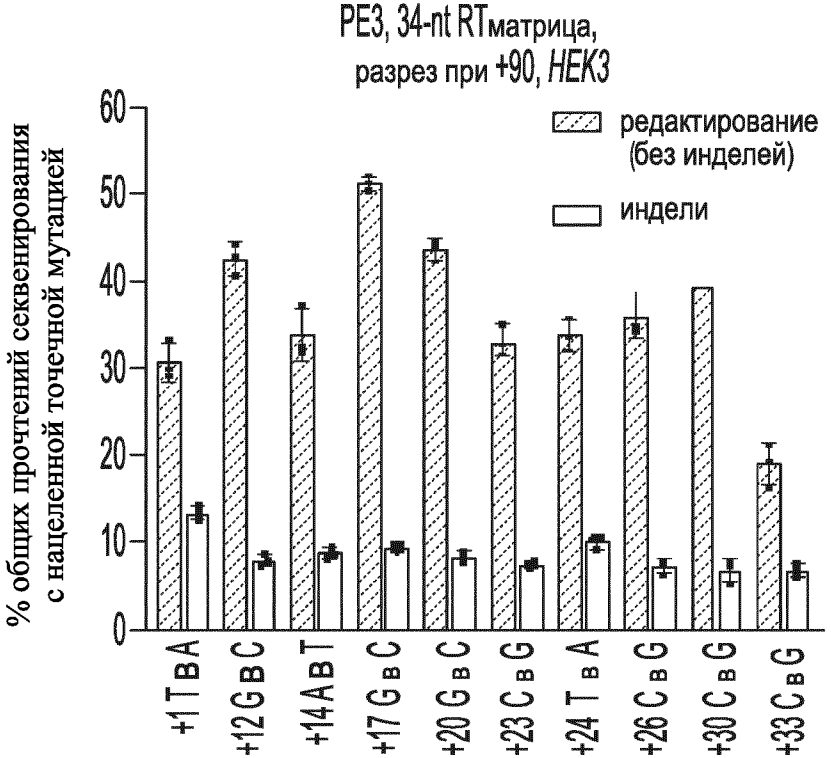
Фиг.40B


% общих прочтений севенирования
со специфическим редактированием или инделями
0
20
40
60
PE2 +4 A я C
PE3 +4 A я C
разрыв +41
PE3b +4 A я C
разрыв +5
PE2 +4 A я G
PE3 +4 A я G
разрыв +41
PE3b +4 A я G
разрыв +5
RNF2
корректное редактирование (без инделей)
индели
% общих прочтений севенирования
со специфическим редактированием или инделями
0
10
20
30
40
50
60
PE2 +5 G я T
PE3 +5 G я T
разрыв +48
PE3b +5 G я T
разрыв +7
PE2 +7 A я C
PE3 +7 A я C
разрыв +48
PE3b +7 A я C
разрыв +7
FANCF
корректное редактирование (без инделей)
индели
% общих прочтений секвенирования
со специфическим редактированием или инделями
0
20
40
60
PE2 +1 T я A
PE3 +1 T я A
разрыв +90
PE3b +1 T я A
разрыв -8
HEK3
корректное редактирование (без инделей)
индели

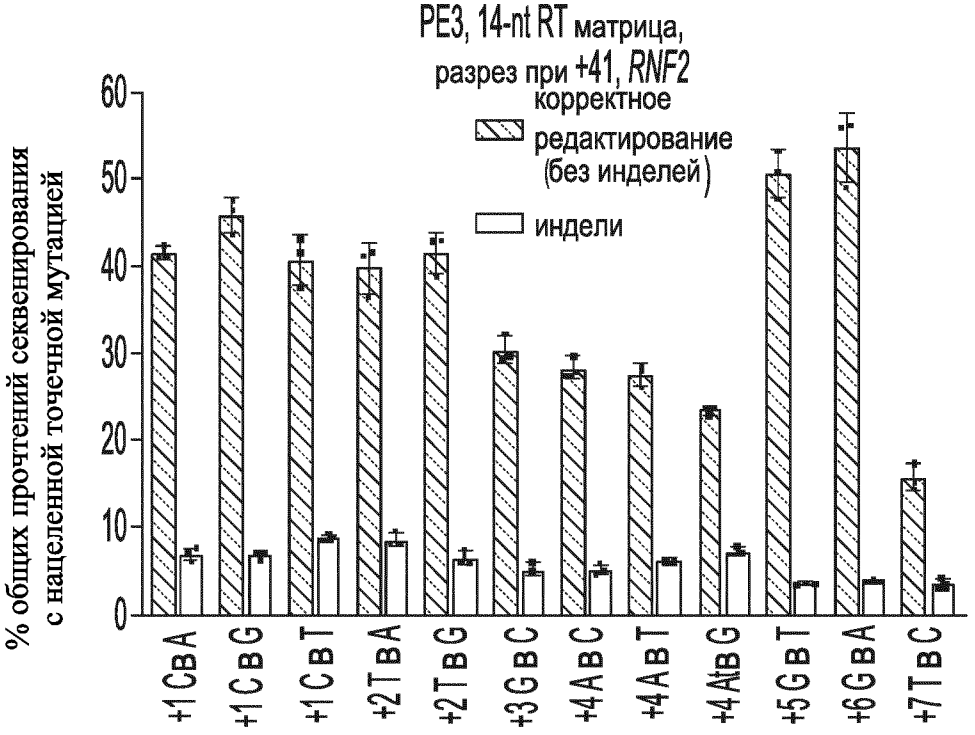
Фиг.40С


PE3, 10-nt RT
матрица, разрез
при +90, HEK3
корректное редактирование (без инделей)
индели
% общих прочтений секвенирования с целевой точечной мутацией
0
10
20
30
40
50
60
+1TB A
+1TB C
+1TB G
+2GB A
+2GB C
+2GB T
+3AB C
+3AB G
+3AB T
+4TB A
+4TB C
+4TB G
+5GB A
+5GB C
+5GB T
+6GB A
+6GB C
+6GB T
+7CB A
+7CB G
+7CB T
+8AB C
+8AB G
+8AB T

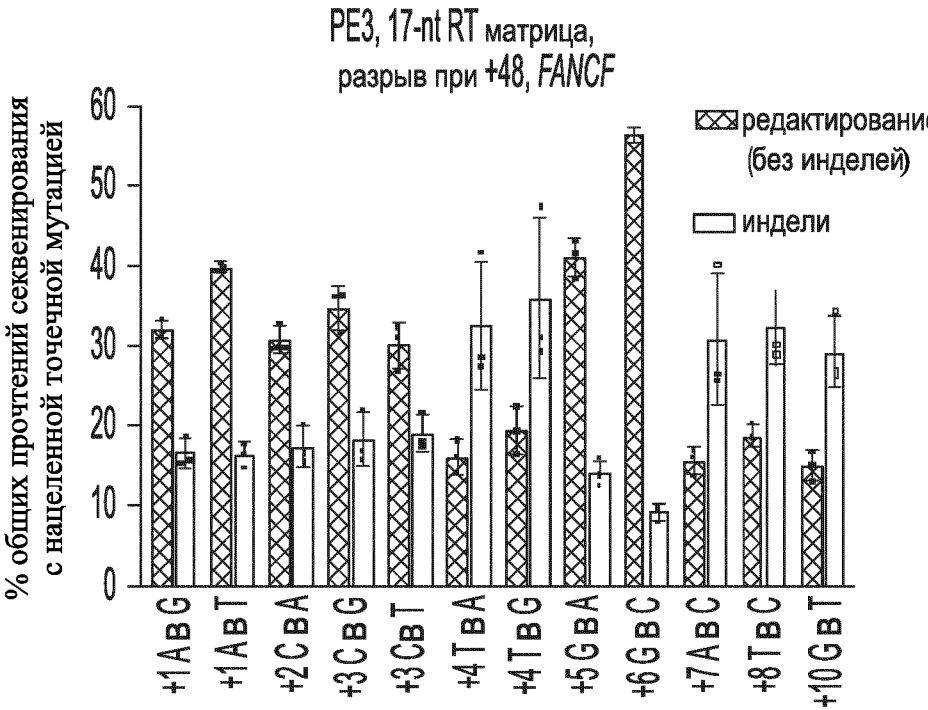
фиг. 41A



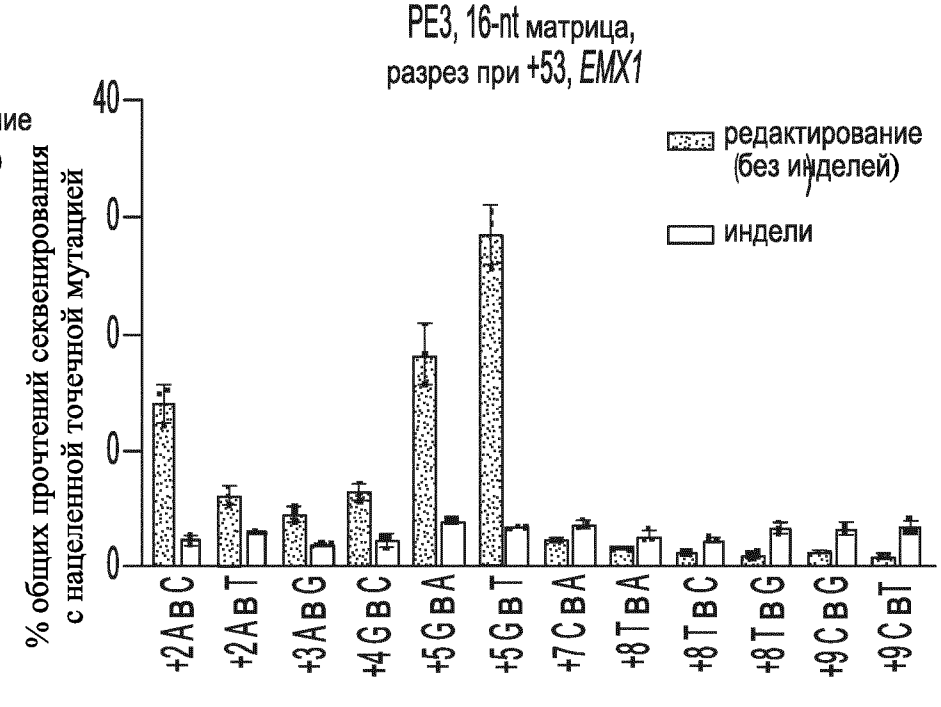
Фиг. 41B



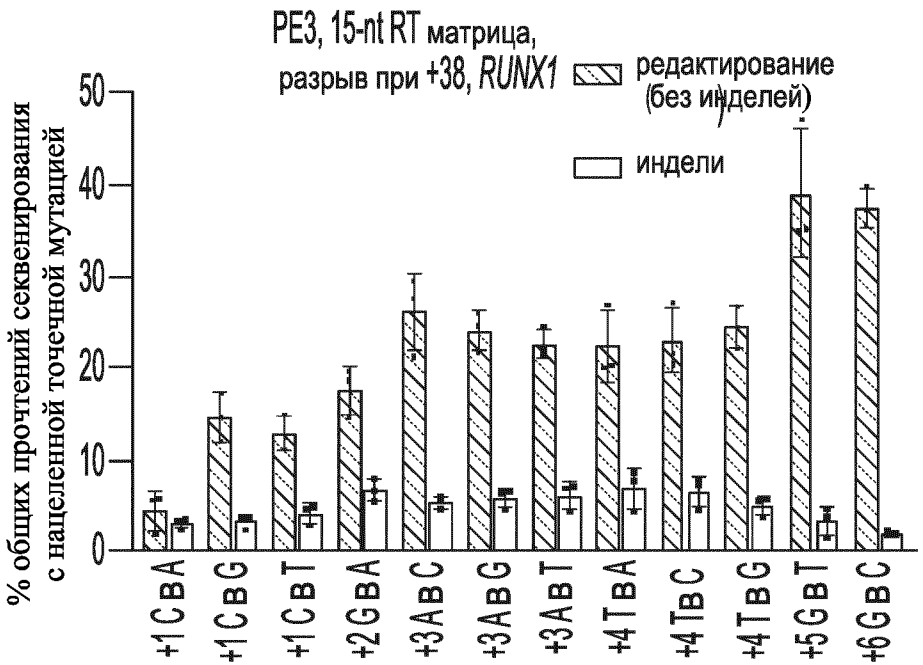
Фиг. 41C



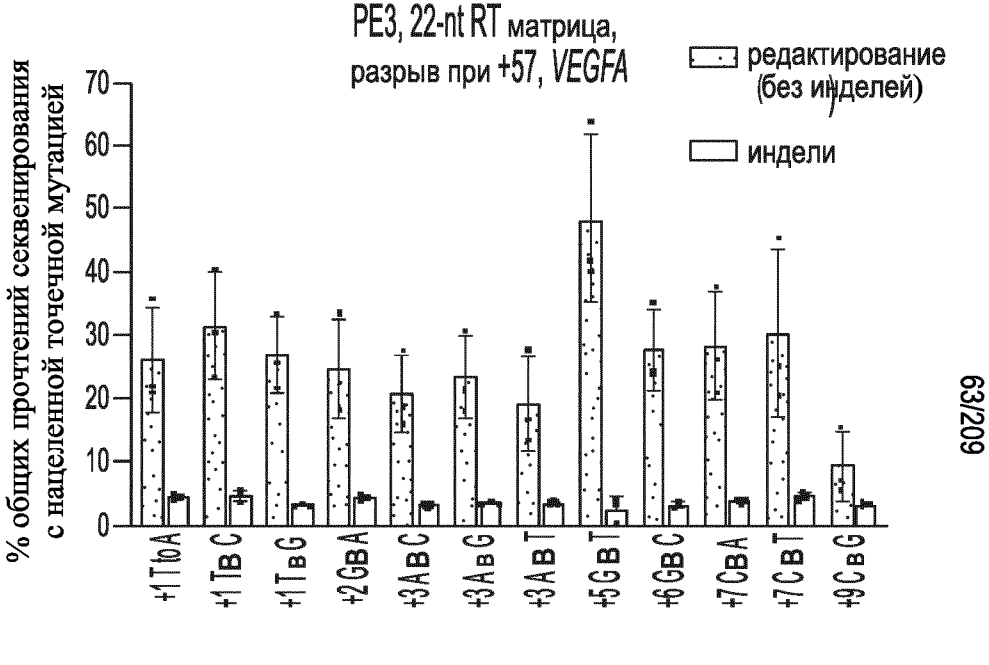
Фиг. 41D



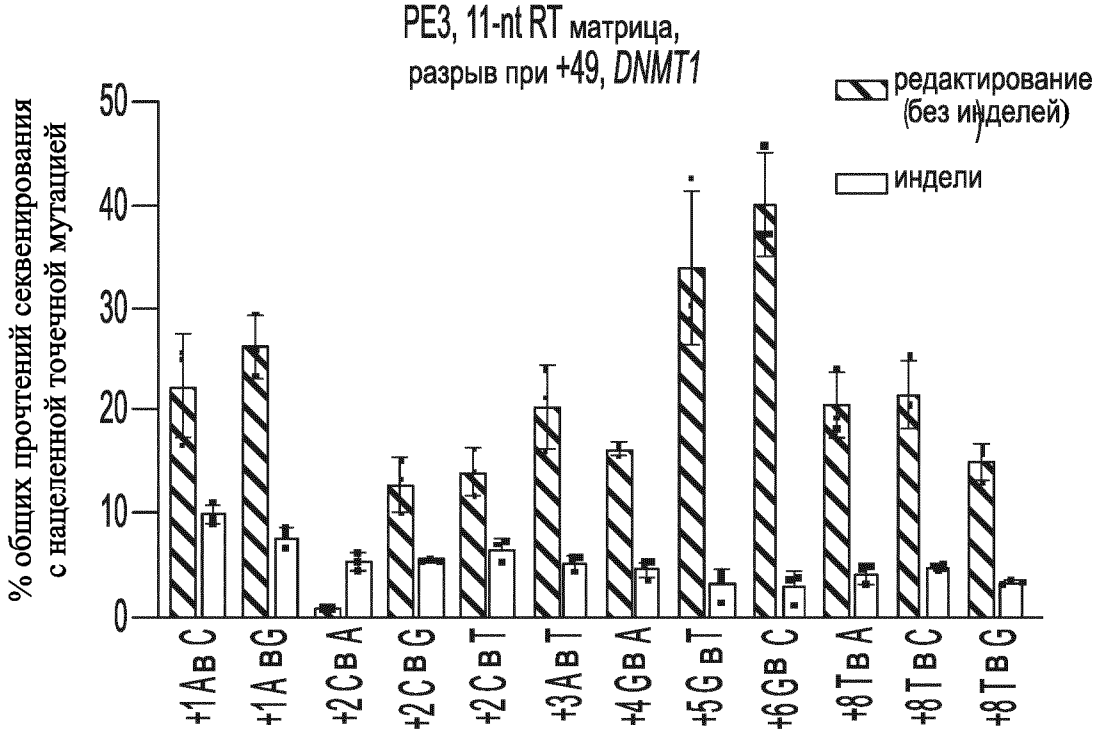
Фиг. 41E



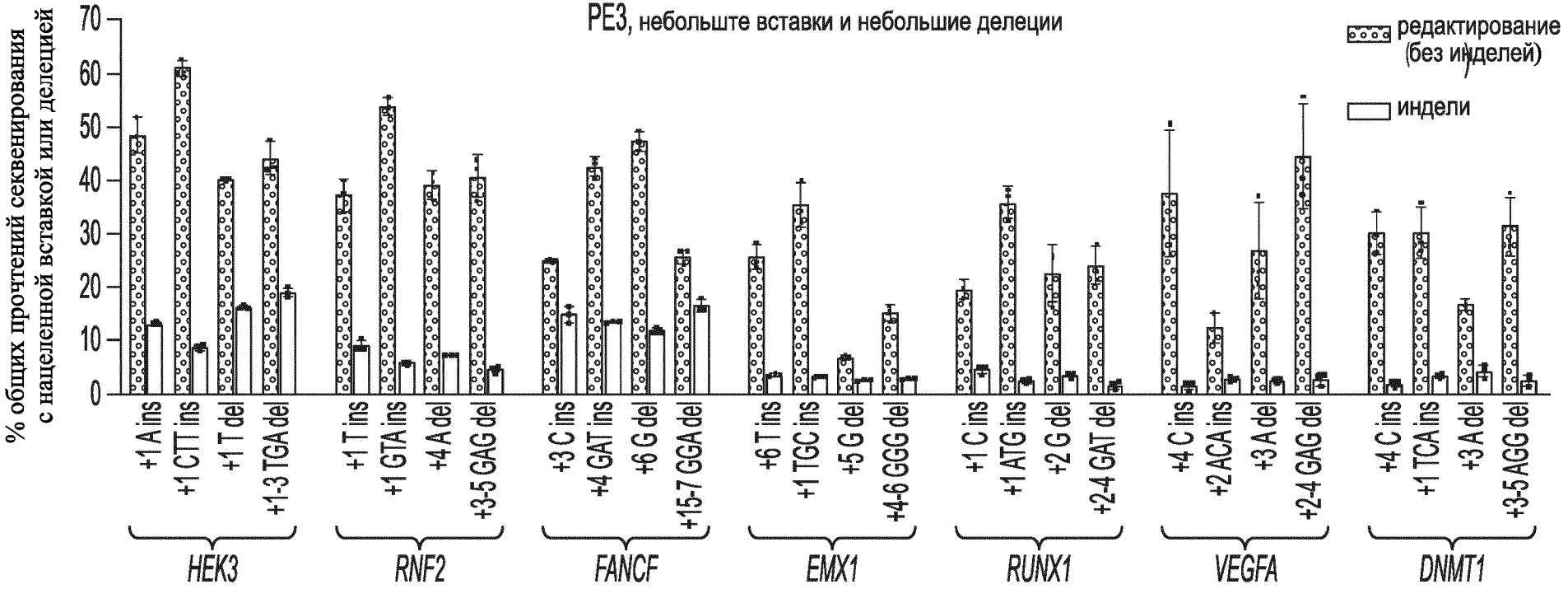
Фиг. 41F



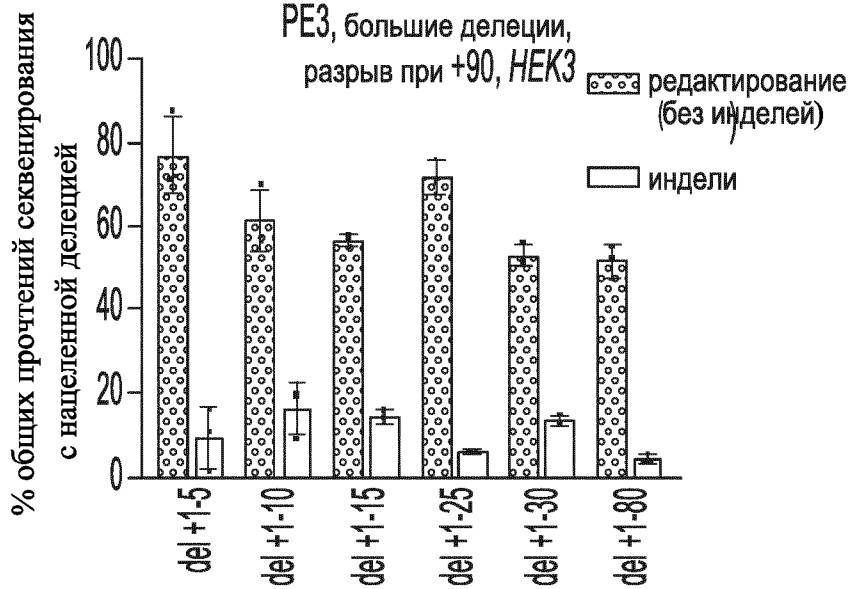
Фиг. 41G


PE3, 11-nt RT матрица,
разрыв при +49, DNMT1
редактирование
(без инделей)
индели
% общих прочтений секвенирования
с нацеленной точечной мутацией
50
40
30
20
10
0
+1A в C
+1A в G
+2C в A
+2C в G
+2C в T
+3A в T
+4G в A
+5G в T
+6G в C
+8T в A
+8T в C
+8T в G

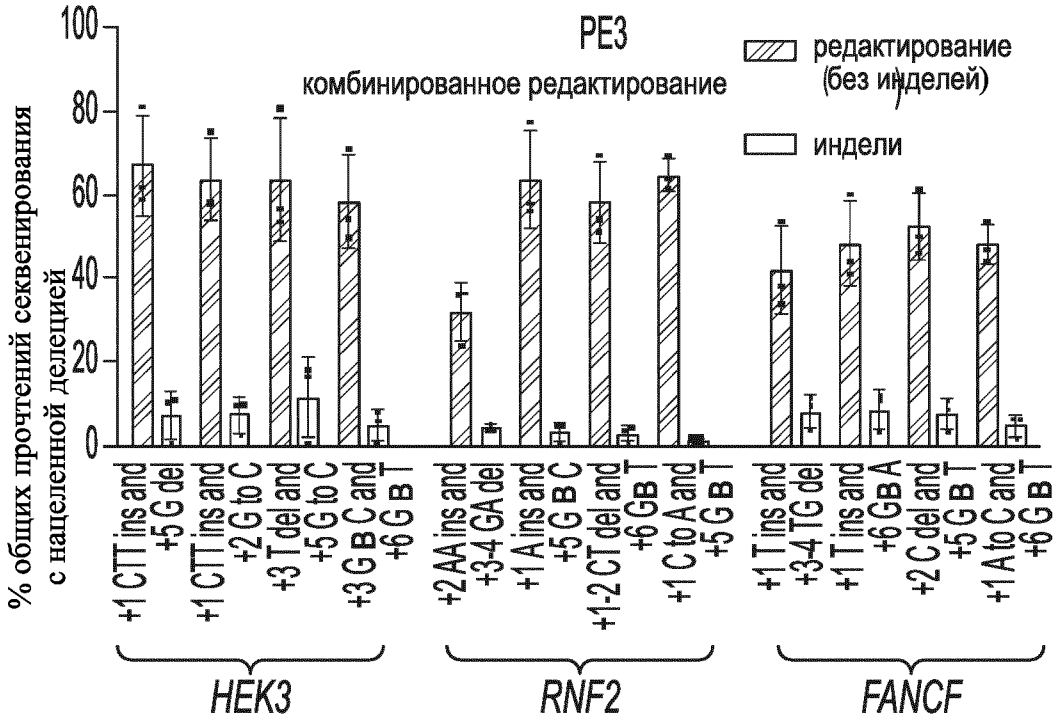
Фиг. 41H


PE3, небольште вставки и небольшие делеции
редактирование (без инделей)
индели
% общих прочтений секвенирования с нацеленной вставкой или делецией
0
10
20
30
40
50
60
70
+1 A ins
+1 CTT ins
+1 T del
+1-3 TGA del
+1 T ins
+1 GTA ins
+4 A del
+3-5 GAG del
+3 C ins
+4 GAT ins
+6 G del
+15-7 GGA del
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
+1 C ins
+1 ATG ins
+2 G del
+2-4 GAT del
+4 C ins
+2 ACA ins
+3 A del
+2-4 GAG del
+4 C ins
+1 TCA ins
+3 A del
+3-5 AGG del
HEK3
RNF2
FANCF
EMX1
RUNX1
VEGFA
DNMT1

Фиг. 41I



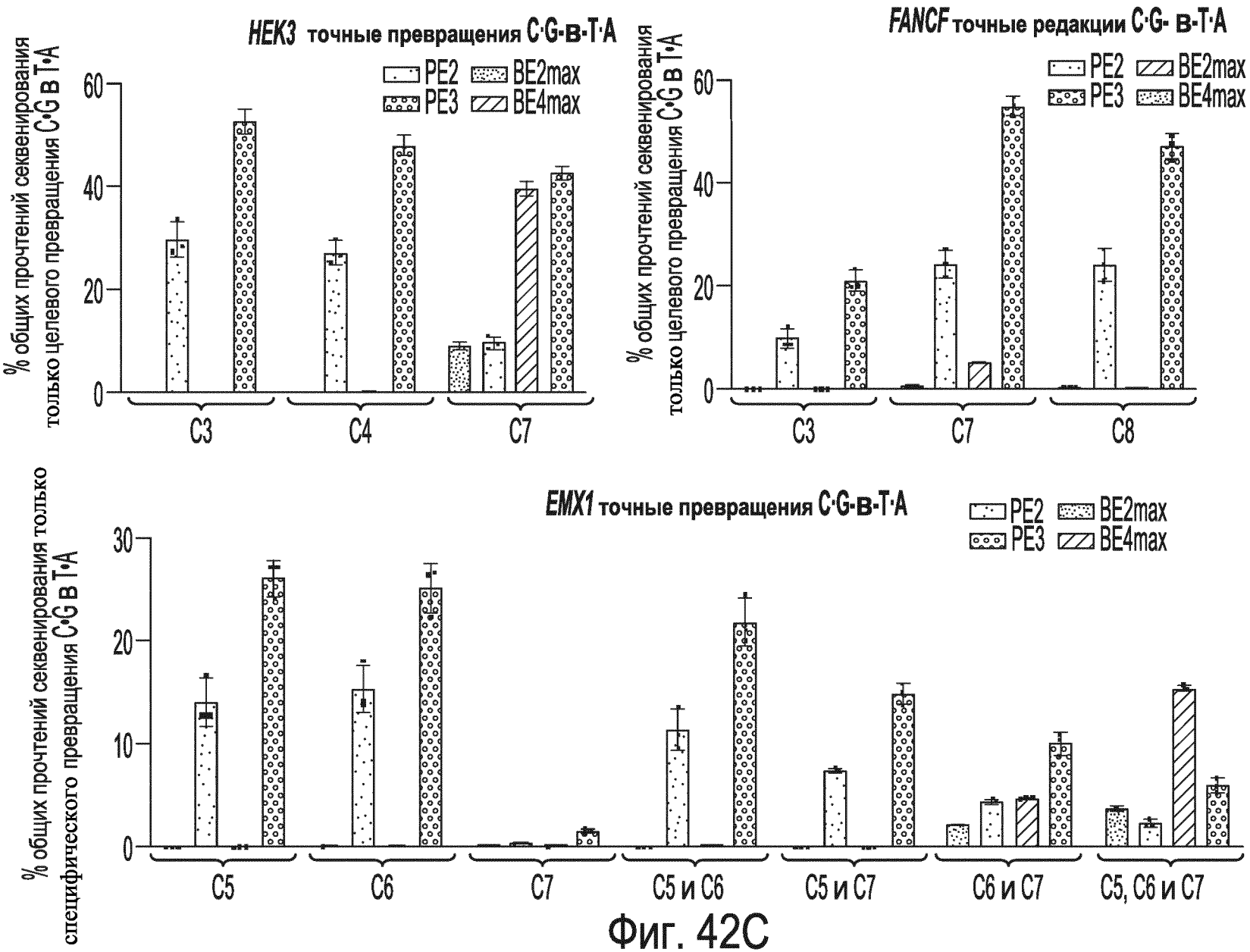
Фиг. 41J



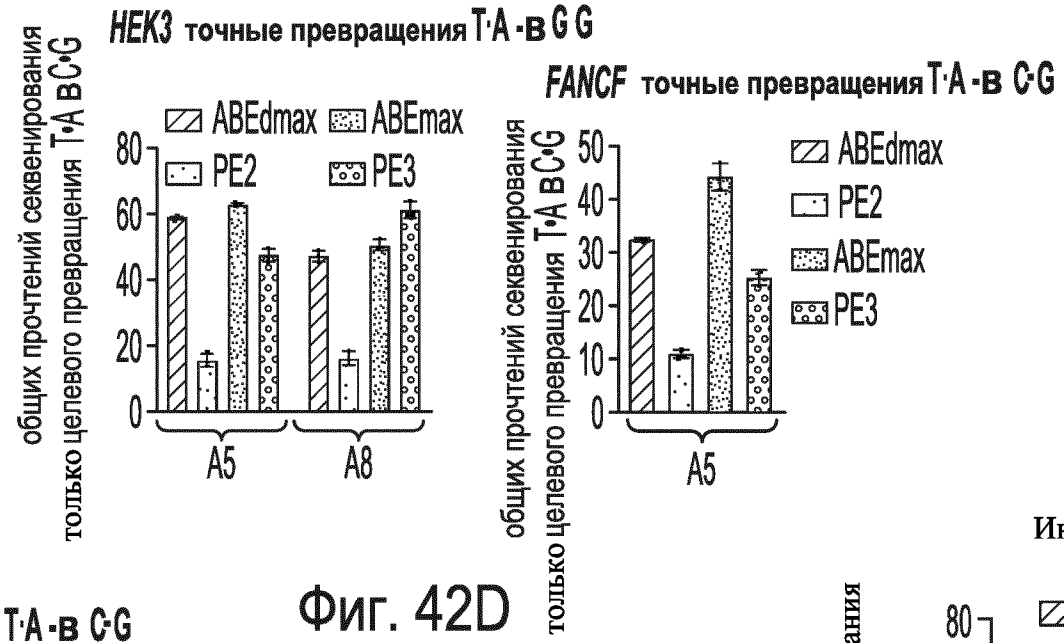
Фиг. 41K

Фиг. 42A

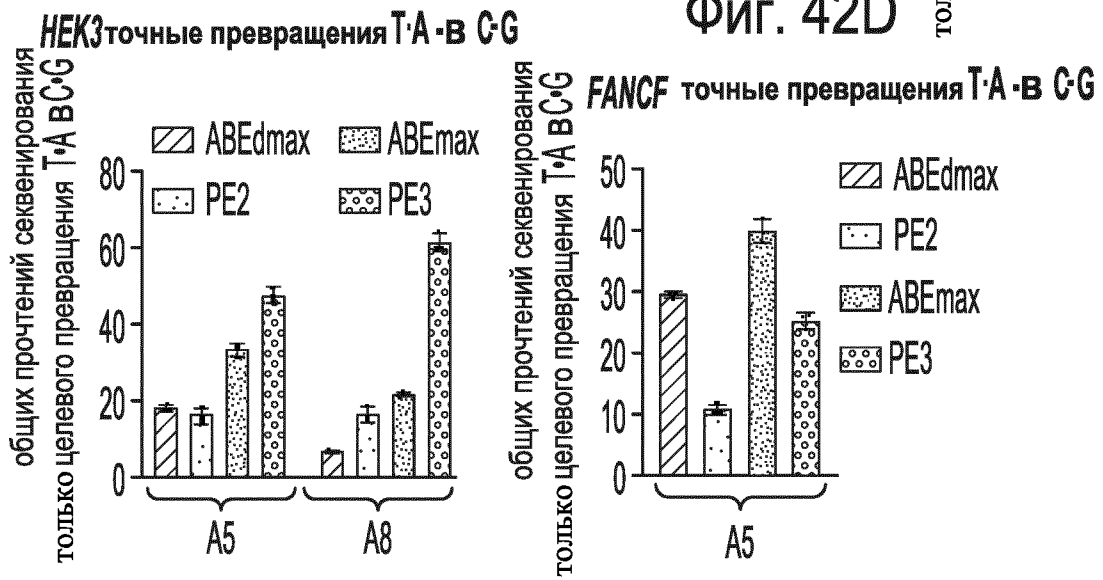
Фиг. 42B


HEK3 точные превращения C·G-в-T·A
PE2
PE3
BE2max
BE4max
% общих прочтений секвенирования только целевого превращения C•G в T•A
60
40
20
0
C3
C4
C7
FANCF точные редакции C·G- в-T·A
PE2
PE3
BE2max
BE4max
% общих прочтений секвенирования только целевого превращения C•G в T•A
60
40
20
0
C3
C7
C8
EMX1 точные превращения C·G-в-T·A
PE2
PE3
BE2max
BE4max
% общих прочтений секвенирования только специфического превращения C•G в T•A
30
20
10
0
C5
C6
C7
C5 и C6
C5 и C7
C6 и C7
C5, C6 и C7

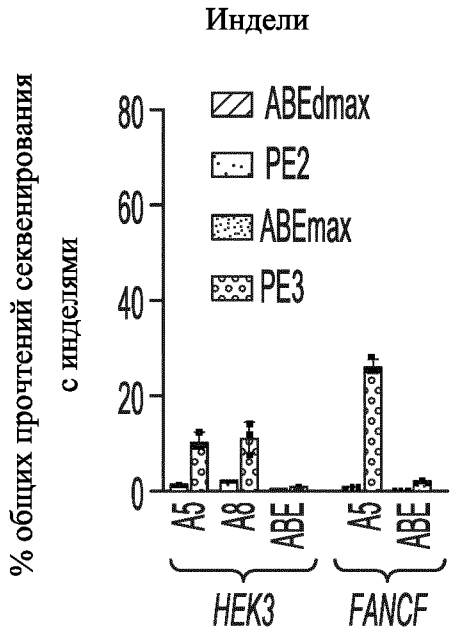
Фиг. 42C



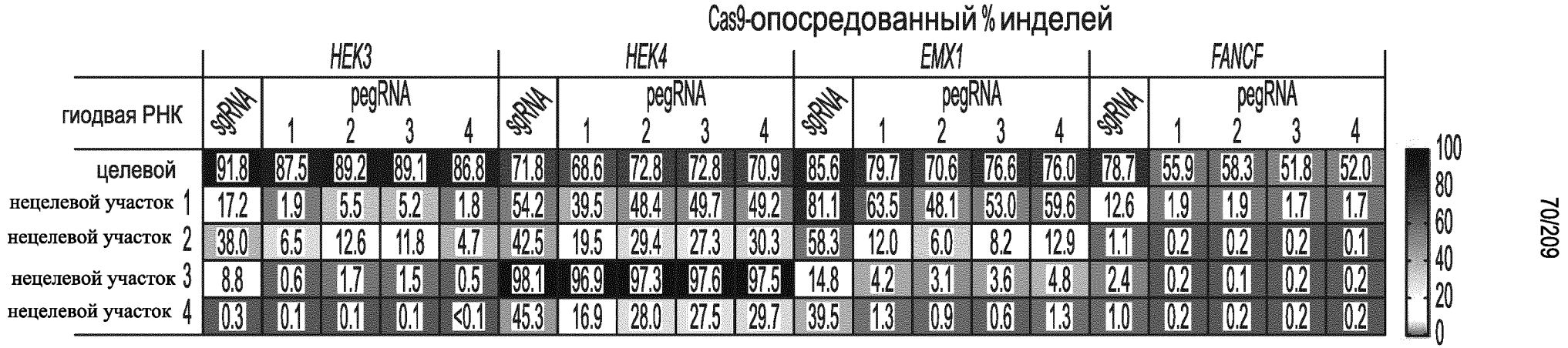
Фиг. 42D



Фиг. 42E



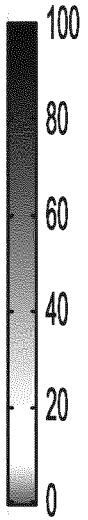
Фиг. 42F


Cas9-опосредованный % инделей
HEK3
HEK4
EMX1
FANCF
гиодвая РНК
sgRNA
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

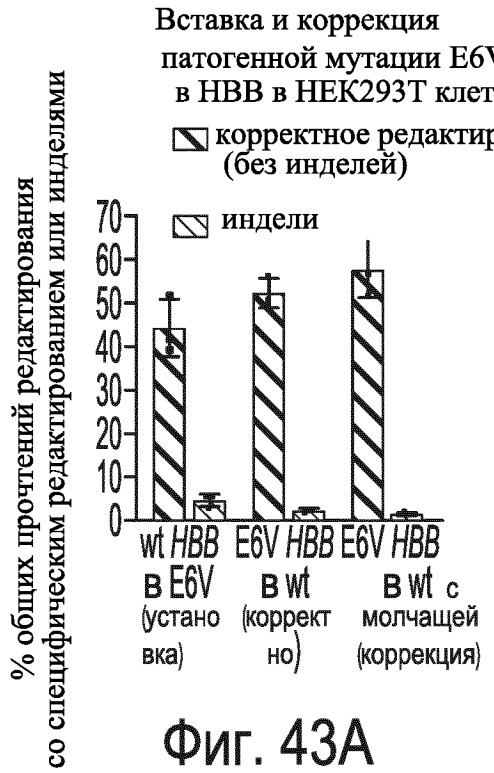
Фиг. 42G

PE-опосредованный % редактирования и %инделей (в паре)

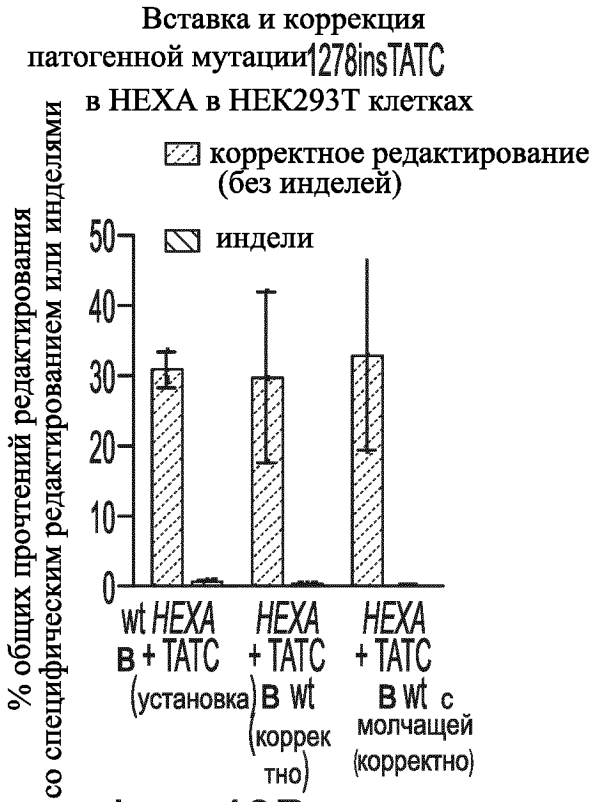
гидовая РНК	*HEK3* (PE3) pegRNA 1	*HEK3* (PE3) pegRNA 2	*HEK3* (PE3) pegRNA 3	*HEK3* (PE3) pegRNA 4	*HEK4* (PE2) pegRNA 1	*HEK4* (PE2) pegRNA 2	*HEK4* (PE2) pegRNA 3	*HEK4* (PE2) pegRNA 4	*EMX1* (PE3) pegRNA 1	*EMX1* (PE3) pegRNA 2	*EMX1* (PE3) pegRNA 3	*EMX1* (PE3) pegRNA 4	*FANCF* (PE3) pegRNA 1	*FANCF* (PE3) pegRNA 2	*FANCF* (PE3) pegRNA 3	*FANCF* (PE3) pegRNA 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)



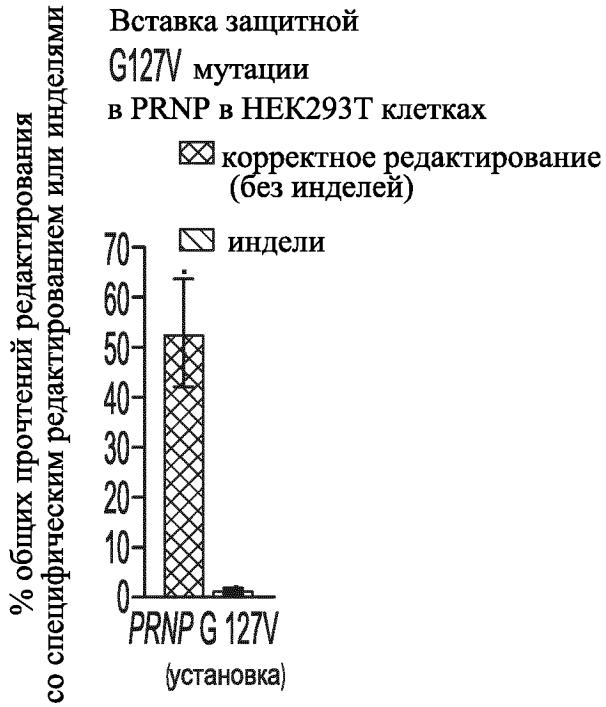
Фиг. 42H



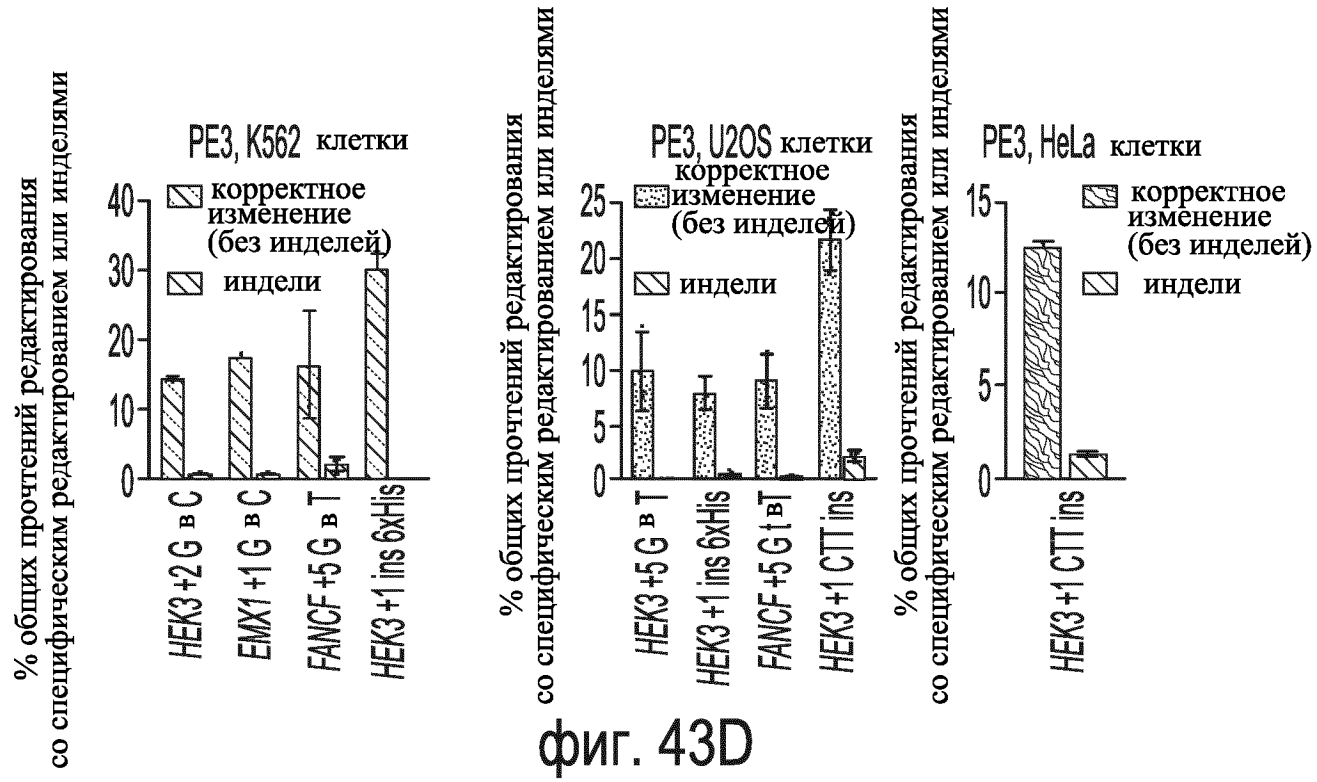
Фиг. 43A



Фиг. 43B



Фиг. 43C

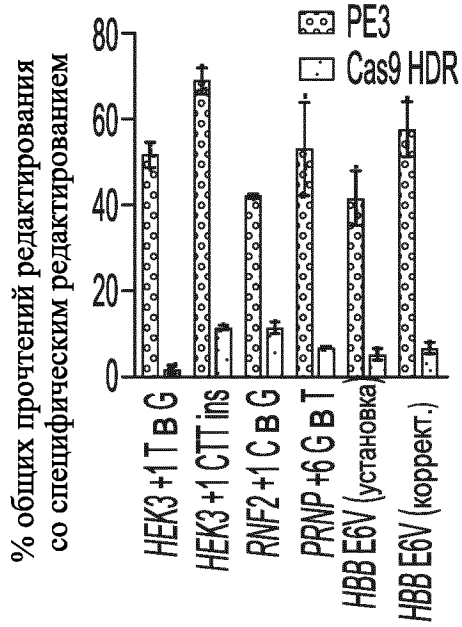


фиг. 43D



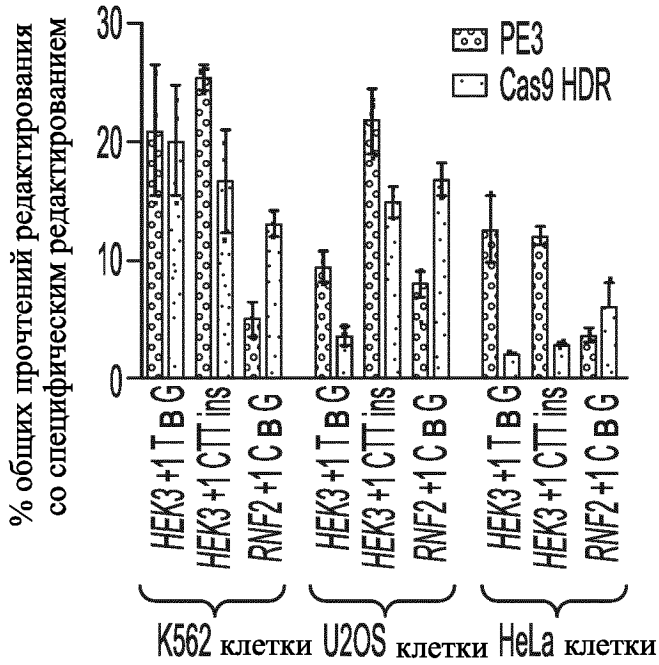
фиг. 43E

Корректное редактирование без инделей для PE3 и Cas9 HDR в HEK293T клетках

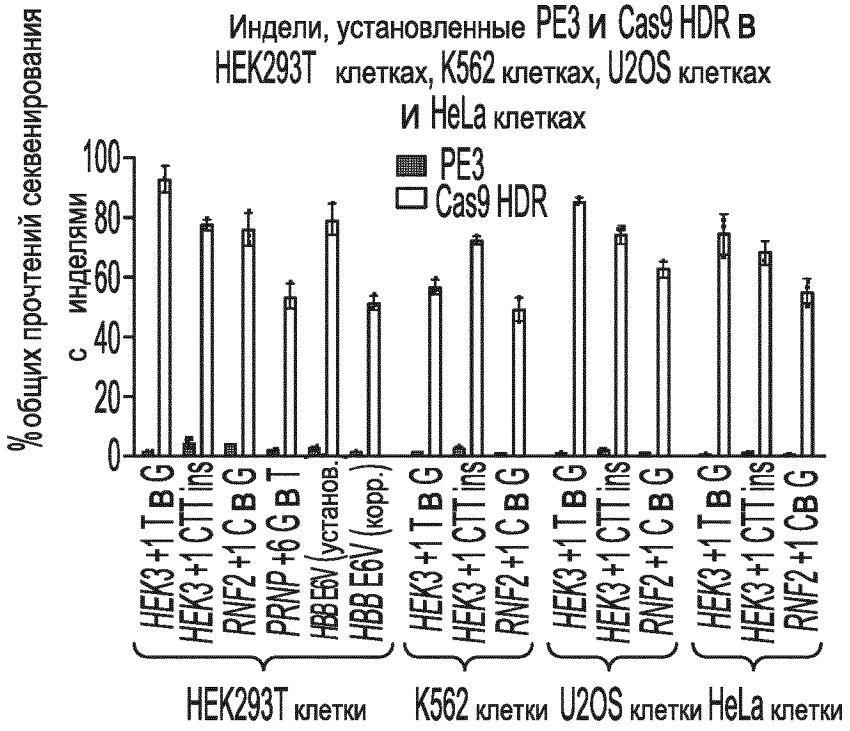


Фиг. 43F

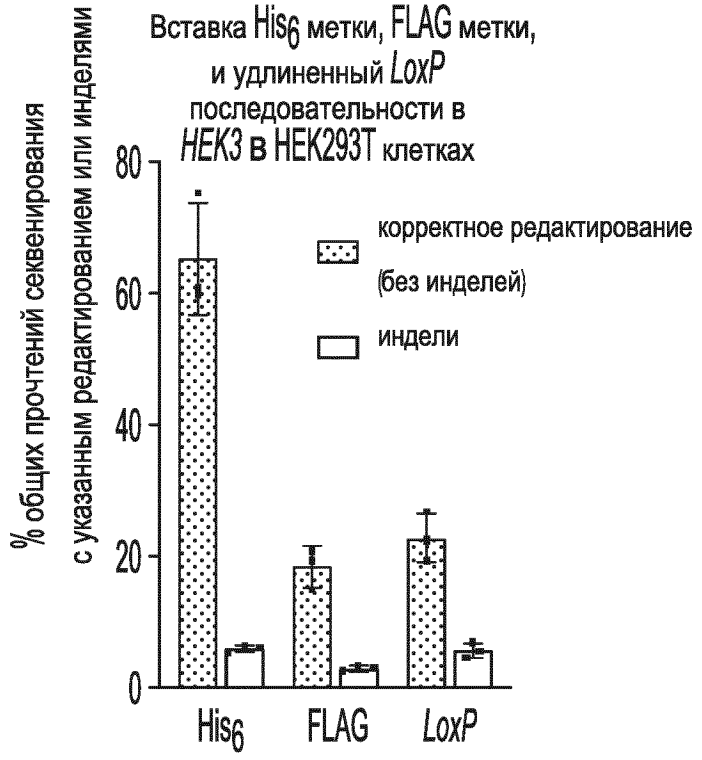
Корректное редактирование без инделей для PE3 и Cas9 HDR в K562 клетках, U2OS клетках и HeLaклетках и



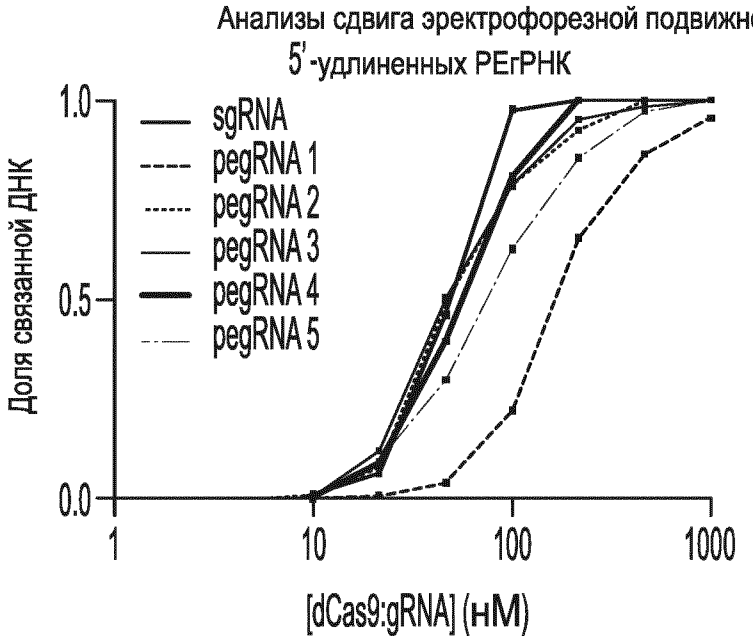
Фиг. 43G



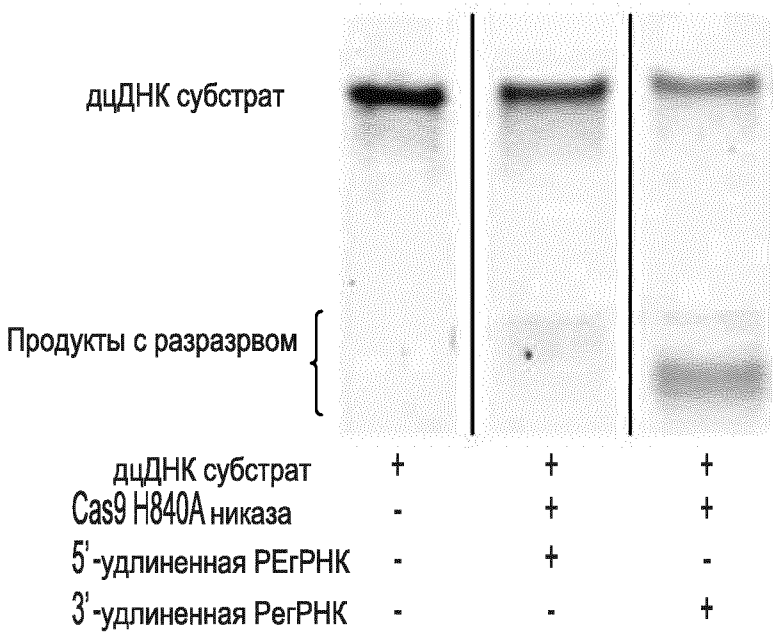
Фиг. 43H



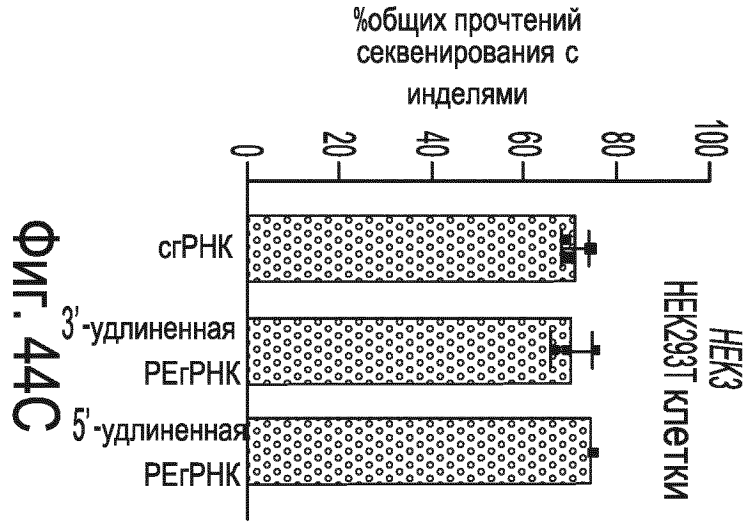
Фиг. 43I



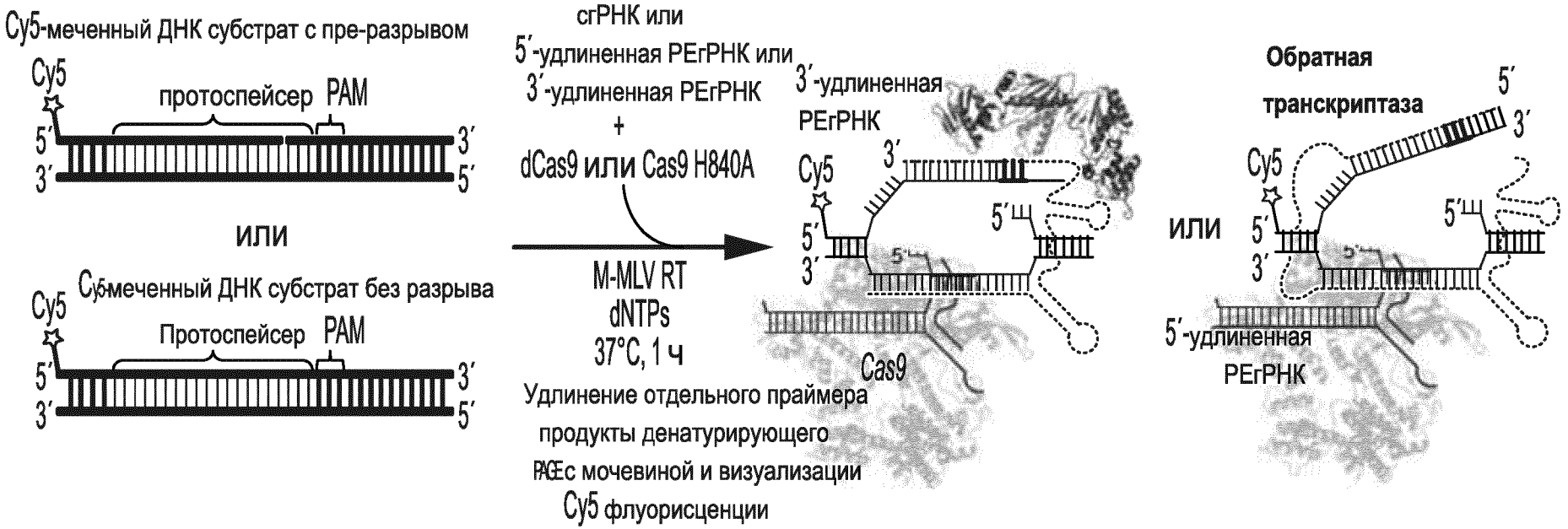
Фиг. 44A



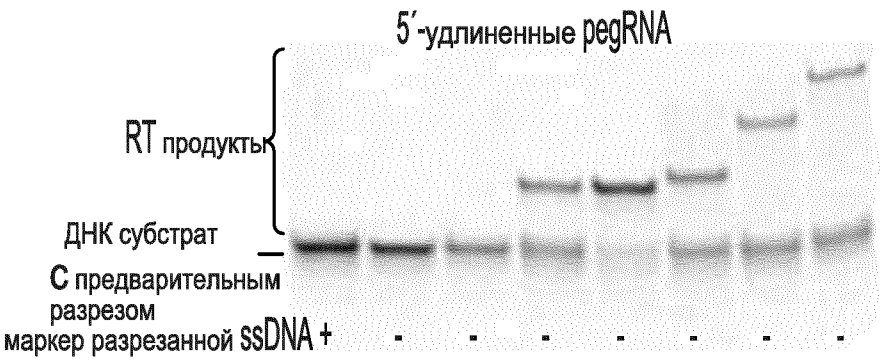
Фиг. 44B


%общих прочтений секвенирования с инделями
0
20
40
60
80
100
сгРНК
3'-удлиненная РЕгРНК
5'-удлиненная РЕгРНК
HEK3
HEK293T клетки

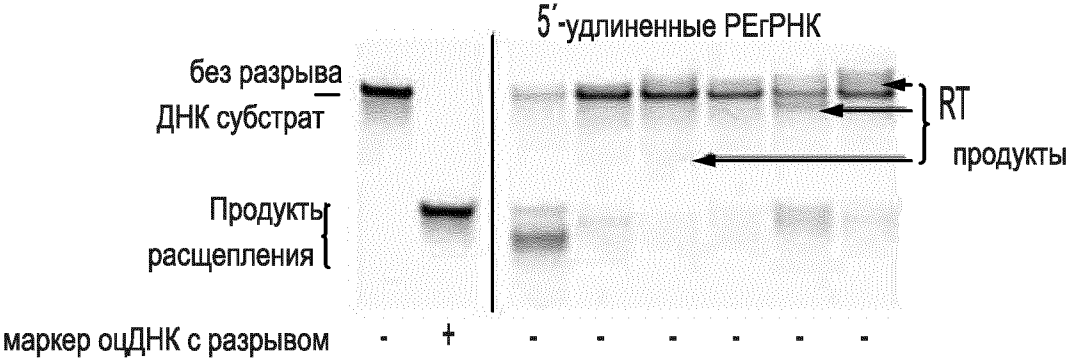
Фиг. 44C


Cy5-меченный ДНК субстрат с пре-разрывом
Cy5
протоспейсер
PAM
5′
3′
3′
5′
или
Cy5-меченный ДНК субстрат без разрыва
Cy5
Протоспейсер
PAM
5′
3′
3′
5′
crРНК или
5′-удлиненная PEгРНК или
3′-удлиненная PEгРНК
+
dCas9 или Cas9 H840A
M-MLV RT
dNTPs
37°C, 1 ч
Удлинение отдельного праймера
продукты денатурирующего
PAGEс мочевиной и визуализации
Cy5 флуоресценции
3′-удлиненная
PEгРНК
Cy5
3′
5′
3′
5′
5′
Cas9
Обратная
транскриптаза
5′
3′
Cy5
5′
3′
5′
5′
или
5′-удлиненная
PEгРНК

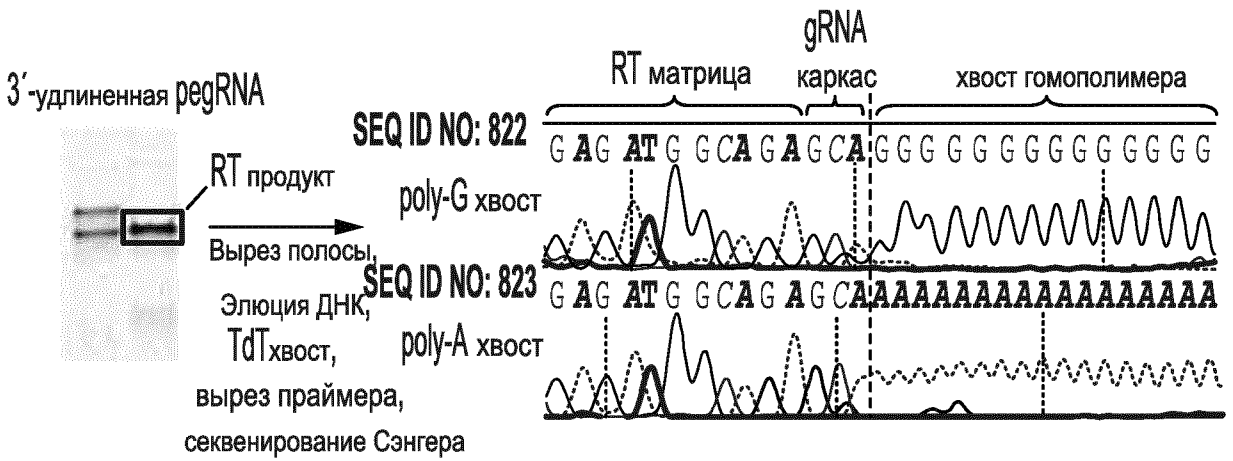
Фиг. 44D



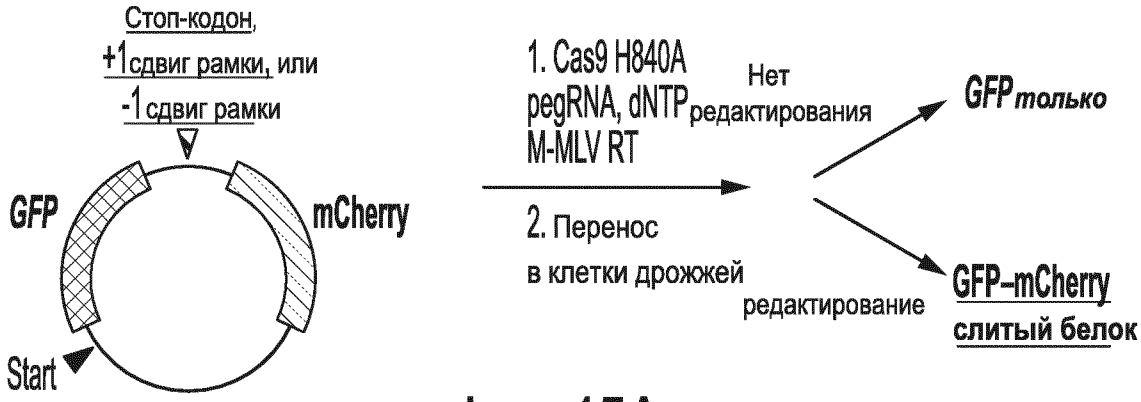
Фиг. 44E



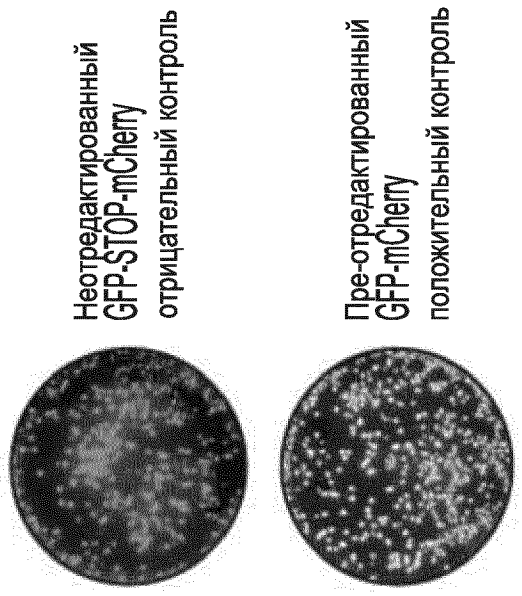
Фиг. 44F



Фиг. 44G

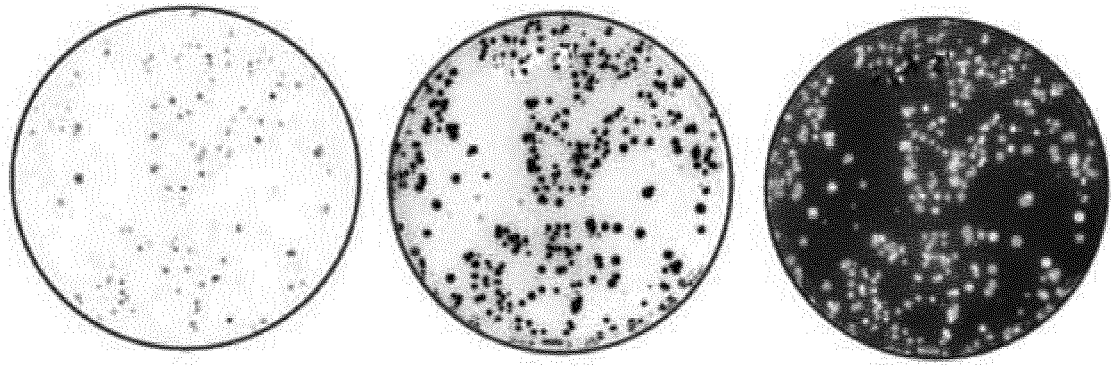


Фиг. 45A


Неотредактированный
GFP-STOP-mCherry
отрицательный контроль
Пре-отредактированный
GFP-mCherry
положительный контроль

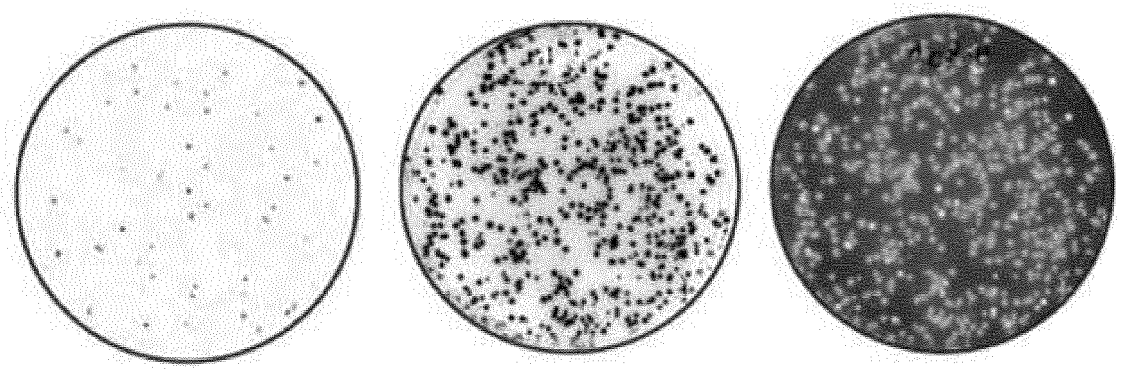
Фиг. 45B

3'-удлиненная pegRNA , коррекция стоп-кодона (T•A в A•T)



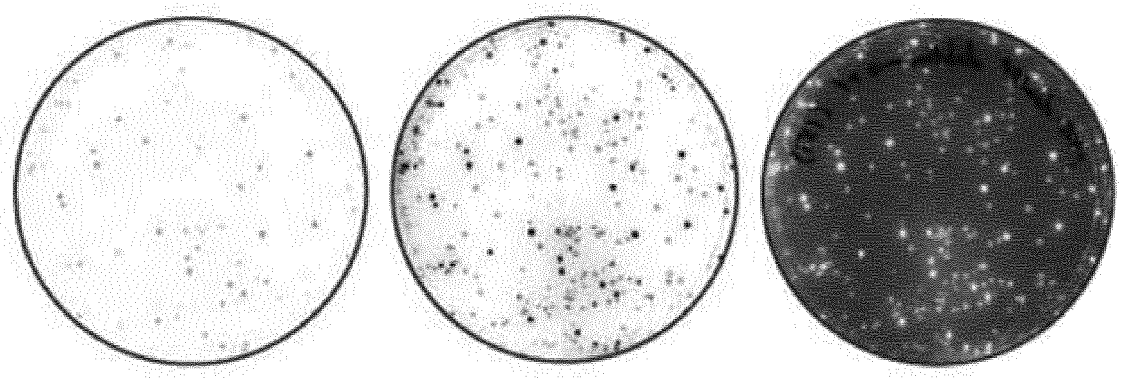
Фиг. 45C

5'-удлиненная pegRNA, коррекция стоп-кодона (T•A в A•T)



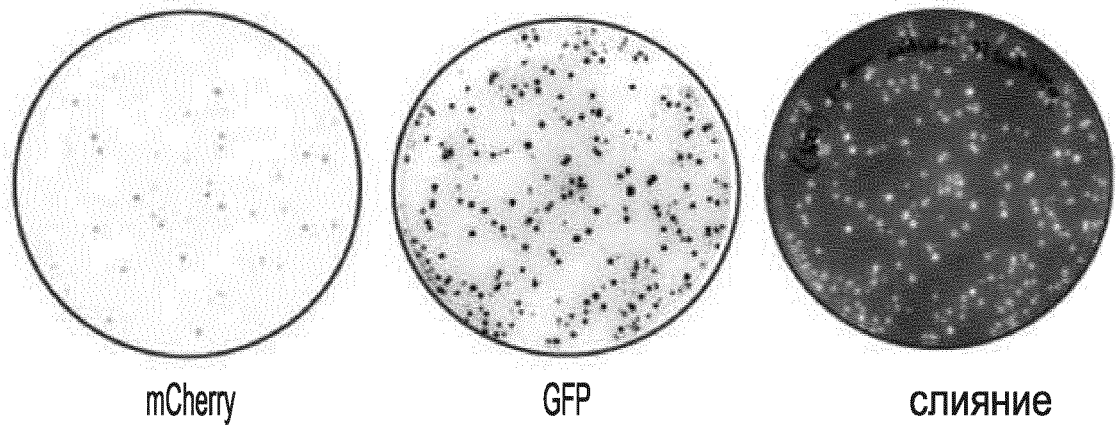
Фиг. 45D

3'-удлиненная , коррекйия pegRNA +1 сдвига рамки (1-п.о. делеция)

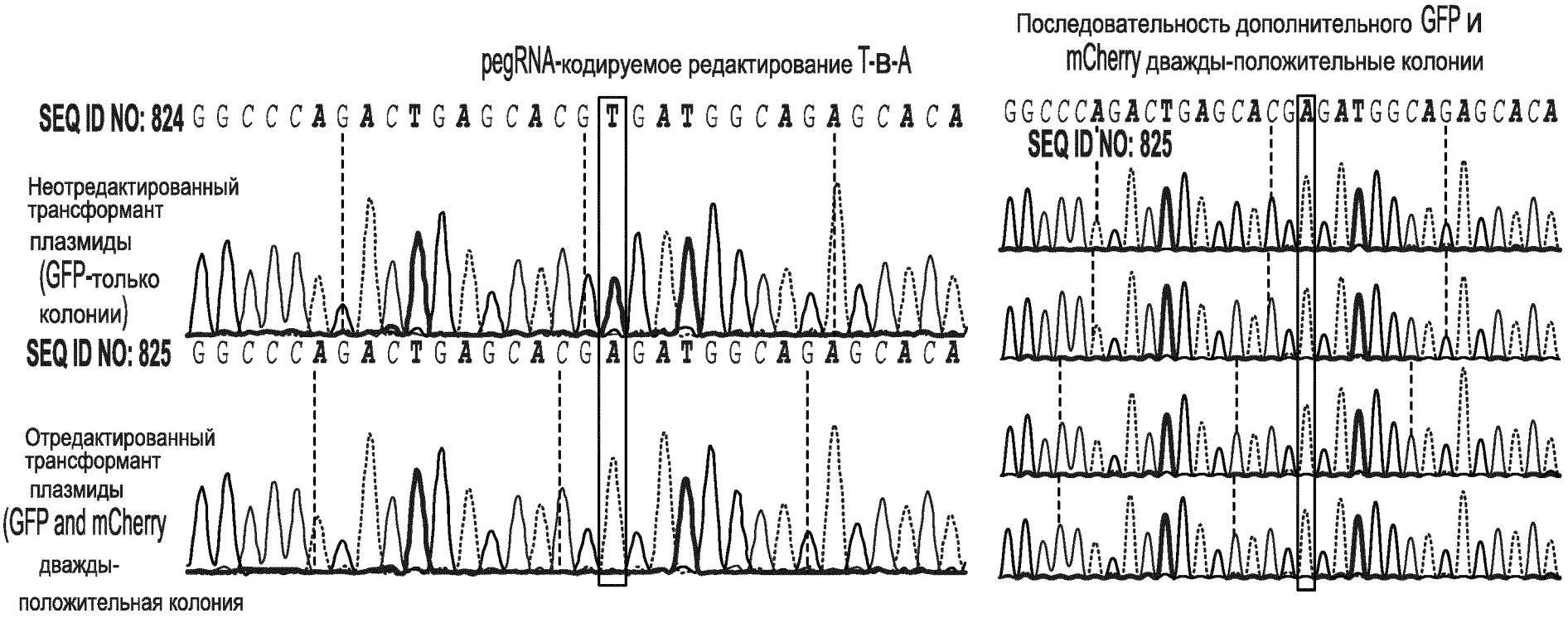


Фиг. 45E

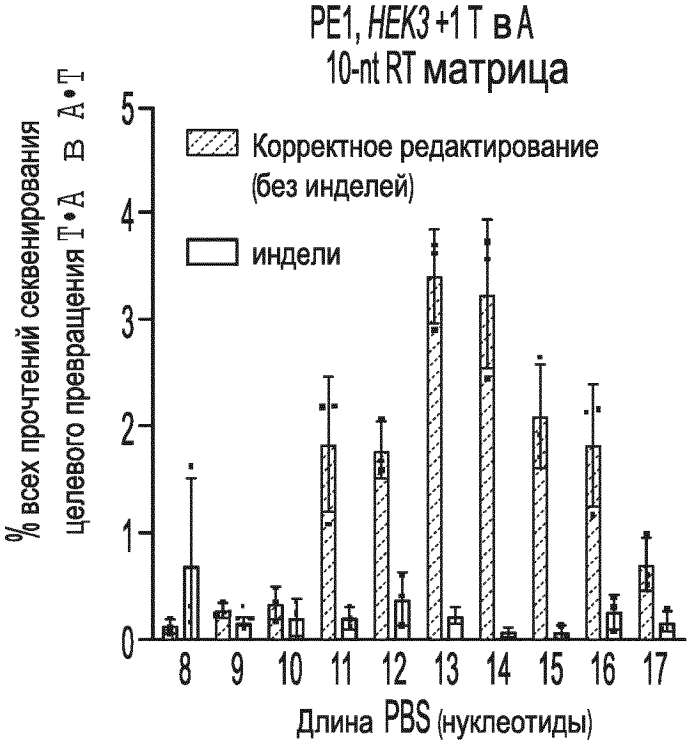
3'-удлиненная pegRNA , коррекция -1 сдвига рамки (1-п.о. вставка)



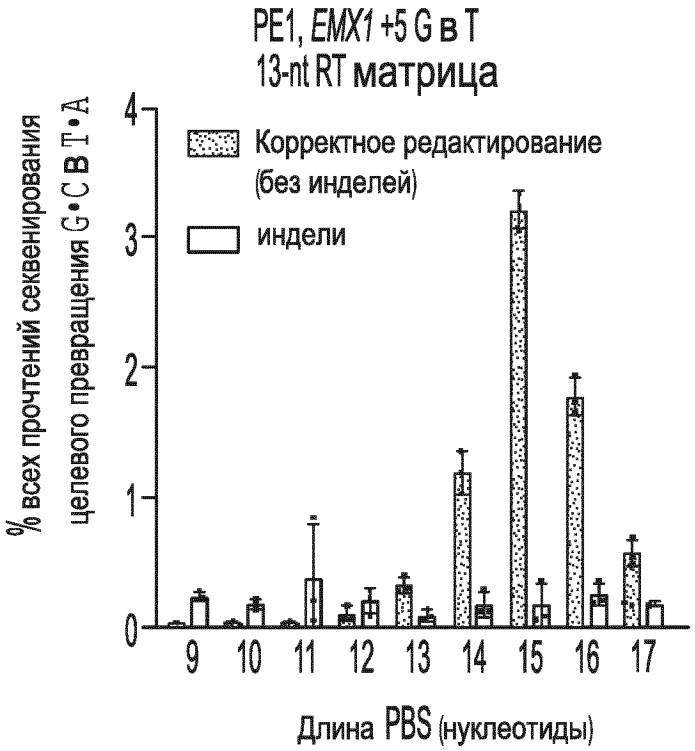
Фиг. 45F


pegRNA-кодируемое редактирование T-в-A
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
Неотредактированный трансформант плазмиды (GFP-только колонии)
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
Отредактированный трансформант плазмиды (GFP and mCherry дважды-положительная колония
Последовательность дополнительного GFP и mCherry дважды-положительные колонии
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

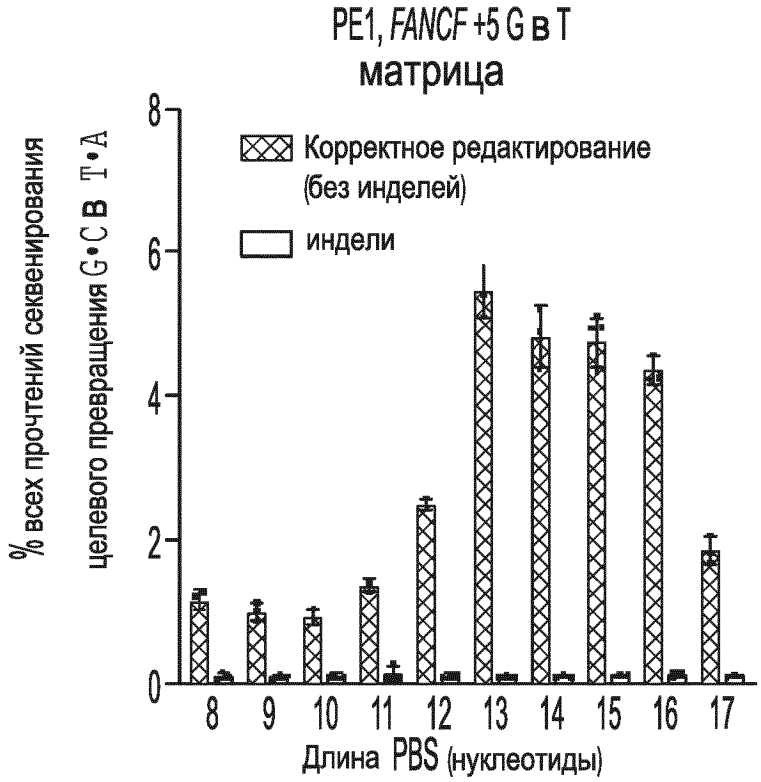
Фиг. 45G


PE1, HEK3 +1 T в A
10-nt RT матрица
% всех прочтений секвенирования целевого превращения T•A в A•T
Корректное редактирование (без инделей)
индели
0
1
2
3
4
5
8
9
10
11
12
13
14
15
16
17
Длина PBS (нуклеотиды)

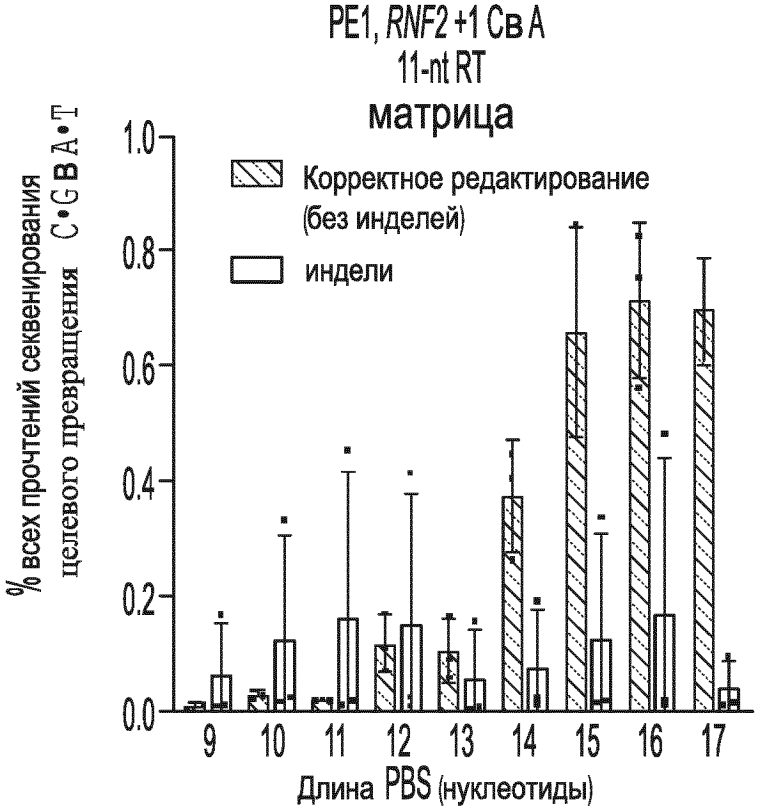
Фиг. 46A



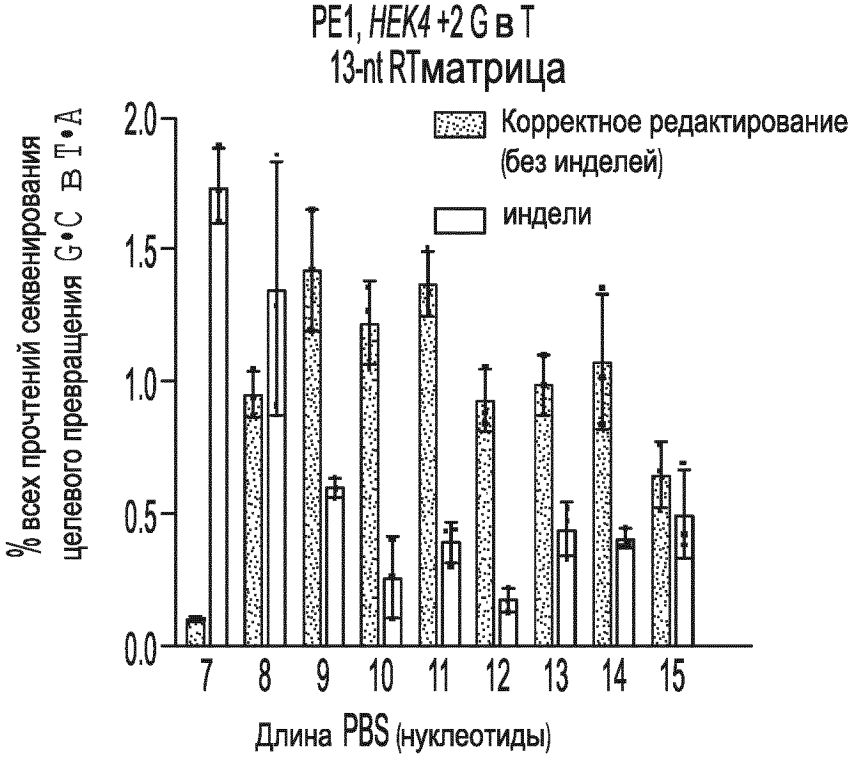
Фиг. 46B



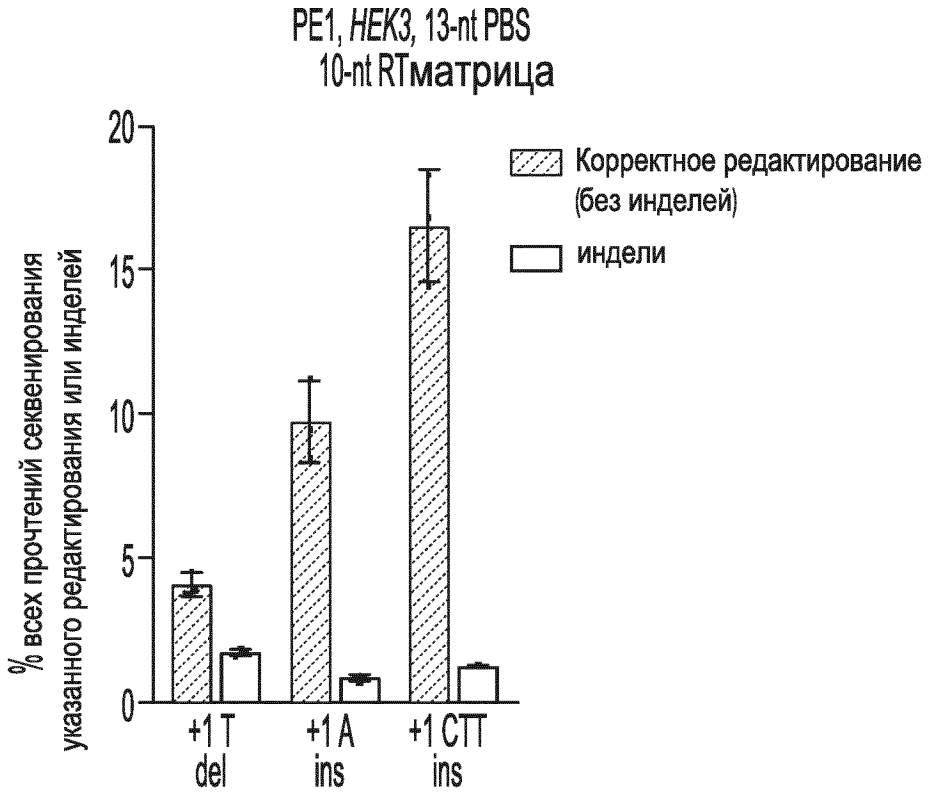
Фиг. 46C



Фиг. 46D



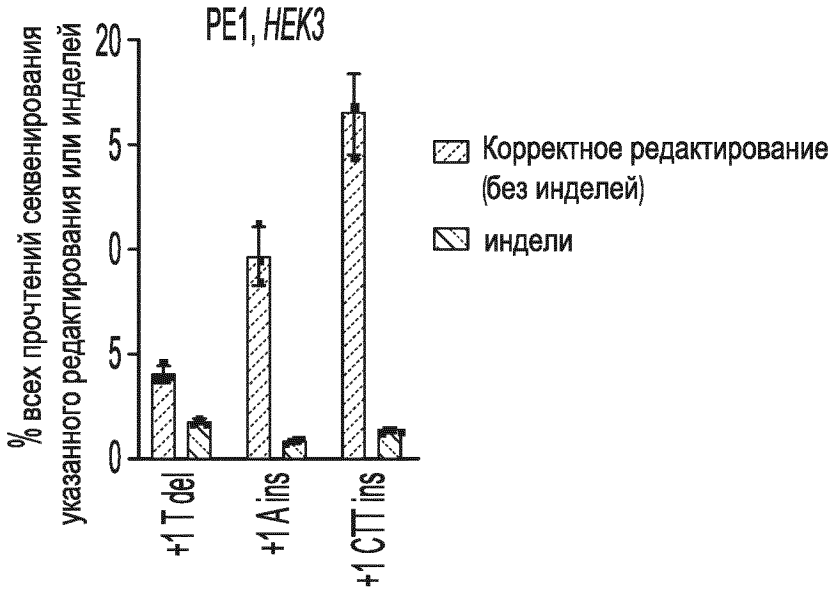
Фиг. 46E



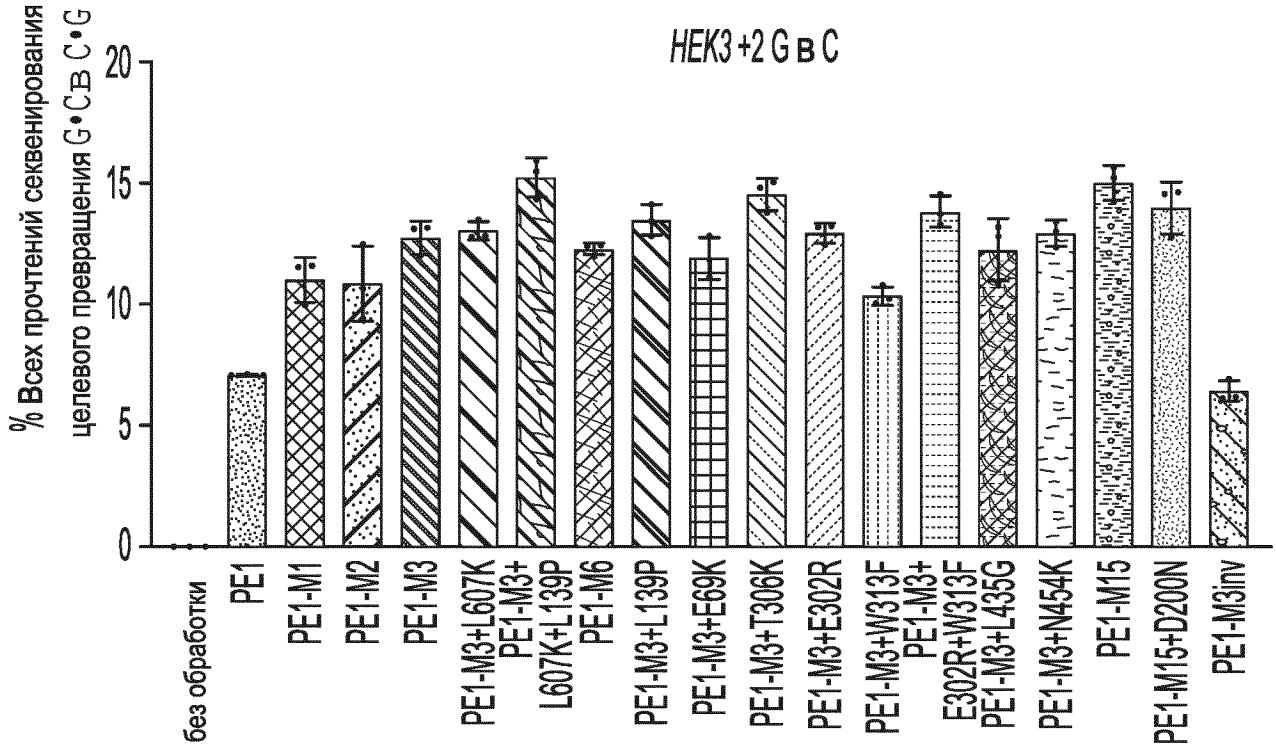
Фиг. 46F

PE1	Cas9 H840A-M-MLV RT (wt)
PE1-M1	Cas9 H840A-M-MLV RT (D200N)
PE1-M2	Cas9 H840A-M-MLV RT (D200N+L603W)
PE1-M3	Cas9 H840A-M-MLV RT (D200N+L603W+T330P)
PE1-M6	Cas9 H840A-M-MLV RT (D200N+L603W+T330P+D524G+E562Q+D583N)
PE1-M15	Cas9 H840A-M-MLV RT (P51L+S67R+E67K+T197A+H204R+E302K+F309N+W1331F+T330P+L345G+ N454K+D524G+D583N+H594Q+D653N)
PE1-M3inv	M-MLV RT (D200N+L603W+T330P)-Cas9 H840A
PE2	Cas9 H840A-M-MLV RT (D200N+L603W+T330P+T306K+W313F)

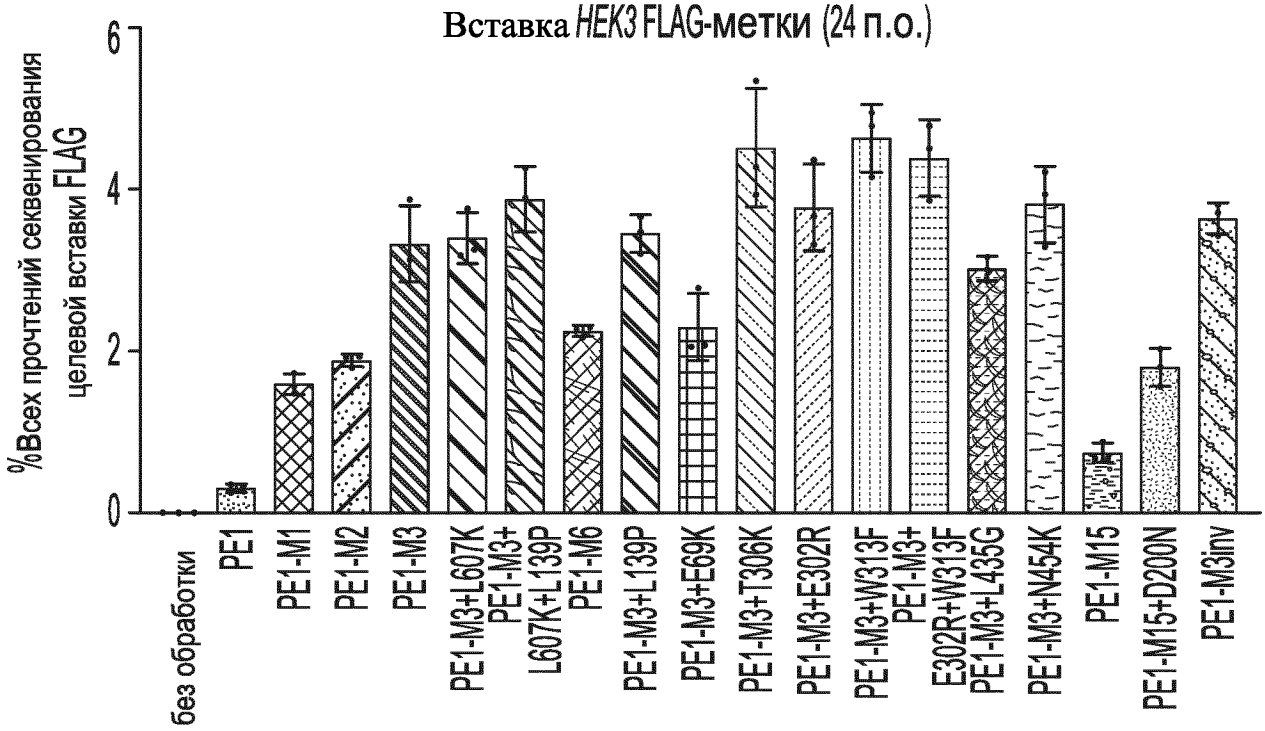
Фиг. 47A



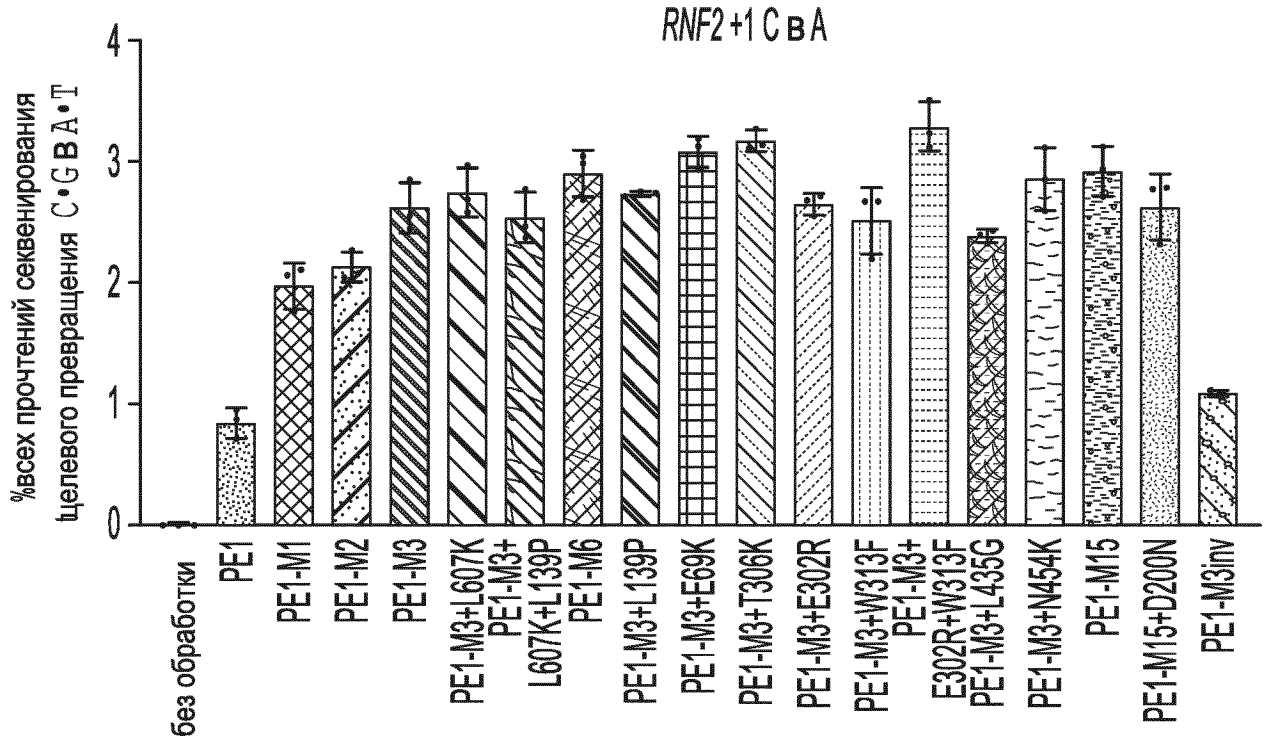
Фиг. 47B



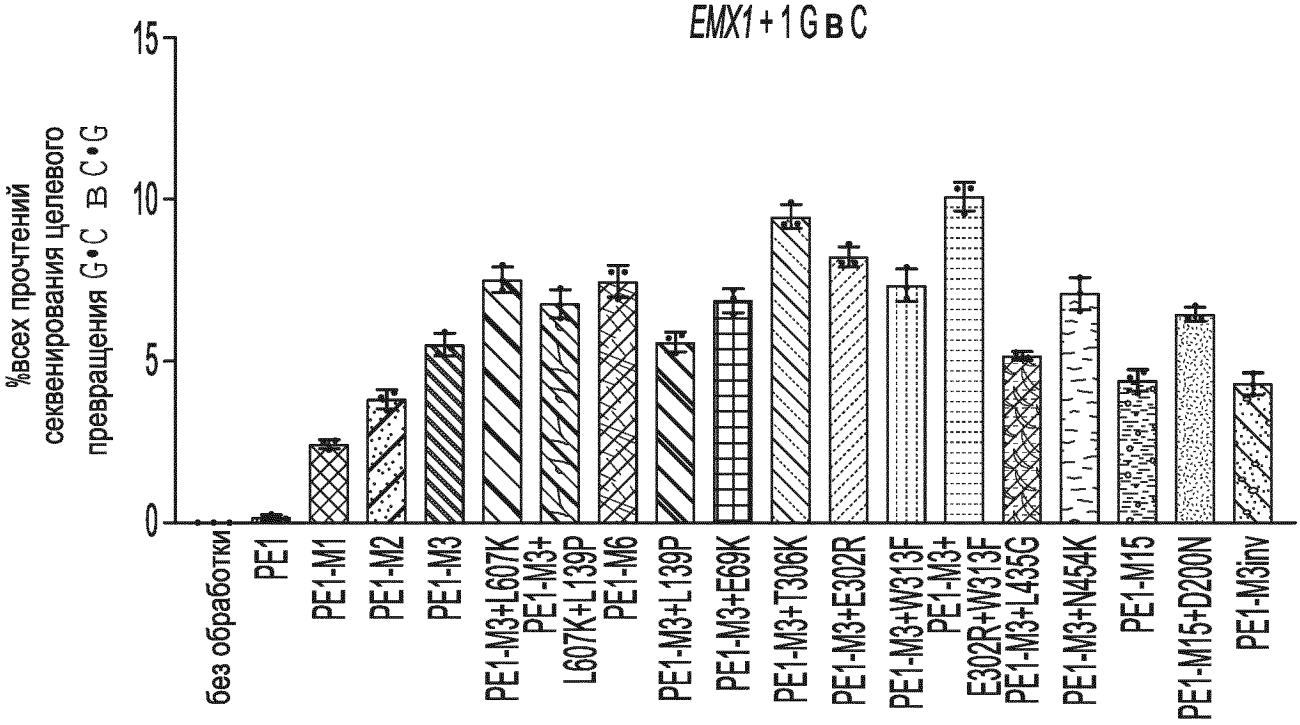
Фиг. 47C



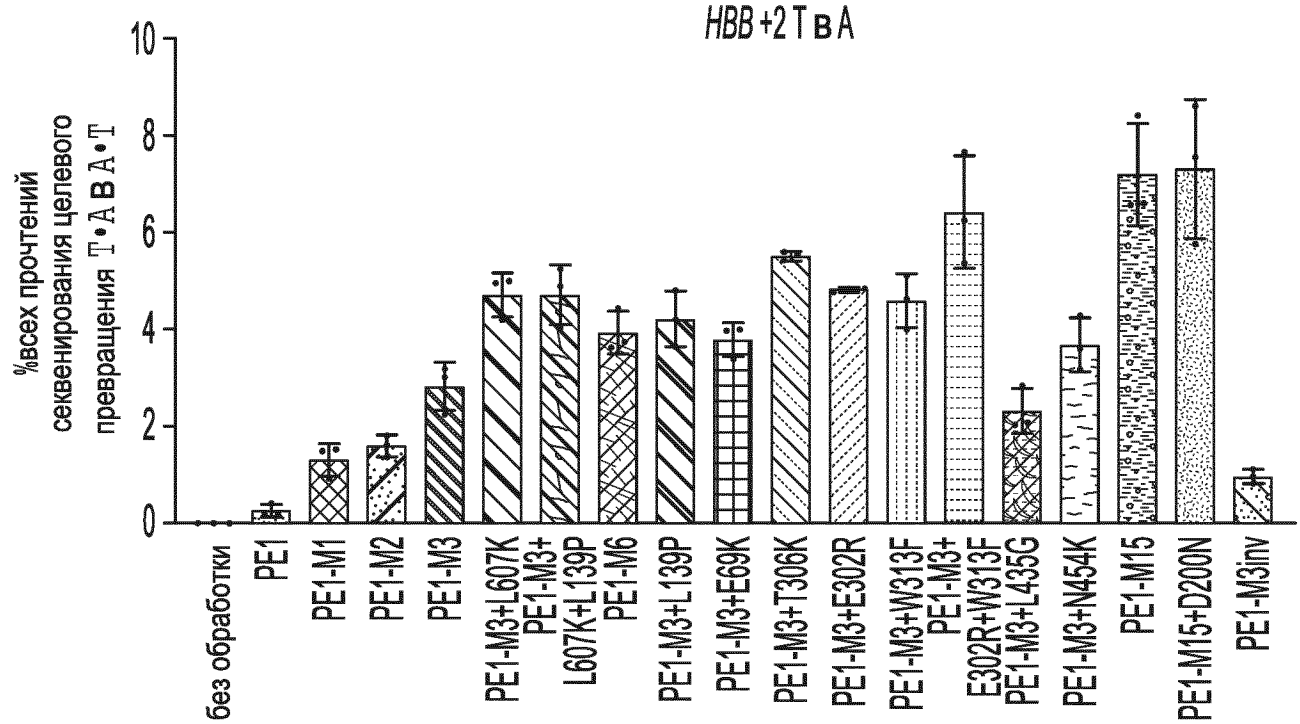
Фиг. 47D



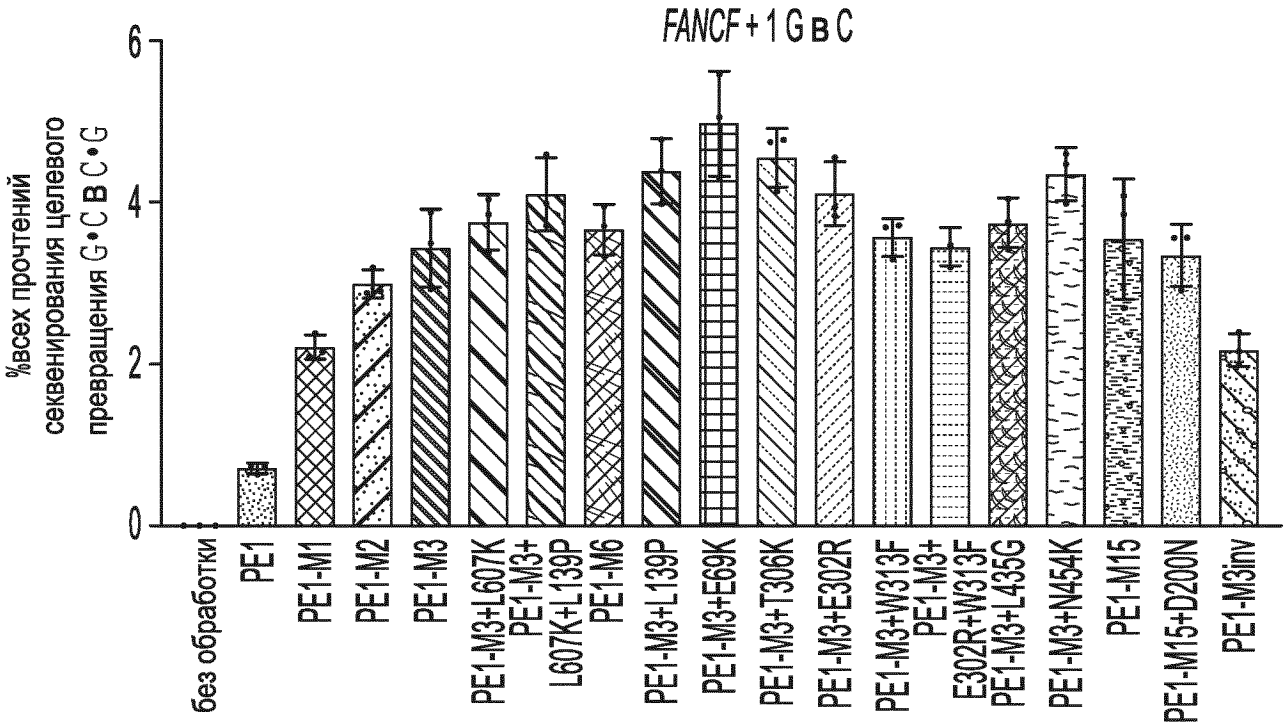
Фиг. 47E



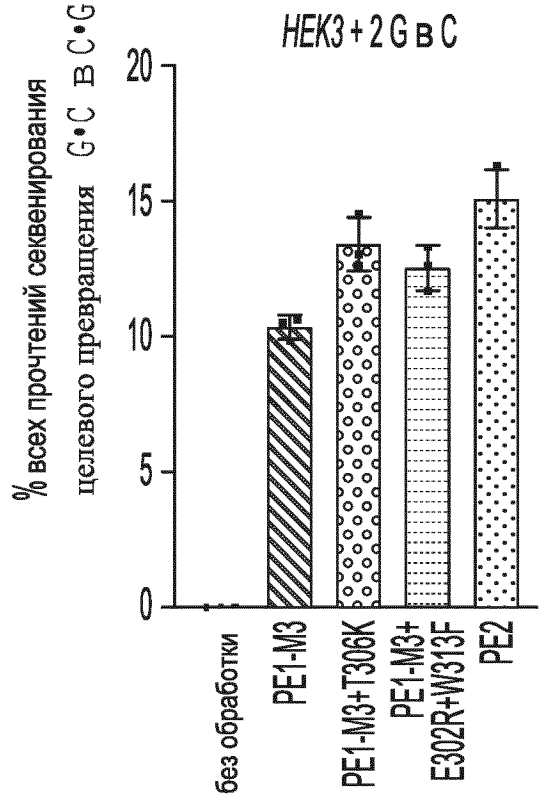
Фиг. 47F



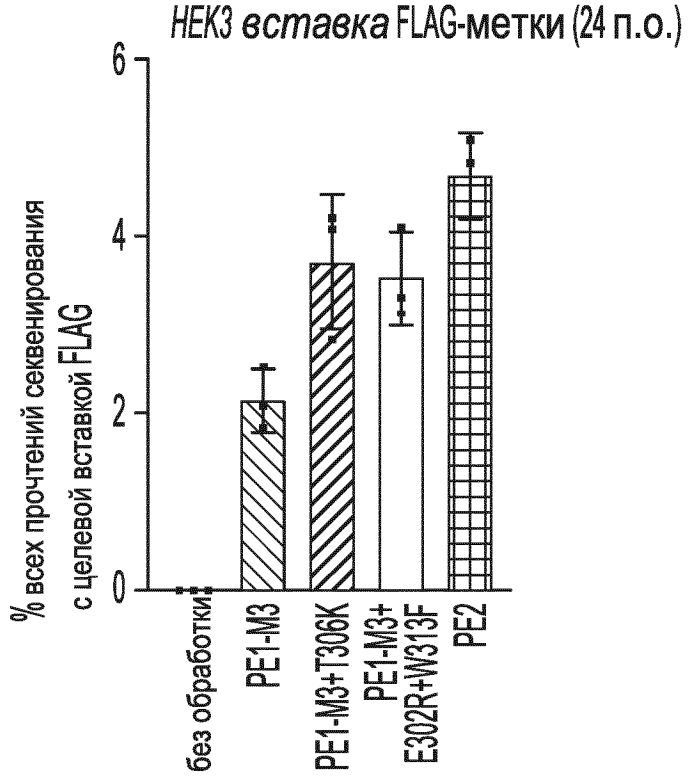
Фиг. 47G



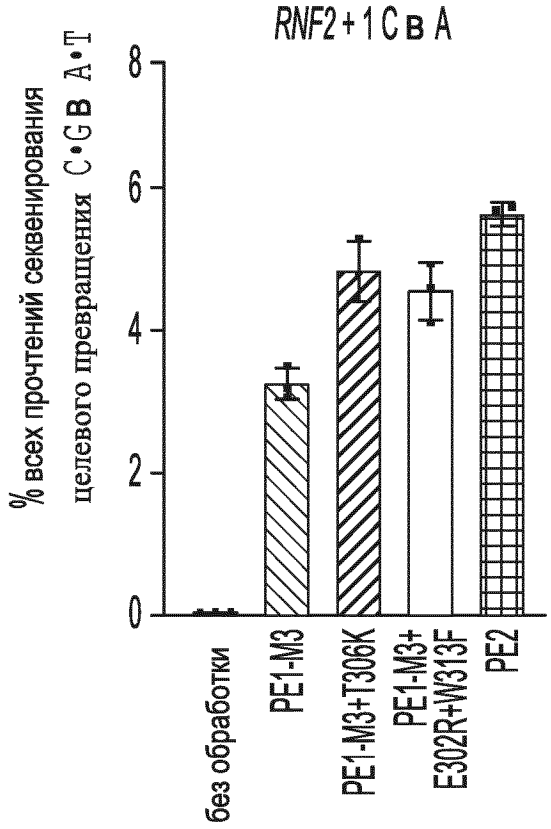
Фиг. 47H



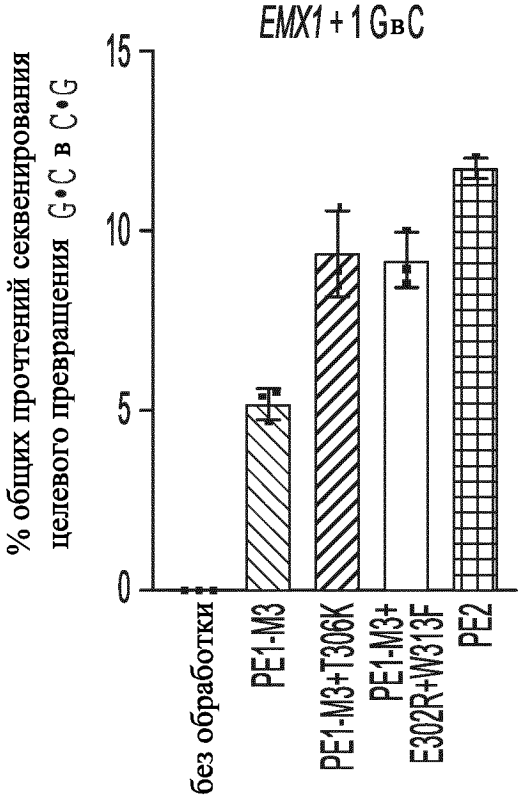
Фиг. 47I



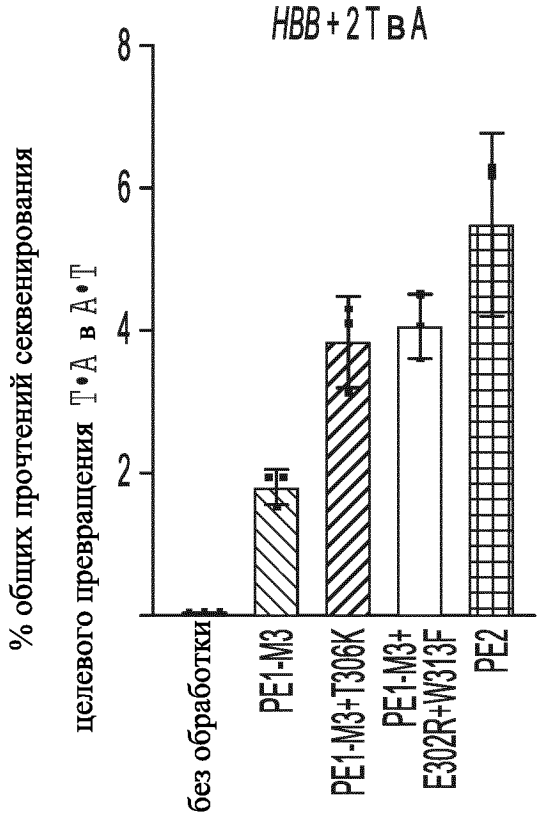
Фиг. 47J



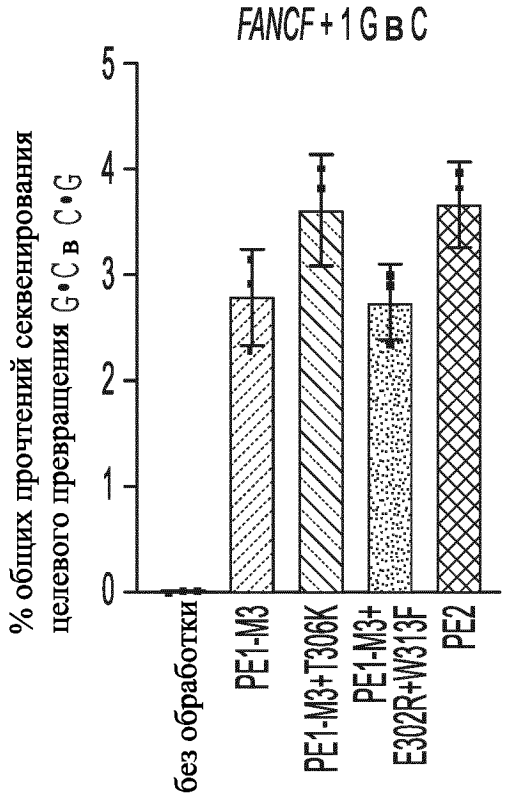
Фиг. 47K



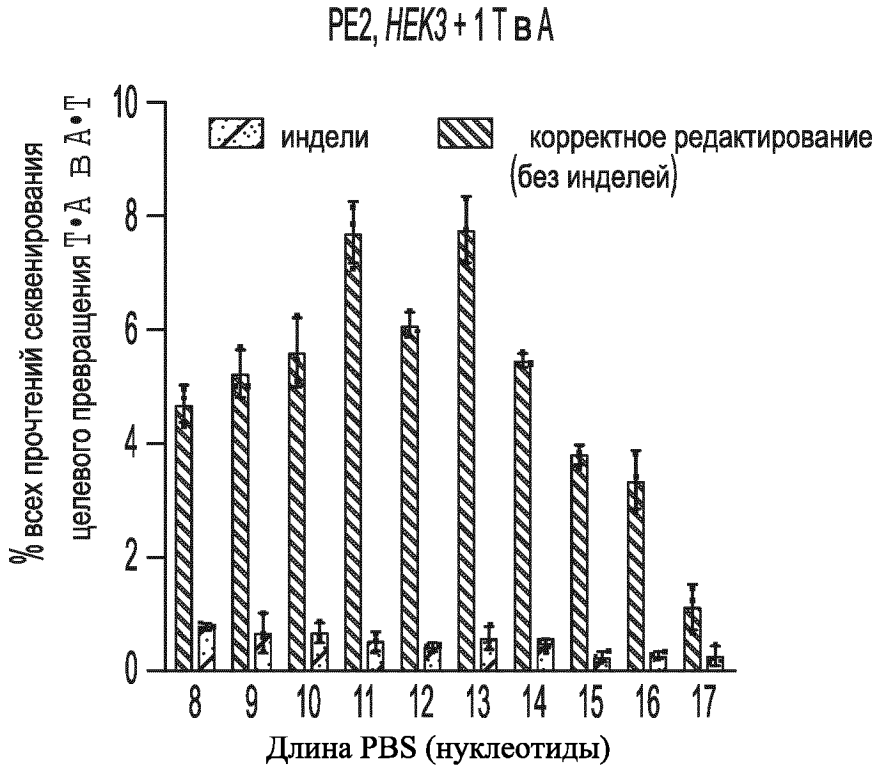
Фиг.47L



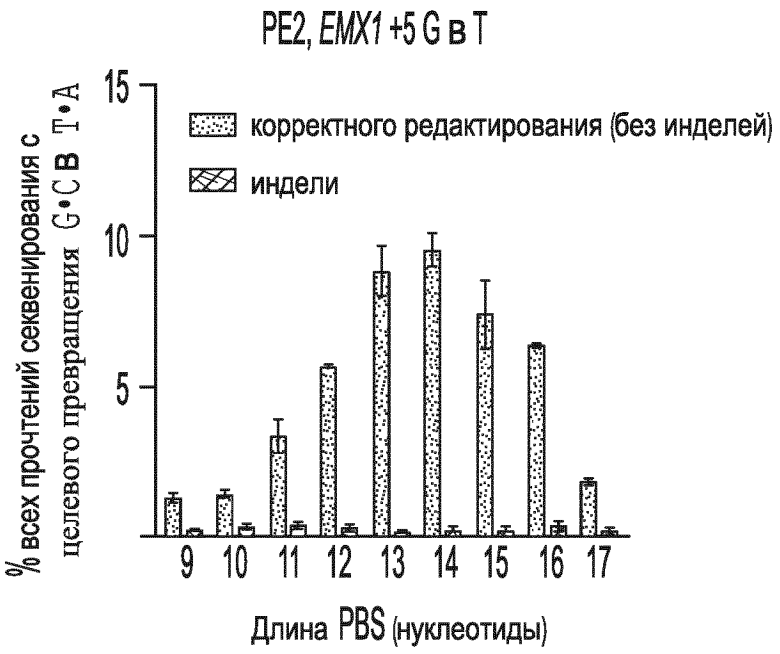
Фиг. 47M



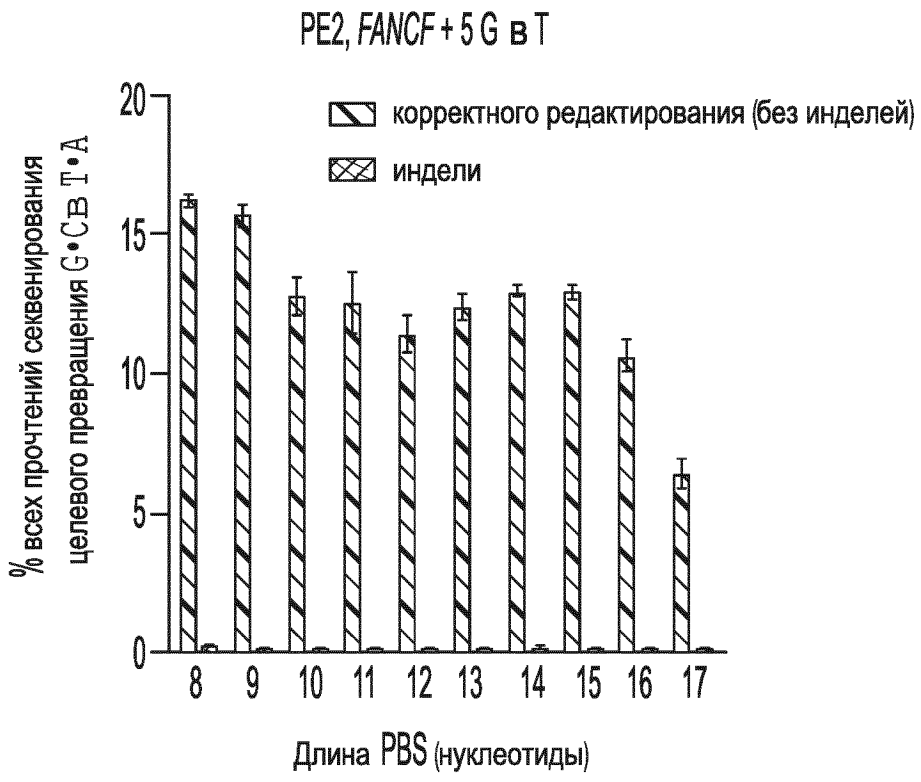
Фиг. 47N



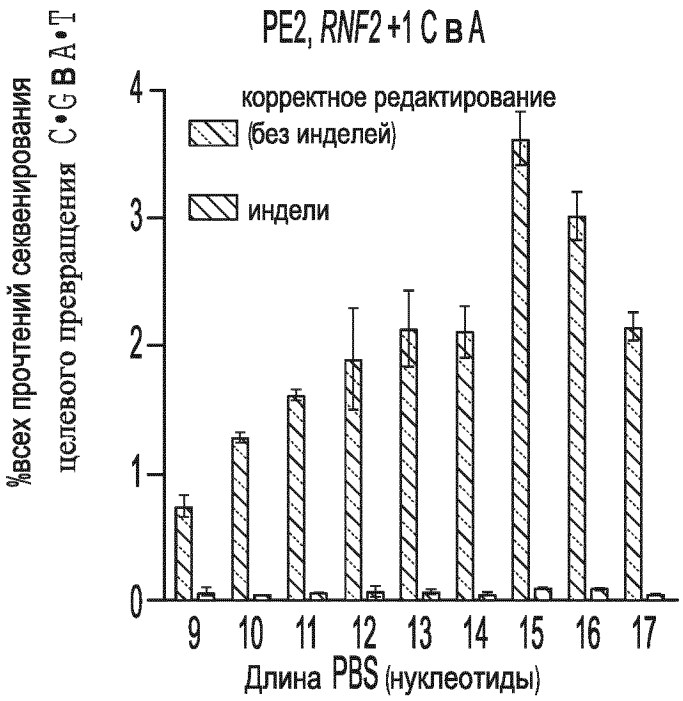
Фиг. 47O



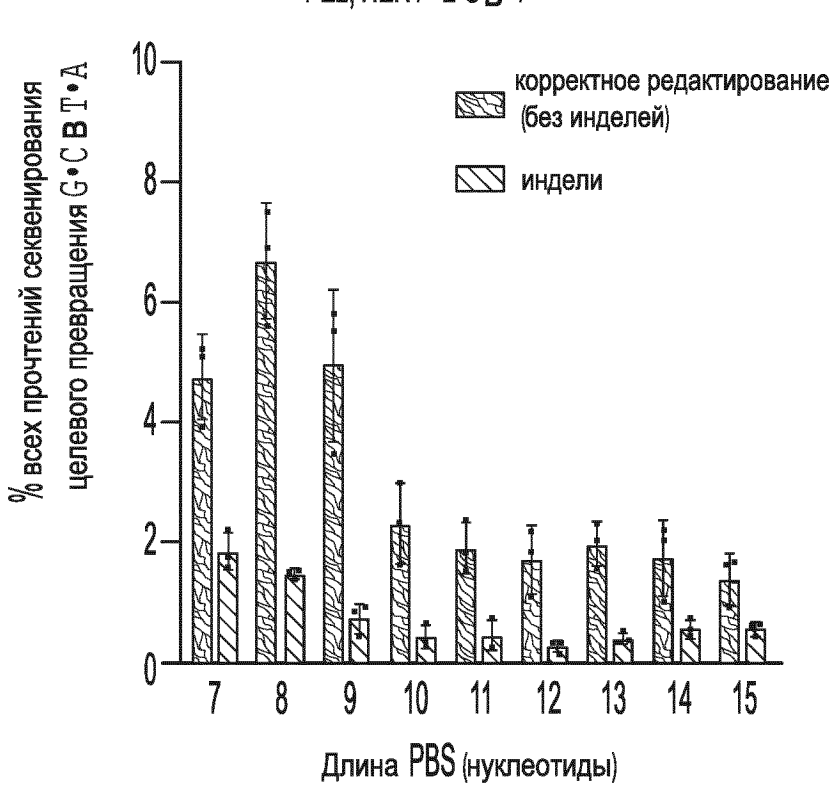
Фиг. 47P



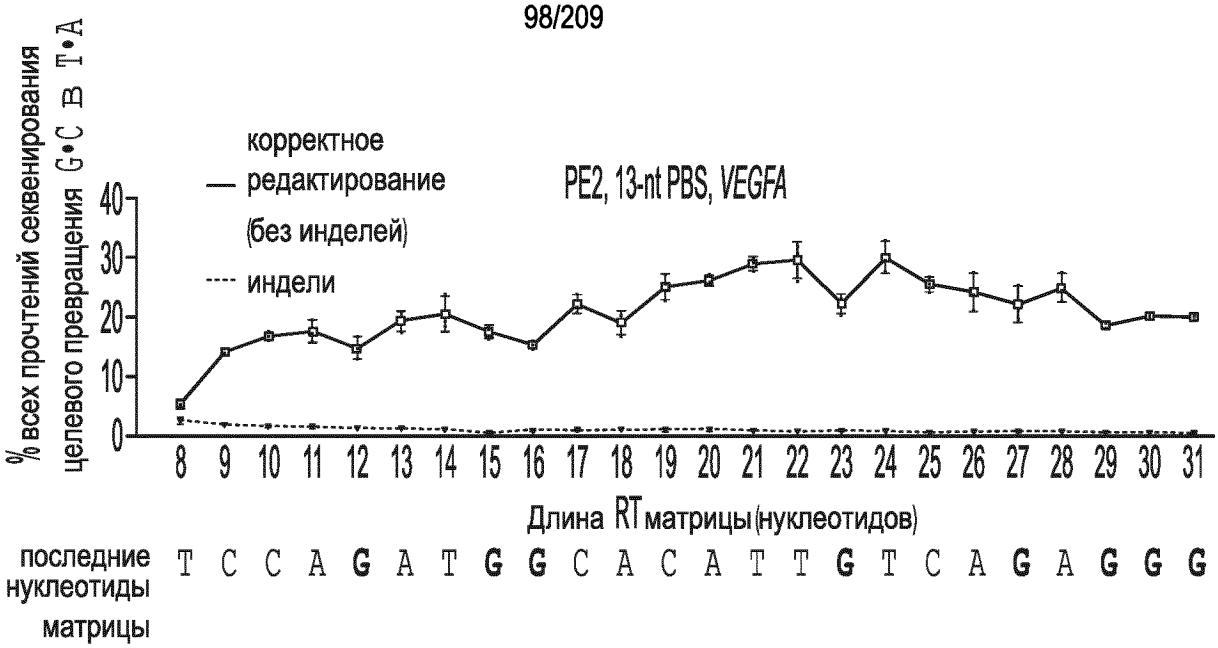
Фиг. 47Q



Фиг. 47R

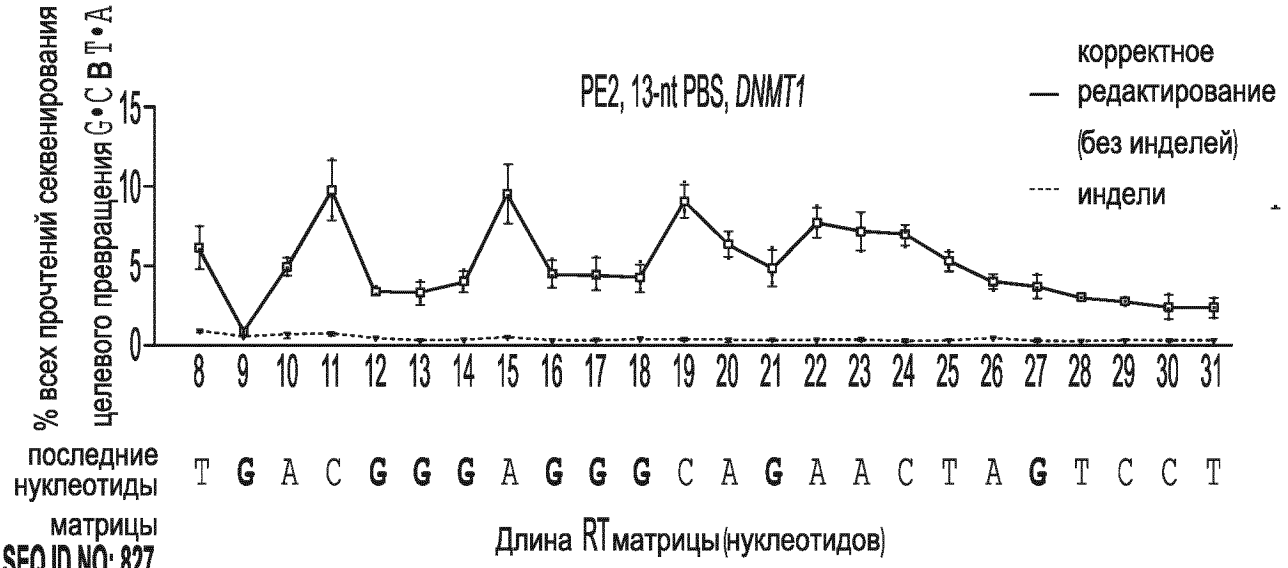


Фиг. 47S



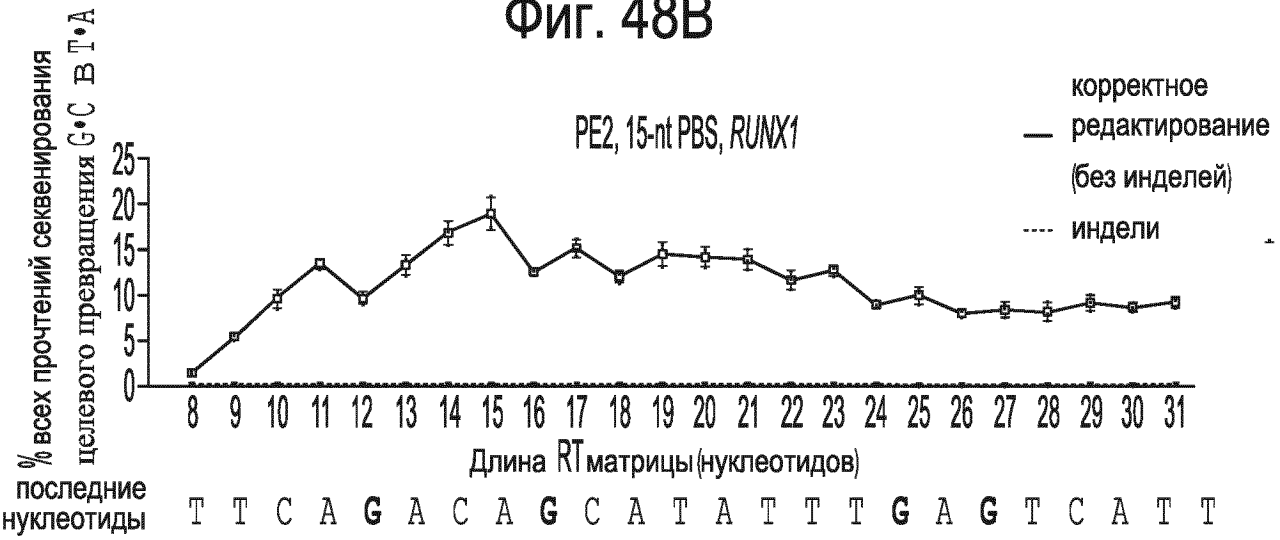
SEQ ID NO: 826

Фиг. 48A



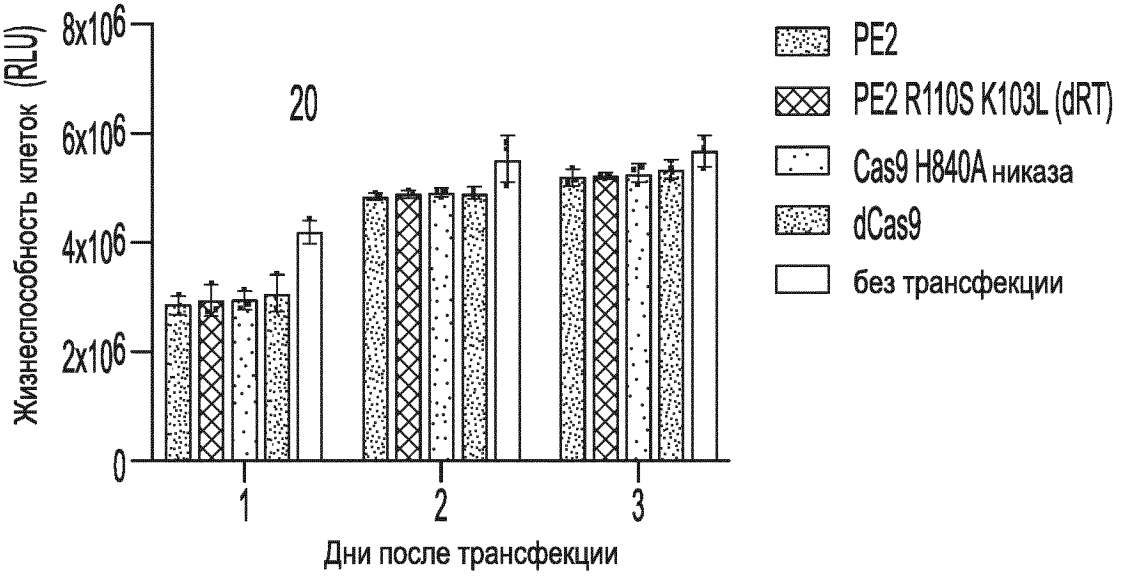
SEQ ID NO: 827

Фиг. 48B

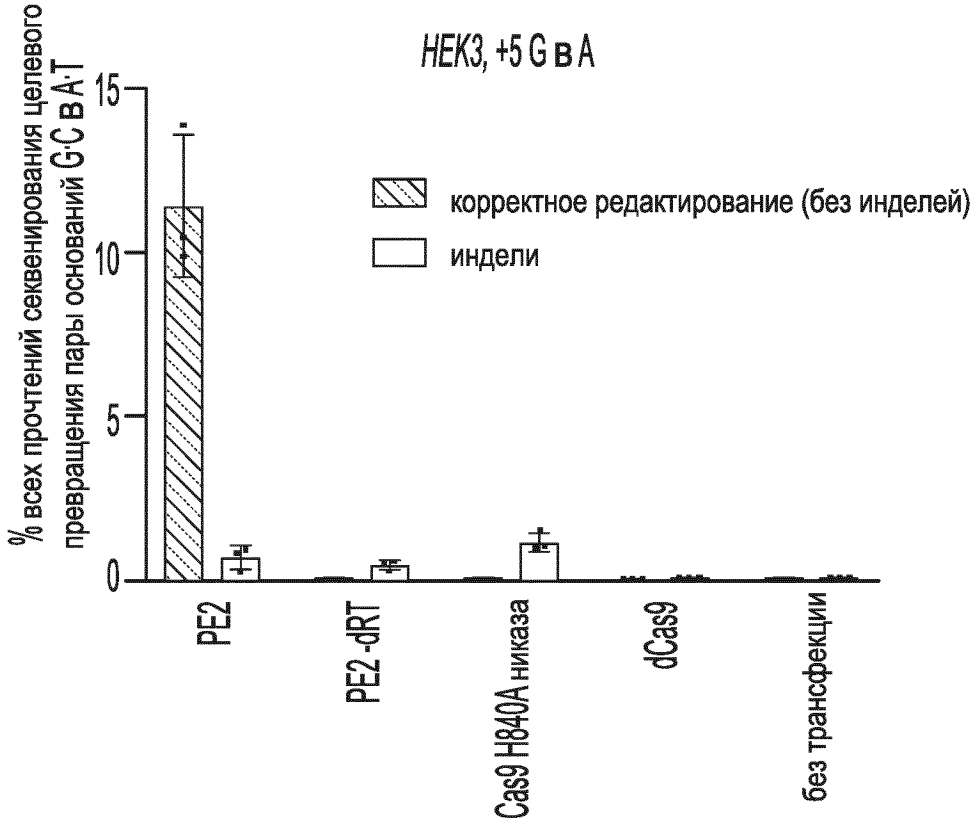


SEQ ID NO: 828

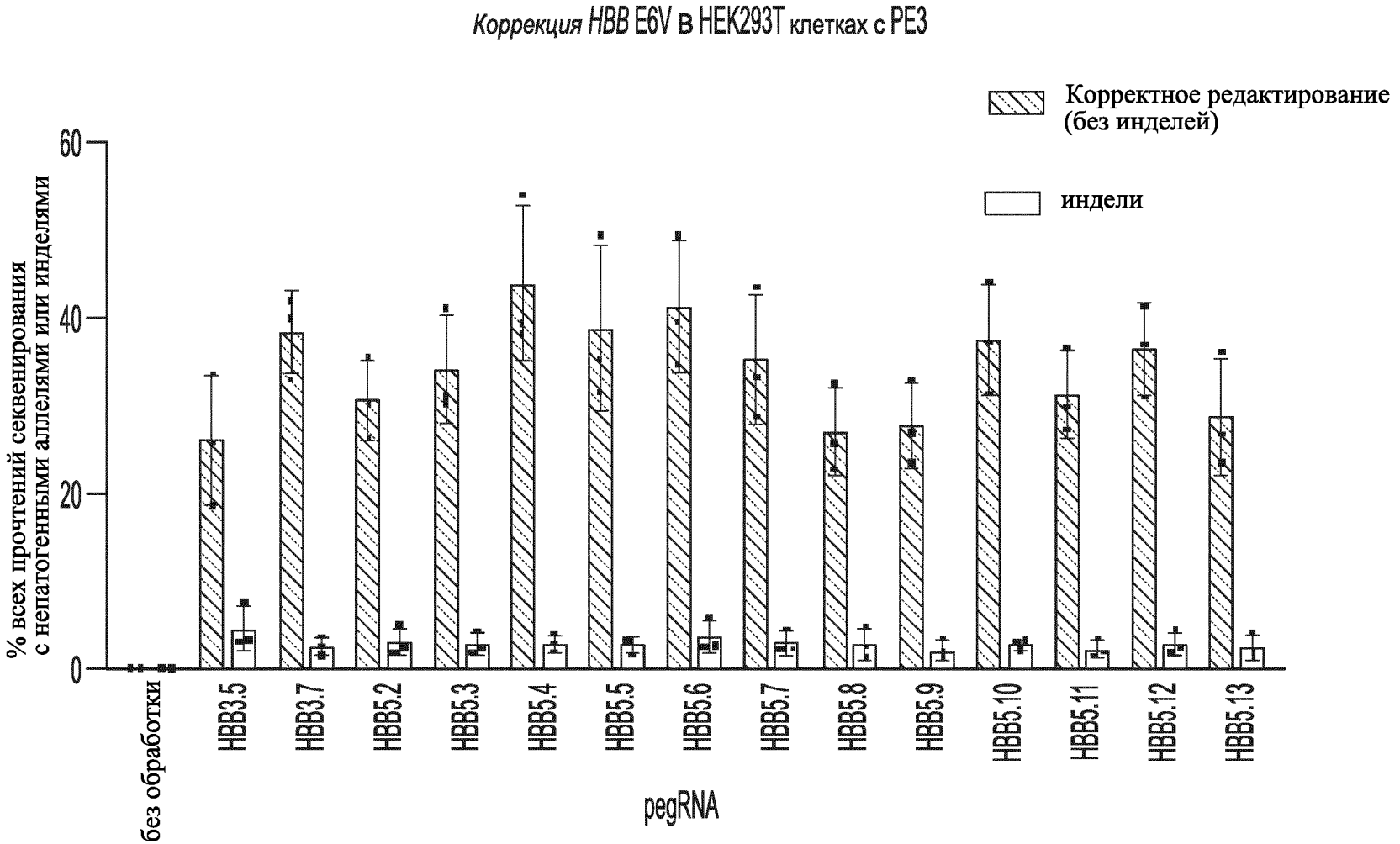
Фиг. 48C



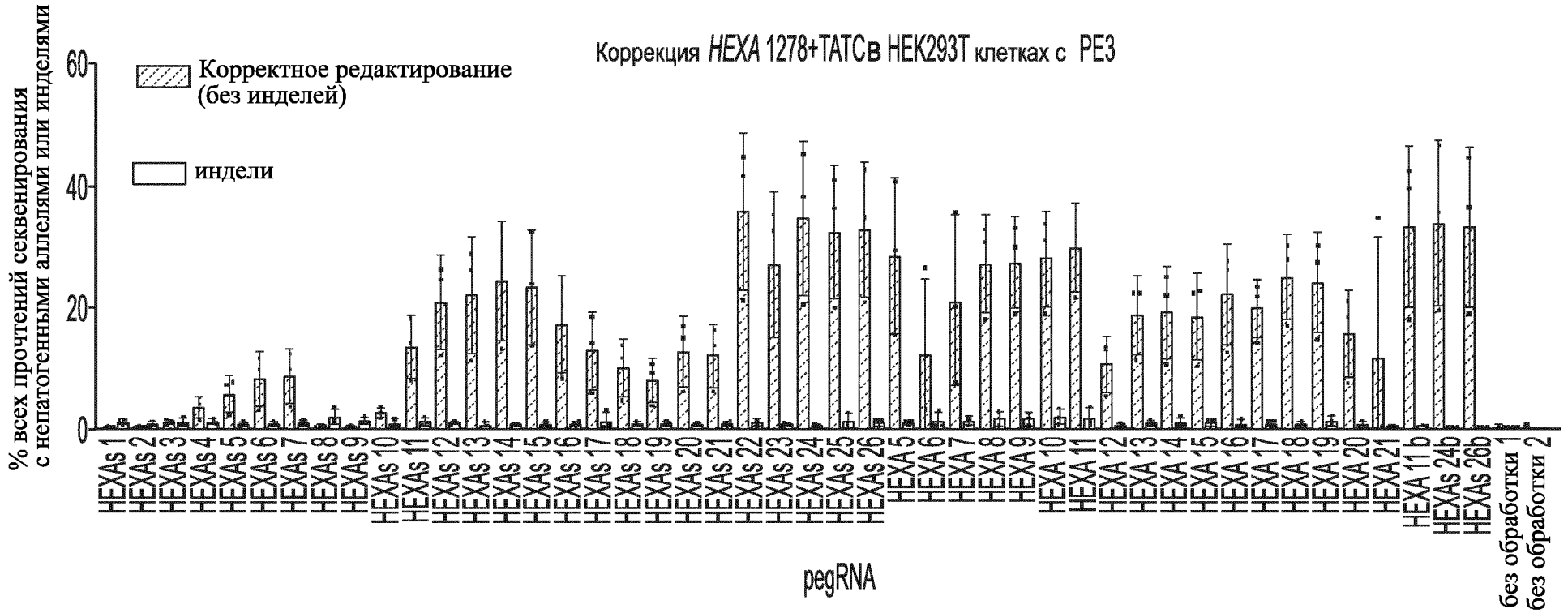
Фиг. 49A



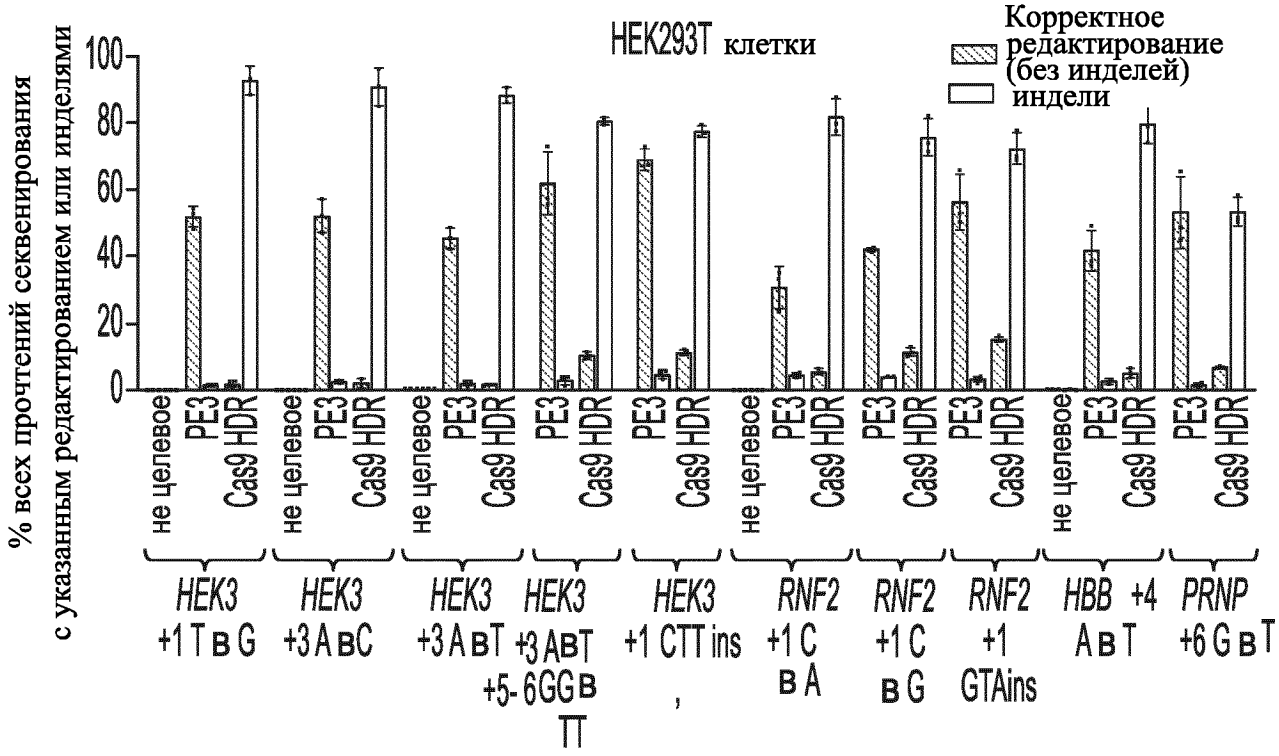
Фиг. 49B


Коррекция HBB E6V в HEK293T клетках с PE3
Корректное редактирование (без инделей)
индели
% всех прочтений секвенирования с непатогенными аллелями или инделями
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

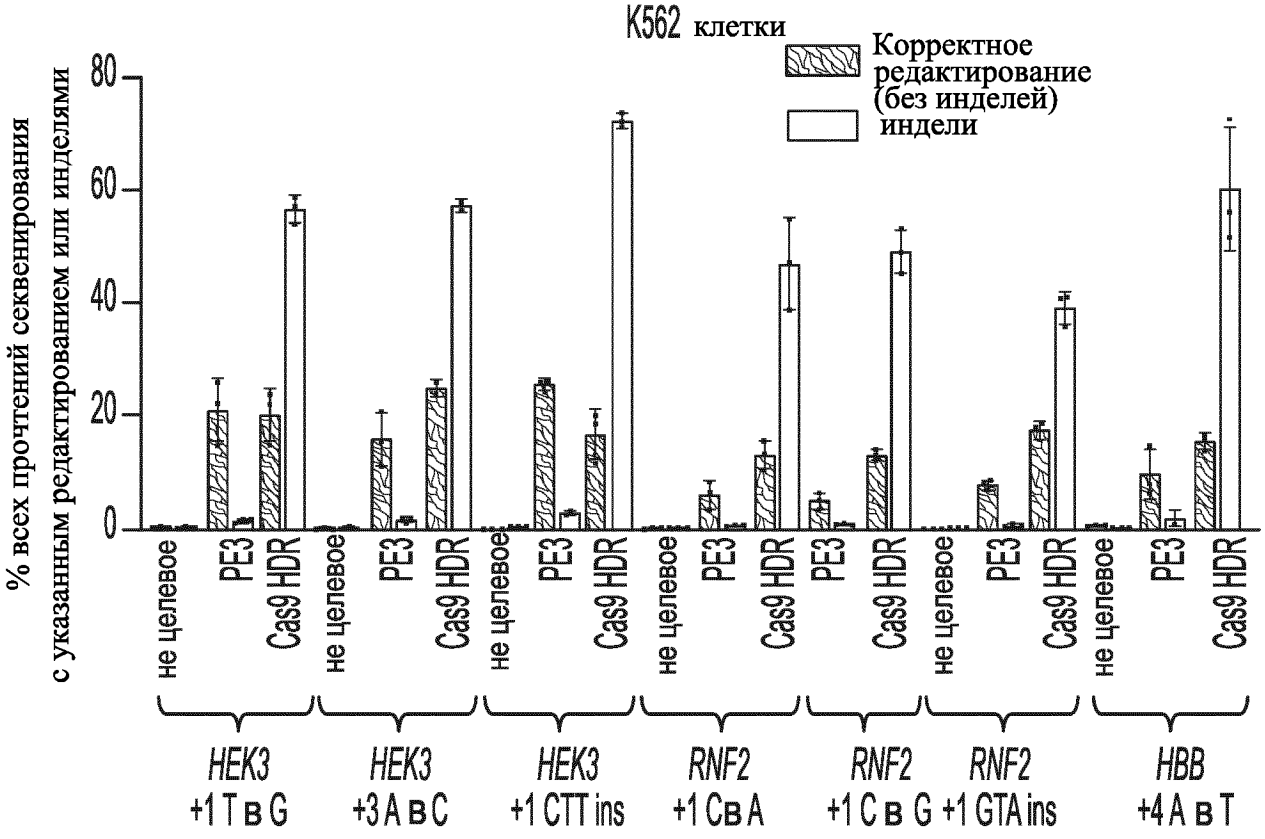
Фиг. 50A


Коррекция HEXA 1278+TATCв HEK293T клетках с PE3
Корректное редактирование (без инделей)
индели
% всех прочтений секвенирования с непатогенными аллелями или инделями
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 11b
HEXAs 24b
HEXAs 26b
без обработки 1
без обработки 2
pegRNA

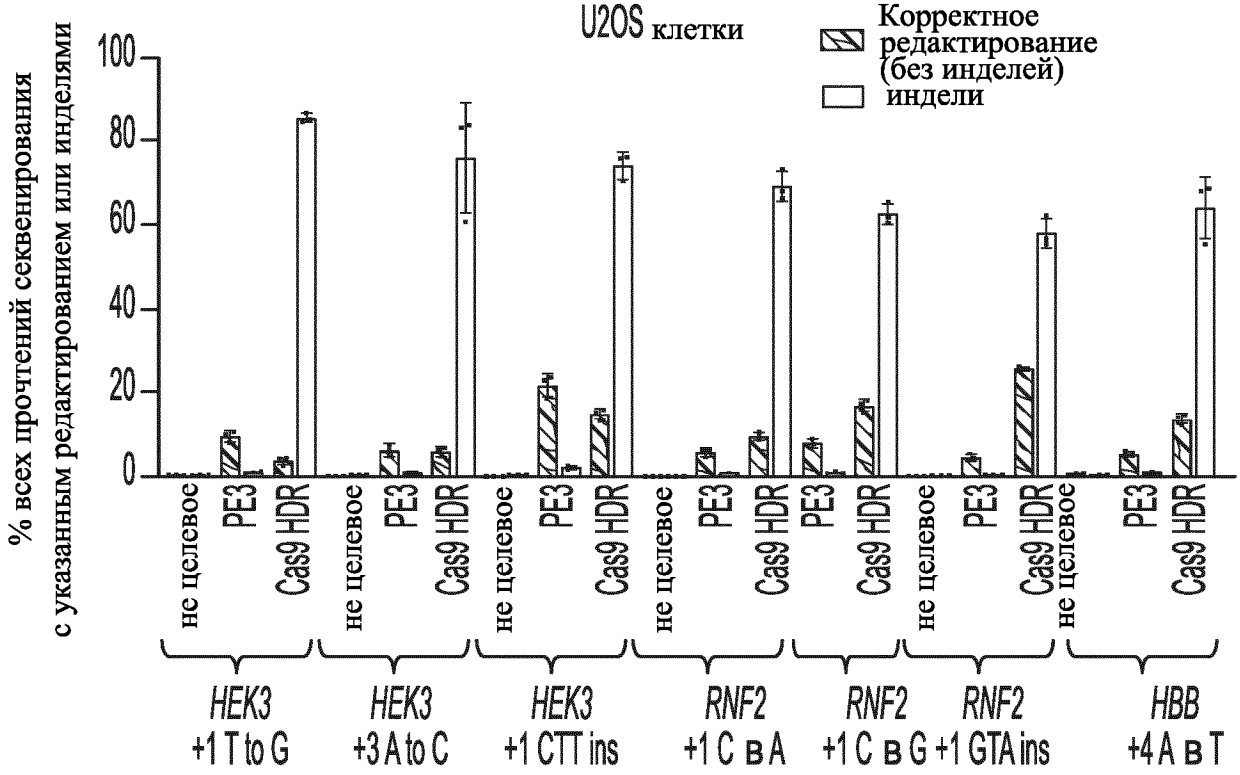
Фиг. 50B



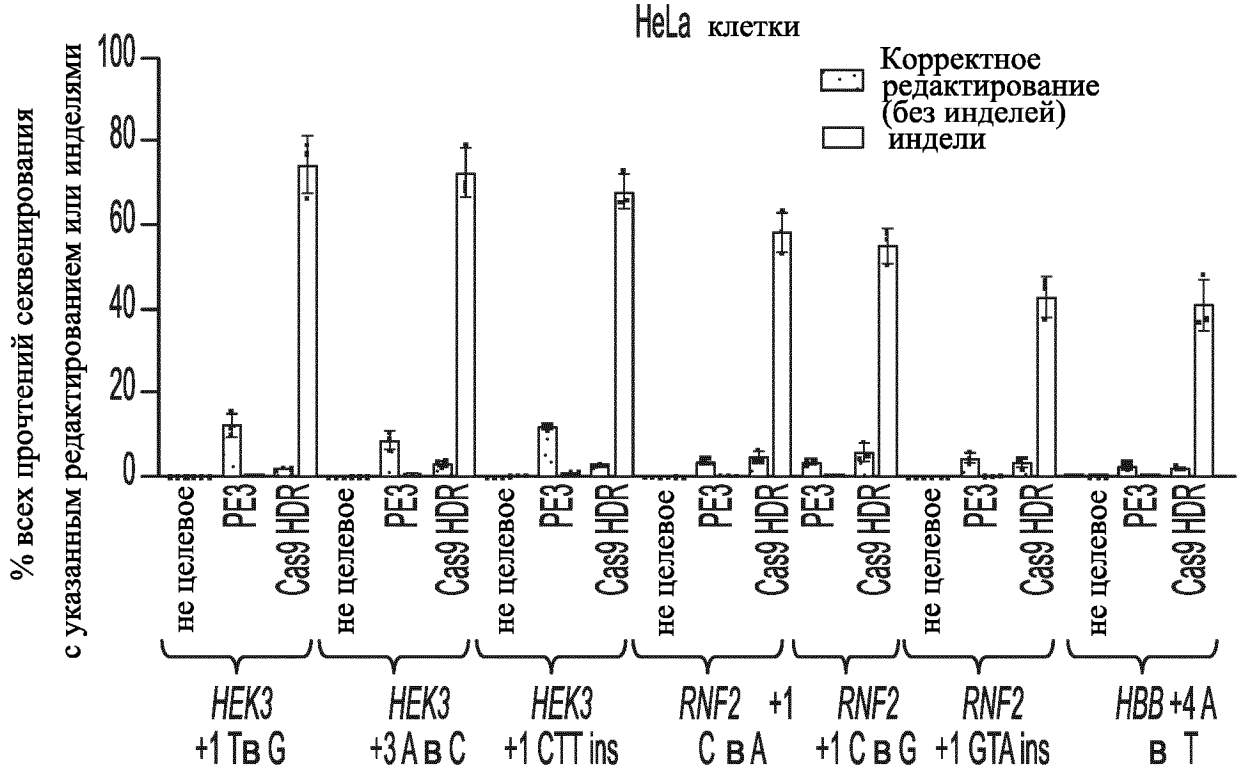
Фиг. 51A



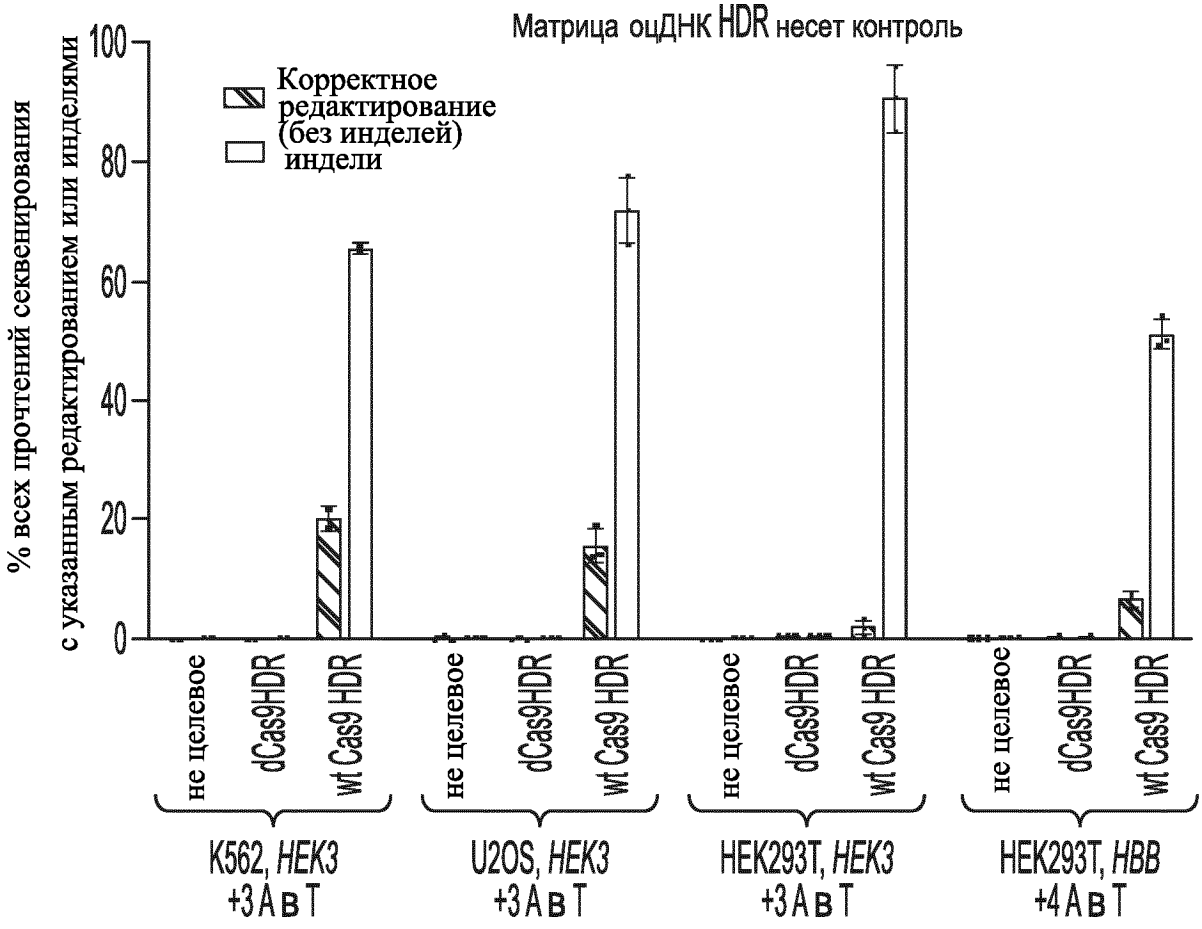
Фиг. 51B



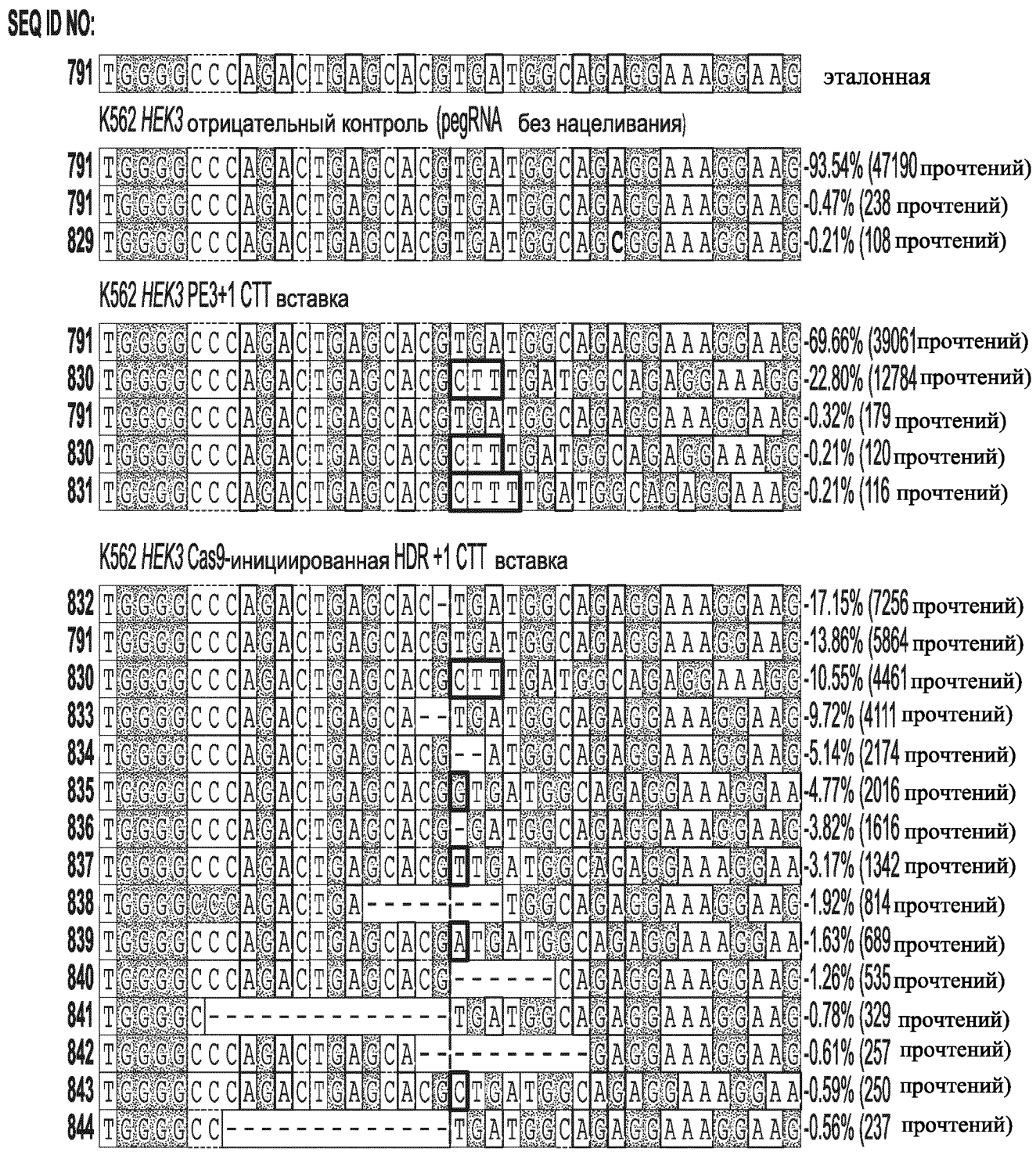
Фиг. 51C



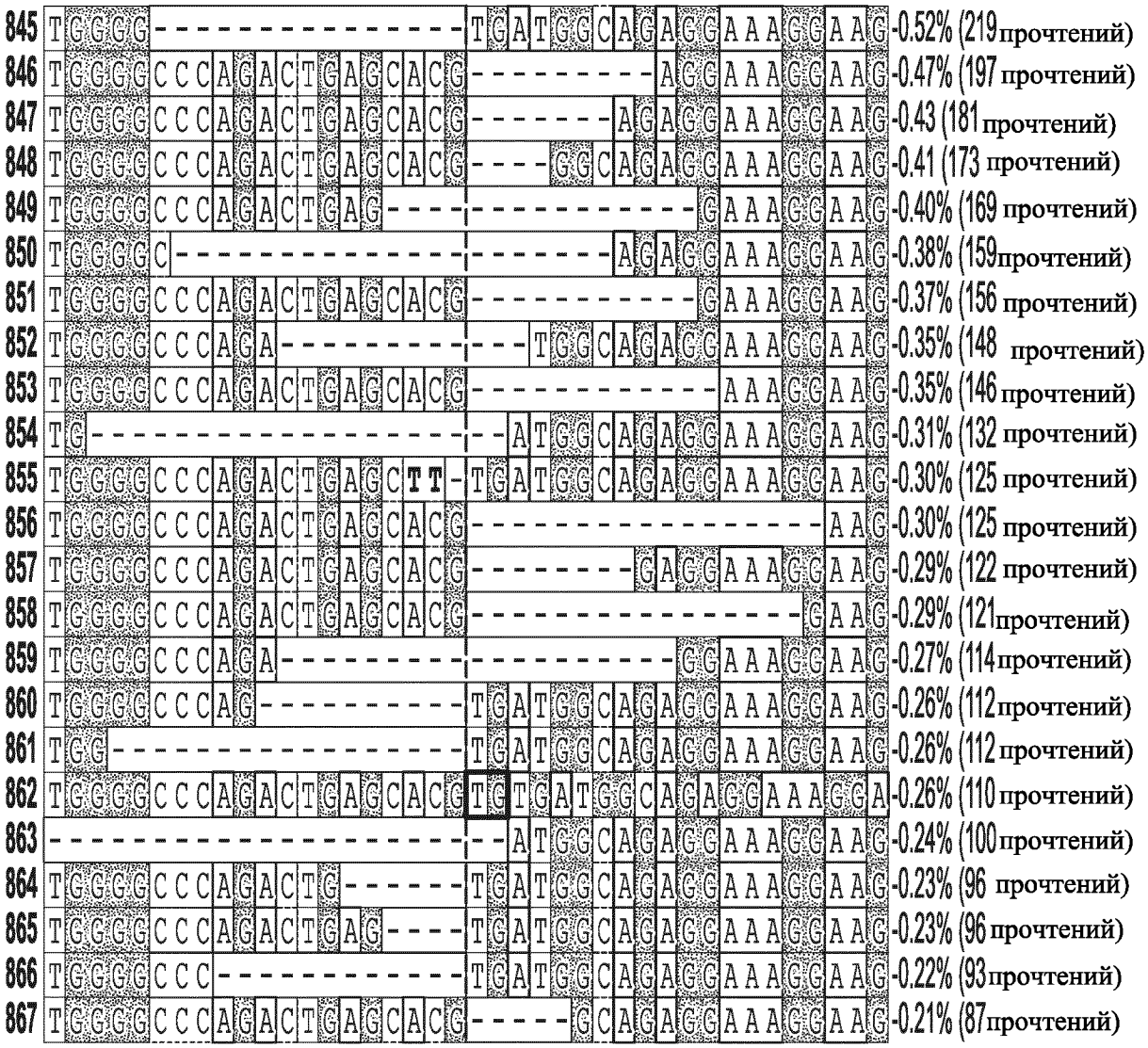
Фиг. 51D


Матрица оцДНК HDR несет контроль
% всех прочтений секвенирования с указанным редактированием или инделями
Корректное редактирование (без инделей)
индели
100
80
60
40
20
0
не целевое
dCas9HDR
wt Cas9 HDR
K562, HEK3 +3 A в T
U2OS, HEK3 +3 A в T
HEK293T, HEK3 +3 A в T
HEK293T, HBB +4 A в T

Фиг. 51E


SEQ ID NO:
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG эталонная
K562 HEK3 отрицательный контроль (pegRNA без нацеливания)
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -93.54% (47190 прочтений)
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -0.47% (238 прочтений)
829 TGGGCCCAGACTGAGCACGTGATGGCAGCGAAAGGAAG -0.21% (108 прочтений)
K562 HEK3 PE3+1 CTT вставка
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -69.66% (39061 прочтений)
830 TGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -22.80% (12784 прочтений)
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -0.32% (179 прочтений)
830 TGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -0.21% (120 прочтений)
831 TGGGCCCAGACTGAGCACGCTTTTGATGGCAGAGGAAAG -0.21% (116 прочтений)
K562 HEK3 Cas9-инициированная HDR +1 CTT вставка
832 TGGGCCCAGACTGAGCAC-TGATGGCAGAGGAAAGGAAG -17.15% (7256 прочтений)
791 TGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -13.86% (5864 прочтений)
830 TGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -10.55% (4461 прочтений)
833 TGGGCCCAGACTGAGCA--TGATGGCAGAGGAAAGGAAG -9.72% (4111 прочтений)
834 TGGGCCCAGACTGAGCACG--ATGGCAGAGGAAAGGAAG -5.14% (2174 прочтений)
835 TGGGCCCAGACTGAGCACGGTGATGGCAGAGGAAAGGAA -4.77% (2016 прочтений)
836 TGGGCCCAGACTGAGCACG-GATGGCAGAGGAAAGGAAG -3.82% (1616 прочтений)
837 TGGGCCCAGACTGAGCACGTTGATGGCAGAGGAAAGGAA -3.17% (1342 прочтений)
838 TGGGCCCAGACTGA-----TGGCAGAGGAAAGGAAG -1.92% (814 прочтений)
839 TGGGCCCAGACTGAGCACGATGATGGCAGAGGAAAGGAA -1.63% (689 прочтений)
840 TGGGCCCAGACTGAGCACG------CAGAGGAAAGGAAG -1.26% (535 прочтений)
841 TGGGGC----------TGATGGCAGAGGAAAGGAAG -0.78% (329 прочтений)
842 TGGGCCCAGACTGAGCA---------GAGGAAAGGAAG -0.61% (257 прочтений)
843 TGGGCCCAGACTGAGCACGCTGATGGCAGAGGAAAGGAA -0.59% (250 прочтений)
844 TGGGCC----------TGATGGCAGAGGAAAGGAAG -0.56% (237 прочтений)

Фиг. 51F



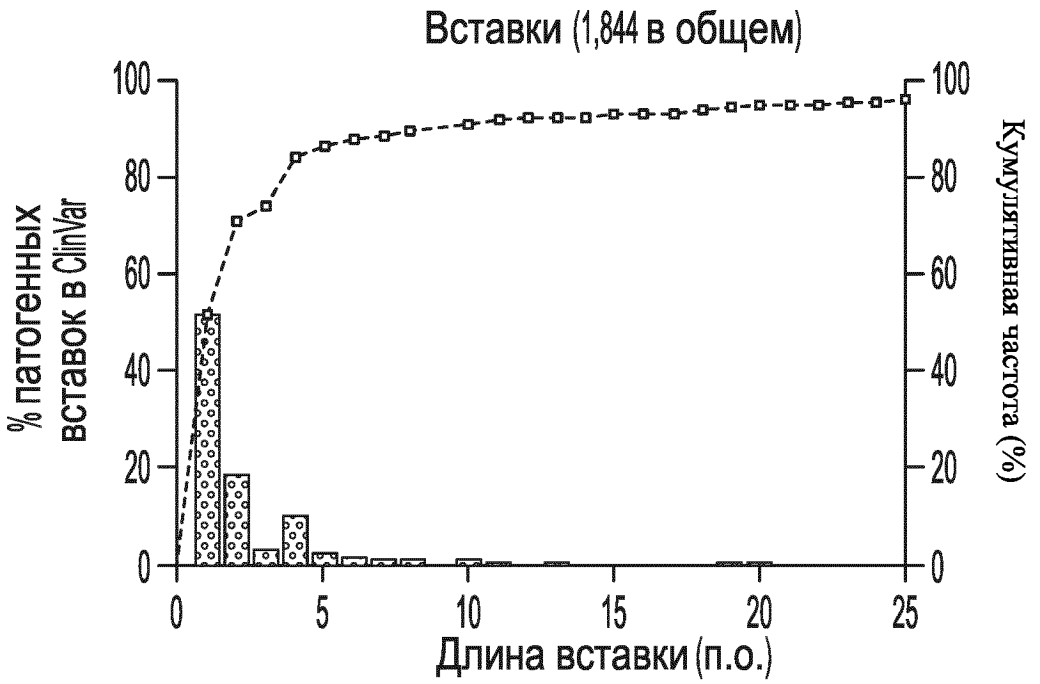
Выделенные замены

вставки

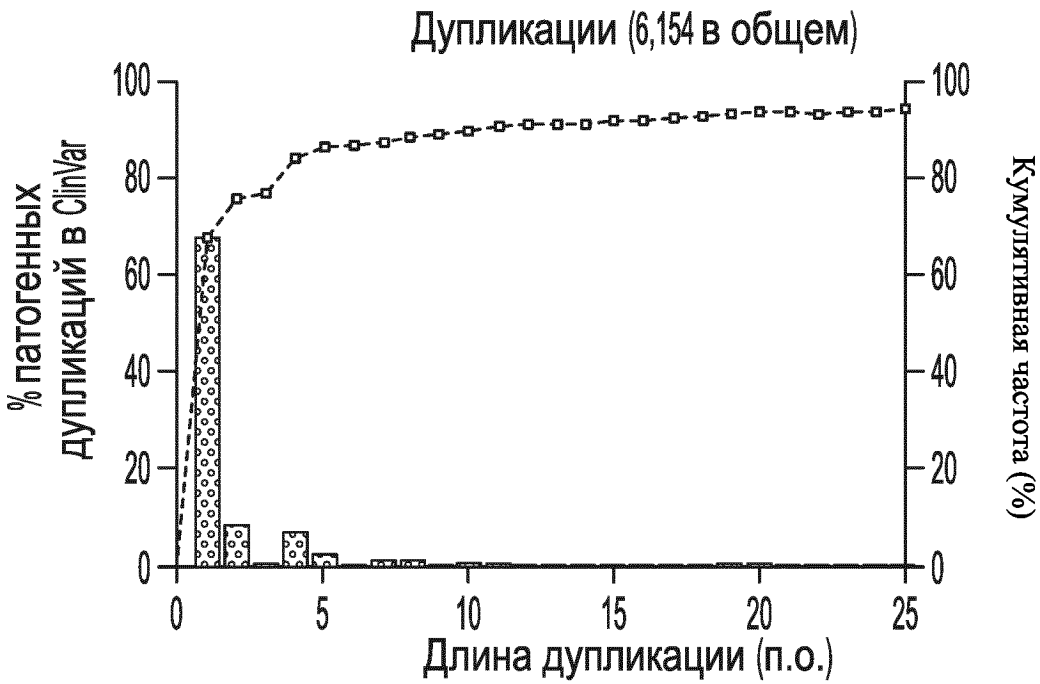
делеции

предсказанное положение расщепления

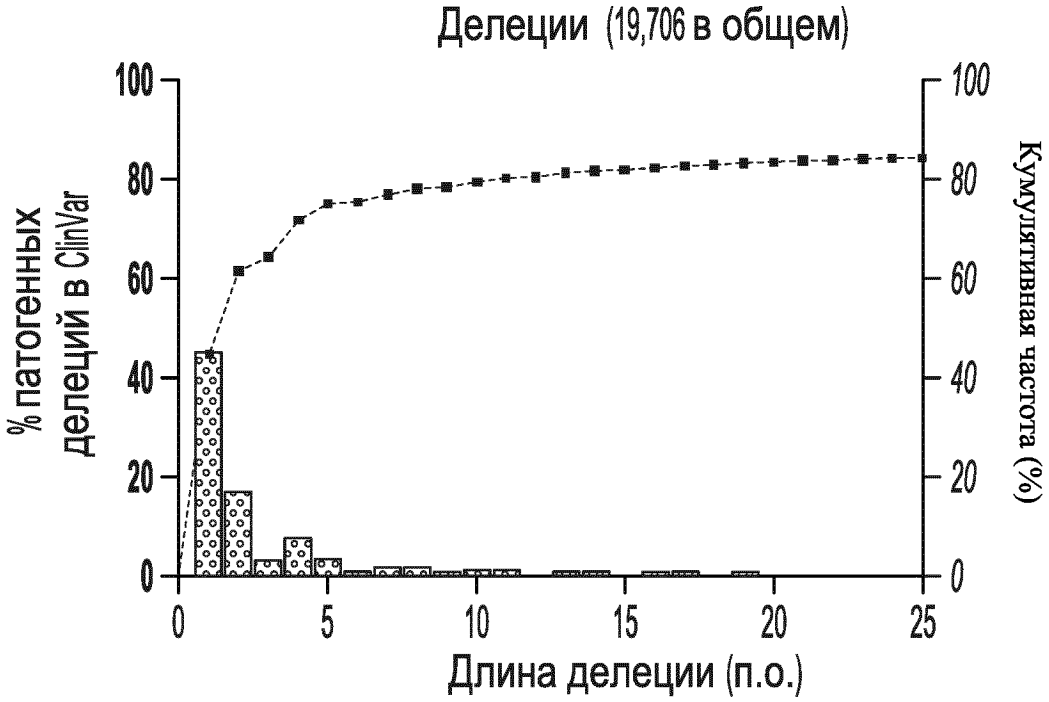
Фиг. 51G



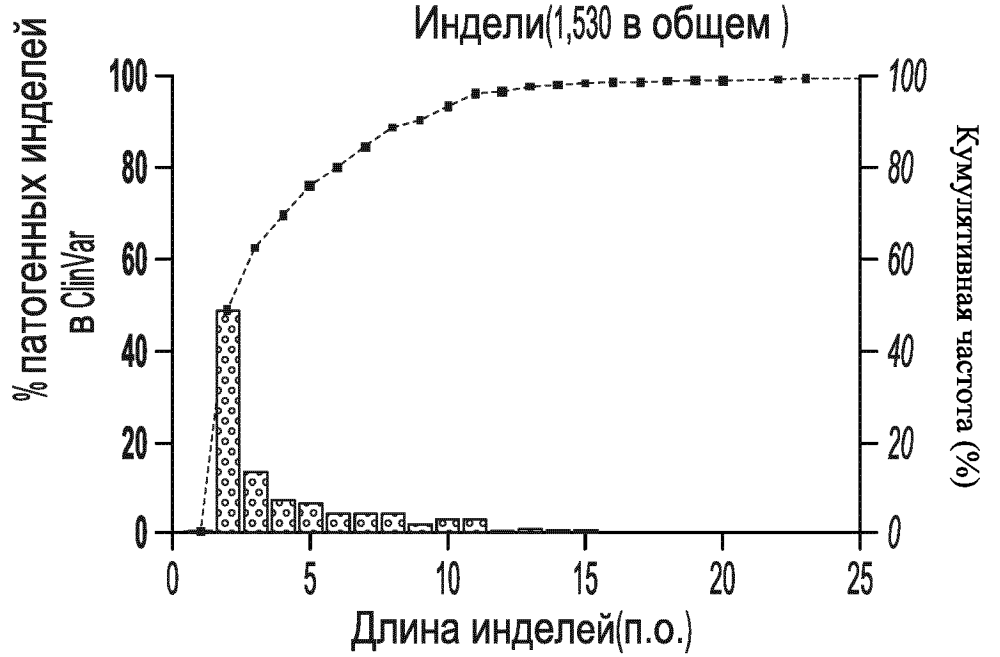
Фиг. 52A



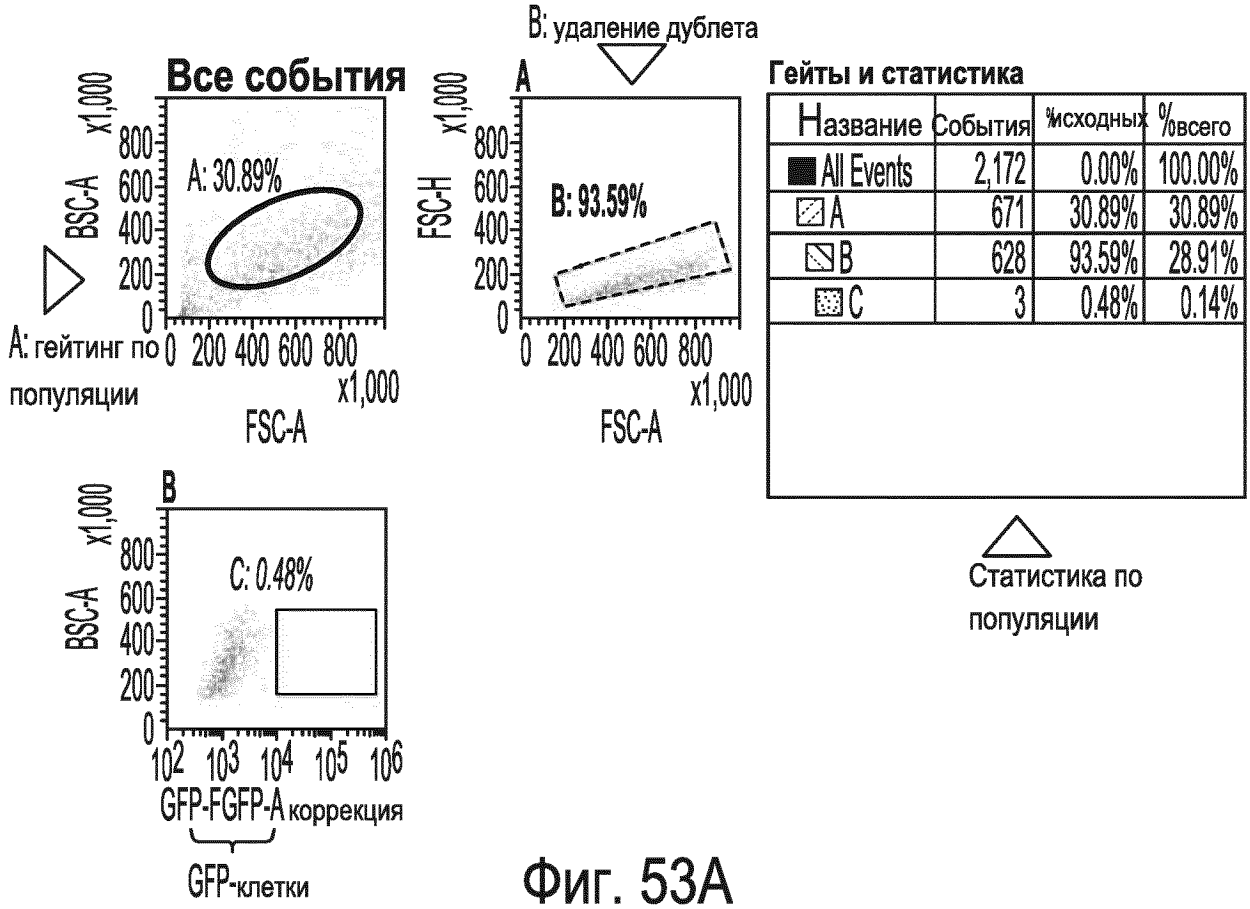
Фиг. 52B



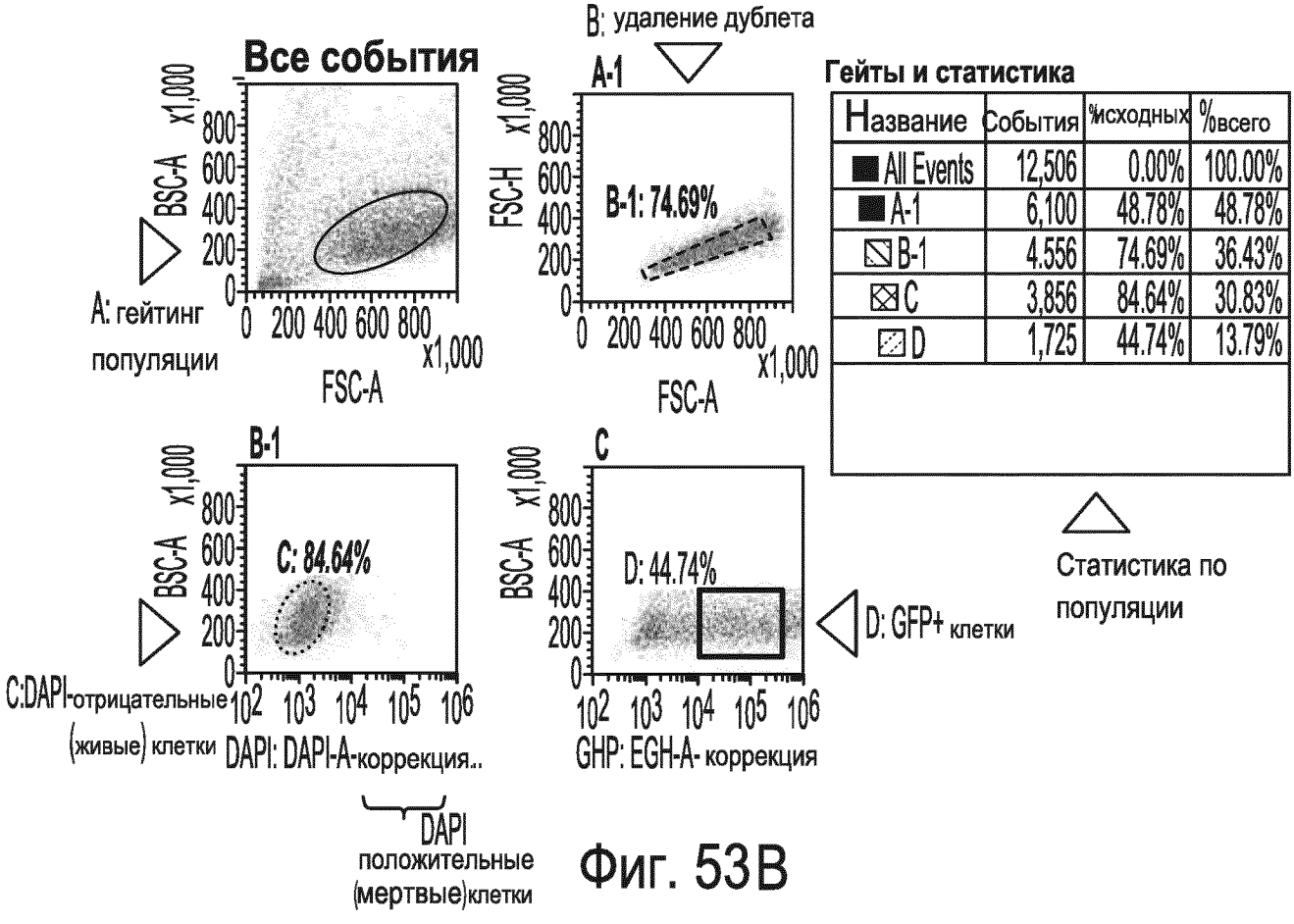
Фиг. 52C



Фиг. 52D

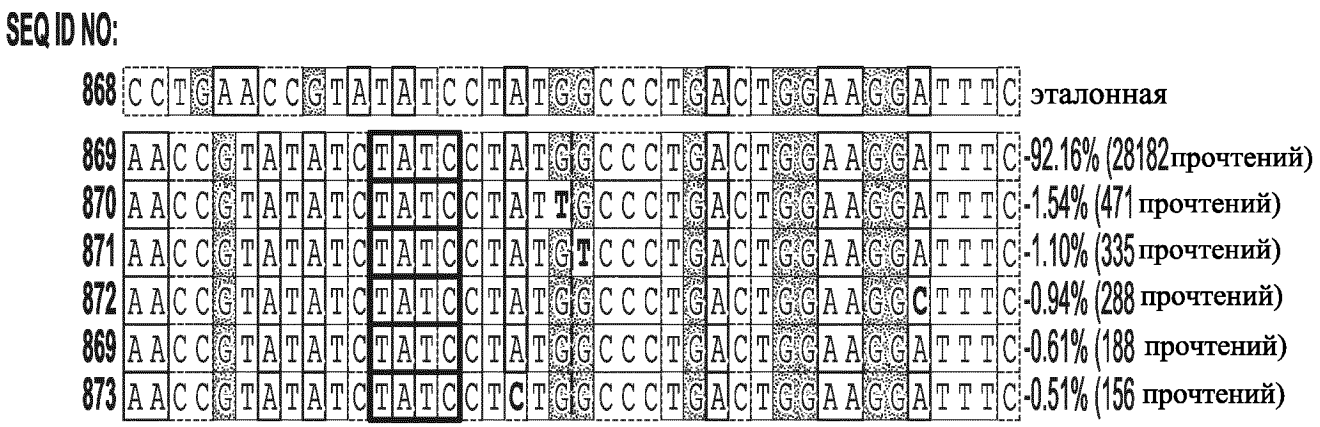


Фиг. 53A

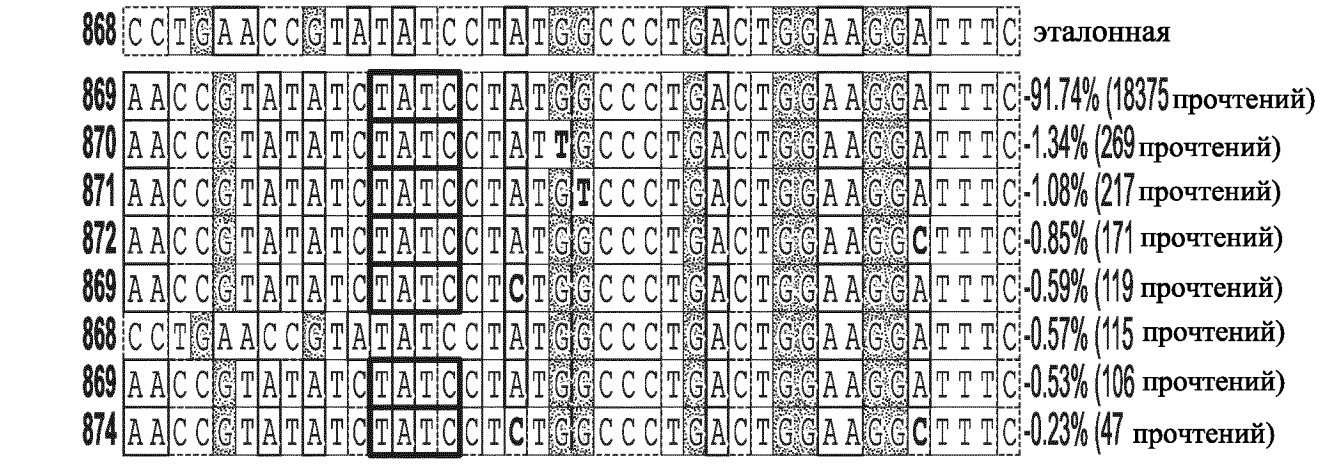


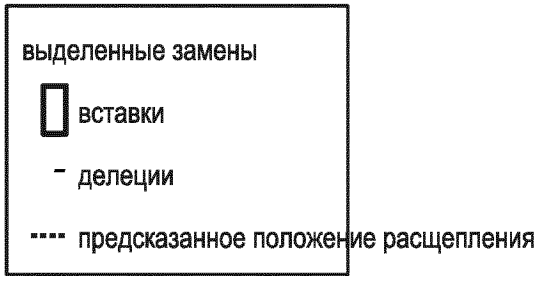
Фиг. 53B

Линия 1

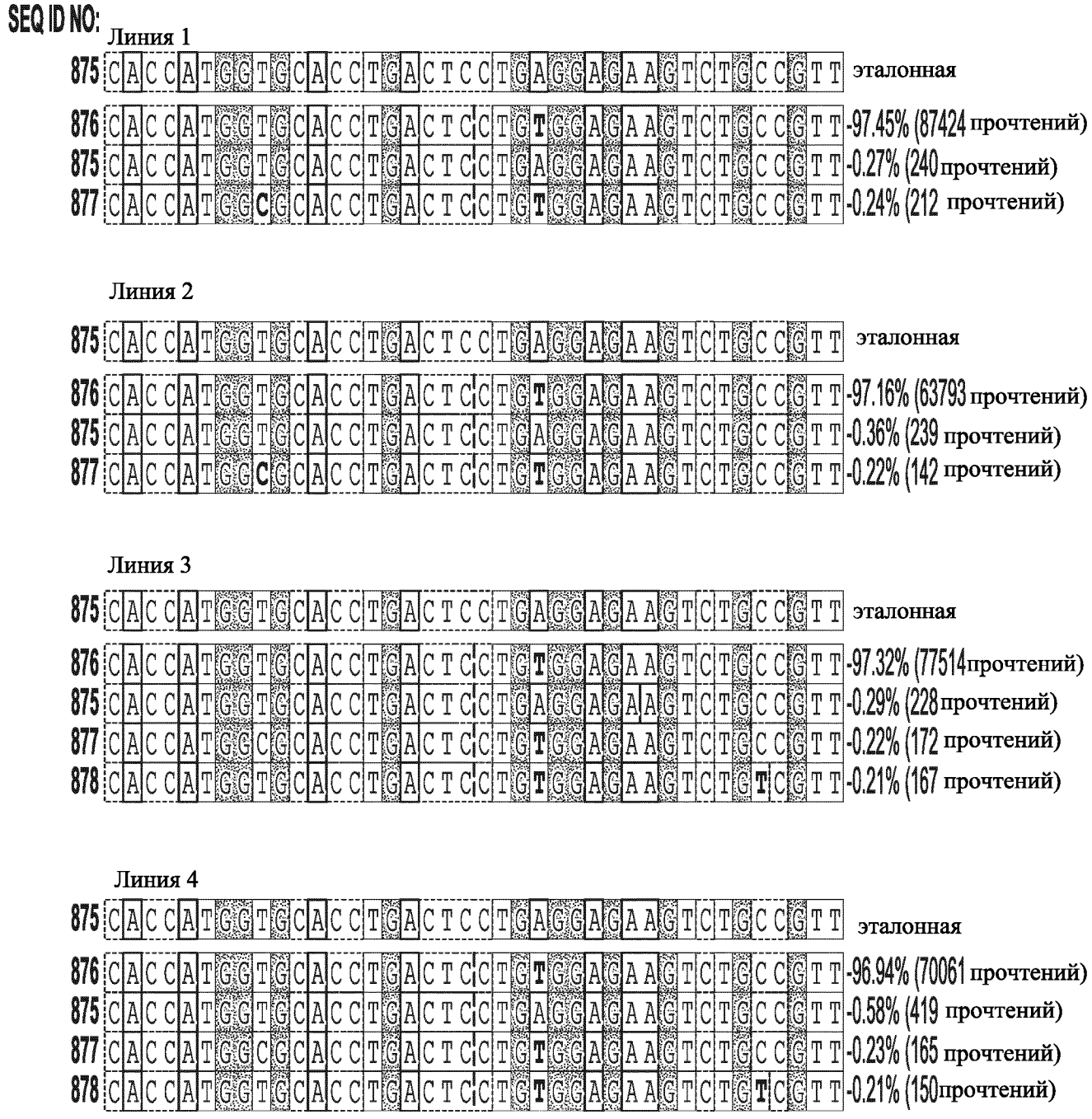


Линия 2

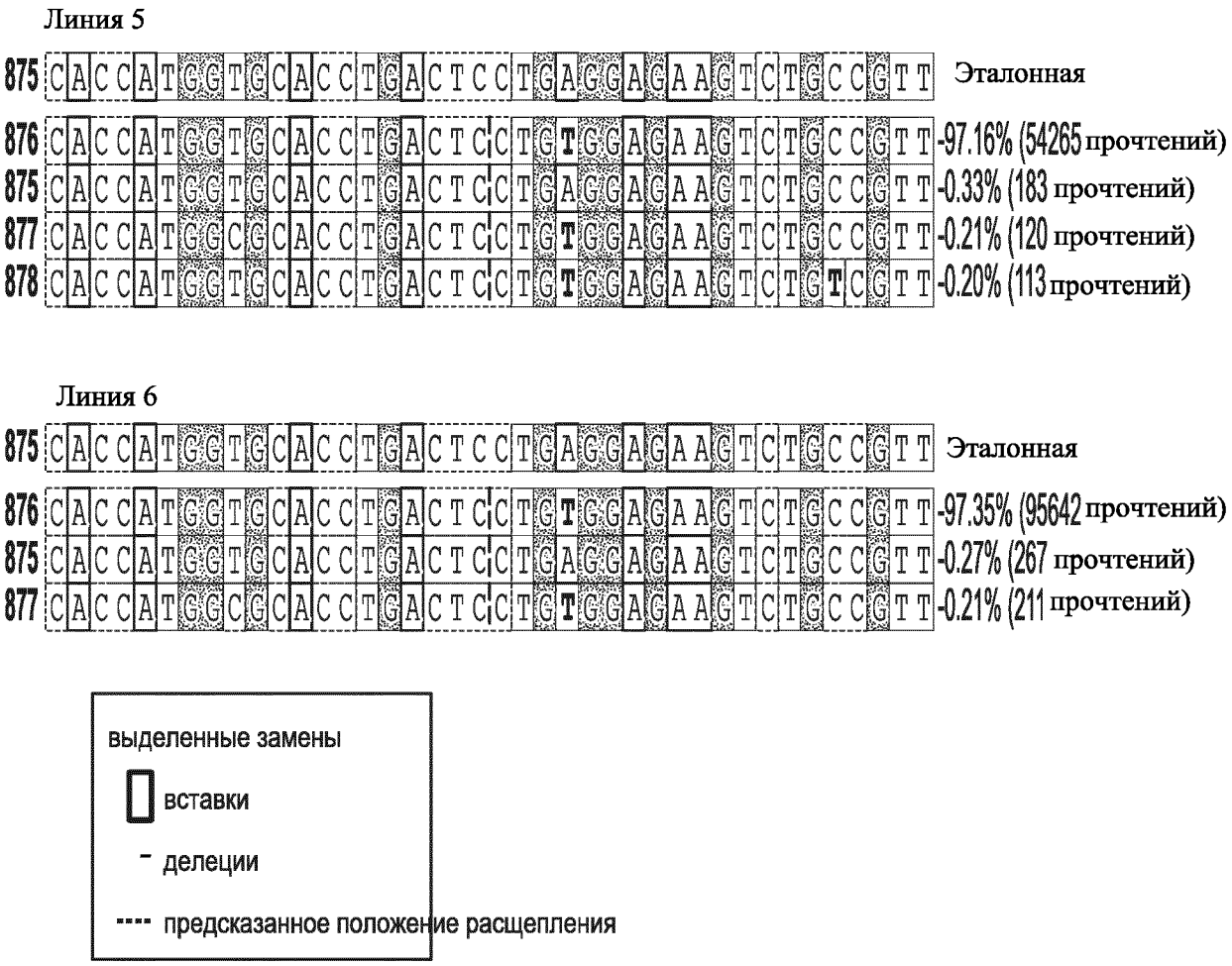




Фиг. 53C


SEQ ID NO:
Линия 1
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.45% (87424 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.27% (240 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.24% (212 прочтений)
Линия 2
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.16% (63793 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.36% (239 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.22% (142 прочтений)
Линия 3
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.32% (77514 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.29% (228 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.22% (172 прочтений)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.21% (167 прочтений)
Линия 4
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -96.94% (70061 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.58% (419 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.23% (165 прочтений)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.21% (150 прочтений)

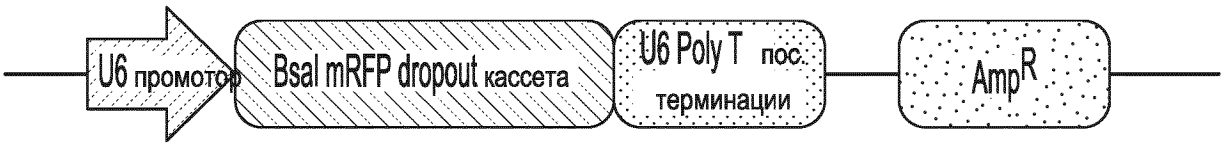
Фиг. 53D


Линия 5
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT Эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.16% (54265 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.33% (183 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.21% (120 прочтений)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.20% (113 прочтений)
Линия 6
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT Эталонная
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.35% (95642 прочтений)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.27% (267 прочтений)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.21% (211 прочтений)
выделенные замены
вставки
делеции
предсказанное положение расщепления

Фиг. 53E

Протокол одновременного клонирования спейсерной последовательности (целевой участок) и 3'-удлинения PEгРНК, содержащее PBS и матрицу RT,
вектор экспрессии PEгРНК с промотором U6

(Компонент 1) схема нерасщепленного вектора **pU6-pegRNA-GG-**:



Плазмида	Описание	признаки	разрез с
pU6-pegRNA-GG-вектор	челpegRNA U6вектор экспрессии	U6_BsaI-mRFP-cassette_U6-Ter	BsaI-HFV2

[Компонент2]порядок олигонуклеотидов с получением (целевой) последовательности спейсера для вектора pegRNA человека: верхние и нижние олго подвергнуты отжигу и собраны в расщепляемый вектор pU6-pegRNA-GG:

SEQ ID NO: 879 5′- CACCNNNNNNNNNNNNNNNNNNNGTTTT -3′
SEQ ID NO: 880 3′- NNNNNNNNNNNNNNNNNNNCAAAATCTC -5′

(показанная последовательность протоспейсера N20 mock может быть любой длины, как необходимо)

[Компонент 3] порядок олигонуклеотидов для каркаса SpCas9 sgRNA для вектора чел. pegRNA:

SEQ ID NO: 882 {Phos}5′- AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG -3′
SEQ ID NO: 881 -3′ GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCACG -5′ {Phos}

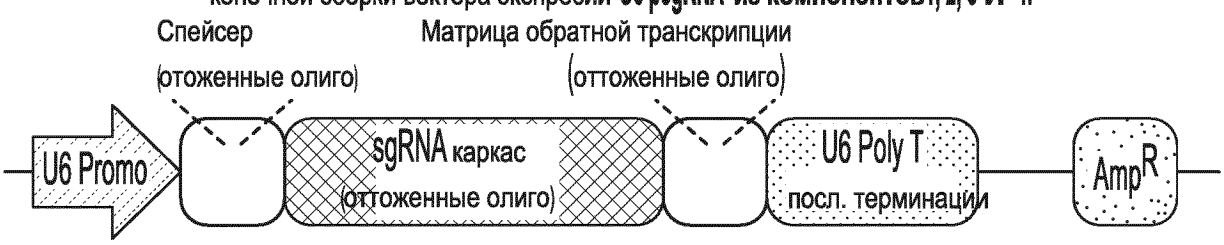
Верхние и нижние олиго подвергнуты отжигу и собраны в расщепляемый вектор pU6-pegRNA-GG

(эти олиго должны быть **5′ фосфорилированы** для сборки вектора pegRNA) [Компонент 4] порядок олигонуклеотидов для 3′ удлинения pegRNA, содержащей PBS и матрицу RT для векторов pegRNA:

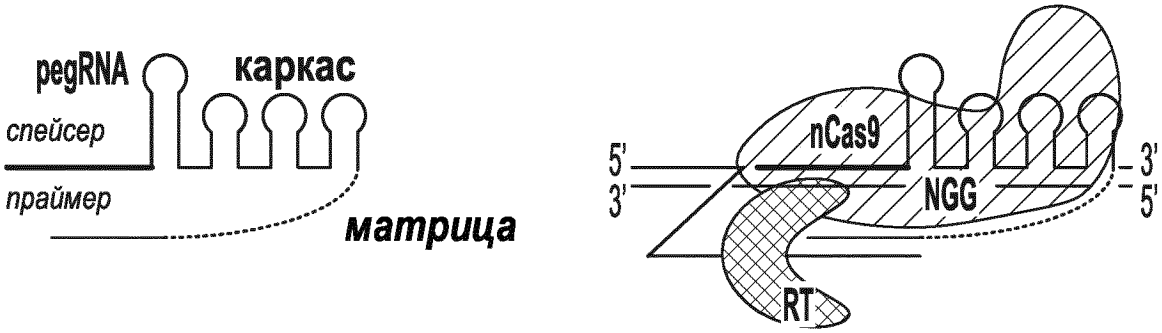
SEQ ID NO: 883 5′ -GTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNN -3′
SEQ ID NO: 884 3′ - NNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAA -5′

Верхние и нижние олиго оттожены и собраны в расщепляемый вектор pU6-pegRNA-GG:

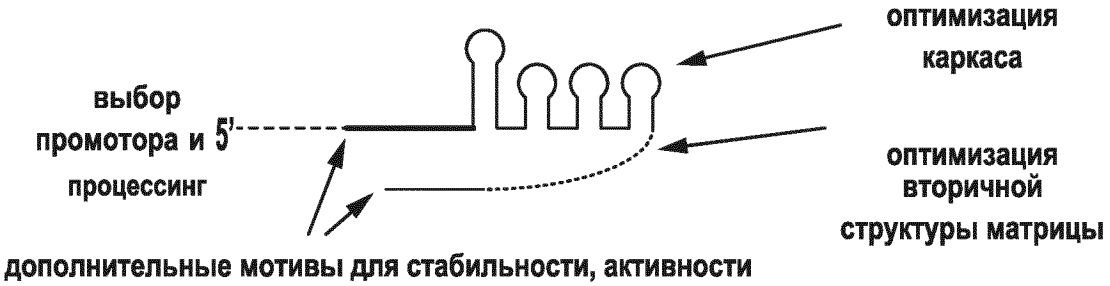
(показанная матрица обратной транскрипции N30 может быть любой длины) Схема конечной сборки вектора экспресии **U6 pegRNA из компонентов 1, 2, 3 и 4**:



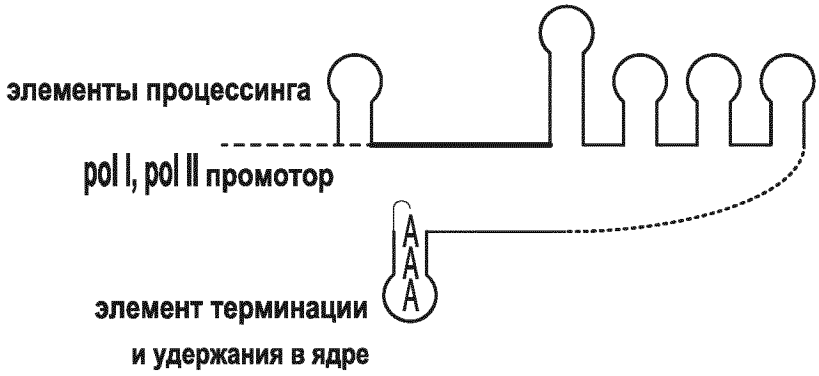
Фиг. 54



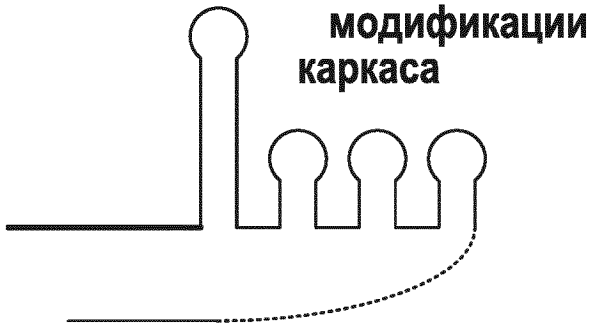
Фиг. 55A



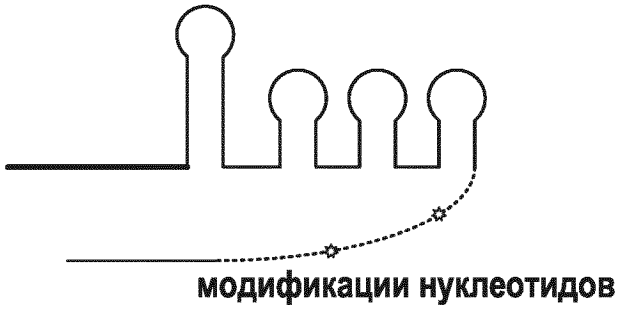
Фиг. 55B



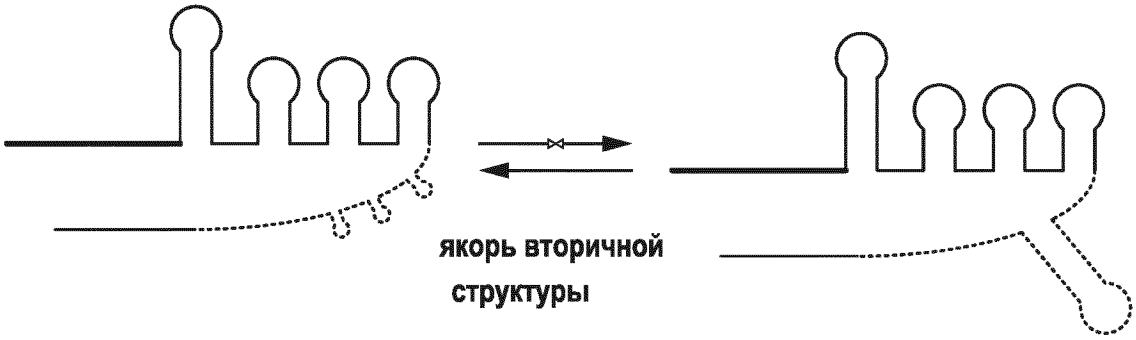
Фиг. 55C



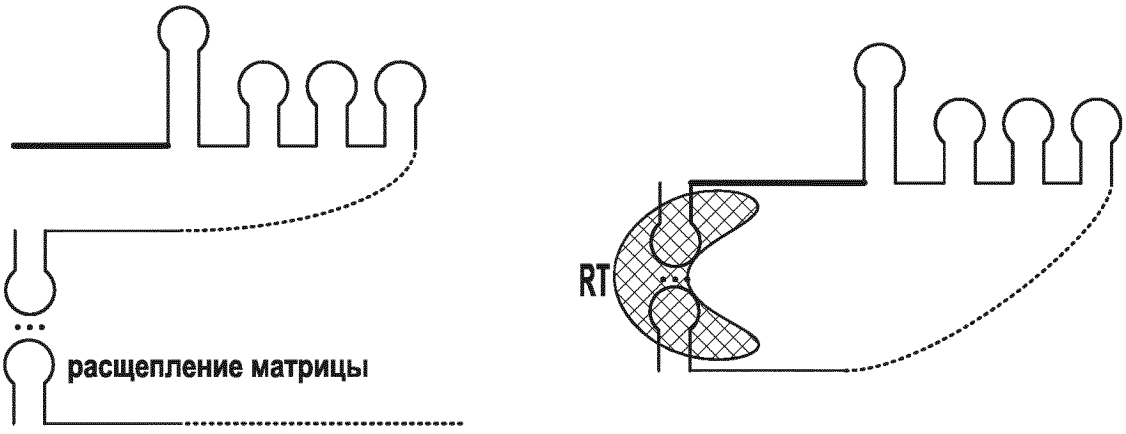
Фиг. 55D



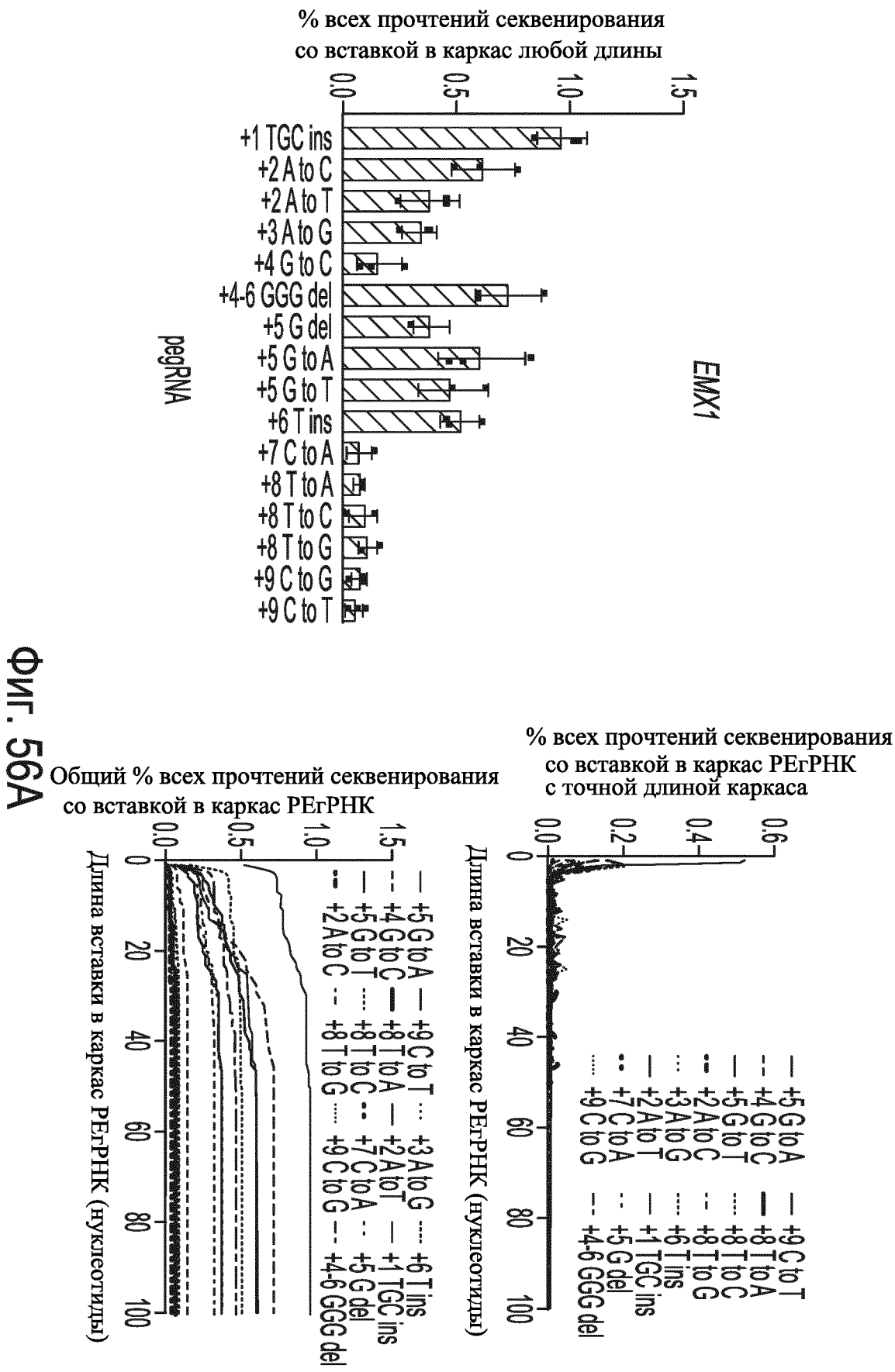
Фиг. 55E



Фиг. 55F



Фиг. 55G

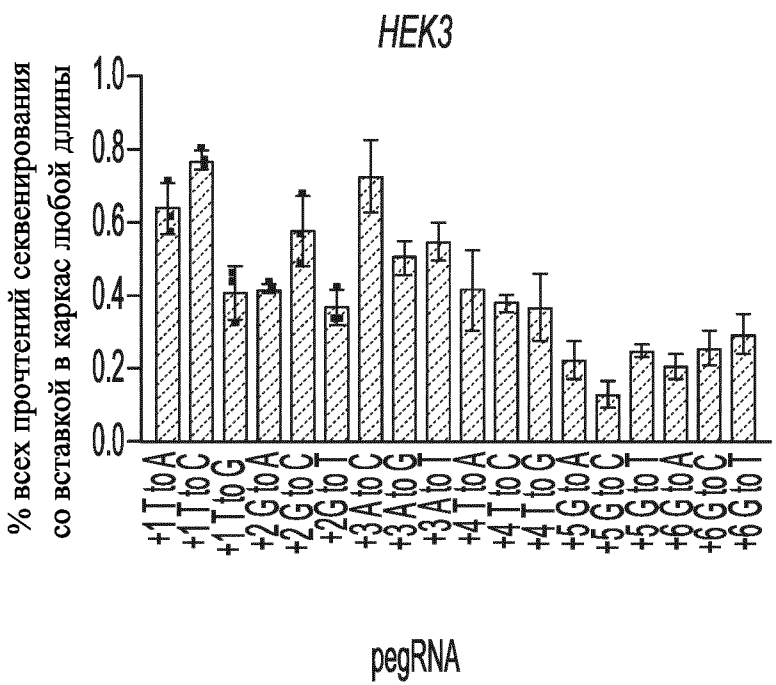

% всех прочтений секвенирования
со вставкой в каркас любой длины
0.0
0.5
1.0
1.5
+1 TGC ins
+2 A to C
+2 A to T
+3 A to G
+4 G to C
+4-6 GGG del
+5 G del
+5 G to A
+5 G to T
+6 T ins
+7 C to A
+8 T to A
+8 T to C
+8 T to G
+9 C to G
+9 C to T
pegRNA
EMX1
% всех прочтений секвенирования
со вставкой в каркас PEгРНК
с точной длиной каркаса
0.0
0.2
0.4
0.6
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+5 G to A
+4 G to C
+5 G to T
+2 A to C
+3 A to G
+2 A to T
+7 C to A
+9 C to G
+9 C to T
+8 T to A
+8 T to C
+8 T to G
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
Общий % всех прочтений секвенирования
со вставкой в каркас PEгРНК
0.0
0.5
1.0
1.5
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+5 G to A
+4 G to C
+5 G to T
+2 A to C
+9 C to T
+8 T to A
+8 T to C
+8 T to G
+3 A to G
+2 A to T
+7 C to A
+9 C to G
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del

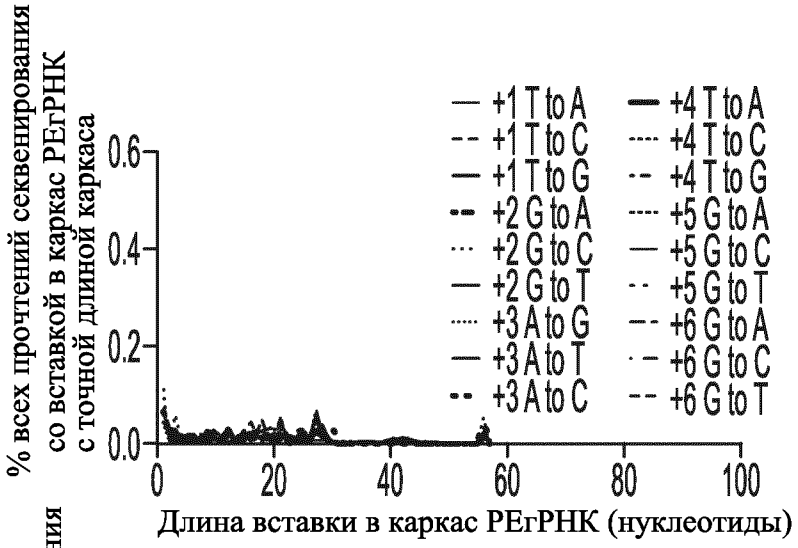
Фиг. 56A

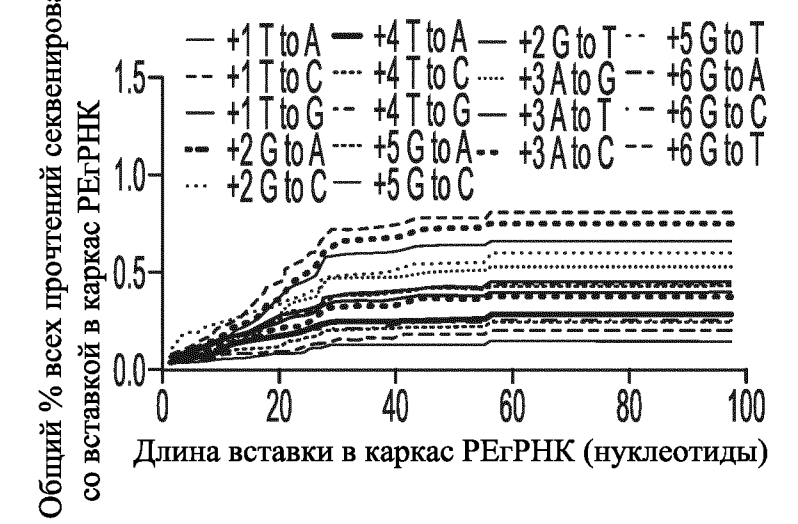
FANCF
% всех прочтений секвенирования со вставкой в каркас любой длины
0
2
4
6
8
+1 A to G
+1 A to T
+2 C to A
+3 C to G
+3 C to T
+3 C ins
+4 T to A
+4 T to G
+4 GAT ins
+5 G to A
+5-7 GGA del
+6 G to C
+6 G del
+7 A to C
+8 T to C
+10 G to T
pegRNA

% всех прочтений секвенирования со вставкой в каркас PEгРНК с точной длиной каркаса
0.0
0.5
1.0
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+5 G to A
+6 G to C
+10 G to T
+7 A to C
+1 A to G
+1 A toT
+2 C to A
+3 C to G
+3 C to T
+4 T to A
+8 T to C
+4 T to G
+3 C ins
+4 GAT ins
+6 G del
+5-7 GGA del
Общий % всех прочтений секвенирования со вставкой в каркас PEгРНК
0
2
4
6
8
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)

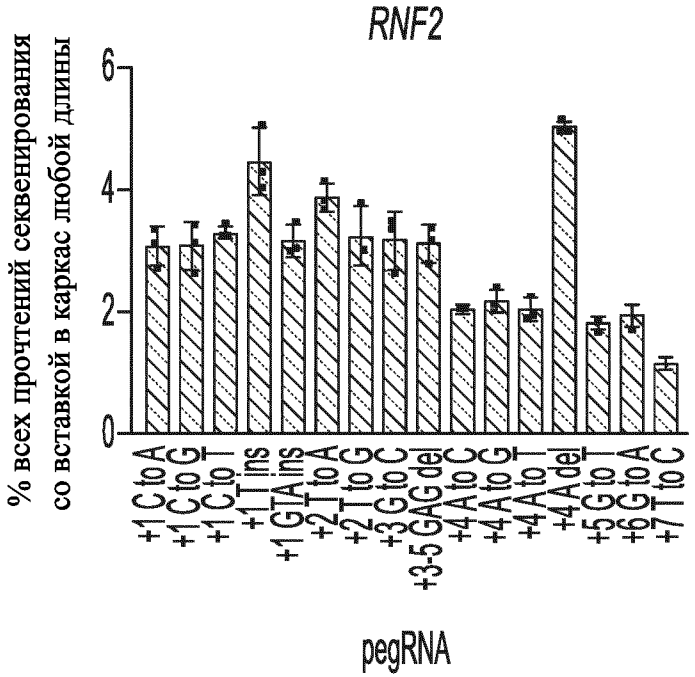
Фиг. 56B

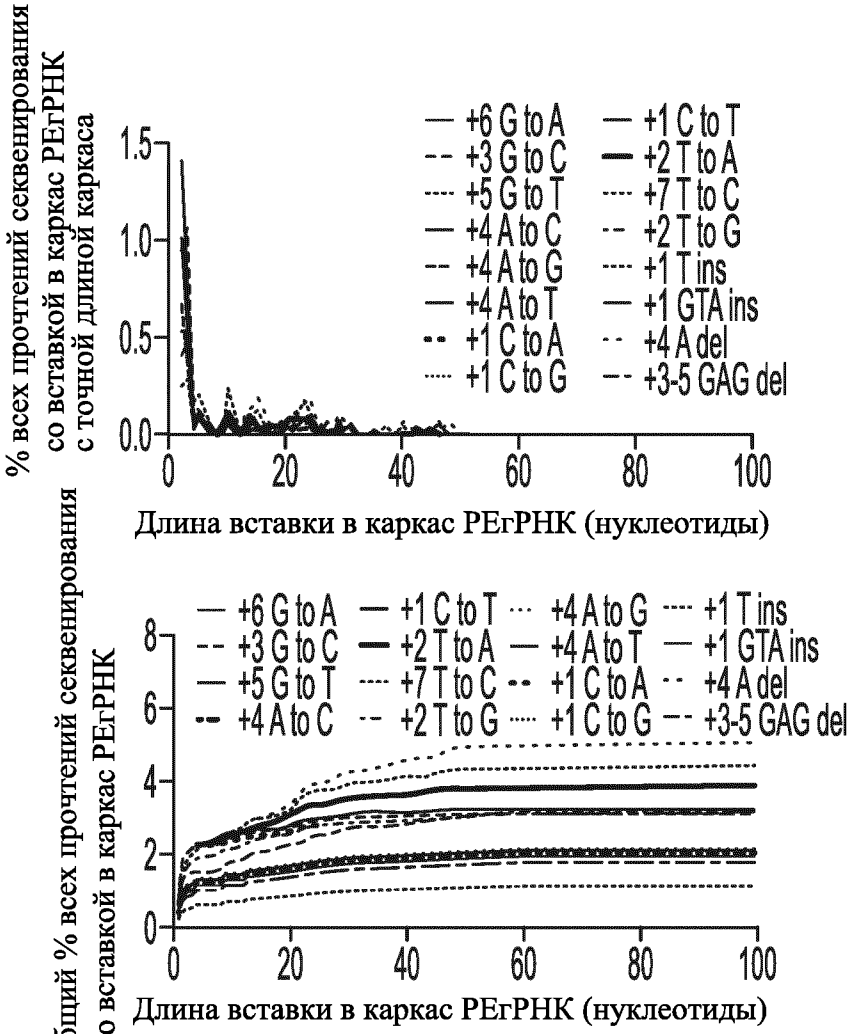

HEK3
% всех прочтений секвенирования со вставкой в каркас любой длины
1.0
0.8
0.6
0.4
0.2
0.0
+1 T to A
+1 T to C
+1 T to G
+2 G to A
+2 G to C
+2 G to T
+3 A to C
+3 A to G
+3 A to T
+4 T to A
+4 T to C
+4 T to G
+5 G to A
+5 G to C
+5 G to T
+6 G to A
+6 G to C
+6 G to T
pegRNA


% всех прочтений секвенирования со вставкой в каркас PEгРНК с точной длиной каркаса
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+1 T to A
+1 T to C
+1 T to G
+2 G to A
+2 G to C
+2 G to T
+3 A to G
+3 A to T
+3 A to C
+4 T to A
+4 T to C
+4 T to G
+5 G to A
+5 G to C
+5 G to T
+6 G to A
+6 G to C
+6 G to T

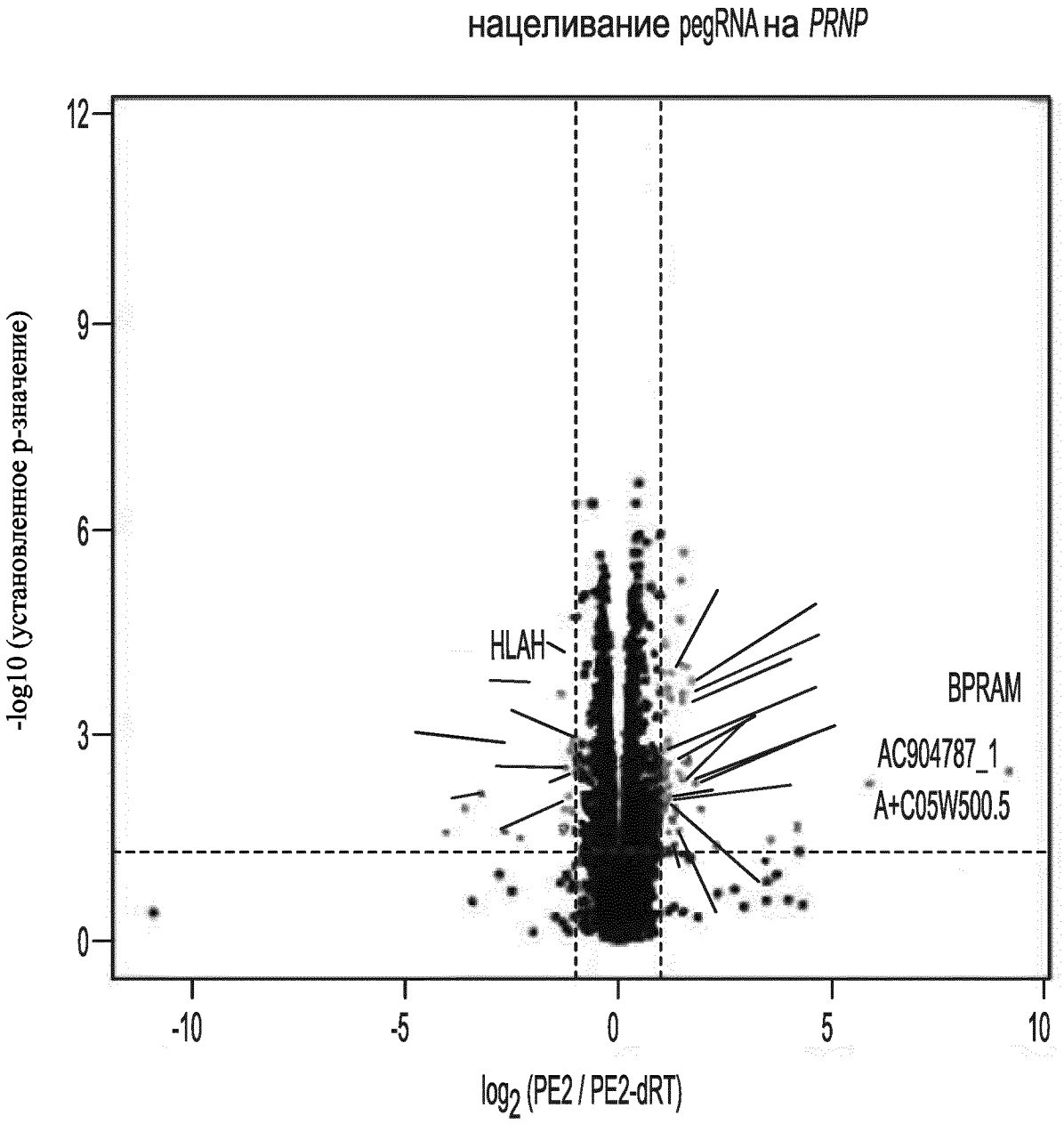

Общий % всех прочтений секвенирования со вставкой в каркас PEгРНК
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+1 T to A
+1 T to C
+1 T to G
+2 G to A
+2 G to C
+4 T to A
+4 T to C
+4 T to G
+5 G to A
+5 G to C
+2 G to T
+3 A to G
+3 A to T
+3 A to C
+5 G to T
+6 G to A
+6 G to C
+6 G to T

Фиг. 56C

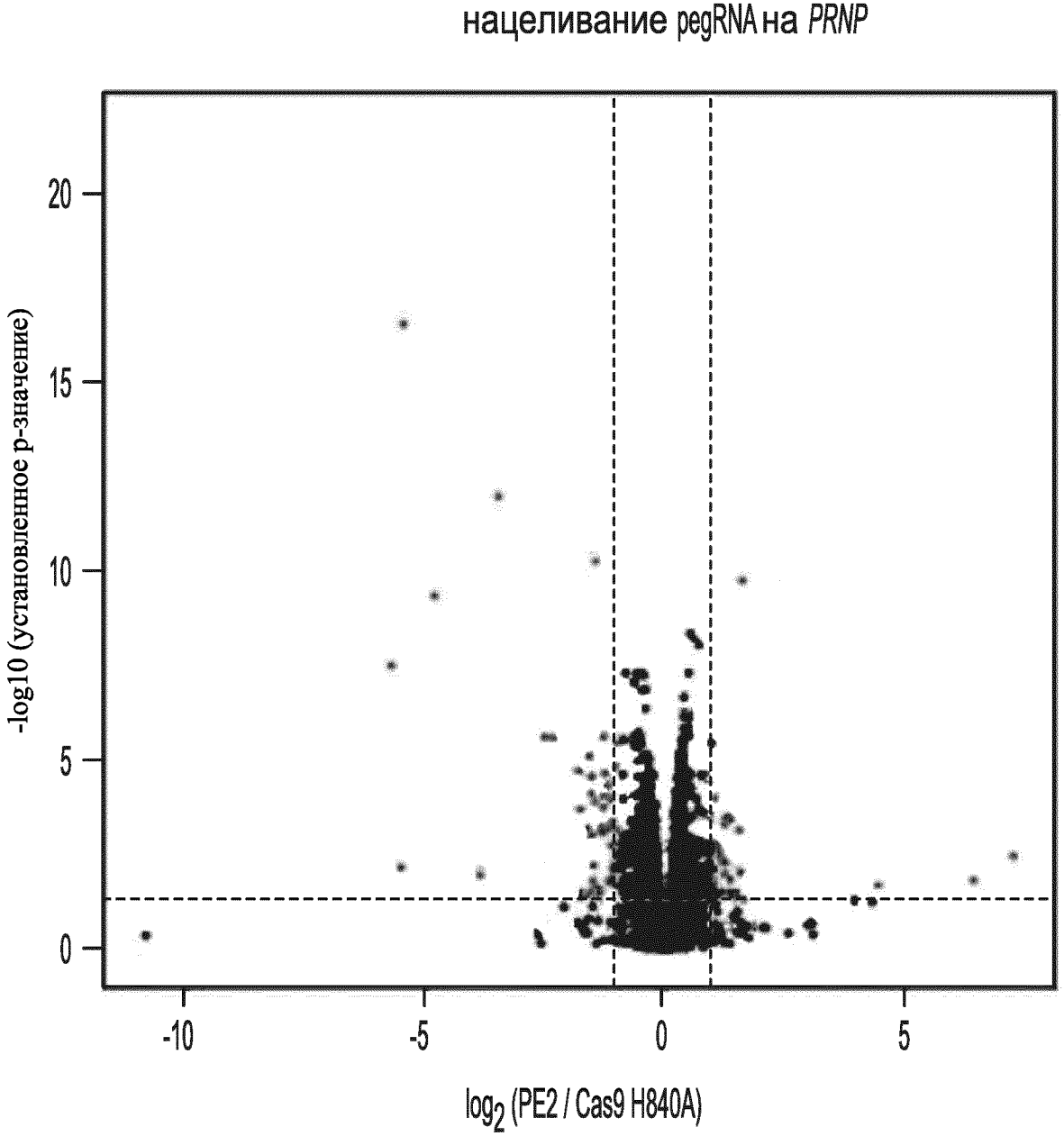

RNF2
% всех прочтений секвенирования со вставкой в каркас любой длины
0
2
4
6
+1 C to A
+1 C to G
+1 C to T
+1 T ins
+1 GTA ins
+2 T to A
+2 T to G
+3 G to C
+3-5 GAG del
+4 A to C
+4 A to G
+4 A to T
+4 A del
+5 G to T
+6 G to A
+7 T to C
pegRNA


% всех прочтений секвенирования со вставкой в каркас PEгРНК с точной длиной каркаса
0.0
0.5
1.0
1.5
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)
+6 G to A
+3 G to C
+5 G to T
+4 A to C
+4 A to G
+4 A to T
+1 C to A
+1 C to G
+1 C to T
+2 T to A
+7 T to C
+2 T to G
+1 T ins
+1 GTA ins
+4 A del
+3-5 GAG del
Общий % всех прочтений секвенирования со вставкой в каркас PEгРНК
0
2
4
6
8
0
20
40
60
80
100
Длина вставки в каркас PEгРНК (нуклеотиды)

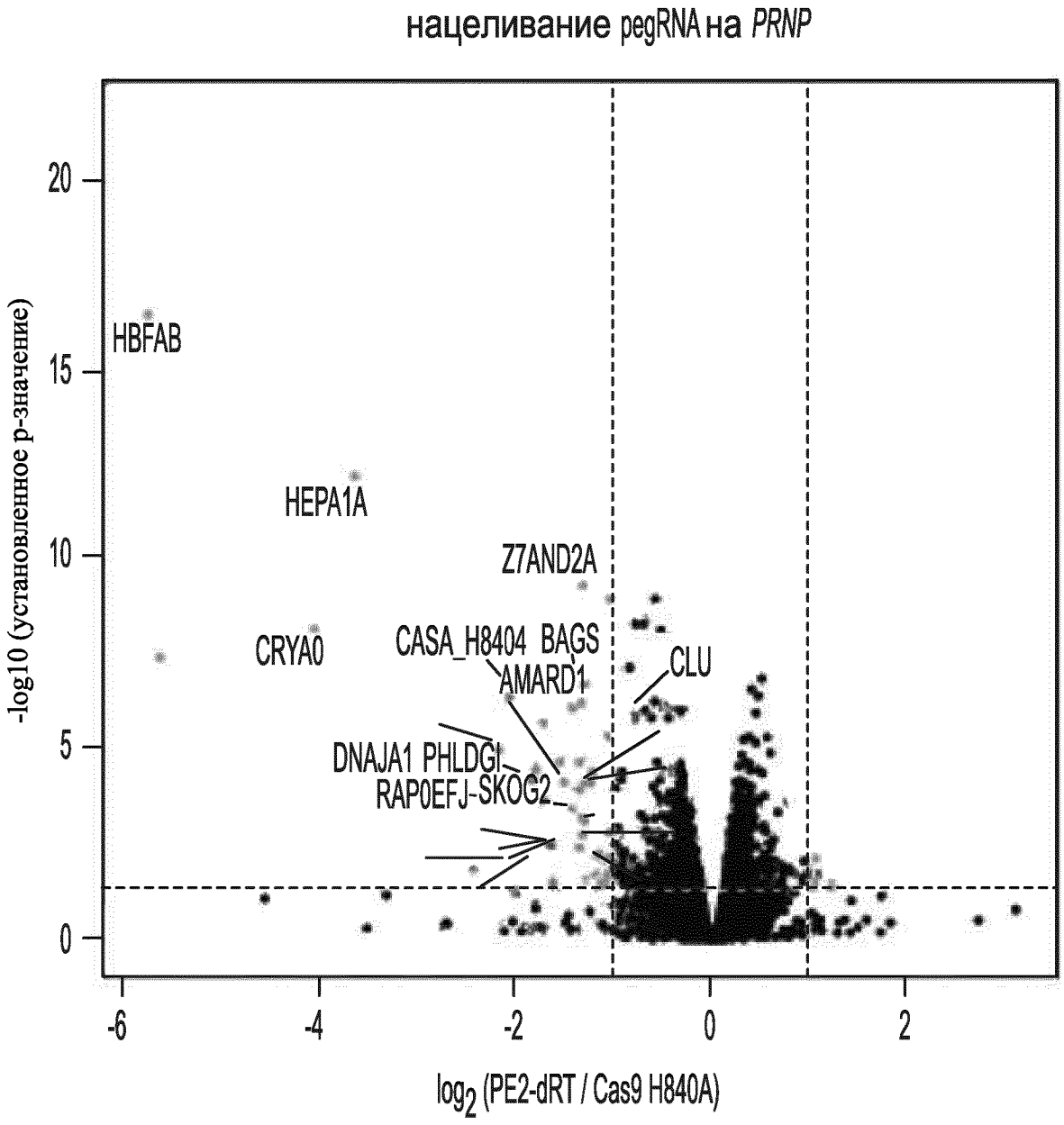
Фиг. 56D


нацеливание pegRNA на PRNP
-log10 (установленное p-значение)
12
9
6
3
0
HLAH
BPRAM
AC904787_1
A+C05W500.5
-10
-5
0
5
10
log2 (PE2 / PE2-dRT)

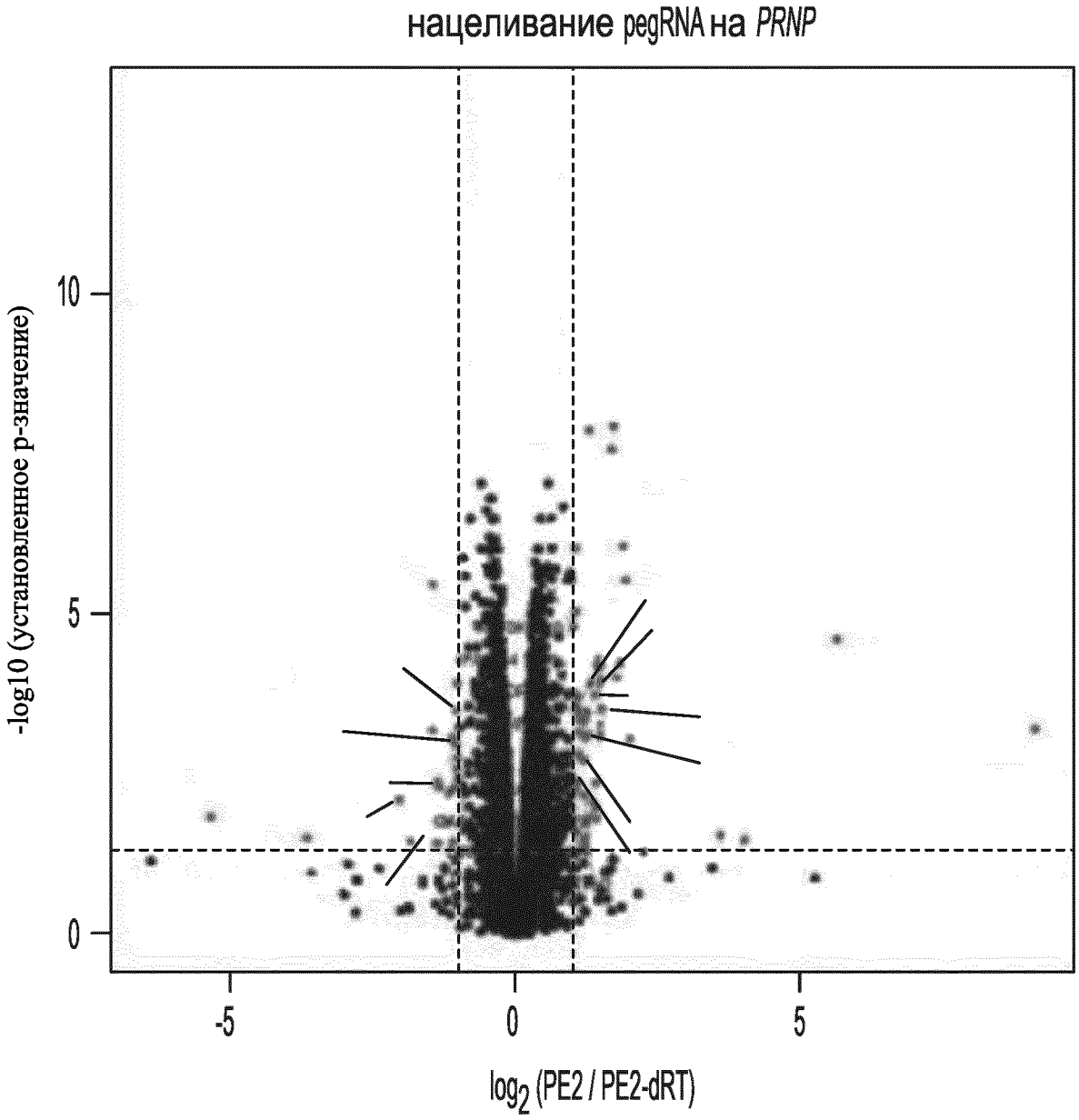
Фиг. 57A


нацеливание pegRNA на PRNP
-log10 (установленное p-значение)
20
15
10
5
0
-10
-5
0
5
log2 (PE2 / Cas9 H840A)

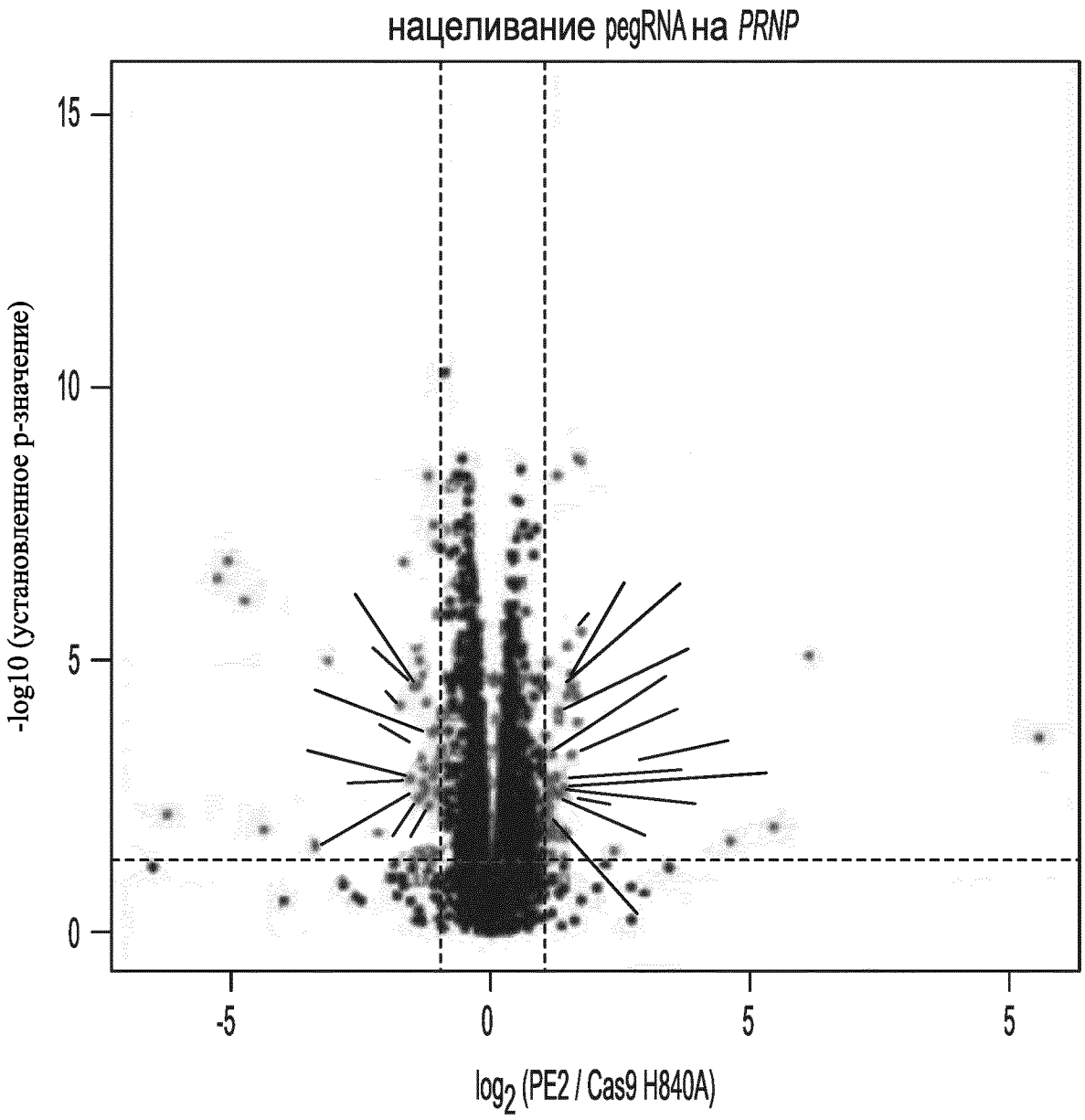
Фиг. 57B


нацеливание pegRNA на PRNP
-log10 (установленное p-значение)
20
15
10
5
0
HBFAB
HEPA1A
Z7AND2A
CRYA0
CASA_H8404
BAGS
AMARD1
CLU
DNAJA1
PHLDGI
RAP0EFJ-SKOG2
-6
-4
-2
0
2
log2 (PE2-dRT / Cas9 H840A)

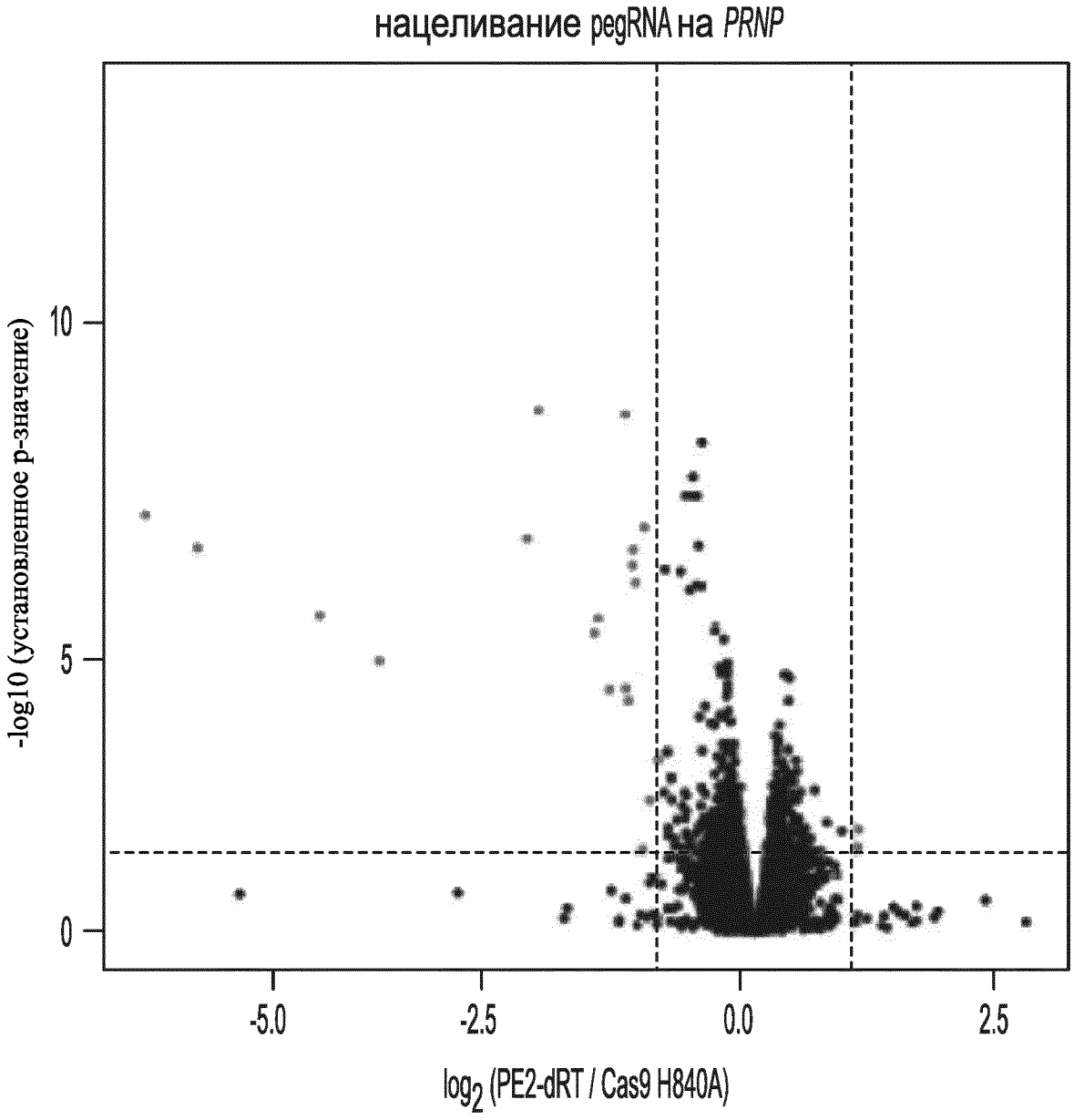
Фиг. 57C


нацеливание pegRNA на PRNP
-log10 (установленное p-значение)
10
5
0
-5
0
5
log2 (PE2 / PE2-dRT)

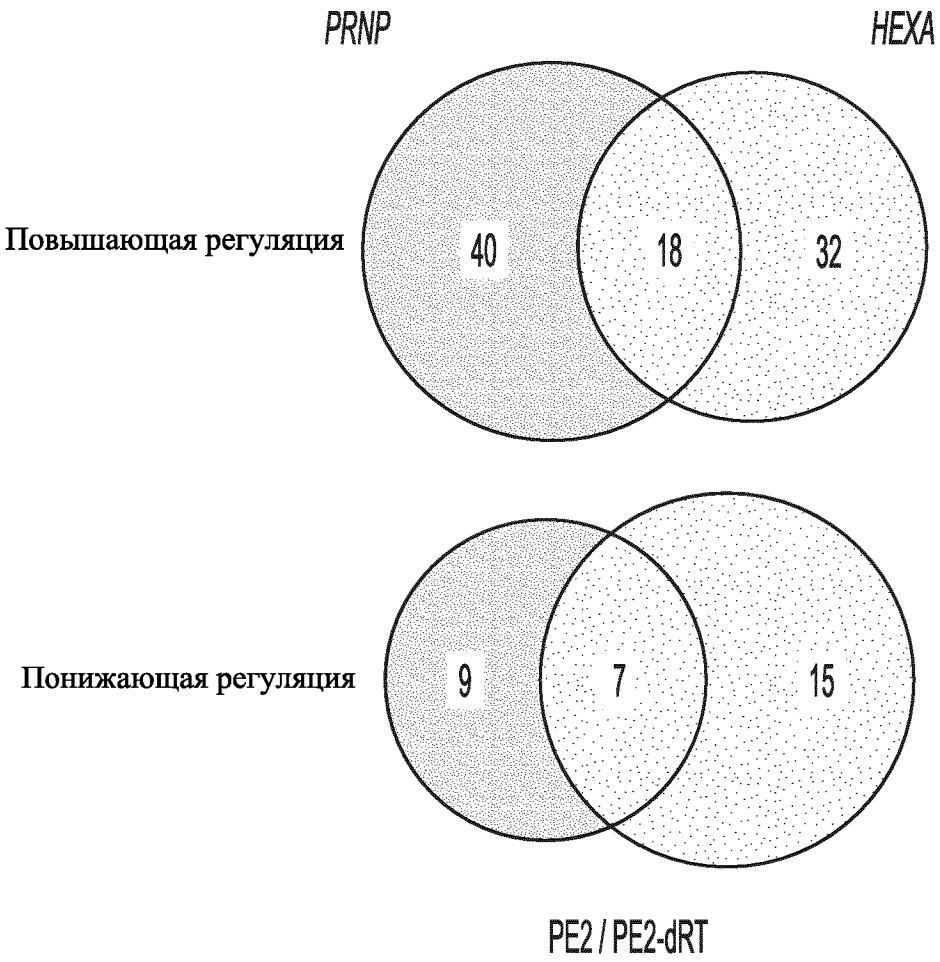
Фиг. 57D


нацеливание pegRNA на PRNP
15
10
5
0
-log10 (установленное p-значение)
-5
0
5
5
log2 (PE2 / Cas9 H840A)

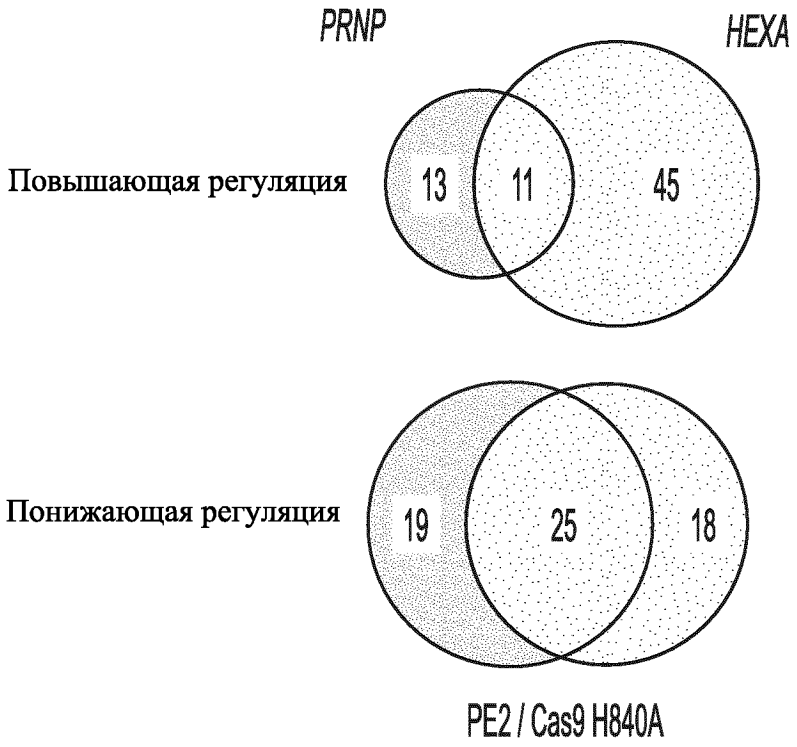
Фиг. 57E


нацеливание pegRNA на PRNP
-log10 (установленное p-значение)
10
5
0
-5.0
-2.5
0.0
2.5
log2 (PE2-dRT / Cas9 H840A)

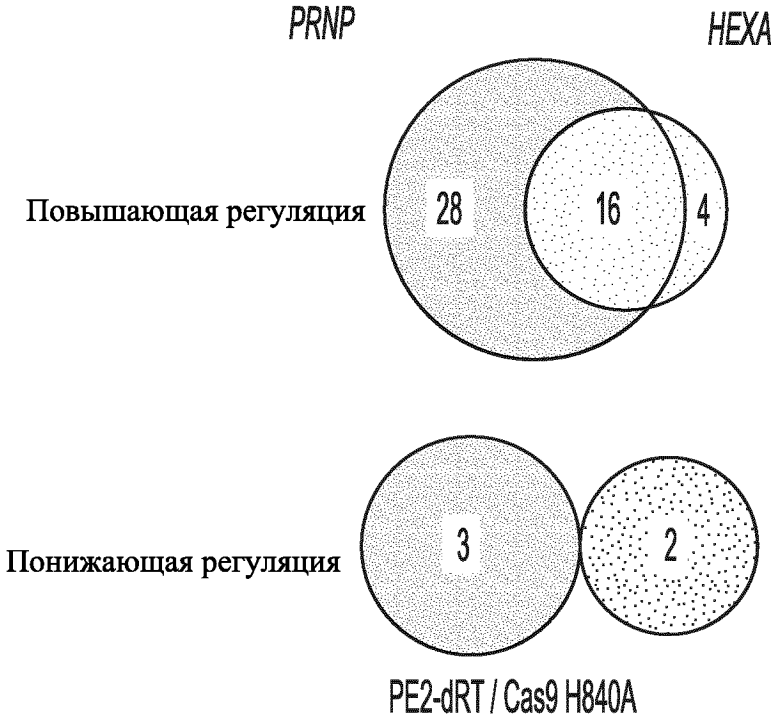
Фиг. 57F


PRNP
HEXA
Повышающая регуляция
40
18
32
Понижающая регуляция
9
7
15
PE2 / PE2-dRT

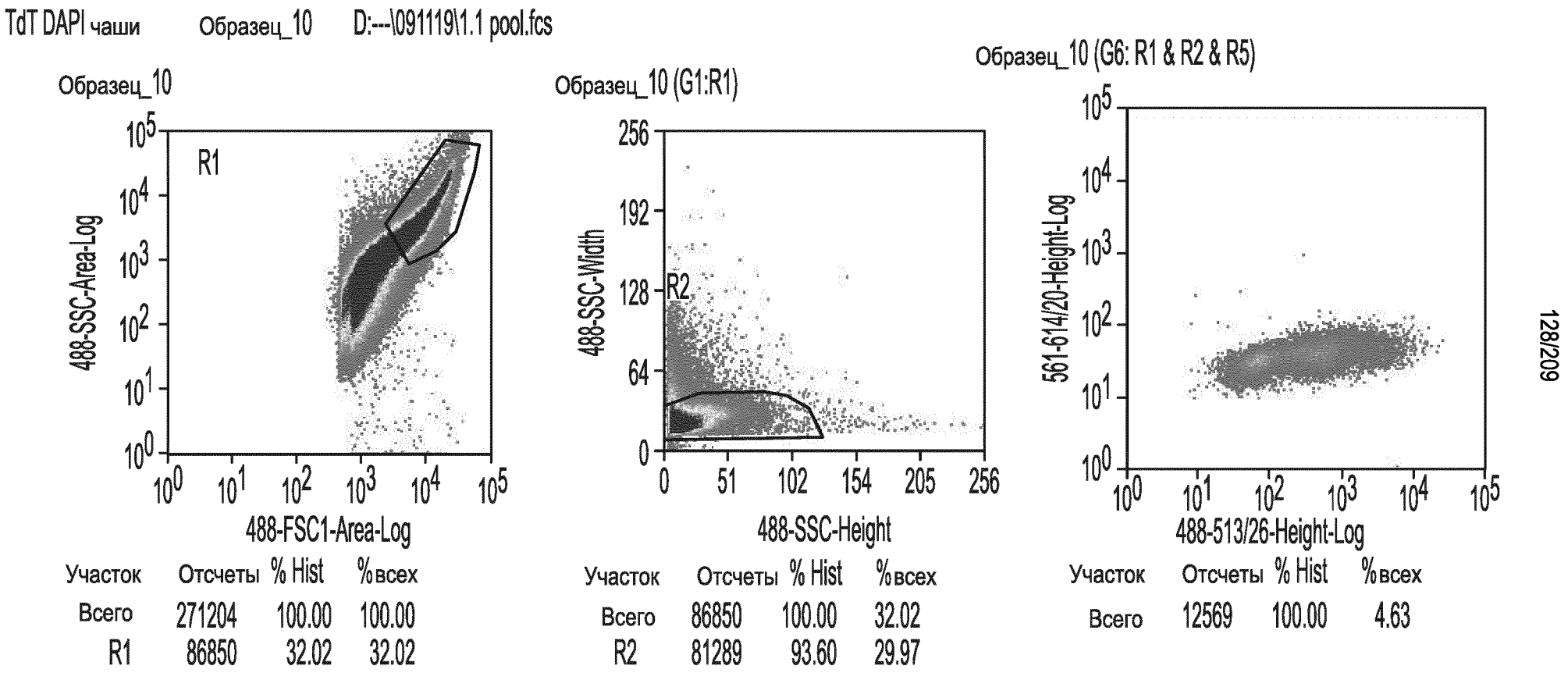
Фиг. 57G



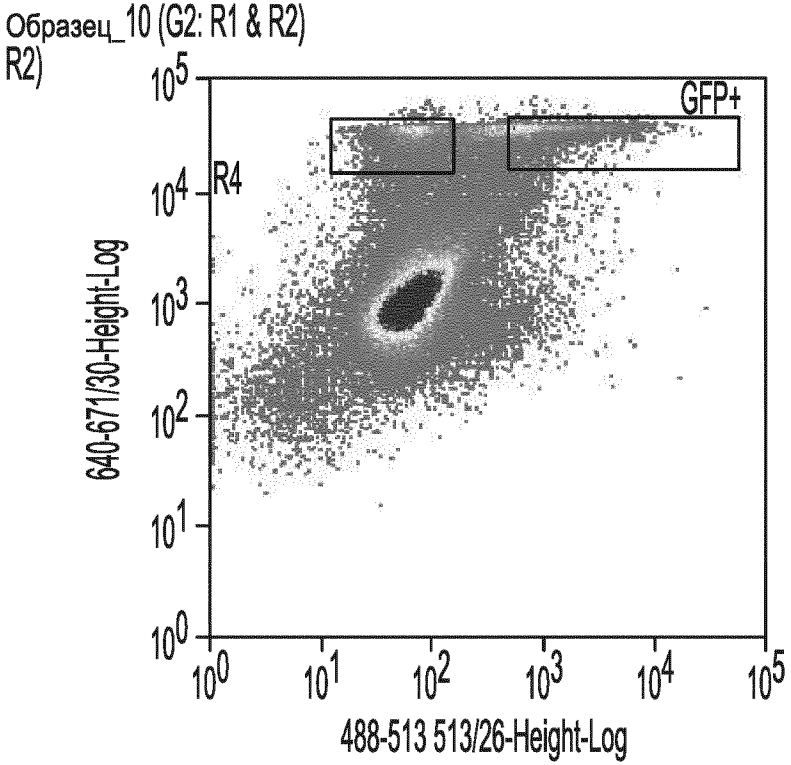
Фиг. 57H



Фиг. 57I

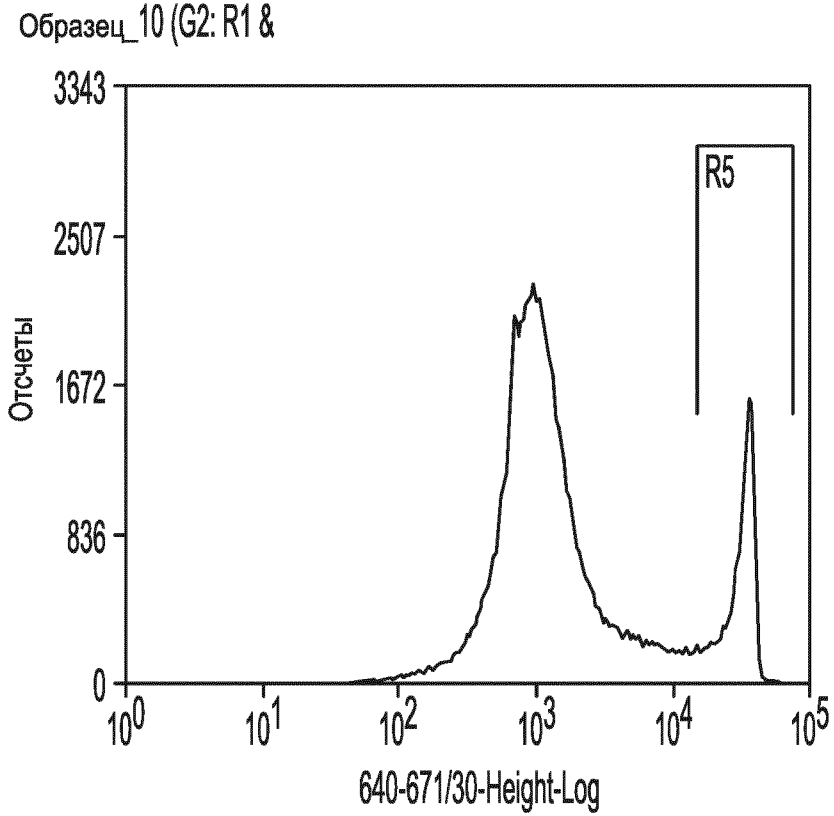

TdT DAPI чаши Образец_10 D:---\091119\1.1 pool.fcs
Образец_10
R1
488-SSC-Area-Log
488-FSC1-Area-Log
Участок Отсчеты % Hist %всех
Всего 271204 100.00 100.00
R1 86850 32.02 32.02
Образец_10 (G1:R1)
R2
488-SSC-Width
488-SSC-Height
Участок Отсчеты % Hist %всех
Всего 86850 100.00 32.02
R2 81289 93.60 29.97
Образец_10 (G6: R1 & R2 & R5)
561-614/20-Height-Log
488-513/26-Height-Log
Участок Отсчеты % Hist %всех
Всего 12569 100.00 4.63

Фиг. 58A



Участок	Отсчеты	% Hist	%всех	Связи	Отсчеты
Всего	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 01:36:4 1PM Page 1



Участок	Отсчеты	% Hist	Среднее	CV
Всего	81289	100.00	6521.68	179.64
R5	12569	15.46	32516.62	24.27

Фиг. 58B

Обзор клонирования

1. **Расщепляемый плазмидный вектор pU6-pegRNA-GG(компонент 1) с** *BsaI*, и выделенный плазмидный фрашмент (~2.2т.о.) содержит ориджин репликации, U6 промотор, U6 poly-T терминация, и ген Amp^R
2. Порядок олигонуклеотидов для:
   a. Требуемый спейсер (нацеленный), фланкированный указанными выступами (компонент 2)
      i. Применение последовательности требуемой мишени 5-3для олигонуклеотидов верхней цепи (включая 5' CACC и 3' GTTTT выступы)и применение обратного комплемента для олигонуклеотидов нижней цепи(включая выступ 5' CTCTAAAAC). Последовательность спейсера должна иметь G нуклеотид для эффективной инициации транскрипции.
   b. Матрица требуемого 3' удлинения pegRNA, фланкированная указанными выступами (компонент 3)
      i. Применение смысловой цепи РНК в качестве олигонуклеотидов верхней цепи (признаки 5 GTCC выступа)и применение обратного этой последовательности для олигонуклеотидов нижней цепи (призаки 5' AAAA выступа).
   c. SpCas9 sgRNA последовательность каркаса, соответствующая выступам golden gate (компонент 4)
      i. Эти олигонуклеотиды не являются полной последовательностью каркаса , так как выступы оставшихся компонентов приводят к нескольким пропущенным нуклеотидам
      ii. Примечание:эти олигонуклеотиды должны быть 5' фосфорилированными. Олигонуклеотиды могут быть 5' фосфорилированы производителем ИЛИ 5' фосфорилированы ферментативно посредством T4 PNK (см. протокол ниже)
3. Отжиг верхних и нижних олигонуклеотидов для компонентов 2, 3 и 4 в отдельных реакциях отжига согласно протоколу ниже.Если SpCas9 sgRNA последовательность каркаса (компонент 4) не фосфорилирована, фосфорилируют с T4 PNK.
4. Сборка Golden Gate **выделенного фрагмента 2.2-т.о. компонента 1 с компонентами 2, 3 и 4**
5. **Трансформация продукта лигирования с** ***E. coli***. Антибиотическая устойчивость, подтвержденная компонентом 1 из плазмидного вектора pU6- pegRNA-GG, представляет собой устойчивость к ампициллину и карбенициллину.
6. Выделение и секвенирование плазмид из полученных клональных трансформантов

Фиг. 59A

Протокол клонирования pegRNA

Стадия 1: Расщепление плазмидного вектора pU6-pegRNA-GG (компонент 1)

Объединение следующего в пробирке для ПЦР:

2000 нг pU6-pegRNA-GG-вектор (компонент 1)	X мкл
Bsa1-HFv2 (NEB)	1.0 мкл
10x Cutsmart буфер	3.0 мкл
H_2O	до 30.0 мкл
Общий объем реакции	

Инкубирование при 37 °C в течение 4-16 часов Выделение фрагмента ~2.2-т.о. из разрезанной плазмиды.	30.0 мкл

Фиг. 59B

Стадии 2 и3: Порядок и отжиг олигонуклеотидных частей (компоненты 2, 3 и 4)

Материалы

Буфер отжига: H_2O добавка с 10 мМTris-Cl pH 8.5 и 50 мМ NaCl
Комплементарные олигонуклетидные пары

Протокол

Объединение следующего в пробирке для ПЦР:

Верхние олигонуклеотиды, 100 мкМ	1.0 мкл
Нижние олигонуклеотиды, 100 мкМ	1.0 мкл
Буфер для отжига (компоненты 2, 3 и 4)	23.0 мкл
общий объем реакции	25.0 мкл

В термоциклере,нагрев при 95 °С в течение 3 минут,затем охлаждение постепенно (0.1 °С/с) до 22 °С
Разбавление отоженных олигонуклеотидов 1:4 путем добавления 75 мкл H_2O. Конечная концентрация каждого олигонуклеотида составит 1 мкМ после этого разбавления. Не разбавлять каркас sgRNA (компонент 4), в случае фосфорилирования PNK на стадии 2.5.

Фиг. 59C

Стадия 2.b.ii.: фосфорилирование каркаса sgRNA (необходимы, если олигонуклеотиды фосфорилированы)

Протокол

Объедините следующее в пробирке для ПЦР:

4 мкМ дуплекса олигонуклеотидов со стадии 1	6.25 мкл
10x T4 буфер лигазы ДНК (NEB)	2.50 мкл
T4 PNK (NEB)	0.50 мкл
H_2O	15.75 мкл
Общий объем реакции	20.0 мкл

В термоциклере, инкубация при 37 °С в течение 60 минут
После фосфорилирования, отоженные олигонуклеотиды каркаса находятся при концентрации 1 мкМ.
Переход к стадии 3.

Фиг. 59D

Стадия 4: сборка peg RNA pegRNA,
Golden Gate реакция сборки

Расщепленный pU6-pegRNA-GG плазмидный вектор - Предварительно разрезанный, выделенный фрагмент 2,2т.о.	1.00 мкл при 30 нг/мкл
Отоженные олигонуклеотиды протоспейсера (компонент 2)	1.00 мкл при 1 мкМ
Отоженные олигонуклеотиды pegRNA 3'-удлинения (компонент 4)	1.00 мкл при 1 мкМ
отоженные олигонуклеотиды каркаса sgRNA (компонент 3) - олигонуклеотиды должны бфть фосфорилированы	1.00 мкл при 1 мкМ
*Bsa*I-HFv2 (NEB)	0.25 мкл
T4 DNA лигаза (NEB)	0.50 мкл
10x T4 буфер ДНК лигазы (NEB)	1.00 мкл
H_2O	4.25 мкл
Общий объем реакции:	10.0 мкл

Инкубация при комнатной температуре в течение 10 мин После инкубации , проводят следующую программу в термоциклере:
5 мин при 37 °C, затем 5мин при 80 °C, затем поддержание 12 °C

Фиг. 59E

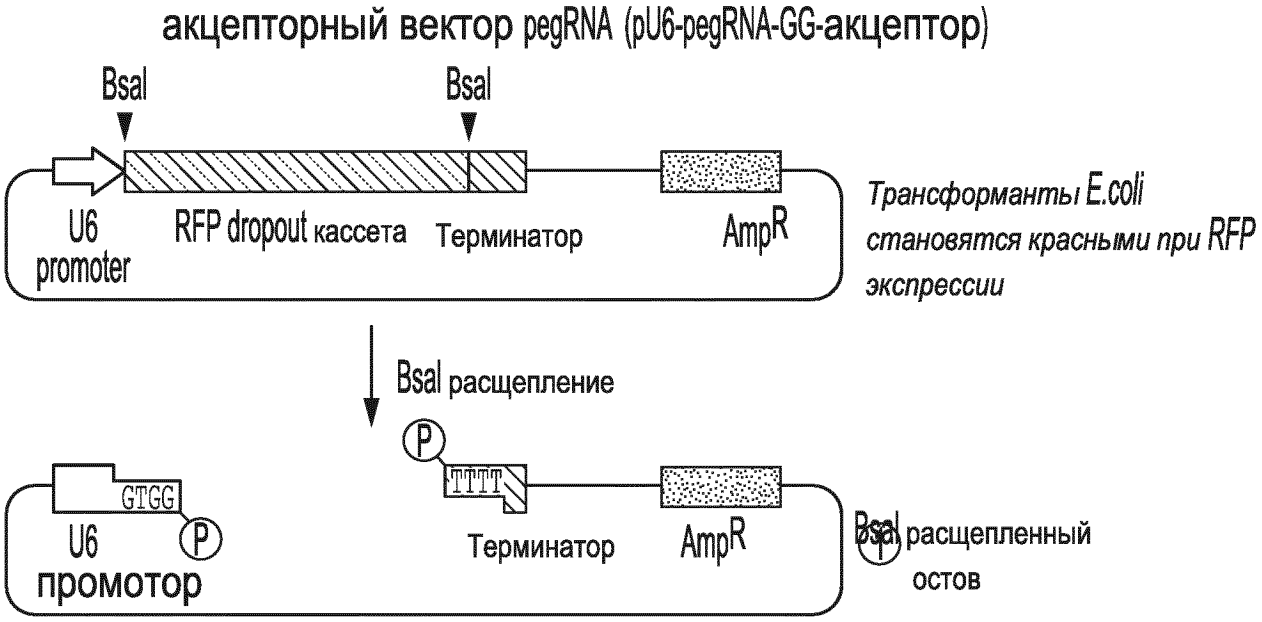
Стадии 5 и 6: Трансформация собранных плазмид

Трансформация 1 мкл из 10 мкл реакции сборки в 10мкл компентентных клеток. Желаемые трансформанты устойчивы к ампициллину и карбенициллину.

Протокол клонирования pegRNA приведен на следующей диаграммеl.

Фиг. 59F

Клонирование pegRNA посредством сборки Golden Gate



② отоженные олиго спейсера pegRNA

SEQ ID NO: 879

5′ CACC*NNNNNNNNNNNNNNNNNNNN**GTTTT*** 3′
3′ *NNNNNNNNNNNNNNNNNNNN**CAAAATCTC*** 5′

SEQ ID NO: 880

③ отоженные олиго pegRNA 3′ удлинения

SEQ ID NO: 883

5′ ***GTGC**NNNNNNNNNNNNNNNNNNNNNNNNNNNN* 3′
3′ *NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAA* 5′

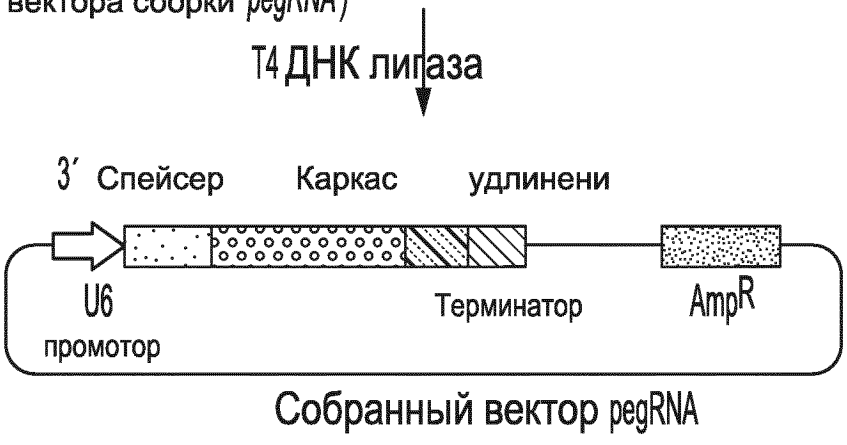
SEQ ID NO: 884

④ Отоженные олиго каркаса pegRNA (5′ фосфорилированные)

SEQ ID NO: 881

5′ [Phos]-***AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG*** 3′
3′ ***GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCACG***-[Phos] 5′

SEQ ID NO: 882 (*эти олигонуклеотиды должны быть* **5′** фосфорилированы для обеспечения вектора сборки *pegRNA*)



Фиг. 59G

```
## sgRNA scaffold sequence search ##

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')

#reads the fastq files into a dictionary with the file names as keys
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))]

#the referenced sequence to be searched for is entered into the following dictionary with
#an appropriate key
scaffdict =
{'HEK3':'CAGAGGACCGACTCGGTCCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACTCACGTGCTCAGTCTGGGCCGGTG', (SEQ ID NO: 885)
'EMX1':'ATCACGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TGCTATTTCTAGCTCTAAAACTTCTTCTTCTGCTCGGACTCGGTG', (SEQ ID NO: 886)
'FANCF':'TTTCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAAC
TTGCTATTTCTAGCTCTAAAACGGTGCTGCAGAAGGGATTCCGGTG', (SEQ ID NO: 887)
'RNF2':'TCGTTGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACCAGGTAATGACTAAGATGACGGTG'} (SEQ ID NO: 888)

#matches and counts iterative slices of the reference string to the appropriate fastq files
#reference key must be contained in the name of the fastq file
#generated values represent cumulative counts for a minimum degree of sgRNA integration
#i.e. a given value x means x reads contain y or more bases of the scaffold
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. 60A

```
                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

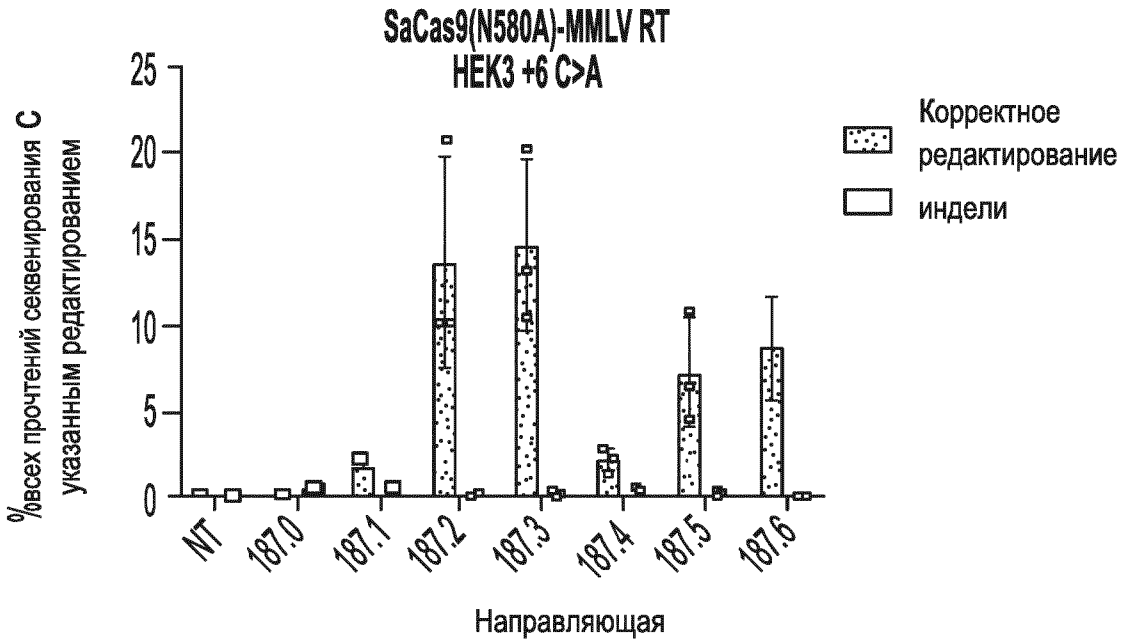
#writes the results into a dataframe indexed from 1
resultdf = pd.DataFrame.from_dict(resultdict)
resultdf = resultdf.reindex(sorted(resultdf.columns), axis=1)
resultdf.index = range(1,len(resultdf)+1)

#converts the cumulative count values into specific counts
#i.e. a given value x means x reads contain exactly y bases of the scaffold
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]

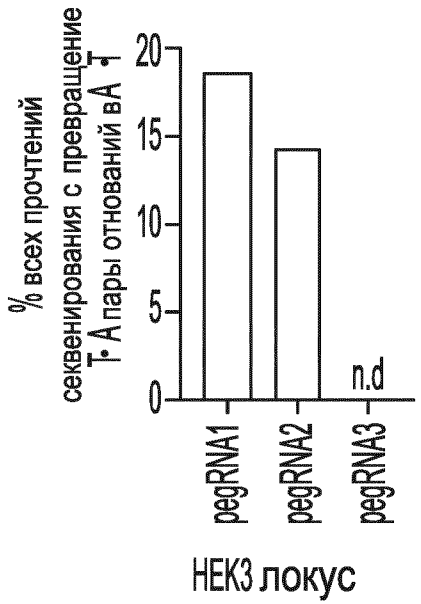
#converts the specific counts values into frequencies
resultdf3=resultdf2.copy()
for entry in resultdf3:
    resultdf3[entry]=resultdf2[entry].div(resultdf[entry][1])*100

#reads the results into excel files
resultdf.to_excel('cumulativecounts.xlsx')
resultdf2.to_excel('specificcounts.xlsx')
resultdf3.to_excel('specificfrequencies.xlsx')
```

Фиг. 60B



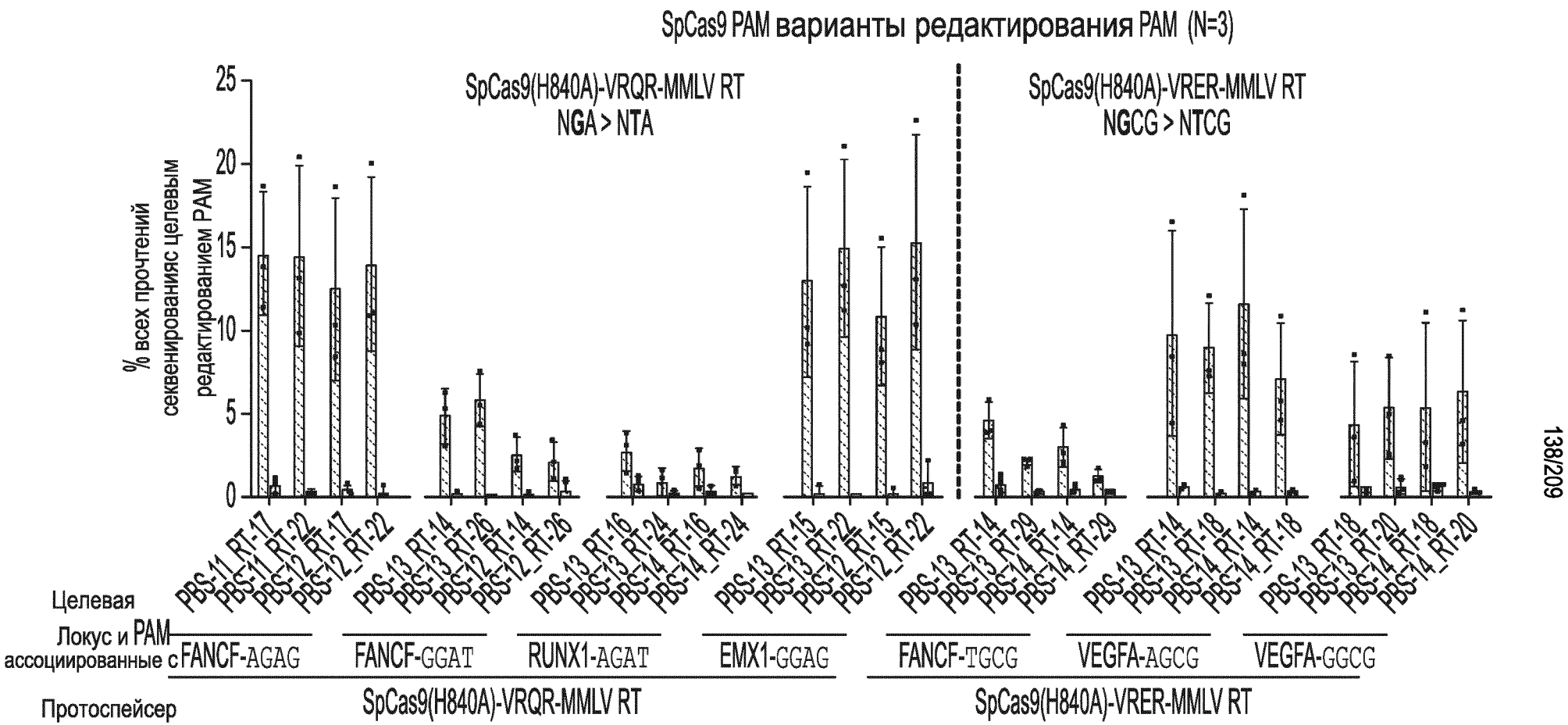
Фиг. 61



Фиг. 62A


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
Целевая пара оснований T·A для редактирования
Протоспейсер B
(SEQ ID NO: 891)
pegRNA 3

Фиг. 62B

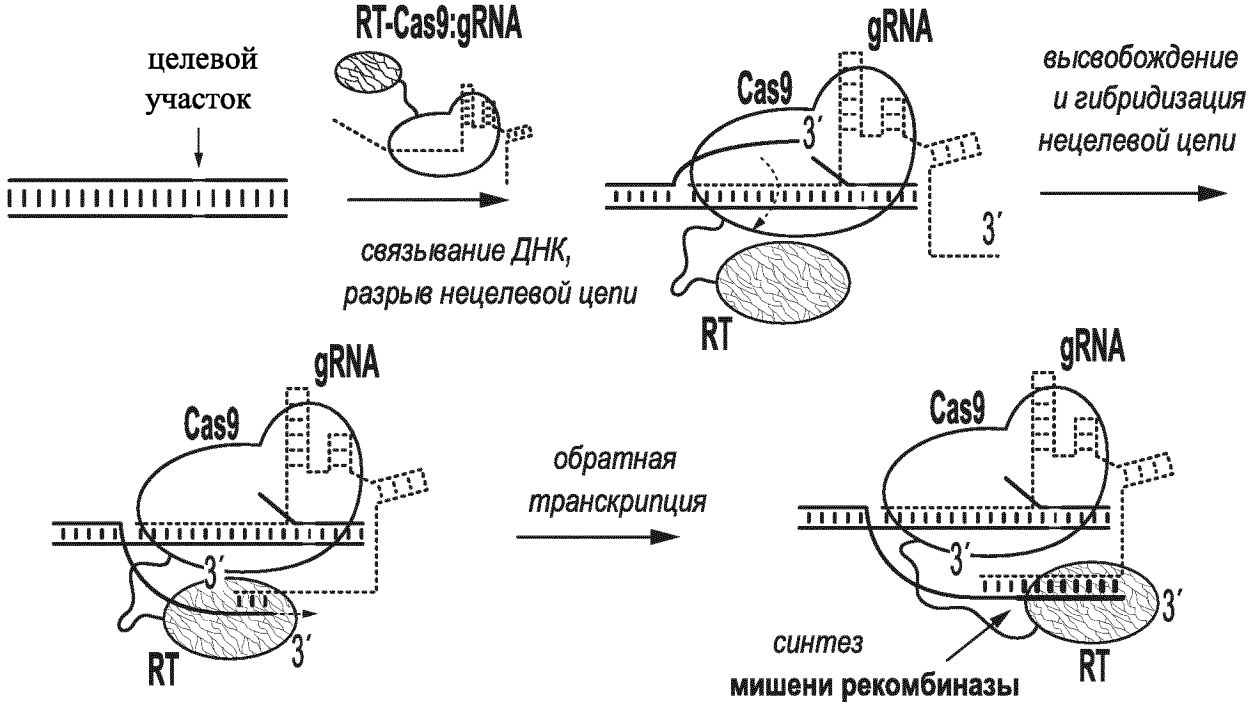


pegRNA , длина участка связывания праймера (PBS) и длина матрицы RT (RT) и используемая система праймированного редактирования

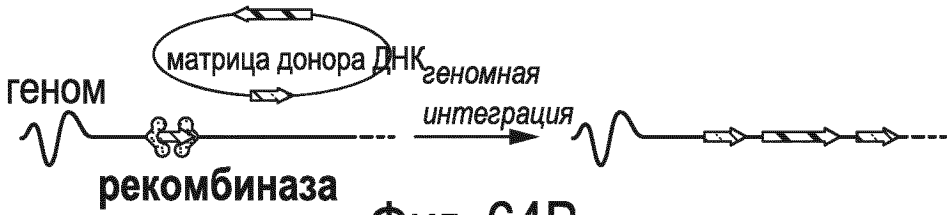
Процент G>T редактирования от общег

Прочтения инделей от общих прочтений

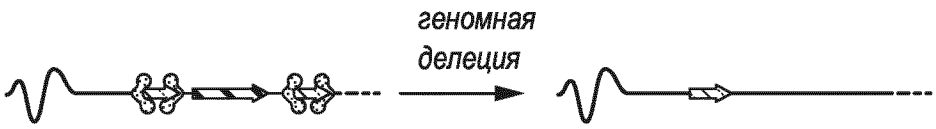
Фиг. 63



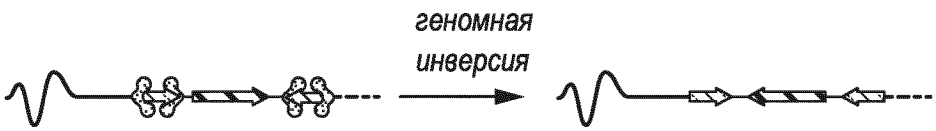
Фиг. 64A



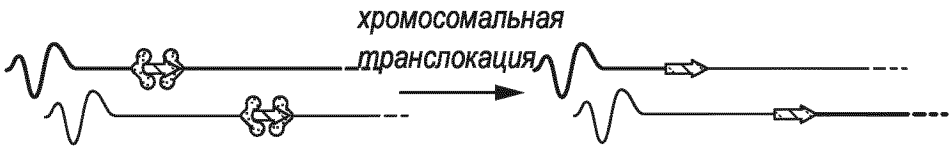
Фиг. 64B



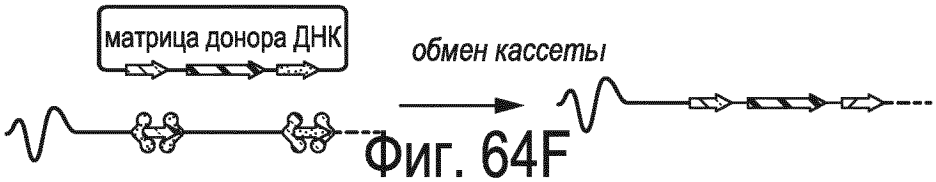
Фиг. 64C



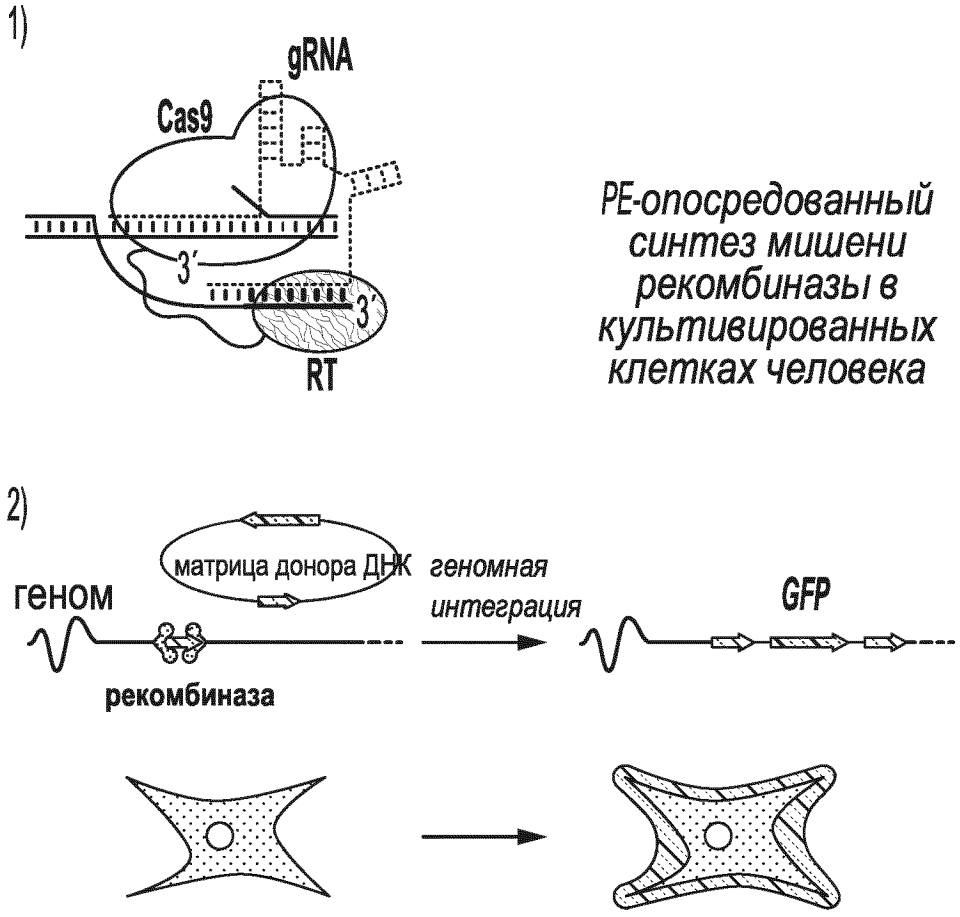
Фиг. 64D



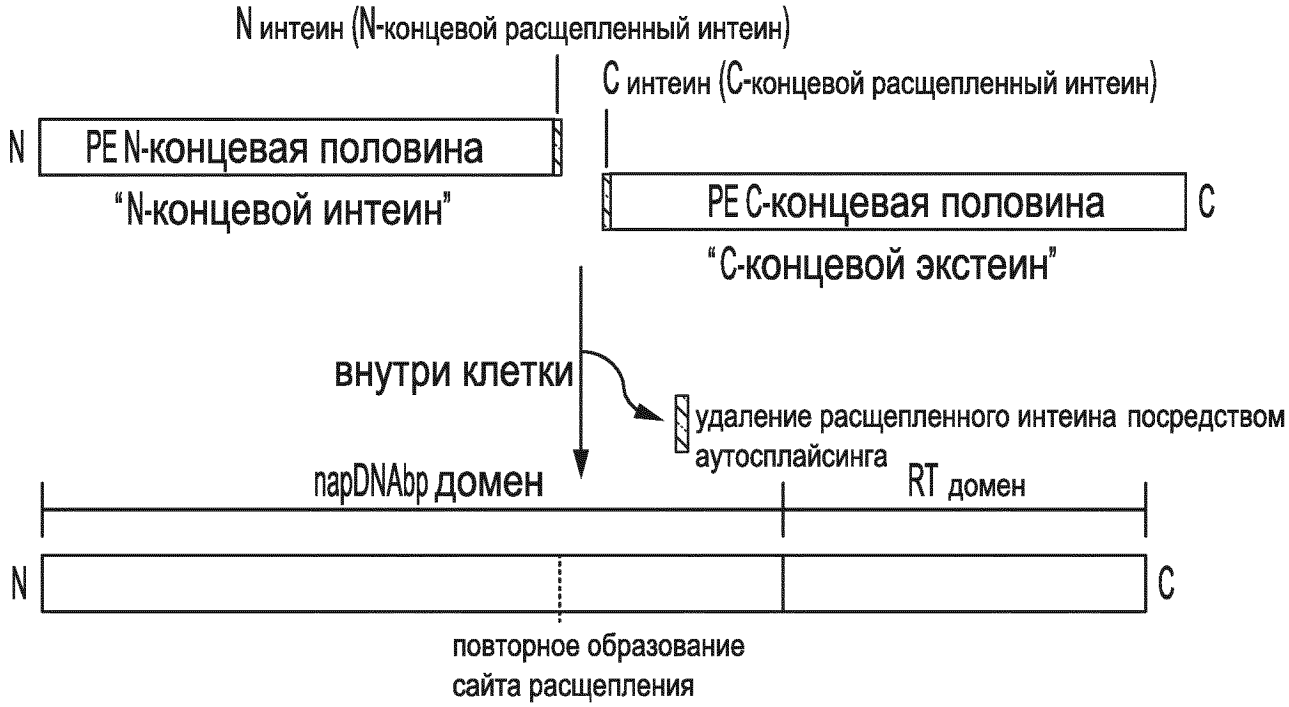
Фиг. 64E



Фиг. 64F

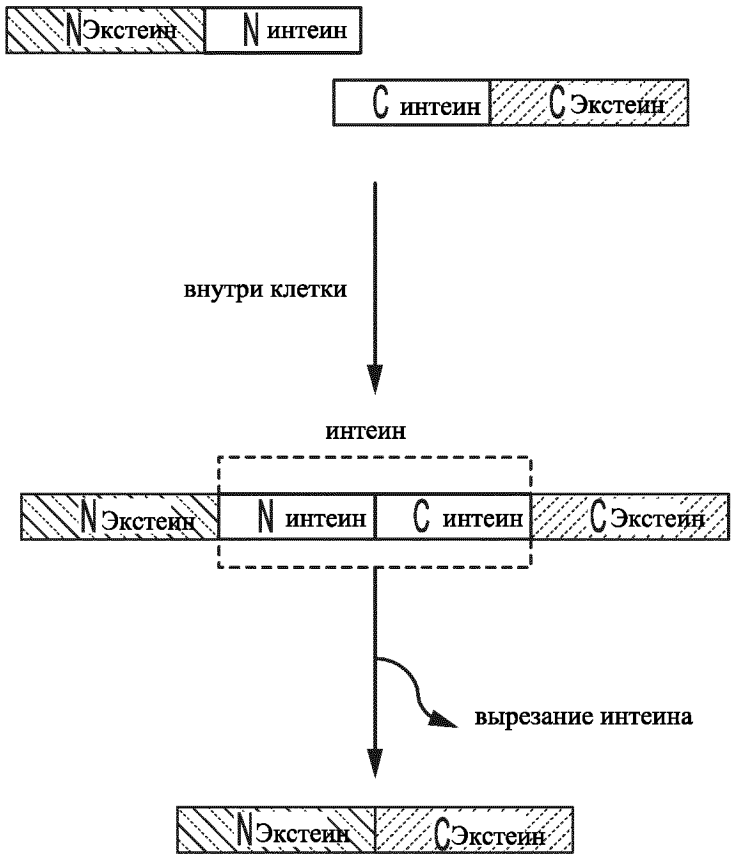

1)
gRNA
Cas9
3′
3′
RT
PE-опосредованный синтез мишени рекомбиназы в культивированных клетках человека
2)
матрица донора ДНК
геномная интеграция
геном
GFP
рекомбиназа

Фиг. 65



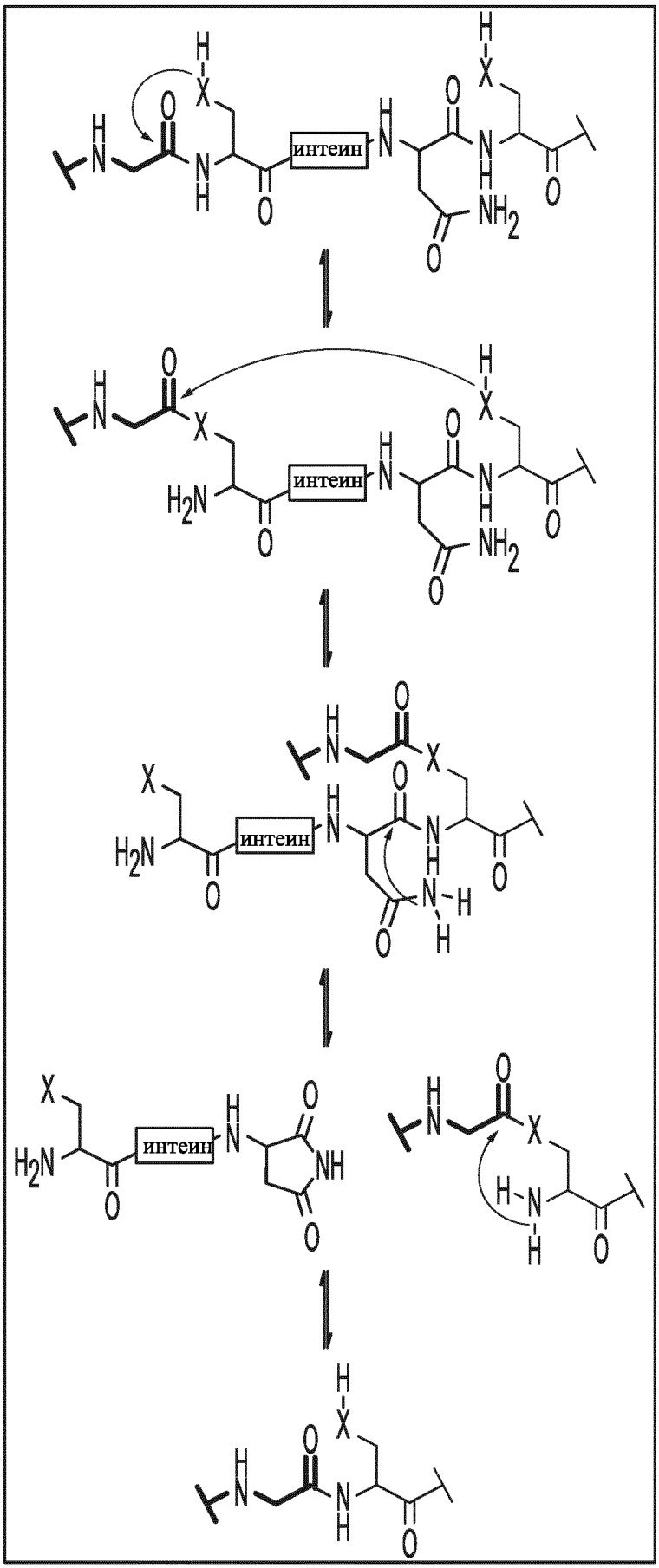
Праймирующий редактор

Фиг. 66

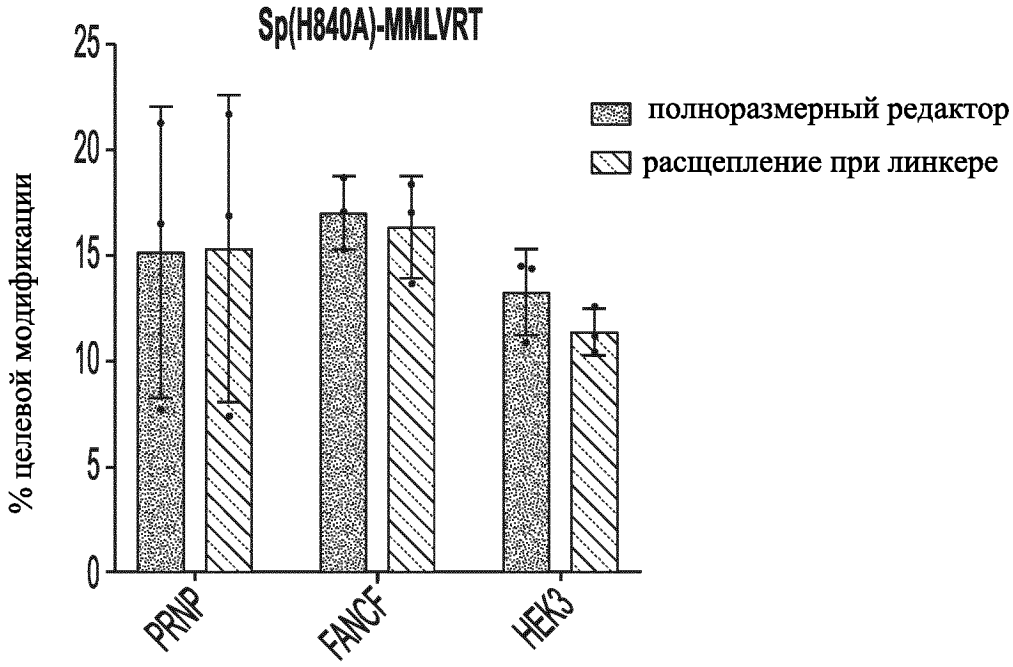

N Экстеин
N интеин
C интеин
C Экстеин
внутри клетки
интеин
N Экстеин
N интеин
C интеин
C Экстеин
вырезание интеина
N Экстеин
C Экстеин

Фиг. 67A

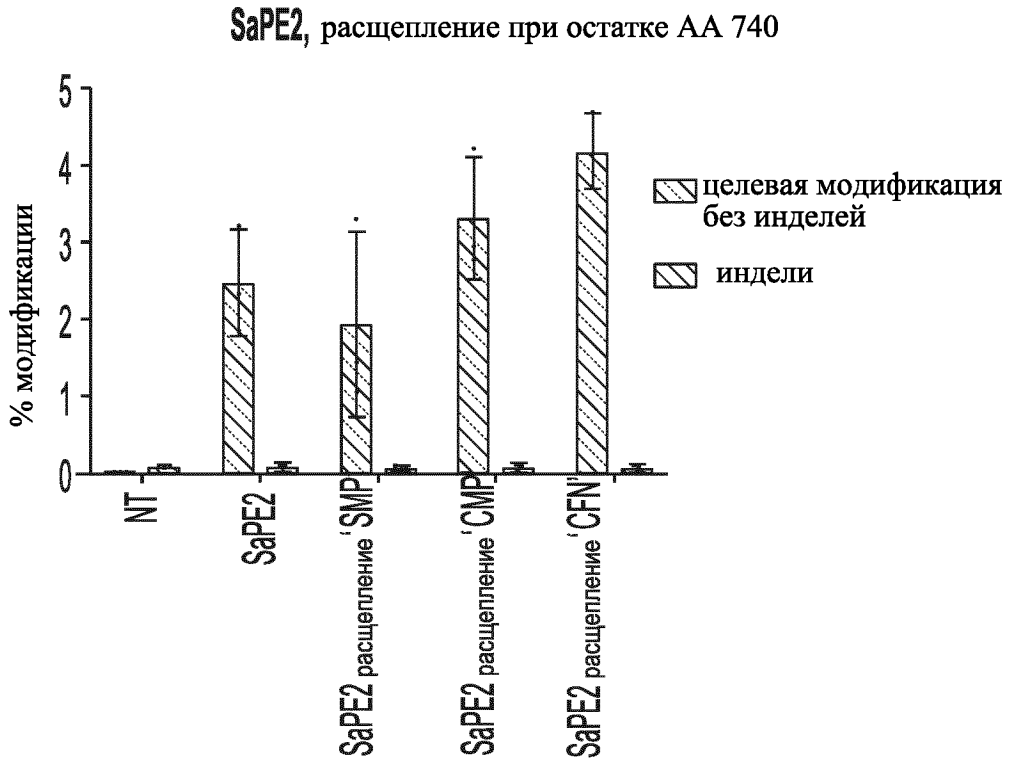
Механизм вырезания интеина



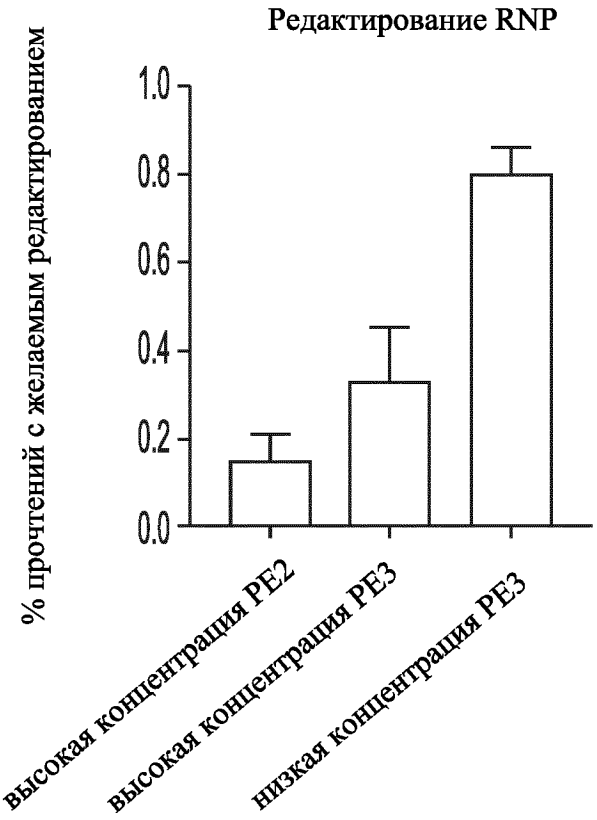
Фиг. 67B



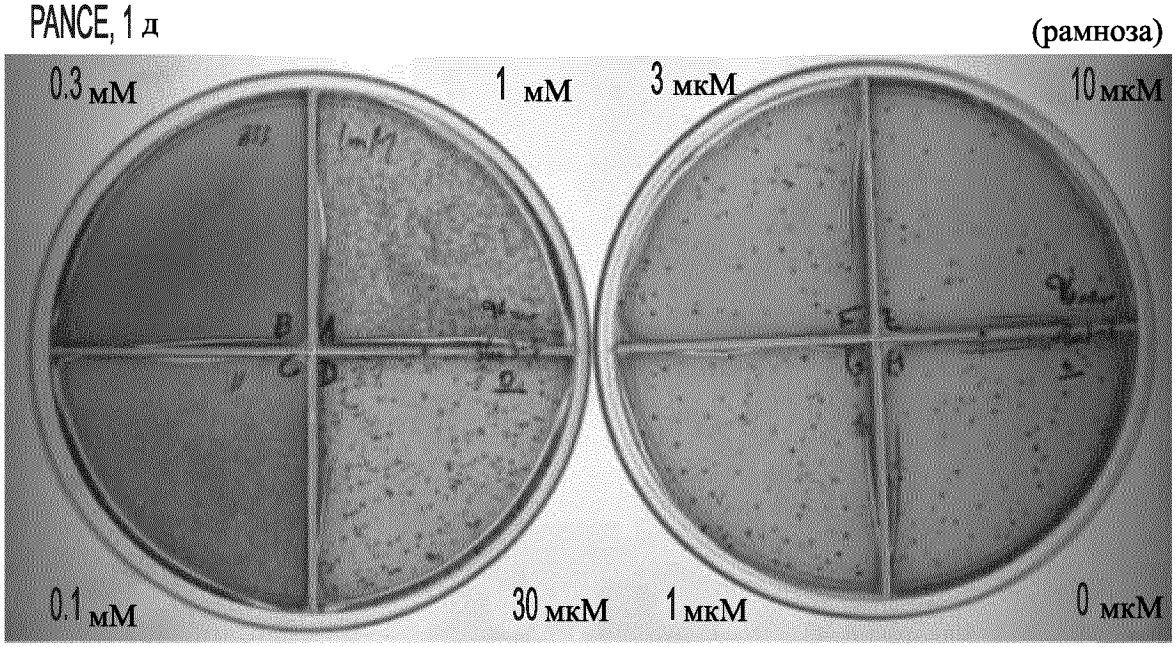
Фиг. 68A



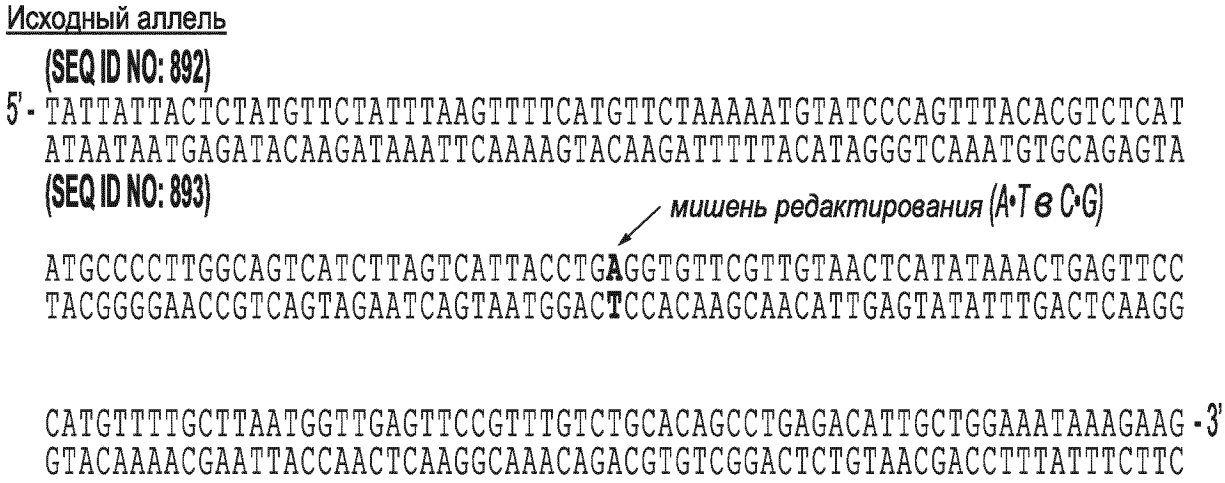
Фиг. 68B


Редактирование RNP
% прочтений с желаемым редактированием
1.0
0.8
0.6
0.4
0.2
0.0
высокая концентрация PE2
высокая концентрация PE3
низкая концентрация PE3

Фиг. 68C



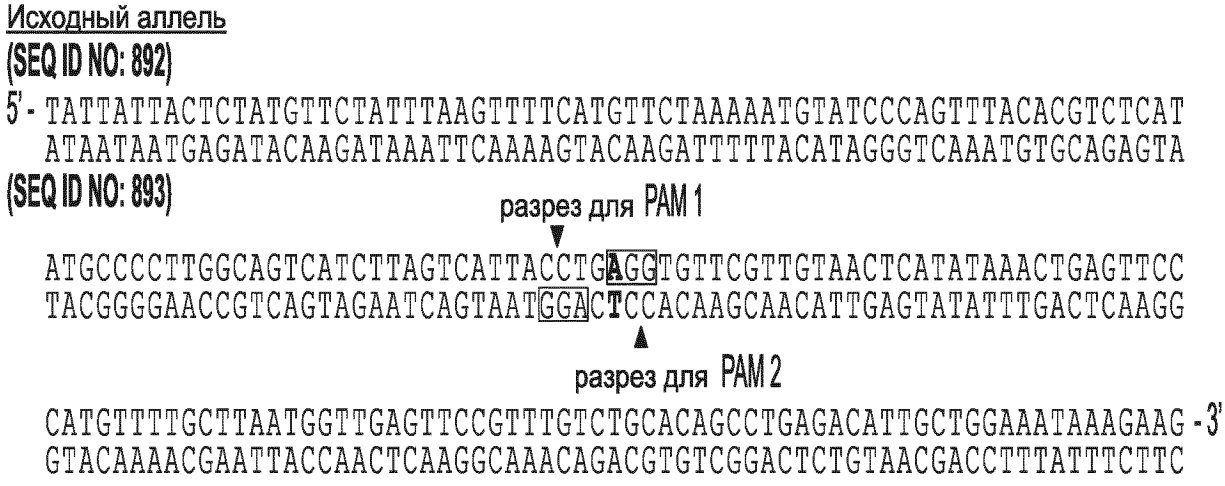
Фиг. 69



Фиг. 70A



Фиг. 70B



Фиг. 70C

Исходный аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

Спейсер

ATGCCCCTTGGCA*GTCATCTTAGTCATTACCTG*AGGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACTCCACAAGCAACATTGAGTATATTTGACTCAAGG

CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG - 3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

Спейсер= GTCATCTTAGTCATTACCTG

Фиг. 70D

Исходный аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

PBS компонент

ATGCCCCTTGGCA*GTCATCTTAGTCATTACCTG*AGGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACTCCACAAGCAACATTGAGTATATTTGACTCAAGG

CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG - 3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

PBS = GTAATGACTAAGATG

Фиг. 70E

Желаемый аллель

(SEQ ID NO: 892)

5' - TATTATTACTCTATGTTCTATTTAAGTTTTCATGTTCTAAAAATGTATCCCAGTTTACACGTCTCAT
ATAATAATGAGATACAAGATAAATTCAAAAGTACAAGATTTTTACATAGGGTCAAATGTGCAGAGTA

(SEQ ID NO: 893)

компонент матрицы RT

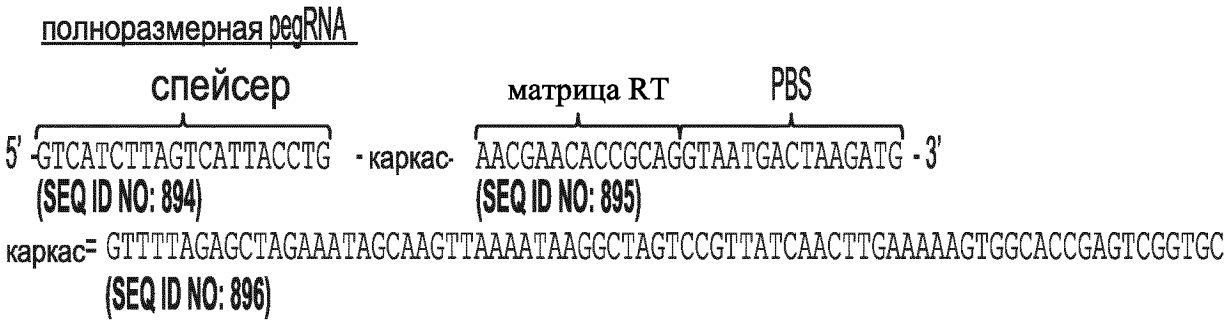
ATGCCCCTTGGCA*GTCATCTTAGTCATTACCTG*CGGTGTTCGTTGTAACTCATATAAACTGAGTTCC
TACGGGGAACCGTCAGTAGAATCAGTAATGGACGCCACAAGCAACATTGAGTATATTTGACTCAAGG

Edit

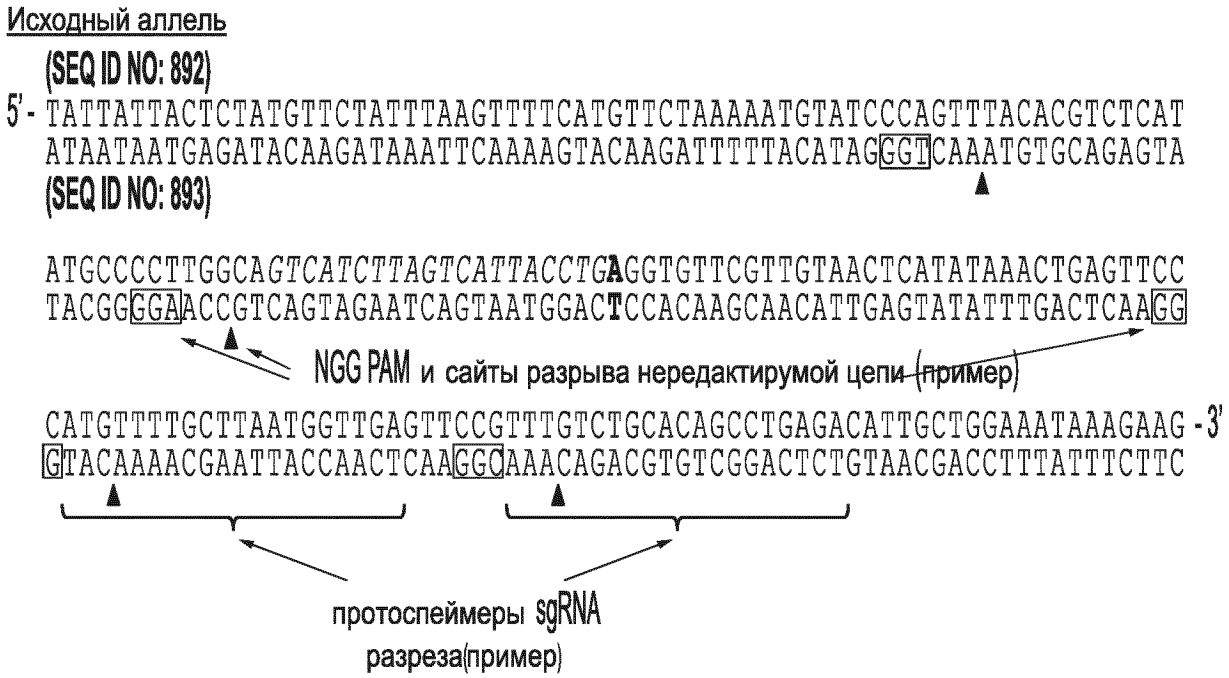
CATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTGGAAATAAAGAAG - 3'
GTACAAAACGAATTACCAACTCAAGGCAAACAGACGTGTCGGACTCTGTAACGACCTTTATTTCTTC

RT матрица= AACGAACACCGCAG

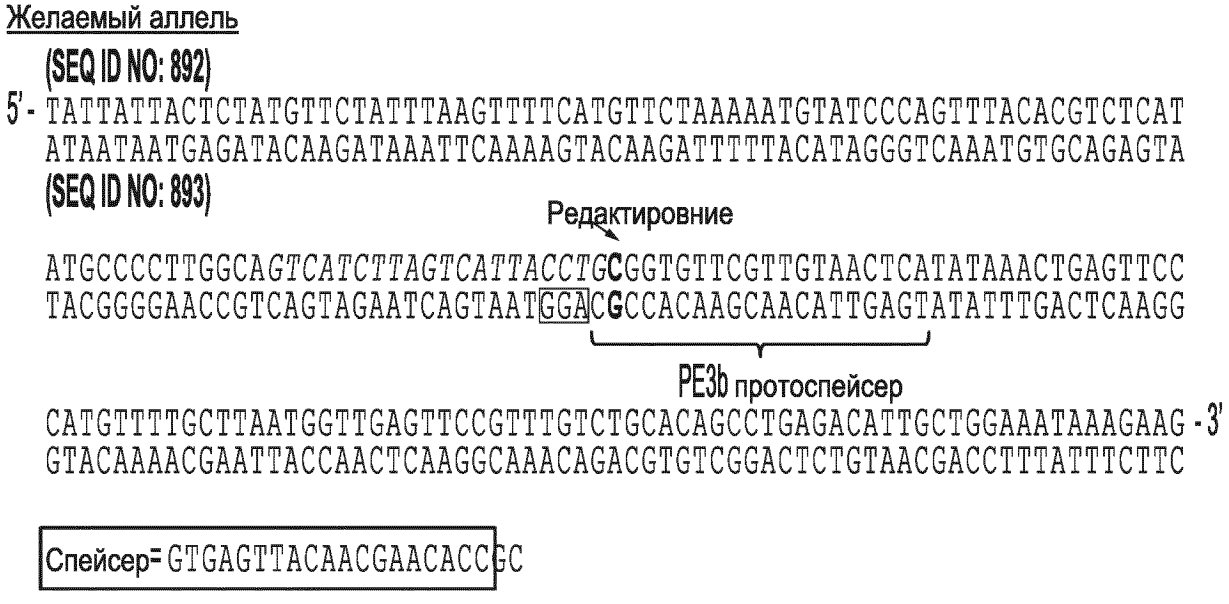
Фиг. 70F



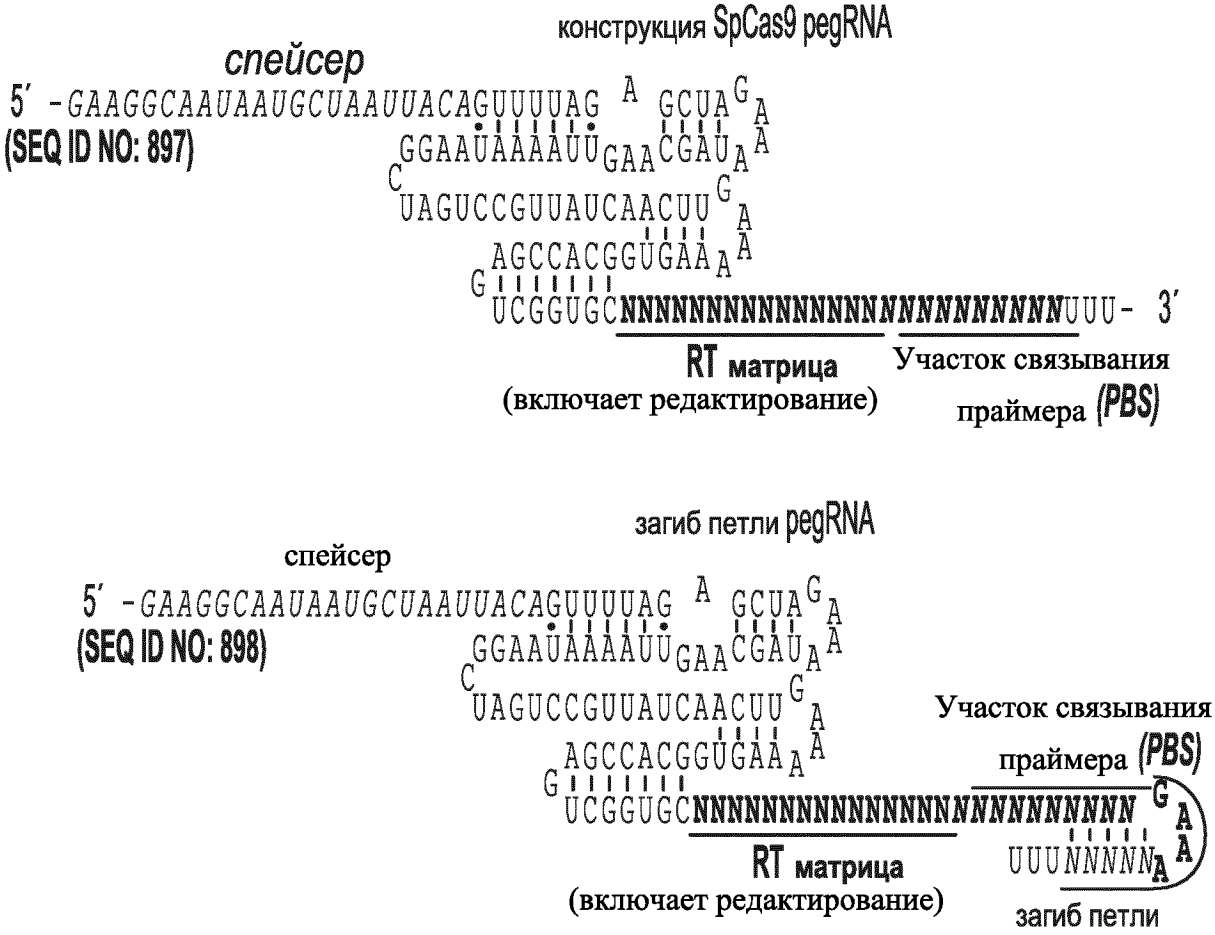
Фиг. 70G



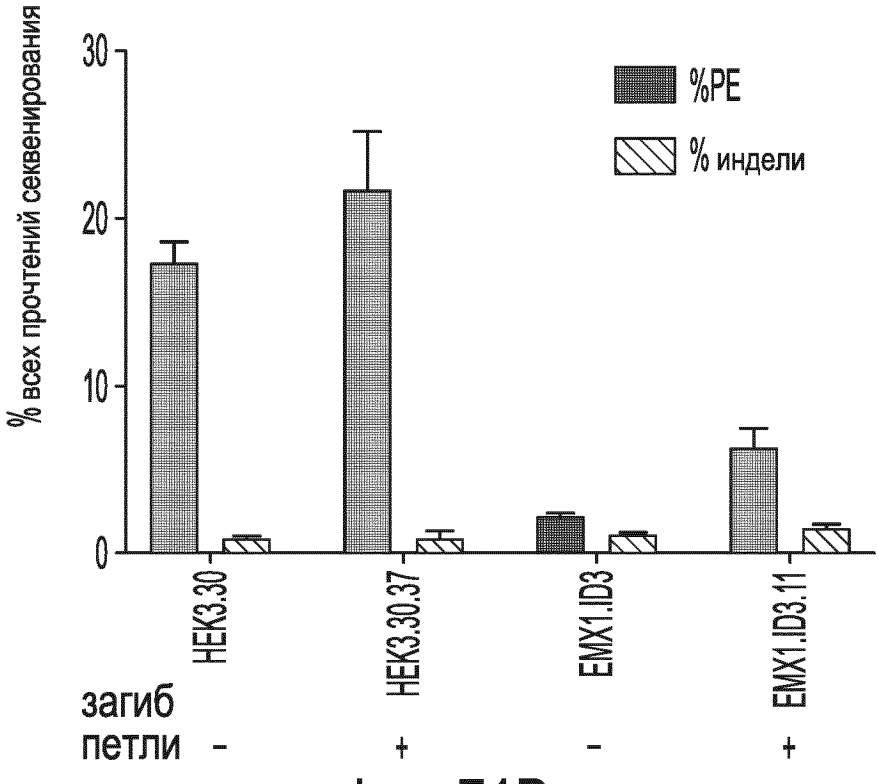
Фиг. 70H



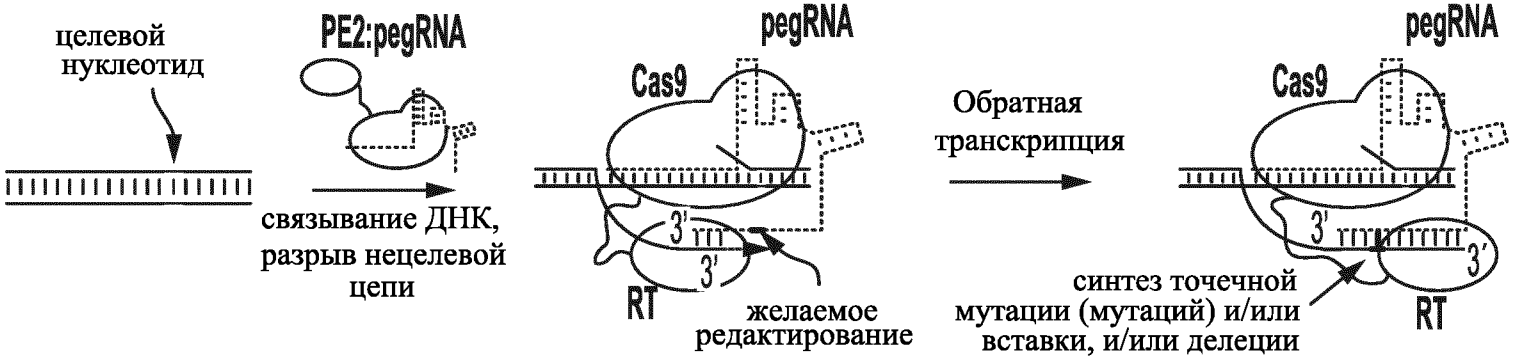
Фиг. 70I



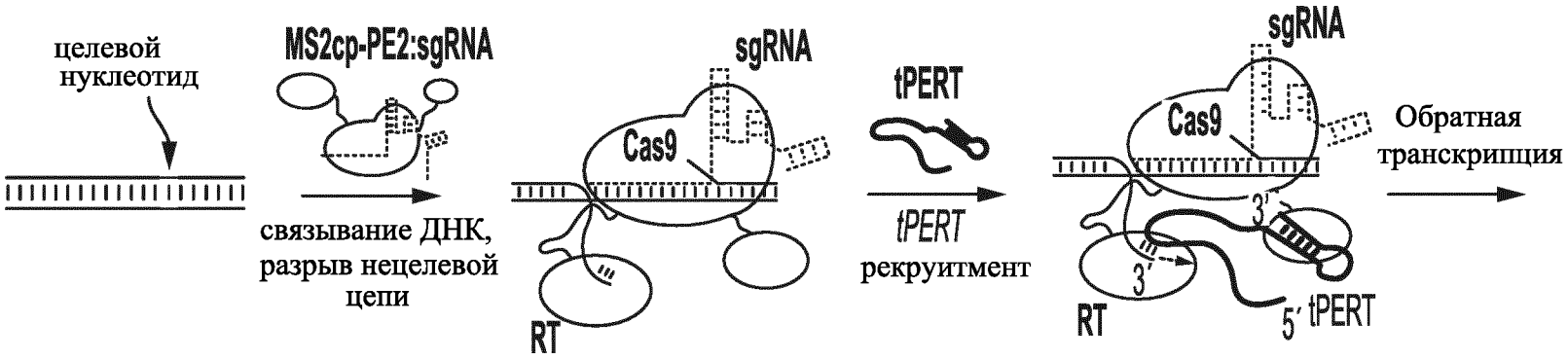
Фиг. 71A



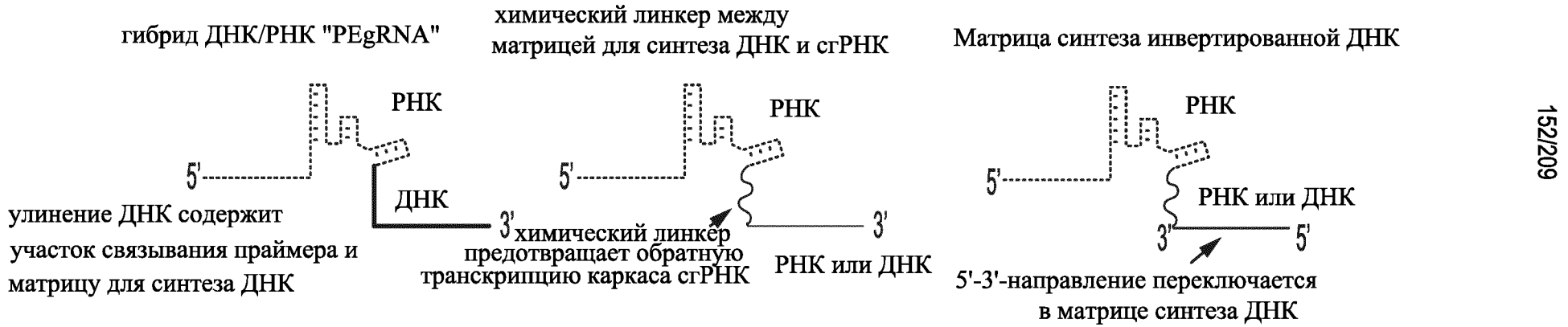
Фиг. 71B



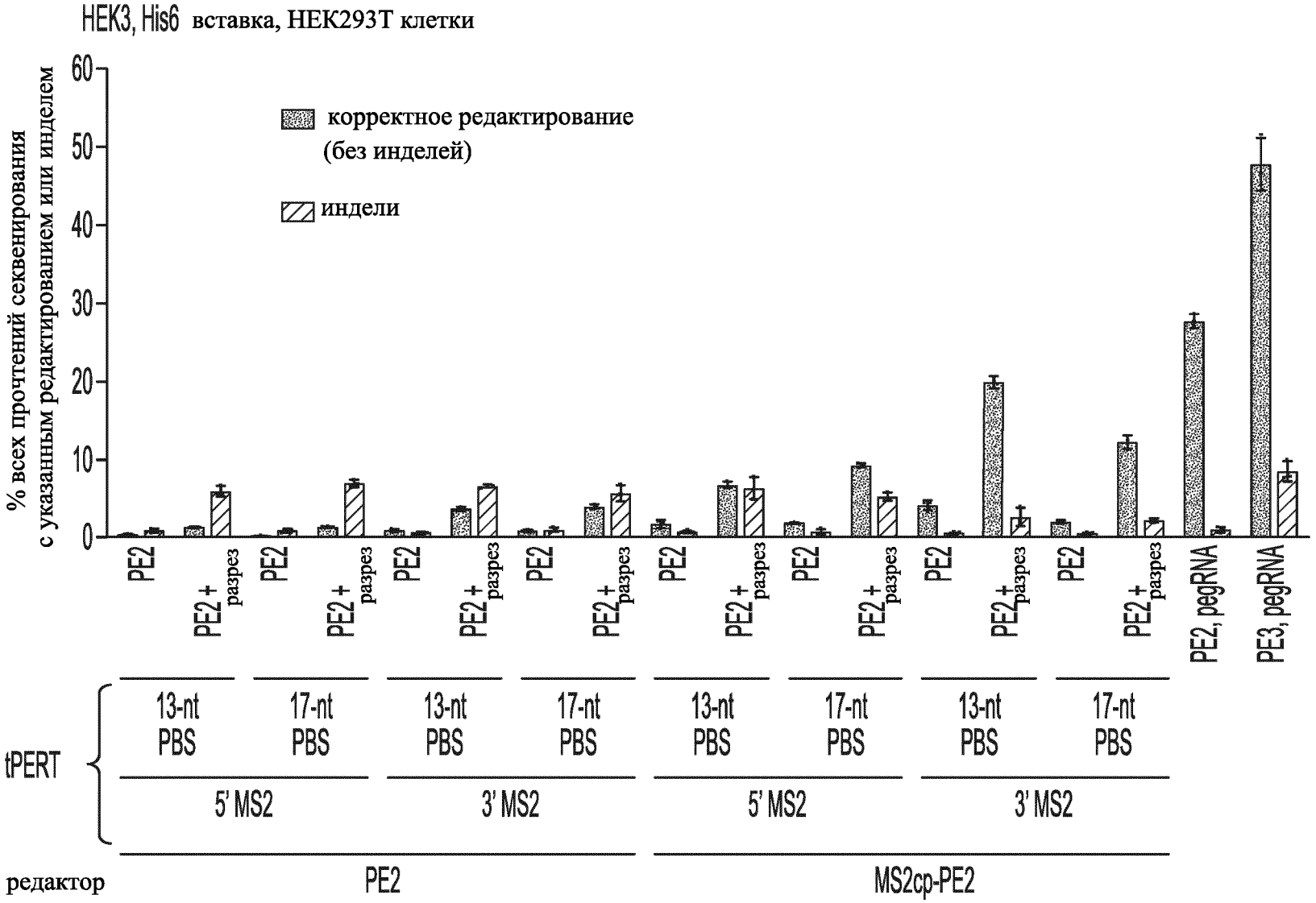
Фиг. 72A



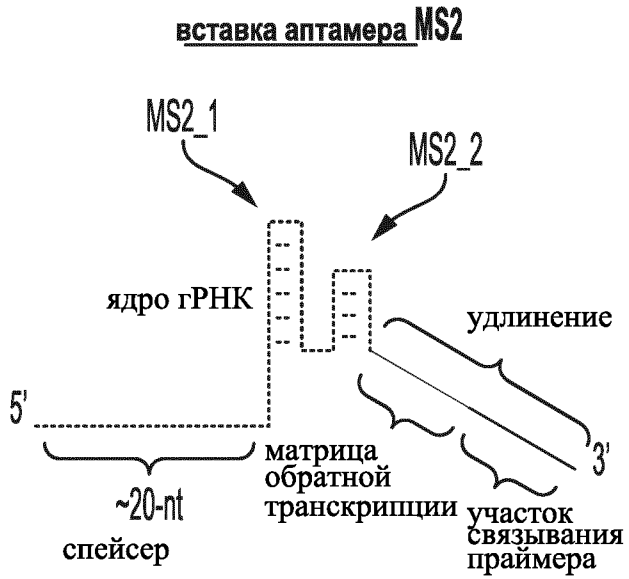
Фиг. 72B


гибрид ДНК/РНК "PEgRNA"
РНК
5'
ДНК
3'
улинение ДНК содержит
участок связывания праймера и
матрицу для синтеза ДНК
химический линкер между
матрицей для синтеза ДНК и сгРНК
РНК
5'
3'
химический линкер
предотвращает обратную
транскрипцию каркаса сгРНК
РНК или ДНК
Матрица синтеза инвертированной ДНК
РНК
5'
РНК или ДНК
3'
5'
5'-3'-направление переключается
в матрице синтеза ДНК
ДНК-зависимая ДНК-полимераза
может использоваться вместо обратной транскриптазы

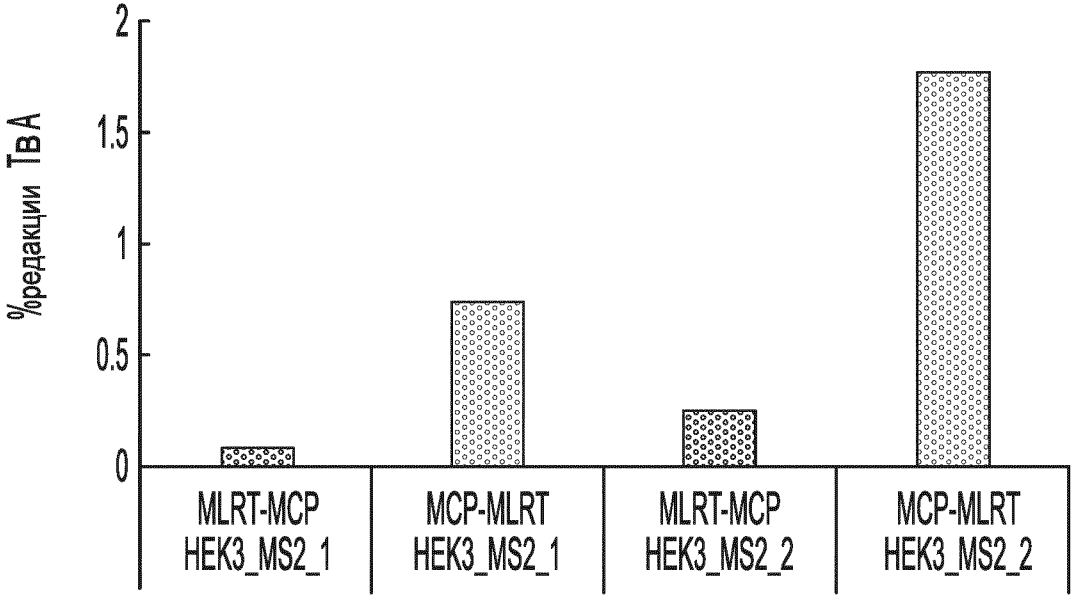
Фиг. 72C


HEK3, His6 вставка, HEK293T клетки
% всех прочтений секвенирования с указанным редактированием или инделем
корректное редактирование (без инделей)
индели
PE2
PE2 + разрез
PE2, pegRNA
PE3, pegRNA
tPERT
13-nt PBS
17-nt PBS
5' MS2
3' MS2
редактор
PE2
MS2cp-PE2

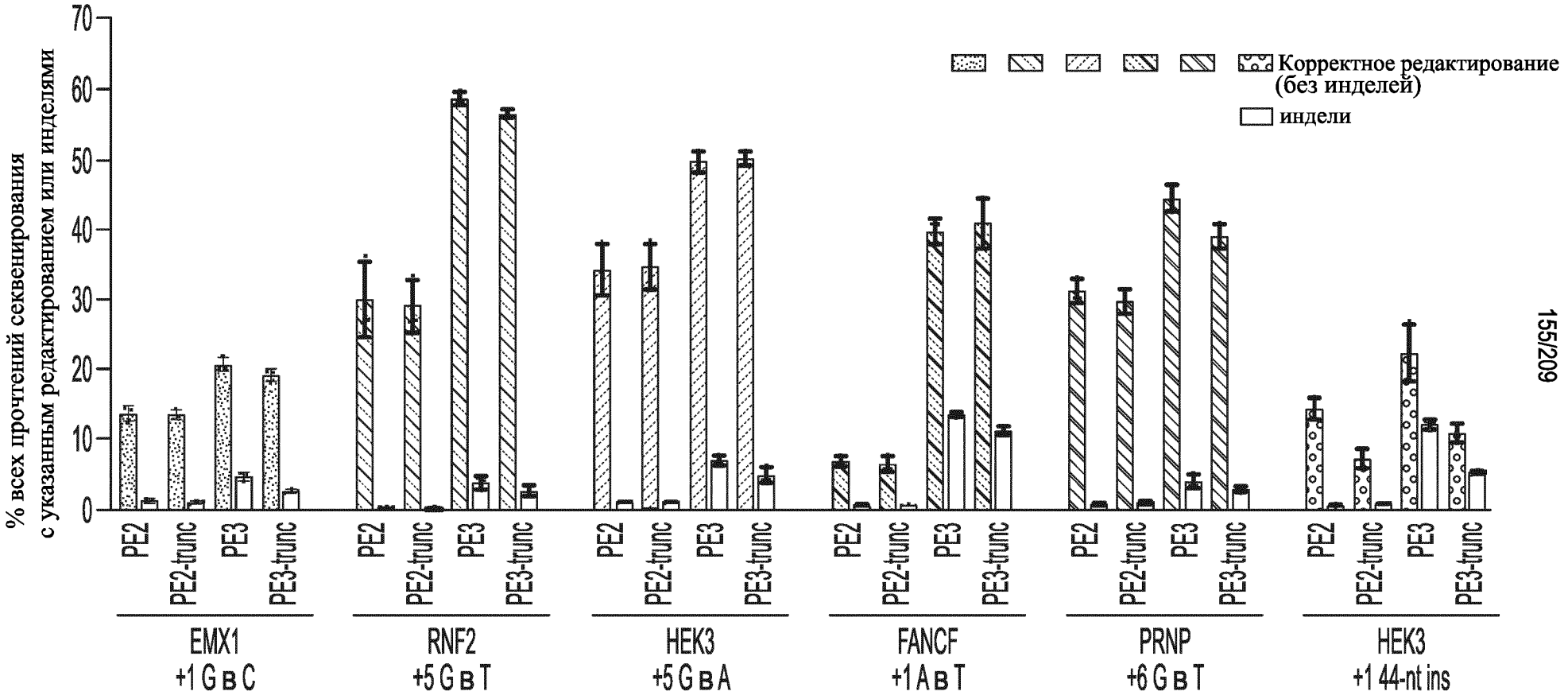
Фиг. 73



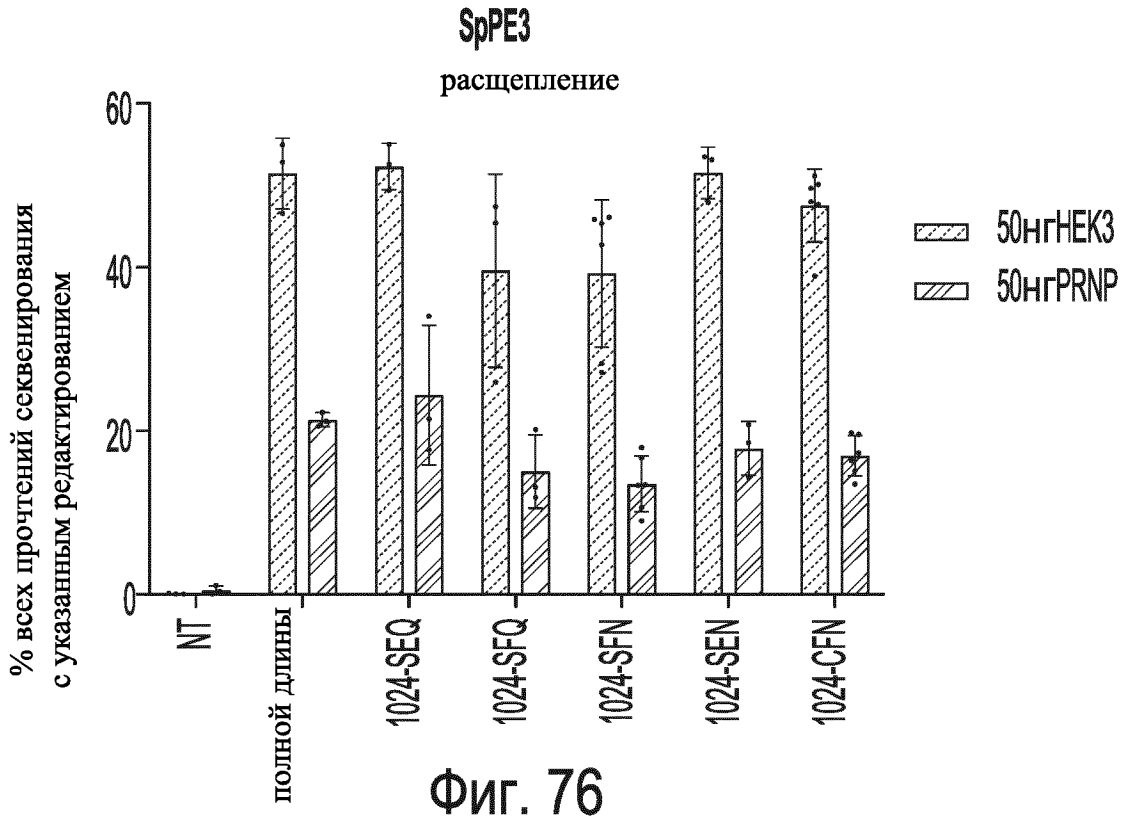
гРНК содержит 12 п.о. PBS



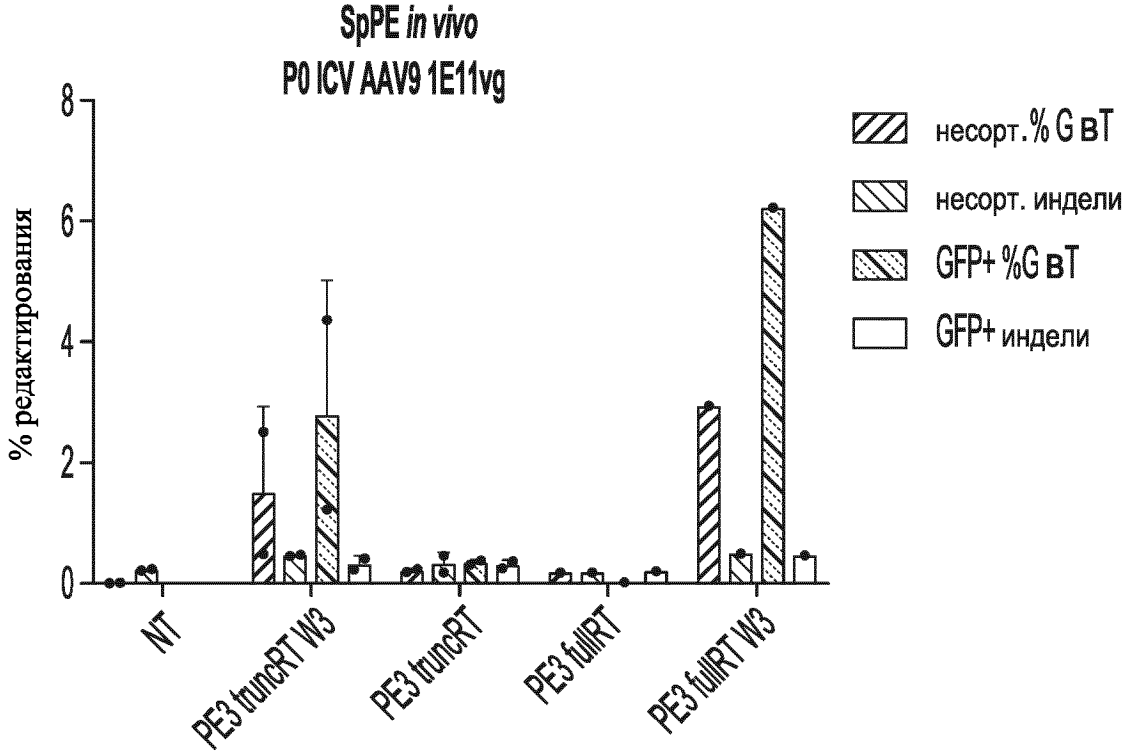
Фиг. 74


% всех прочтений секвенирования
с указанным редактированием или инделями
70
60
50
40
30
20
10
0
Корректное редактирование
(без инделей)
индели
PE2
PE2-trunc
PE3
PE3-trunc
EMX1
+1 G в C
RNF2
+5 G в T
HEK3
+5 G в A
FANCF
+1 A в T
PRNP
+6 G в T
HEK3
+1 44-nt ins

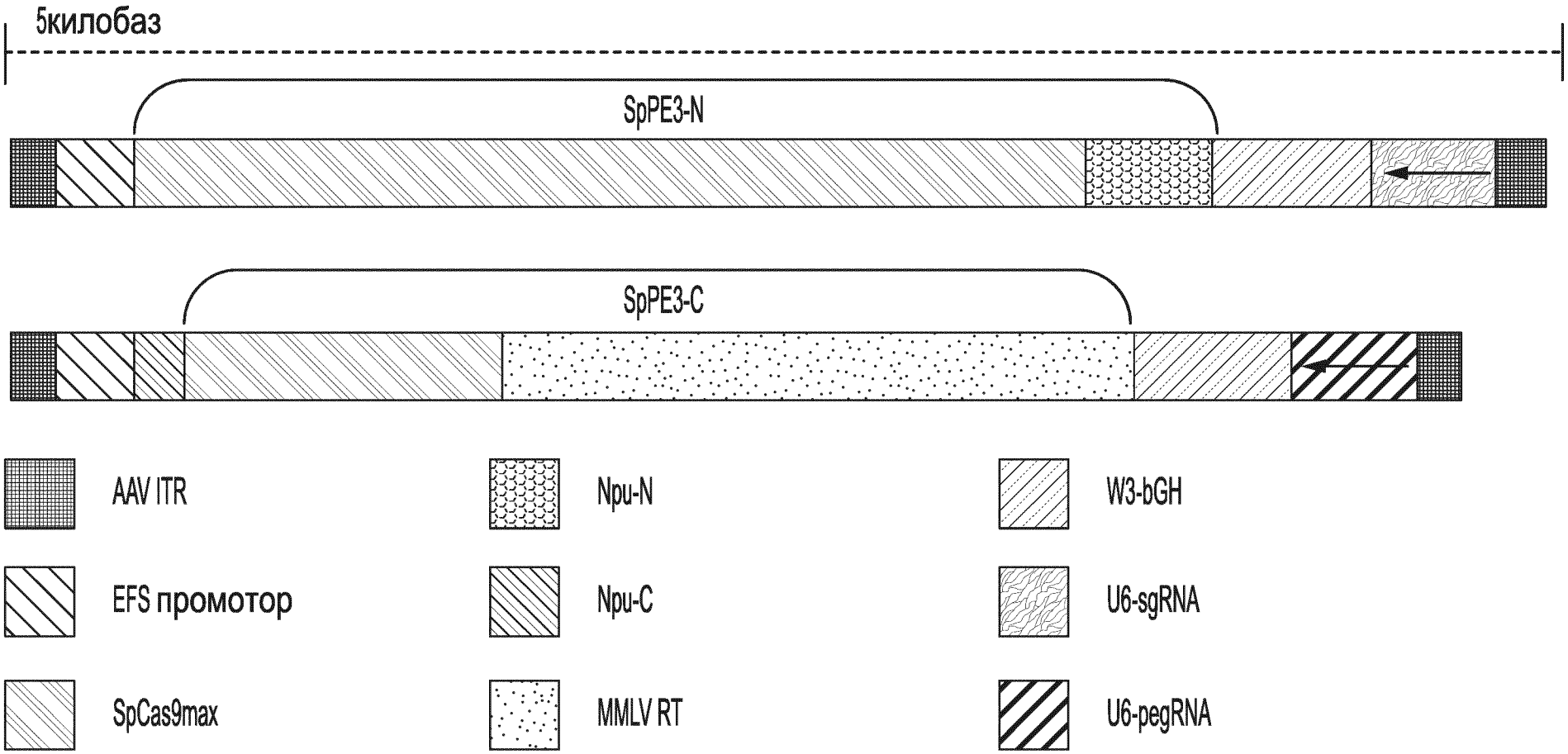
Фиг. 75


SpPE3
расщепление
% всех прочтений секвенирования с указанным редактированием
60
40
20
0
50нгHEK3
50нгPRNP
NT
полной длины
1024-SEQ
1024-SFQ
1024-SFN
1024-SEN
1024-CFN

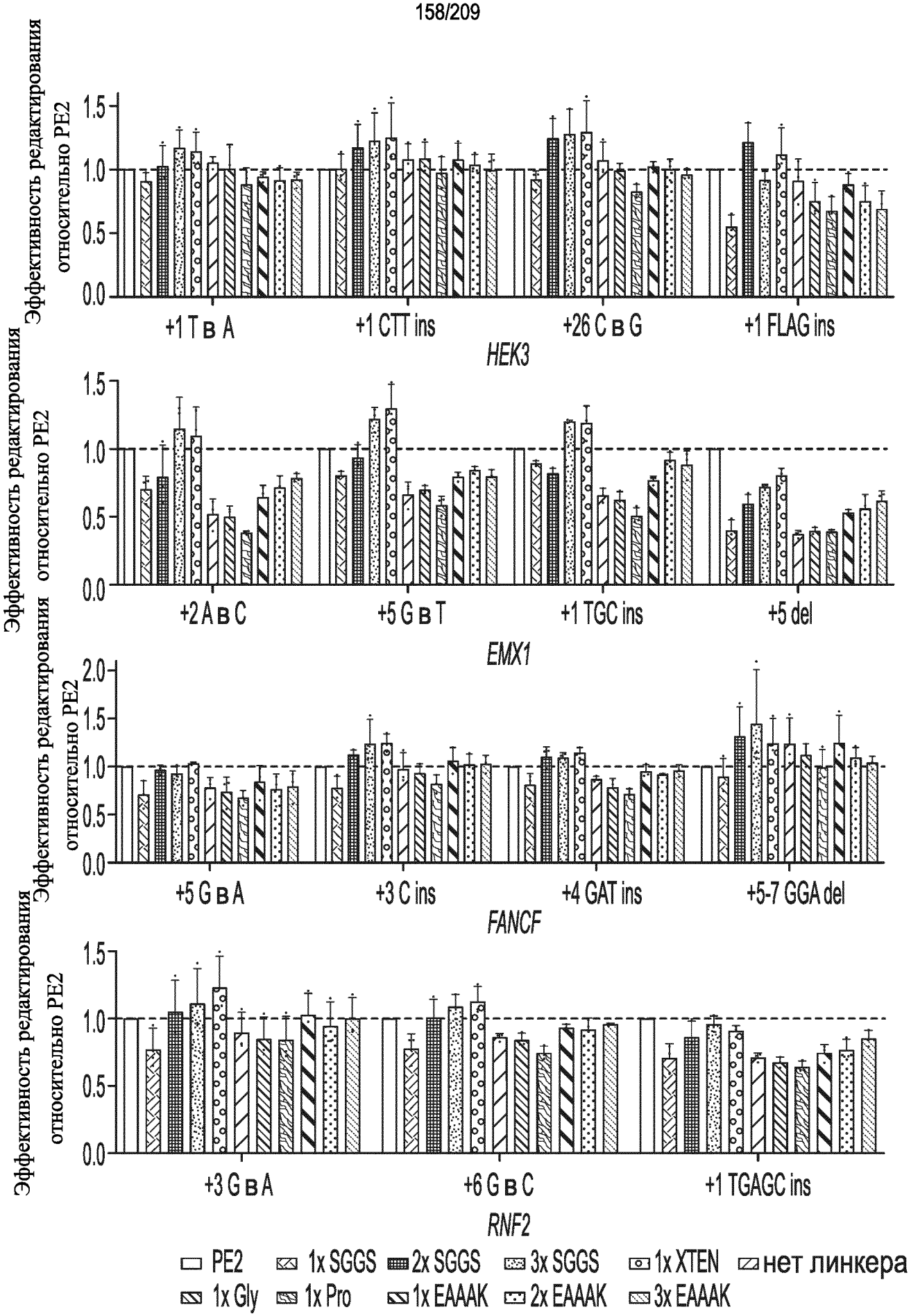
Фиг. 76


SpPE in vivo
P0 ICV AAV9 1E11vg
% редактирования
8
6
4
2
0
несорт.% G вТ
несорт. индели
GFP+ %G вТ
GFP+ индели
NT
PE3 truncRT W3
PE3 truncRT
PE3 fullRT
PE3 fullRT W3

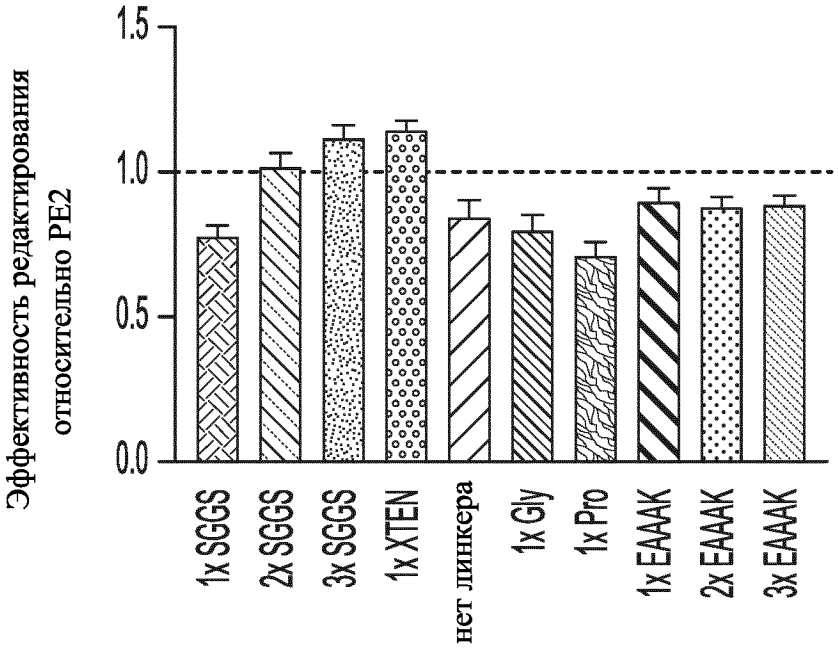
Фиг. 77


5килобаз
SpPE3-N
SpPE3-C
AAV ITR
EFS промотор
SpCas9max
Npu-N
Npu-C
MMLV RT
W3-bGH
U6-sgRNA
U6-pegRNA

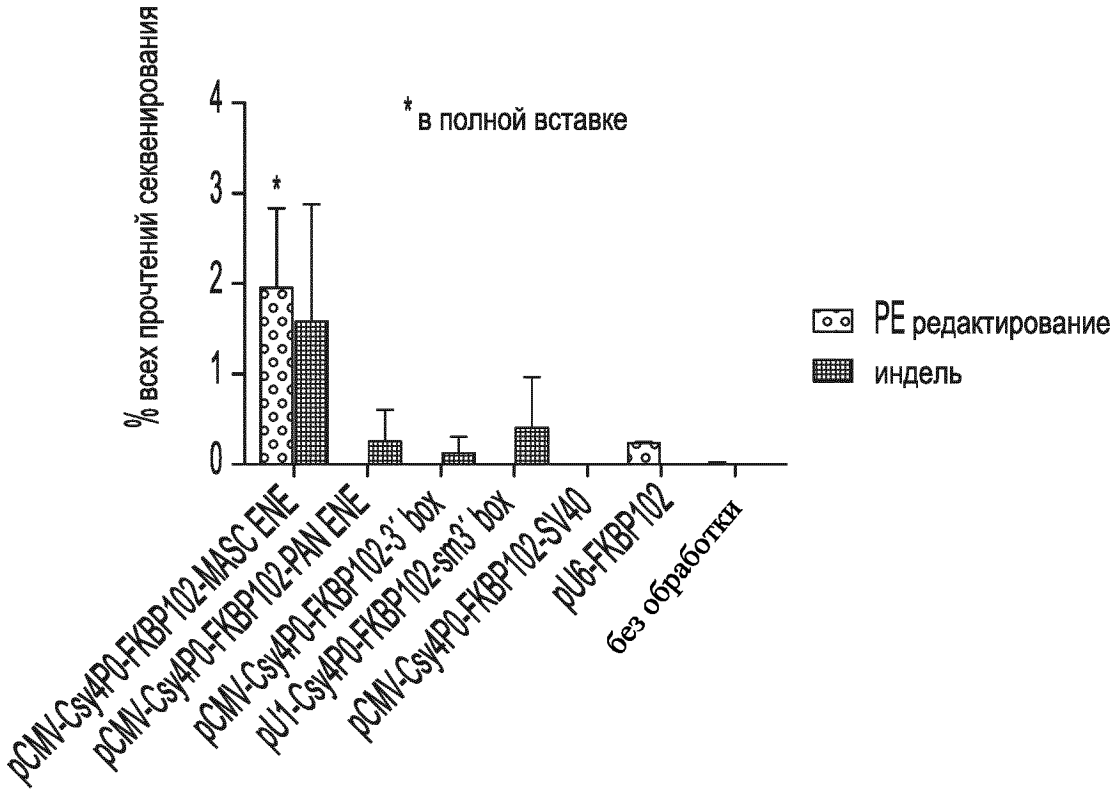
Фиг. 78


Эффективность редактирования относительно PE2
1.5
1.0
0.5
0.0
+1 T в A
+1 CTT ins
+26 C в G
+1 FLAG ins
HEK3
+2 A в C
+5 G в T
+1 TGC ins
+5 del
EMX1
2.0
+5 G в A
+3 C ins
+4 GAT ins
+5-7 GGA del
FANCF
+3 G в A
+6 G в C
+1 TGAGC ins
RNF2
PE2
1x SGGS
2x SGGS
3x SGGS
1x XTEN
нет линкера
1x Gly
1x Pro
1x EAAAK
2x EAAAK
3x EAAAK

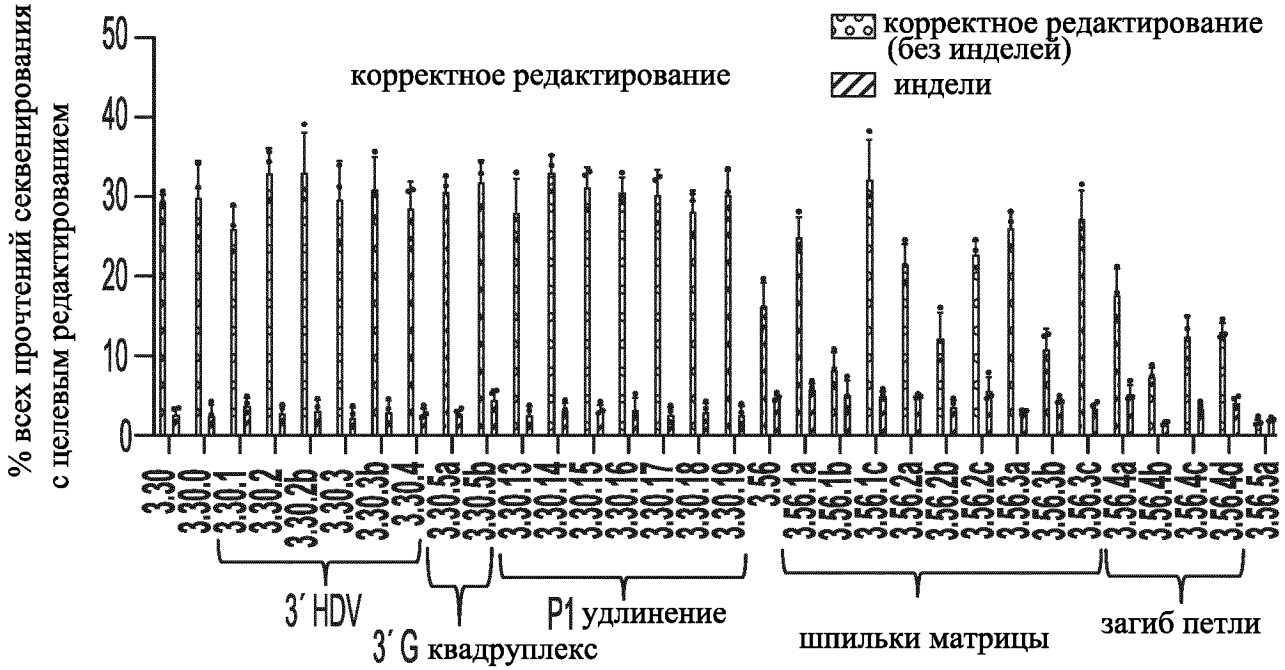
Фиг. 79


Эффективность редактирования
относительно PE2
1.5
1.0
0.5
0.0
1x SGGS
2x SGGS
3x SGGS
1x XTEN
нет линкера
1x Gly
1x Pro
1x EAAAK
2x EAAAK
3x EAAAK

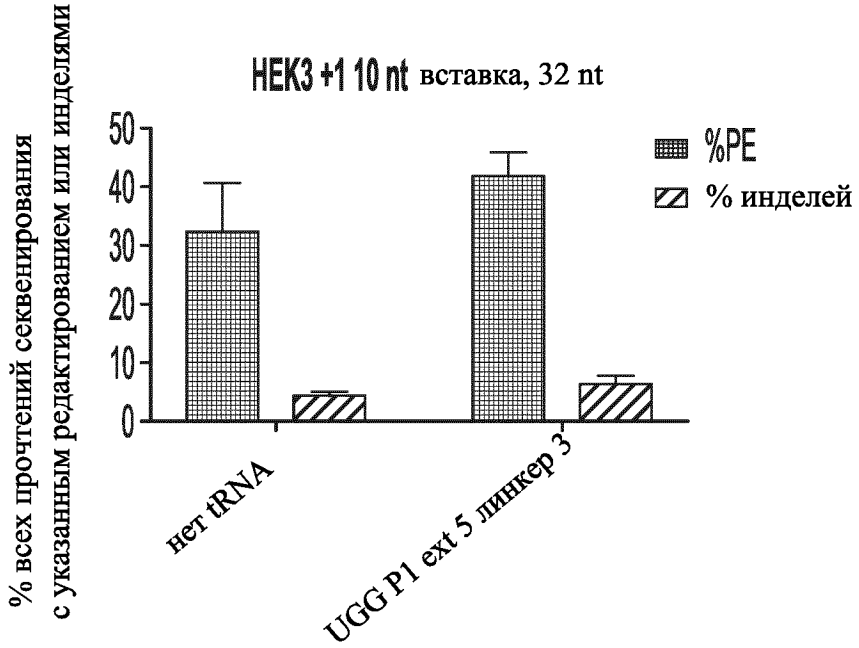
Фиг. 80


% всех прочтений секвенирования
4
3
2
1
0
* в полной вставке
*
PE редактирование
индель
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 прочтений секвенирования
с целевым редактированием
50
40
30
20
10
0
корректное редактирование
корректное редактирование
(без инделей)
индели
3.30
3.30.0
3.30.1
3.30.2
3.30.2b
3.30.3
3.30.3b
3.30.4
3.30.5a
3.30.5b
3.30.13
3.30.14
3.30.15
3.30.16
3.30.17
3.30.18
3.30.19
3.56
3.56.1a
3.56.1b
3.56.1c
3.56.2a
3.56.2b
3.56.2c
3.56.3a
3.56.3b
3.56.3c
3.56.4a
3.56.4b
3.56.4c
3.56.4d
3.56.5a
3´ HDV
3´ G квадруплекс
P1 удлинение
шпильки матрицы
загиб петли

Фиг. 82


HEK3 +1 10 nt вставка, 32 nt
% всех прочтений секвенирования
с указанным редактированием или инделями
50
40
30
20
10
0
%PE
% инделей
нет tRNA
UGG P1 ext 5 линкер 3

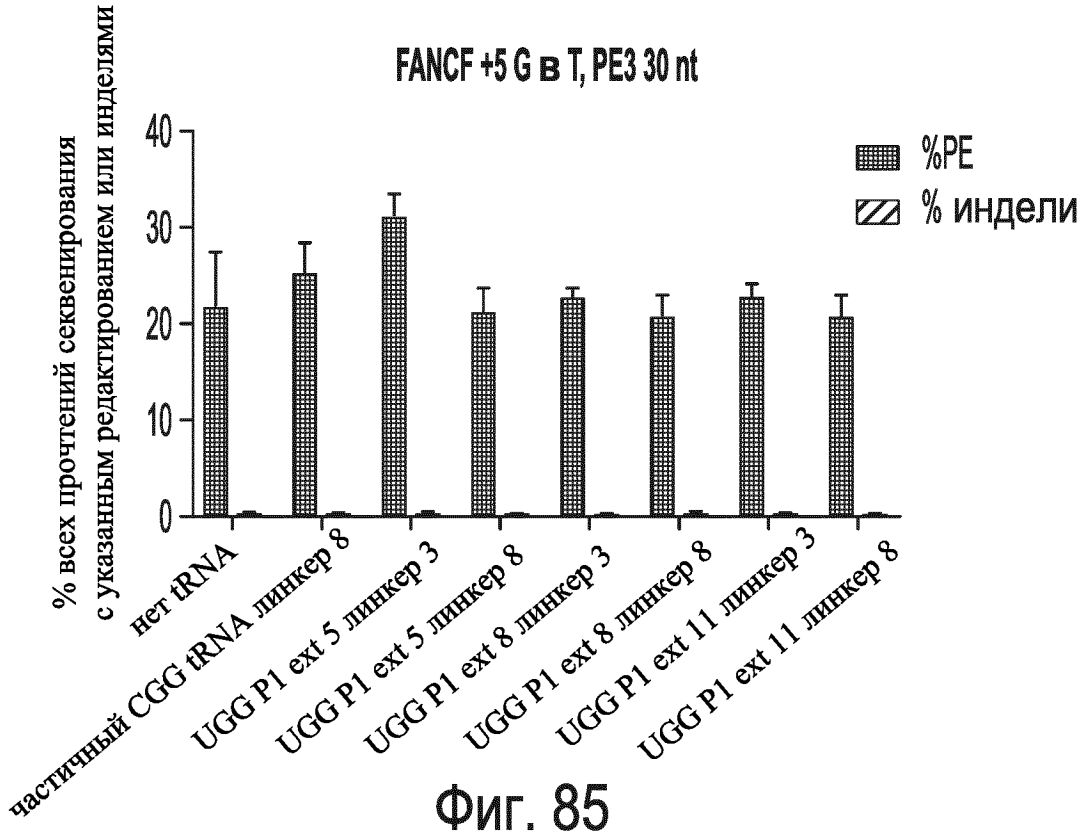
Фиг. 83



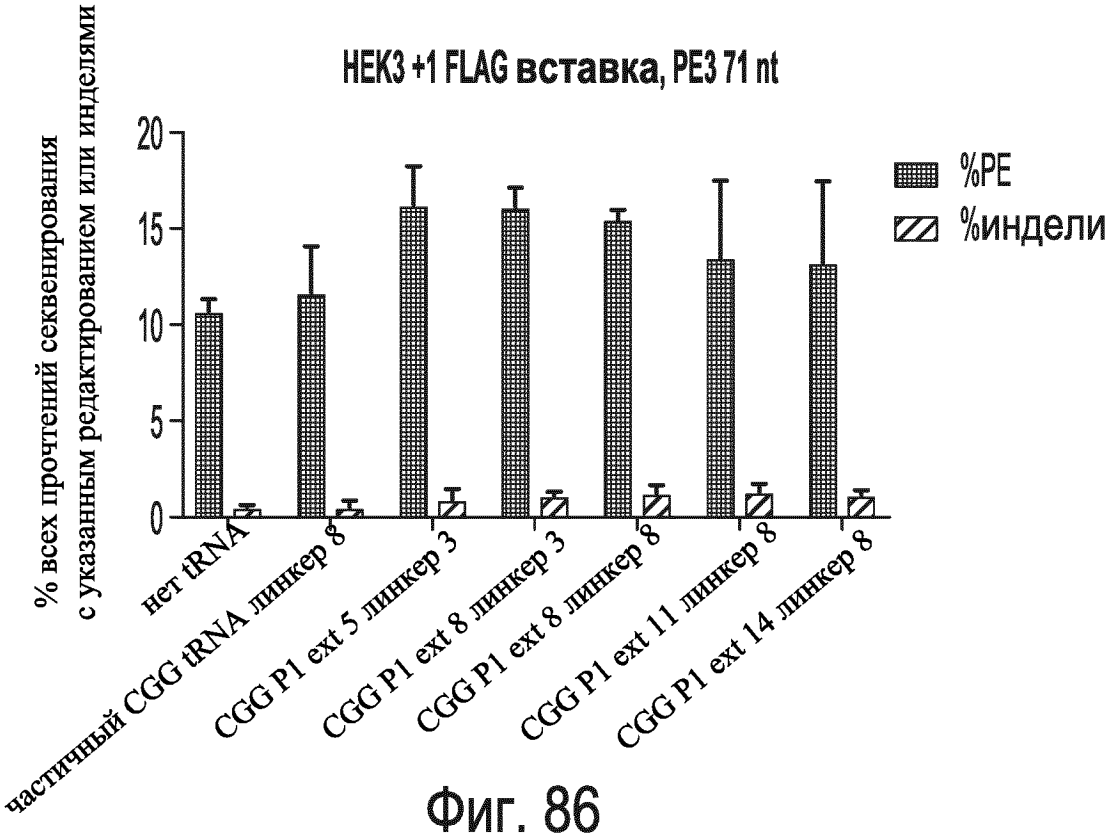
Фиг. 84A



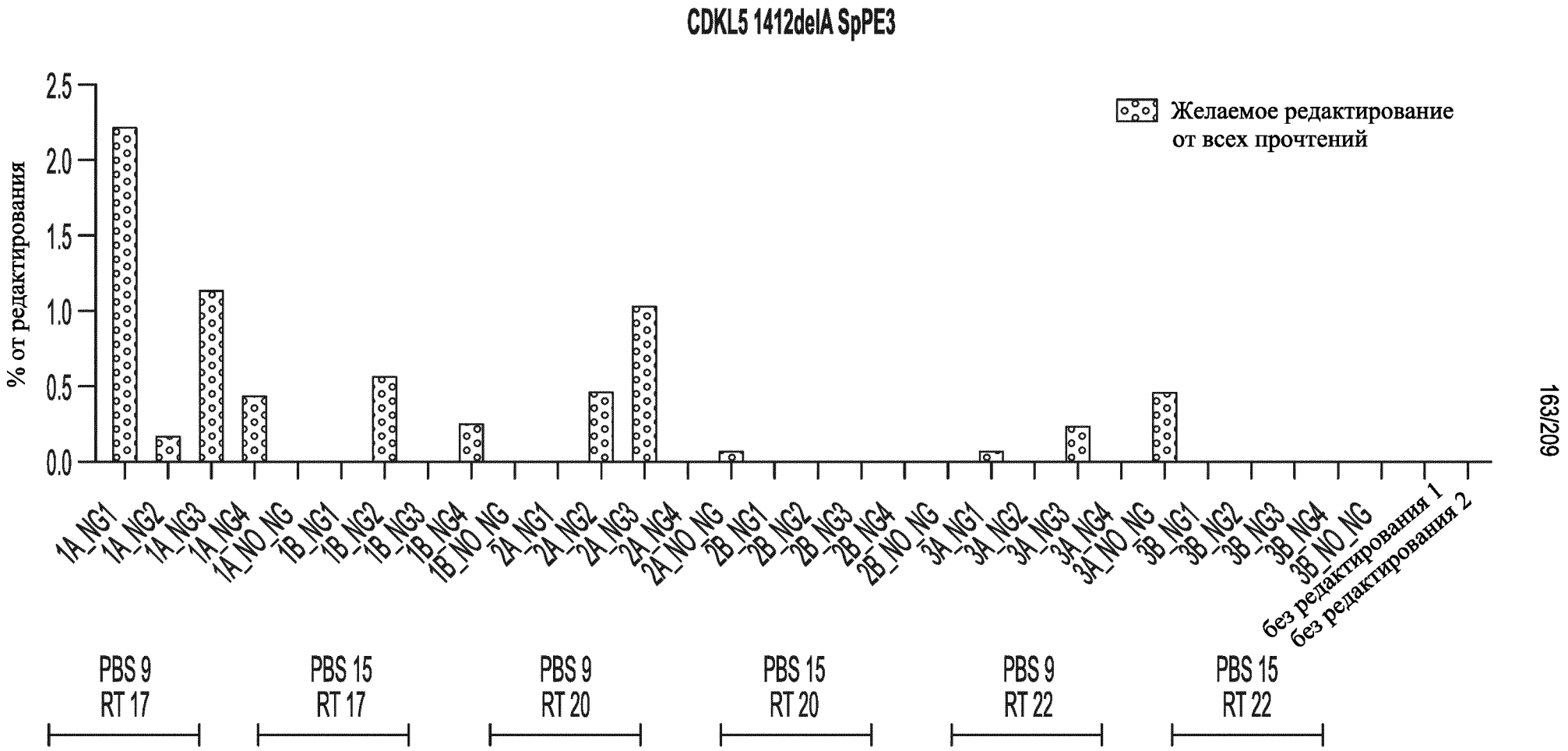
Фиг. 84B


FANCF +5 G в T, PE3 30 nt
% всех прочтений секвенирования с указанным редактированием или инделями
40
30
20
10
0
%PE
% индели
нет tRNA
частичный CGG tRNA линкер 8
UGG P1 ext 5 линкер 3
UGG P1 ext 5 линкер 8
UGG P1 ext 8 линкер 3
UGG P1 ext 8 линкер 8
UGG P1 ext 11 линкер 3
UGG P1 ext 11 линкер 8

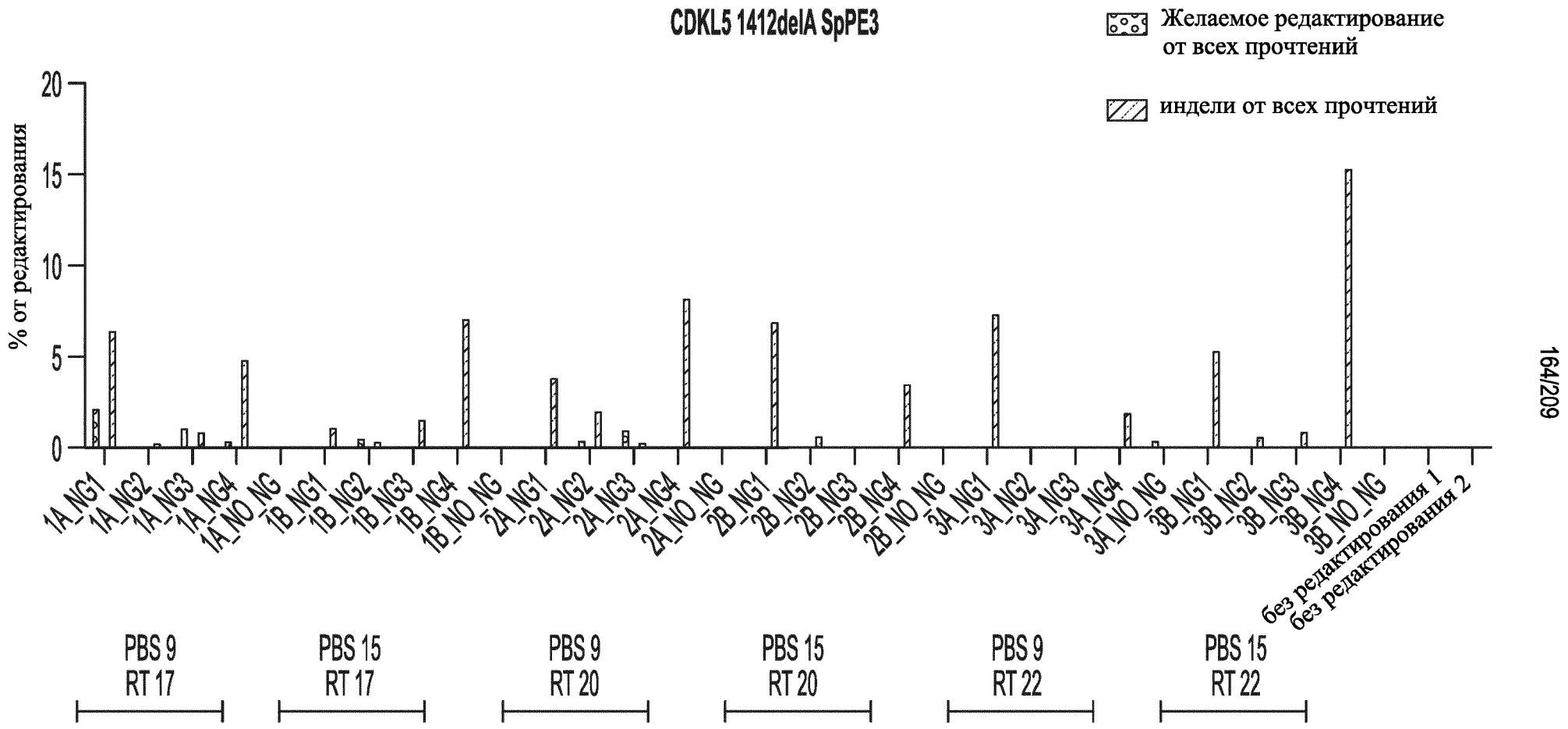
Фиг. 85


HEK3 +1 FLAG вставка, PE3 71 nt
% всех прочтений секвенирования с указанным редактированием или инделями
20
15
10
5
0
%PE
%индели
нет tRNA
частичный CGG tRNA линкер 8
CGG P1 ext 5 линкер 3
CGG P1 ext 8 линкер 3
CGG P1 ext 8 линкер 8
CGG P1 ext 11 линкер 8
CGG P1 ext 14 линкер 8

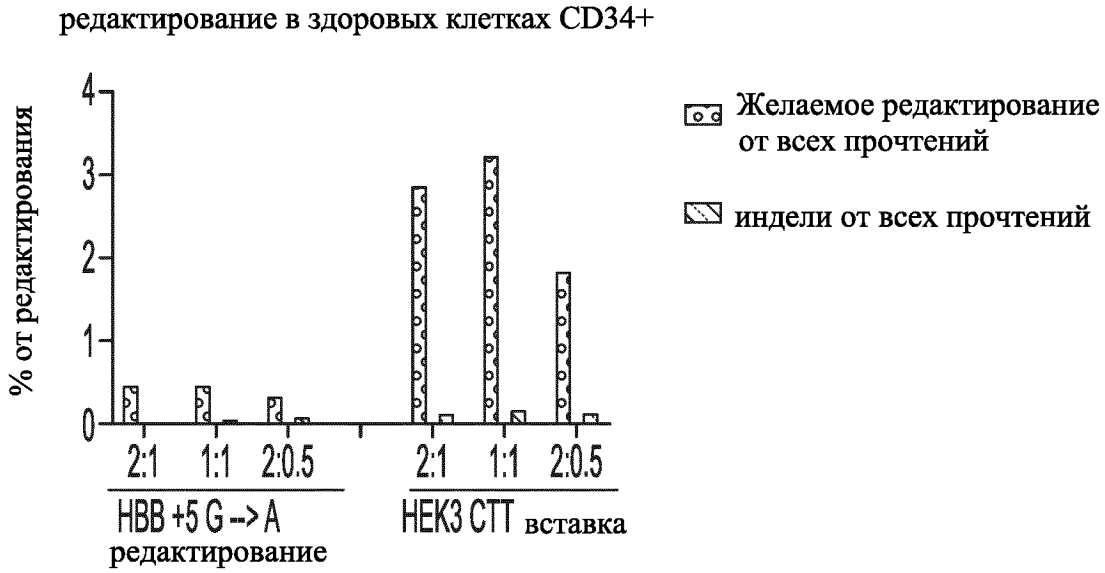
Фиг. 86


CDKL5 1412delA SpPE3
Желаемое редактирование от всех прочтений
% от редактирования
2.5
2.0
1.5
1.0
0.5
0.0
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
без редактирования 1
без редактирования 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

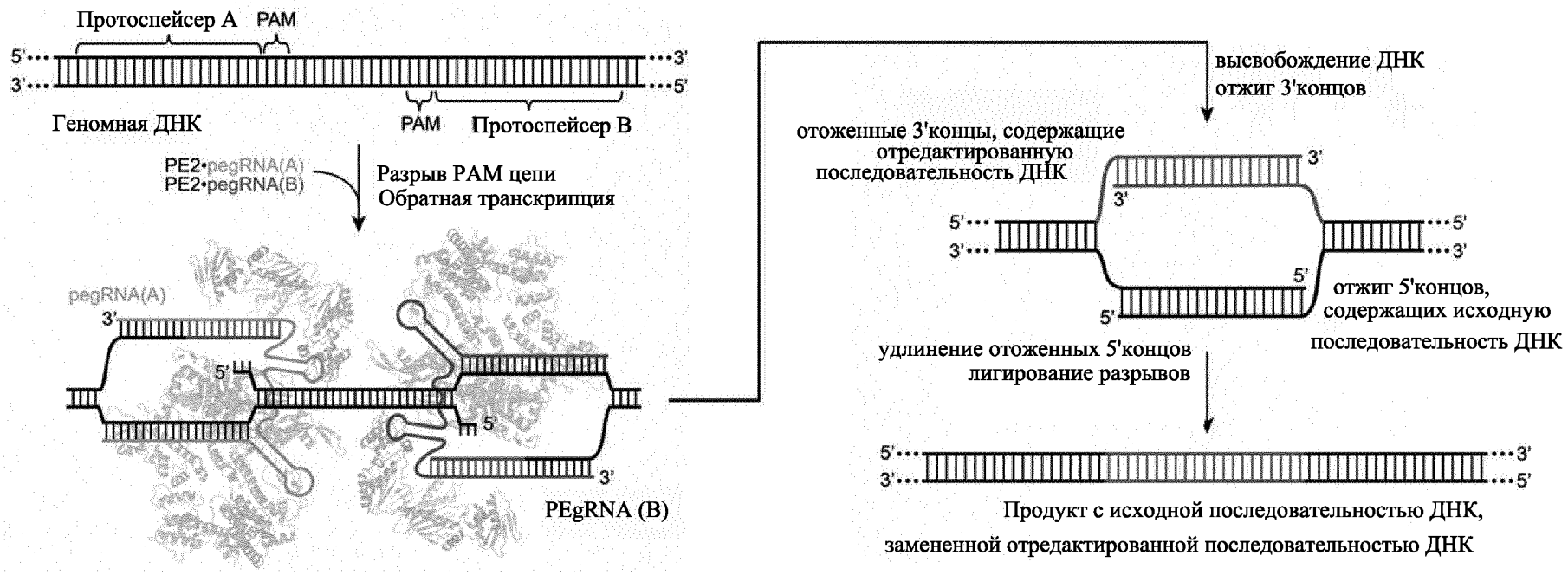
Фиг. 87


CDKL5 1412delA SpPE3
Желаемое редактирование от всех прочтений
индели от всех прочтений
% от редактирования
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
без редактирования 1
без редактирования 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

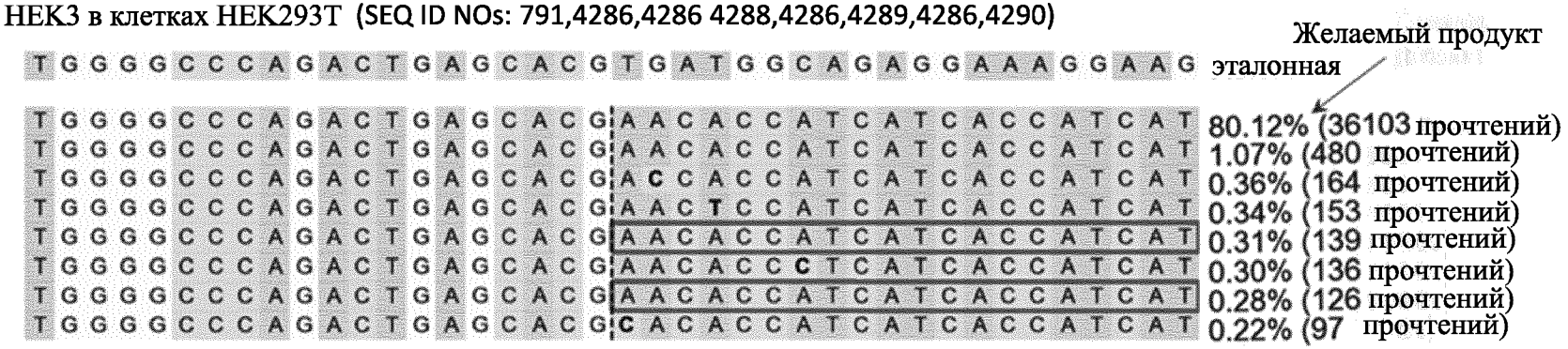

редактирование в здоровых клетках CD34+
% от редактирования
4
3
2
1
0
2:1
1:1
2:0.5
2:1
1:1
2:0.5
HBB +5 G --> A
редактирование
HEK3 CTT вставка
Желаемое редактирование от всех прочтений
индели от всех прочтений
Редактор (мкг):направляющая (мкг)

Фиг. 89


Протоспейсер A
PAM
5'
3'
Геномная ДНК
PAM
Протоспейсер B
PE2•pegRNA(A)
PE2•pegRNA(B)
Разрыв PAM цепи
Обратная транскрипция
pegRNA(A)
3'
5'
5'
3'
PEgRNA (B)
высвобождение ДНК
отжиг 3'концов
отоженные 3'концы, содержащие отредактированную последовательность ДНК
3'
3'
5'
5'
3'
5'
5'
отжиг 5'концов, содержащих исходную последовательность ДНК
удлинение отоженных 5'концов
лигирование разрывов
5'
3'
3'
5'
Продукт с исходной последовательностью ДНК, замененной отредактированной последовательностью ДНК

Фиг. 90

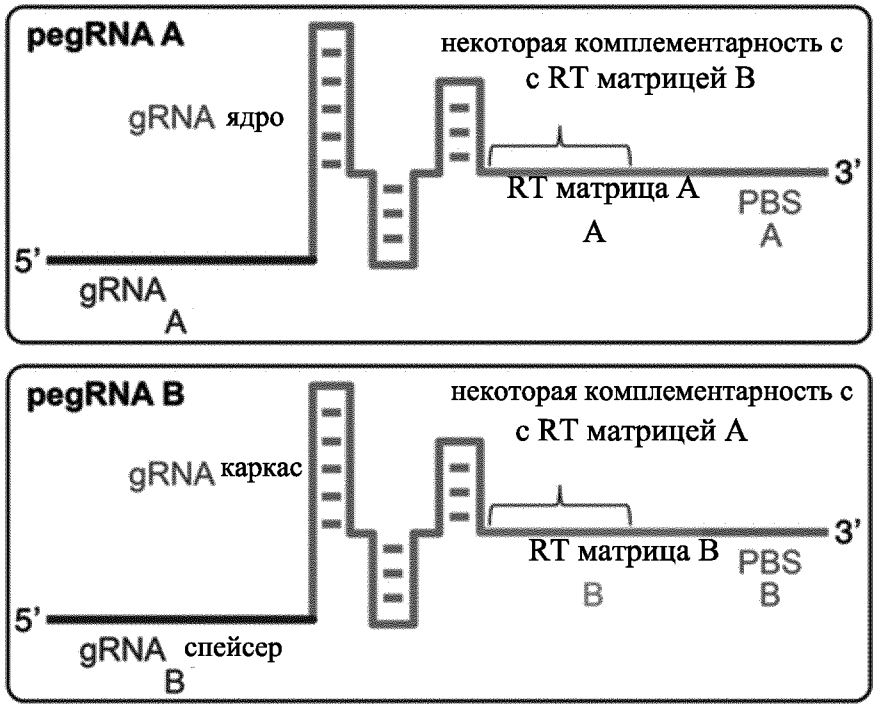
HEK3 в клетках HEK293T (SEQ ID NOs: 791,4286,4286 4288,4286,4289,4286,4290)



Фиг. 91A

pegRNA1: (SEQ ID NO: 4291)
5’ – GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC
CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAAATGATGGTGATGATGGTGTTCGTGCTCAGTCTG – 3’
pegRNA2: (SEQ ID NO: 4292)
5’ – GTCAACCAGTATCCCGGTGCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC
CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAACACCATCATCACCATCATTTCCGGGATACTGG – 3’

Фиг. 91B



Примеры возможных RT матриц в PEgRNA A и B

RT матрица A Комплементарность RT матрице A

5' - CACUGGAUCGAUCGAUGCUGACUGUAGCUACAGCUCA - 3'

(SEQ ID NO: 4293)

RT матрица B Комплементарность RT матрице B

5' - GUCAGCAUCGAUCGAUCCAGUGUGUAGCUACAGCUCA - 3'

(SEQ ID NO: 4294)

RT продукт A Комплементарность RT продукту A

5' - TGAGCTGTAGCTACAGTCAGCATCGATCGATCCAGTG - 3'

(SEQ ID NO: 4295)

RT продукт B Комплементарность RT продукту B

5' - TGAGCTGTAGCTACACACTGGATCGATCGATGCTGAC - 3'

(SEQ ID NO: 4296)

Предпочтительные лоскуты из RT продукотв A и B

(SEQ ID NO: 4295) Область комплементарности между лоскутами для облегчения отжига

5' - TGAGCTGTAGCTACAGTCAGCATCGATCGATCCAGTG - 3'

3' - CAGTCGTAGCTAGCTAGGTCACTGAGCTGTAGCTACA - 5'

(SEQ ID NO: 4297)

Фиг. 92

pegRNA

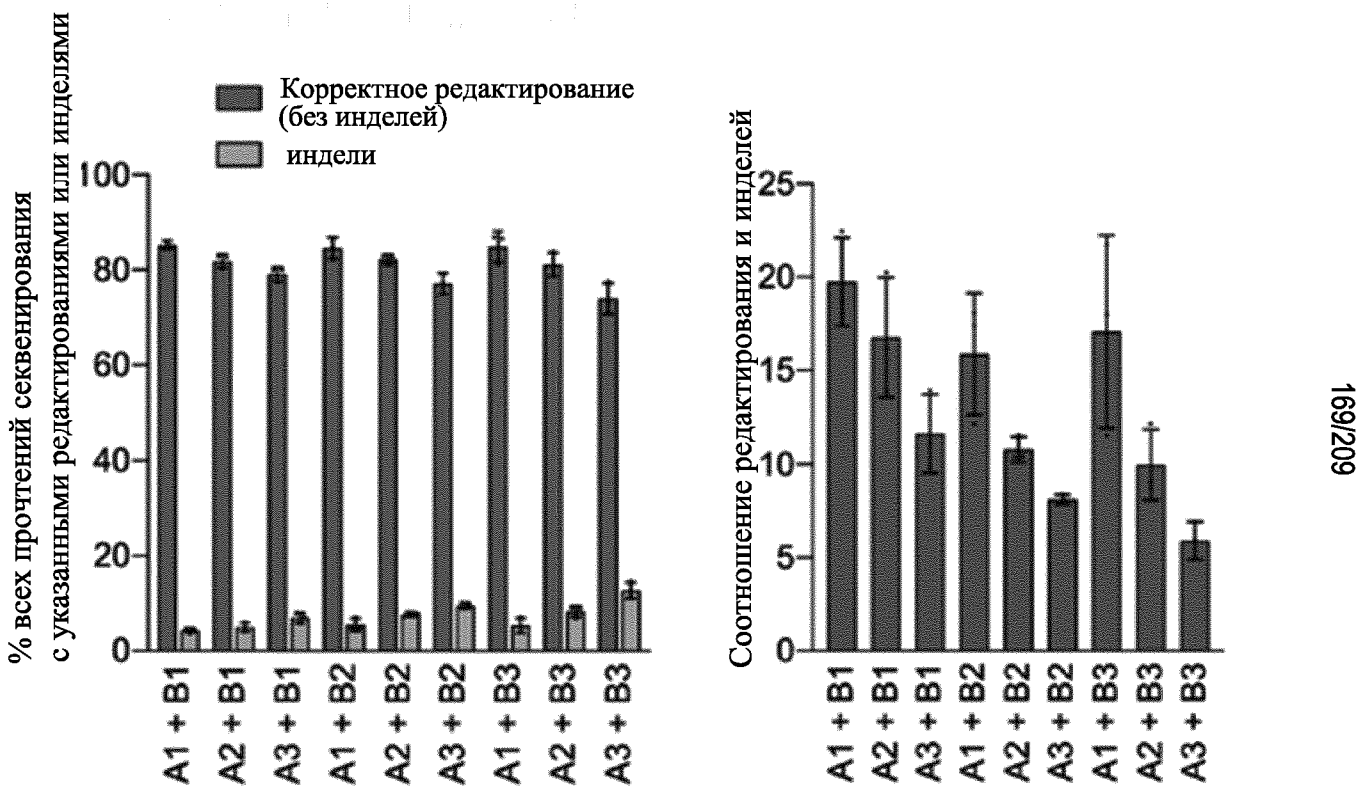
A1 – Bxbl *attB* (+) 1 до 38
A2 – Bxbl *attB* (+) 1 до 34
A3 – Bxbl *attB* (+) 1 до 30

B1 – Bxbl *attB* (-) 1 до 38
B2 – Bxbl *attB* (-) 5 до 38
B3 – Bxbl *attB* (-) 9 до 38

Перекрытие между лоскутами (nt)

	A1	A2	A3
B1	38	34	30
B2	34	30	26
B3	30	26	22

Вставка сайта attB Bxbl (38 п.о.) и делеция 90 п.о.



Фиг. 93A

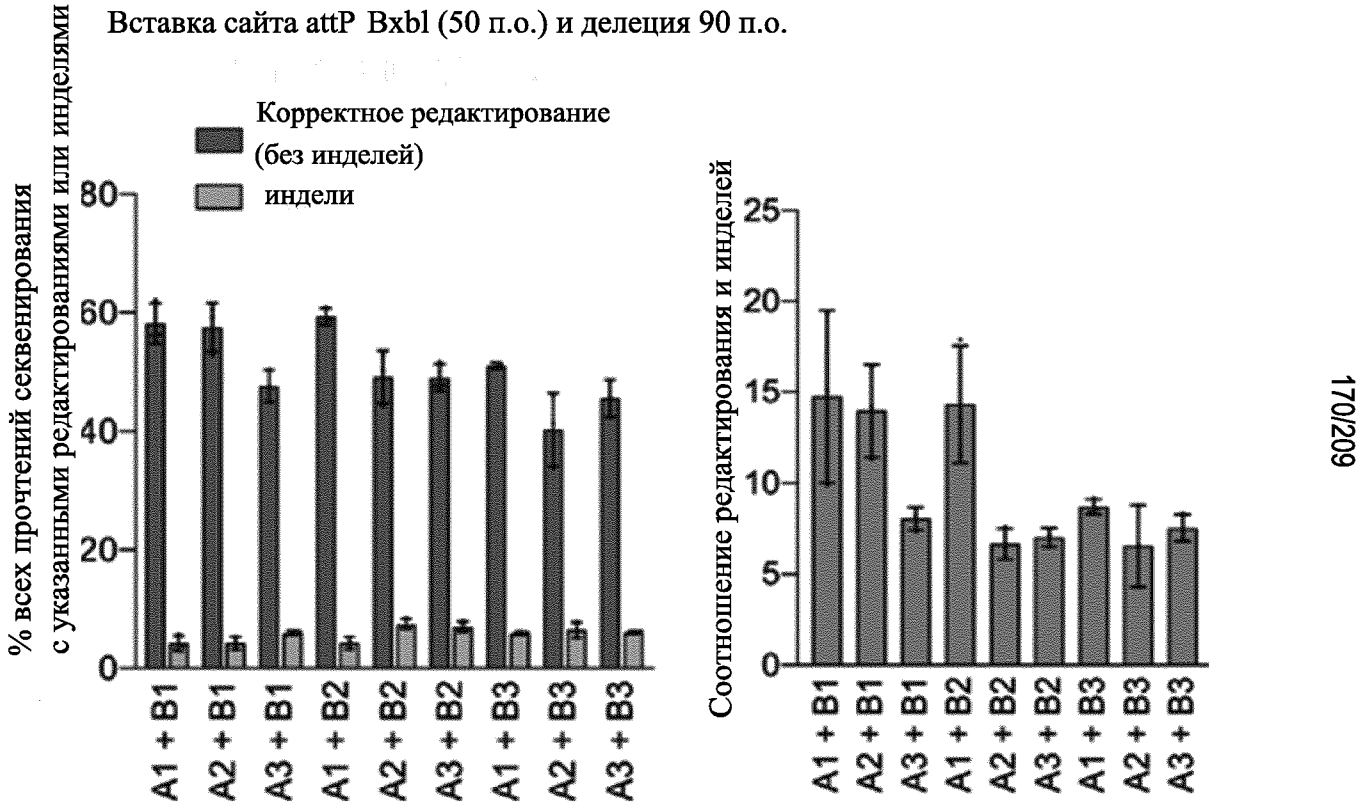
pegRNA

A1 – Bxbl *attP* (+) 1 до 43
A2 – Bxbl *attP* (+) 1 до 39
A3 – Bxbl *attP* (+) 1 до 35

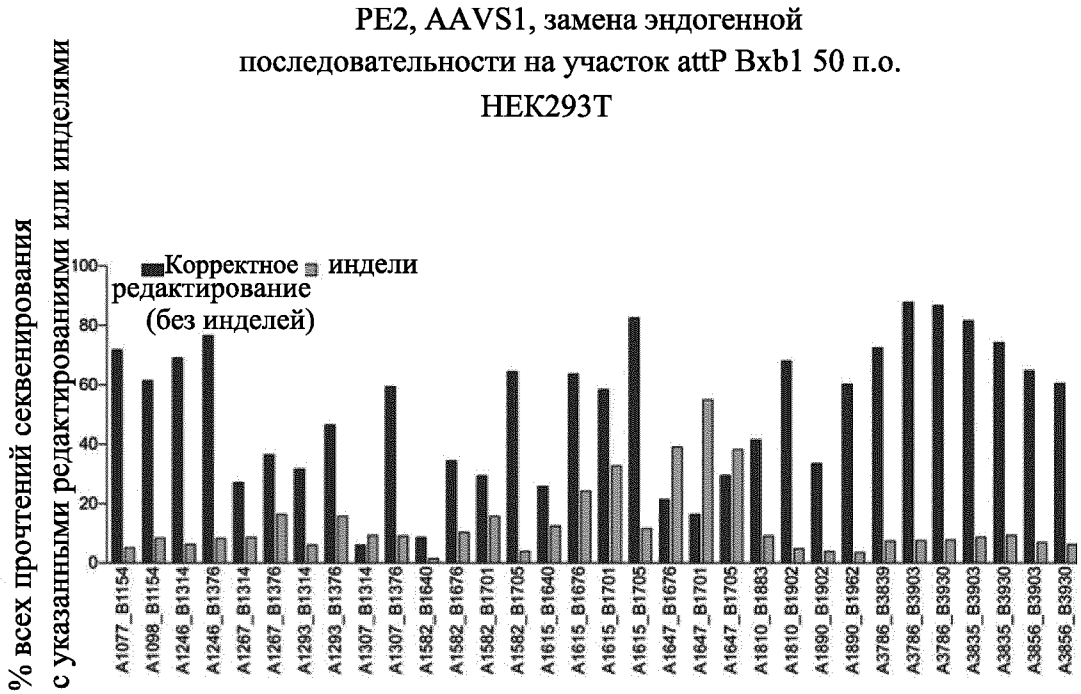
B1 – Bxbl *attP* (-) 7 до 50
B2 – Bxbl *attP* (-) 11 до 50
B3 – Bxbl *attP* (-) 15 до 50

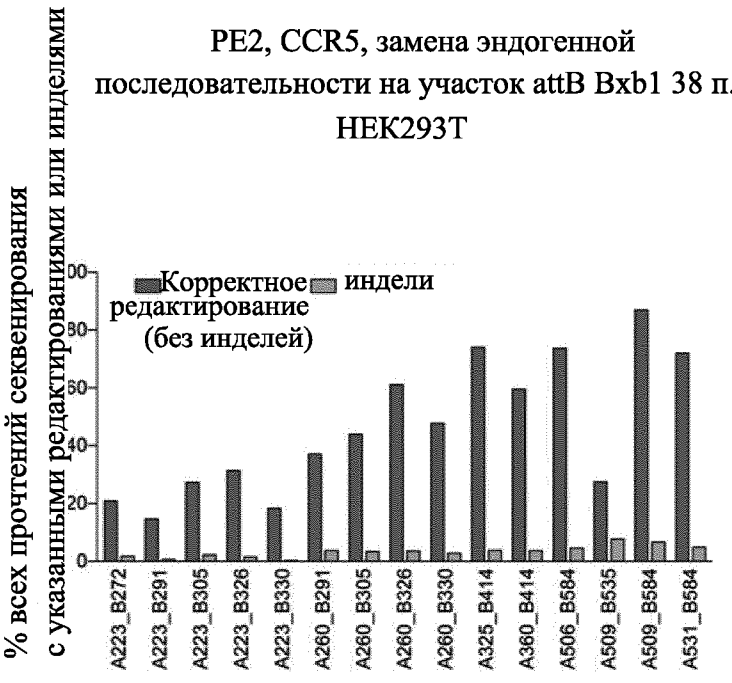
Перекрытие между лоскутами (nt)

	A1	A2	A3
B1	37	33	29
B2	33	29	25
B3	29	25	21

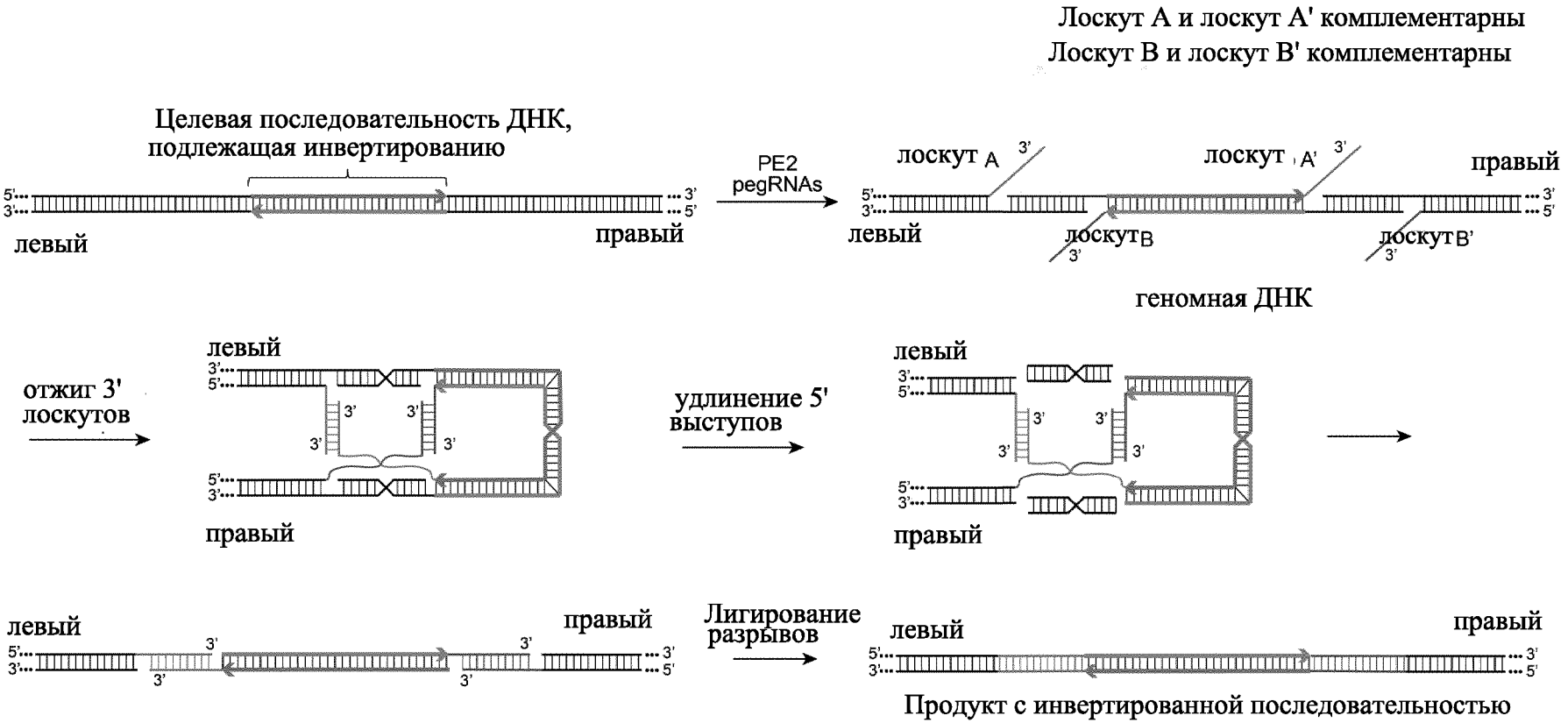


Фиг. 93B


PE2, AAVS1, замена эндогенной последовательности на участок attP Bxb1 50 п.о.
HEK293T
% всех прочтений секвенирования с указанными редактированиями или инделями
Корректное редактирование (без инделей)
индели
A1077_B1154
A1098_B1154
A1246_B1314
A1246_B1376
A1267_B1314
A1267_B1376
A1293_B1314
A1293_B1376
A1307_B1314
A1307_B1376
A1582_B1640
A1582_B1676
A1582_B1701
A1582_B1705
A1615_B1640
A1615_B1676
A1615_B1701
A1615_B1705
A1647_B1676
A1647_B1701
A1647_B1705
A1810_B1883
A1810_B1902
A1890_B1902
A1890_B1962
A3786_B3839
A3786_B3903
A3786_B3930
A3835_B3903
A3835_B3930
A3856_B3903
A3856_B3930


PE2, CCR5, замена эндогенной последовательности на участок attB Bxb1 38 п.о.
HEK293T
% всех прочтений секвенирования с указанными редактированиями или инделями
Корректное редактирование (без инделей)
индели
A223_B272
A223_B291
A223_B305
A223_B326
A223_B330
A260_B291
A260_B305
A260_B326
A260_B330
A325_B414
A360_B414
A506_B584
A509_B535
A509_B584
A531_B584

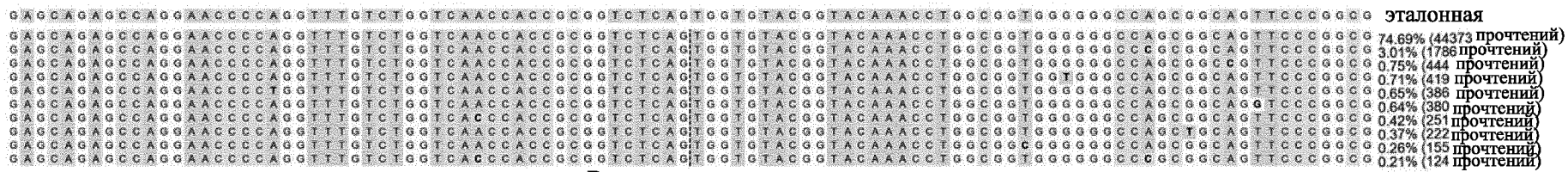
Фиг. 94


Лоскут A и лоскут A' комплементарны
Лоскут B и лоскут B' комплементарны
Целевая последовательность ДНК, подлежащая инвертированию
PE2
pegRNAs
лоскут A
лоскут A'
правый
левый
лоскут B
лоскут B'
геномная ДНК
отжиг 3' лоскутов
удлинение 5' выступов
Лигирование разрывов
Продукт с инвертированной последовательностью

Фиг. 95

HTS ампликонов ПЦР: (SEQ ID NOs: 4298,4298-4307)

Прочтения, выровненные с ожидаемым соединением A (содержит сайт attP) - 90% всех выровненных прочтений, 8-10% инделей

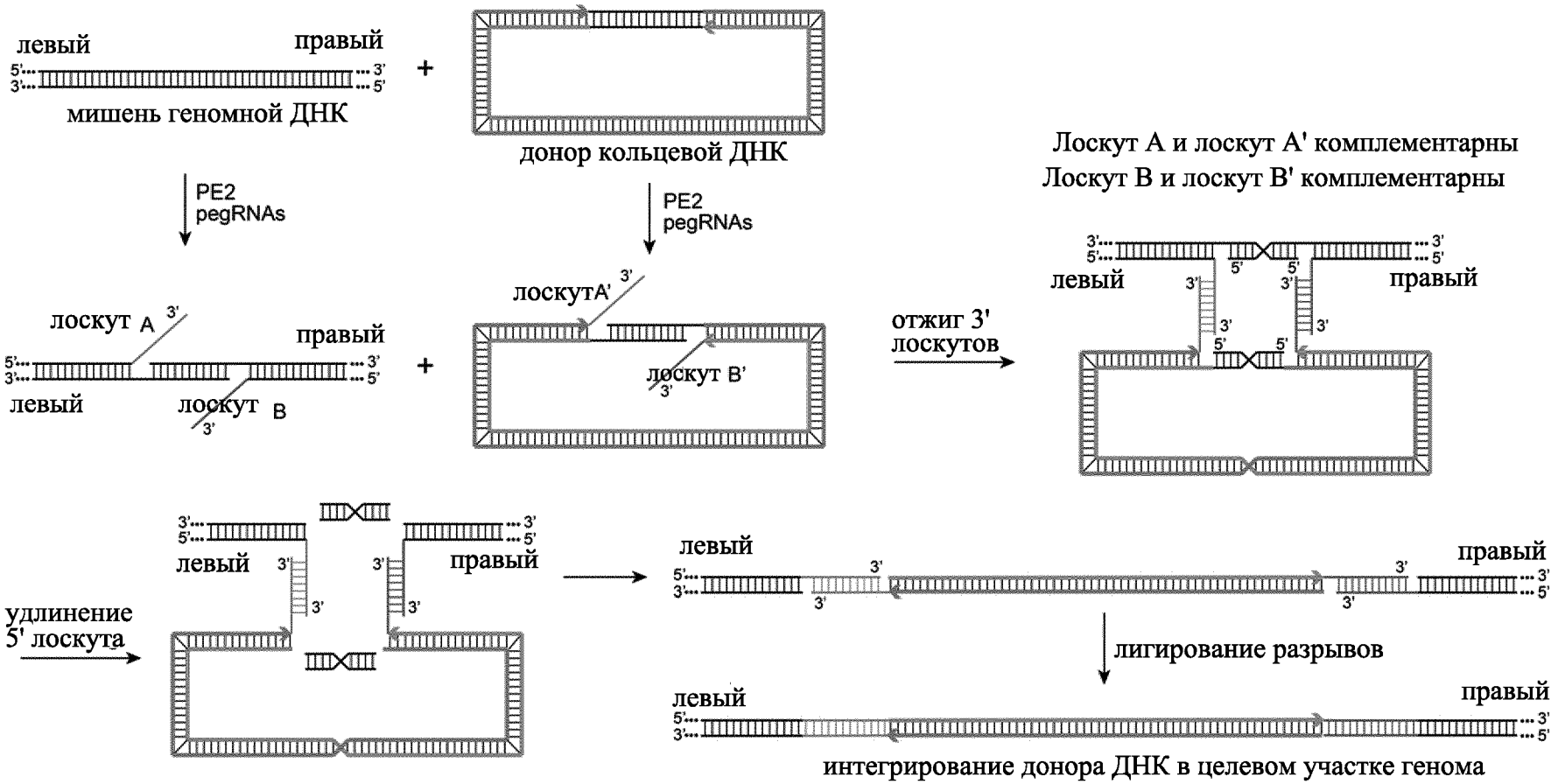


(SEQ ID NOs: 4308,4308-4317)

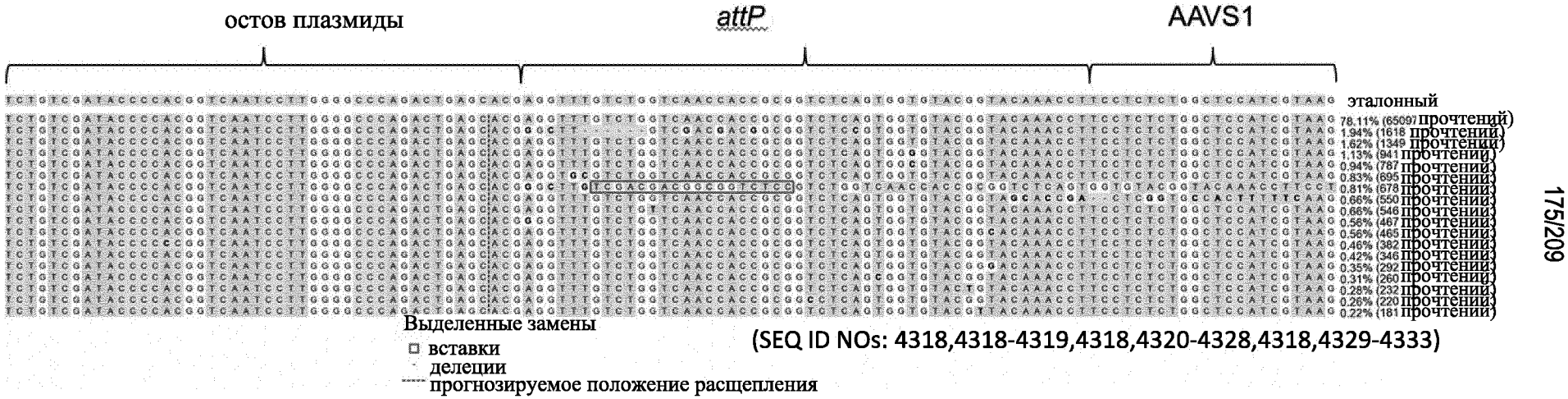
Прочтения, выровненные с ожидаемым соединением B (содержит сайт attB) - 93% всех выровненных прочтений, 2,5% инделей



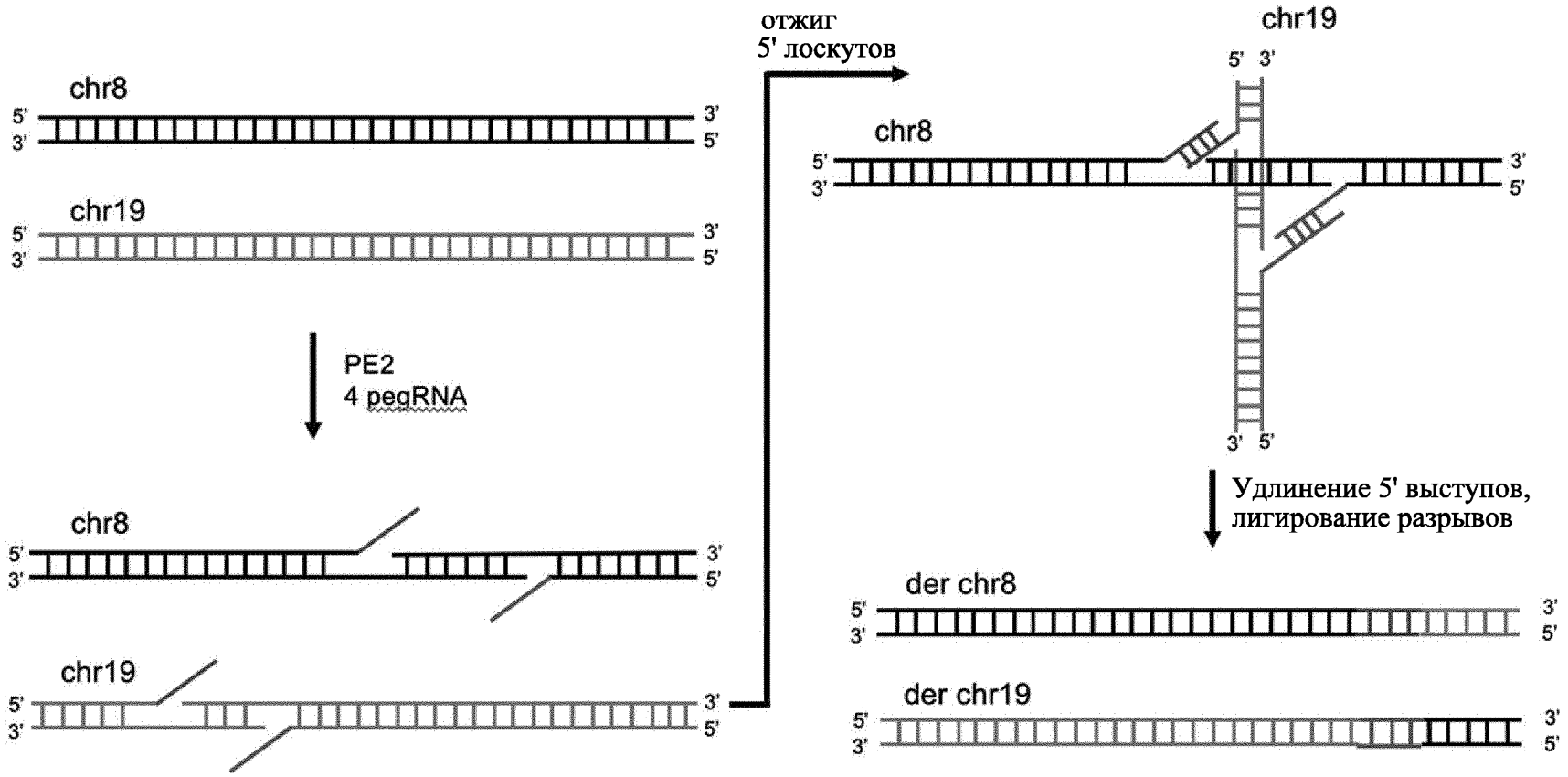
Фиг. 96


левый
правый
мишень геномной ДНК
донор кольцевой ДНК
PE2
pegRNAs
PE2
pegRNAs
локут A
правый
левый
локут B
локутA'
локут B'
Локут A и локут A' комплементарны
Локут B и локут B' комплементарны
отжиг 3'
локутов
левый
правый
удлинение
5' локута
левый
правый
левый
правый
лигирование разрывов
левый
правый
интегрирование донора ДНК в целевом участке генома

Фиг. 97A

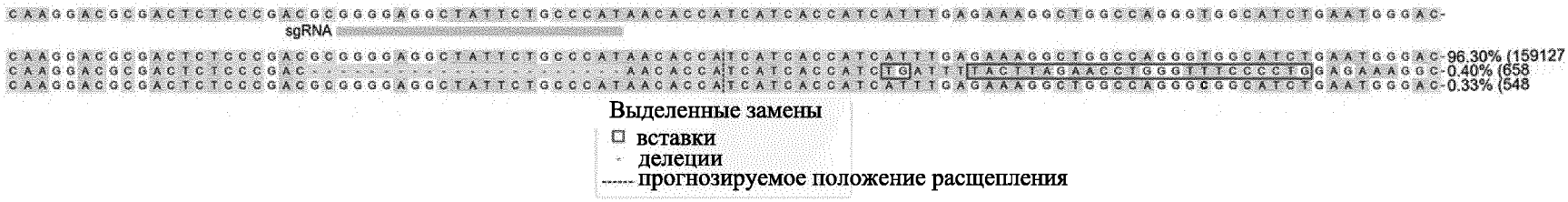

остов плазмиды
attP
AAVS1
эталонный
78.11% (65097 прочтений)
1.94% (1618 прочтений)
1.62% (1349 прочтений)
1.13% (941 прочтений)
0.94% (787 прочтений)
0.83% (695 прочтений)
0.81% (678 прочтений)
0.66% (550 прочтений)
0.66% (546 прочтений)
0.56% (467 прочтений)
0.56% (465 прочтений)
0.46% (382 прочтений)
0.42% (346 прочтений)
0.35% (292 прочтений)
0.31% (260 прочтений)
0.28% (232 прочтений)
0.26% (220 прочтений)
0.22% (181 прочтений)
Выделенные замены
вставки
делеции
прогнозируемое положение расщепления
(SEQ ID NOs: 4318,4318-4319,4318,4320-4328,4318,4329-4333)

Фиг. 97B


chr8
5'
3'
3'
5'
chr19
5'
3'
3'
5'
PE2
4 pegRNA
chr8
5'
3'
3'
5'
chr19
5'
3'
3'
5'
отжиг
5' лоскутов
chr19
5' 3'
chr8
5'
3'
3'
5'
3' 5'
Удлинение 5' выступов,
лигирование разрывов
der chr8
5'
3'
3'
5'
der chr19
5'
3'
3'
5'

Фиг. 98A

HTS амплмконов ПЦР (pegRNA набор №5 rep2):
Прочтения, выровненные с ожидаемым соединением 1 (содержит метку His6) -99.5% всех выровненных прочтений
(SEQ ID NOs: 4334,4334-4336)

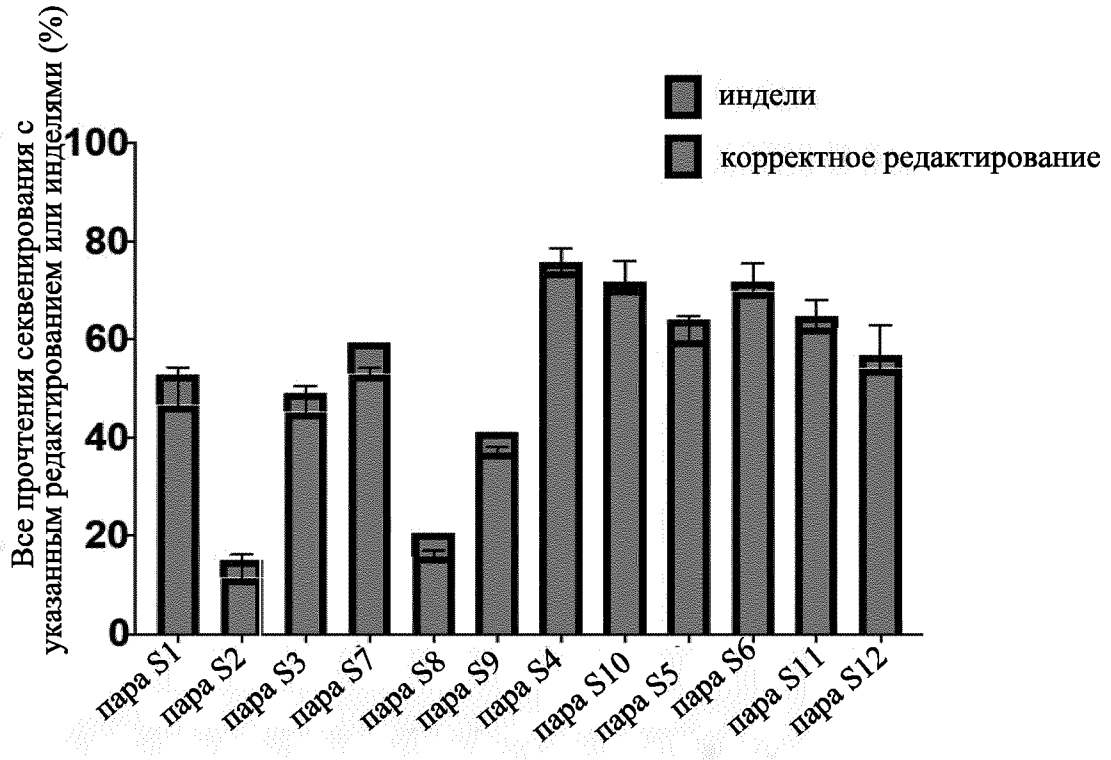


HTS амплмконов ПЦР (pegRNA набор №5 rep2):
Прочтения, выровненные с ожидаемым соединением 2 (содержит случайную последовательность 22nt) _-97.9% всех выровненных прочтений



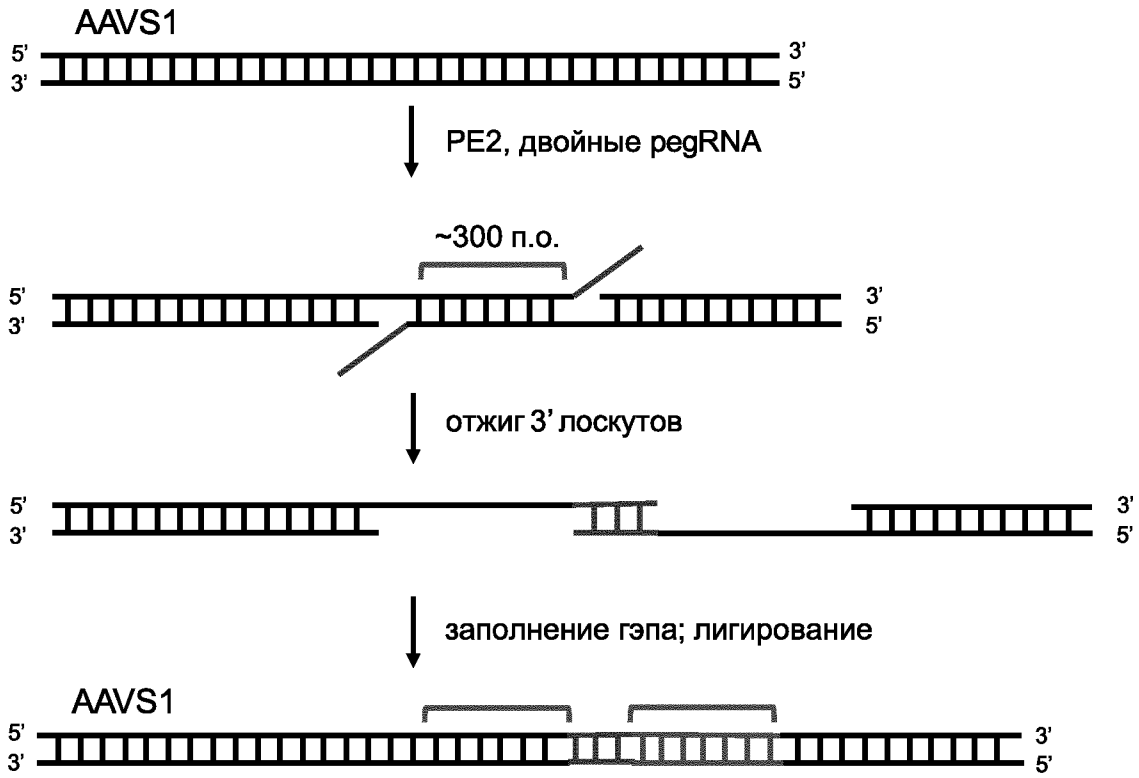
(SEQ ID NOs: 4337,4337-4341)

Фиг. 98B


индели
корректное редактирование
Все прочтения секвенирования с указанным редактированием или инделями (%)
100
80
60
40
20
0
пара S1
пара S2
пара S3
пара S7
пара S8
пара S9
пара S4
пара S10
пара S5
пара S6
пара S11
пара S12

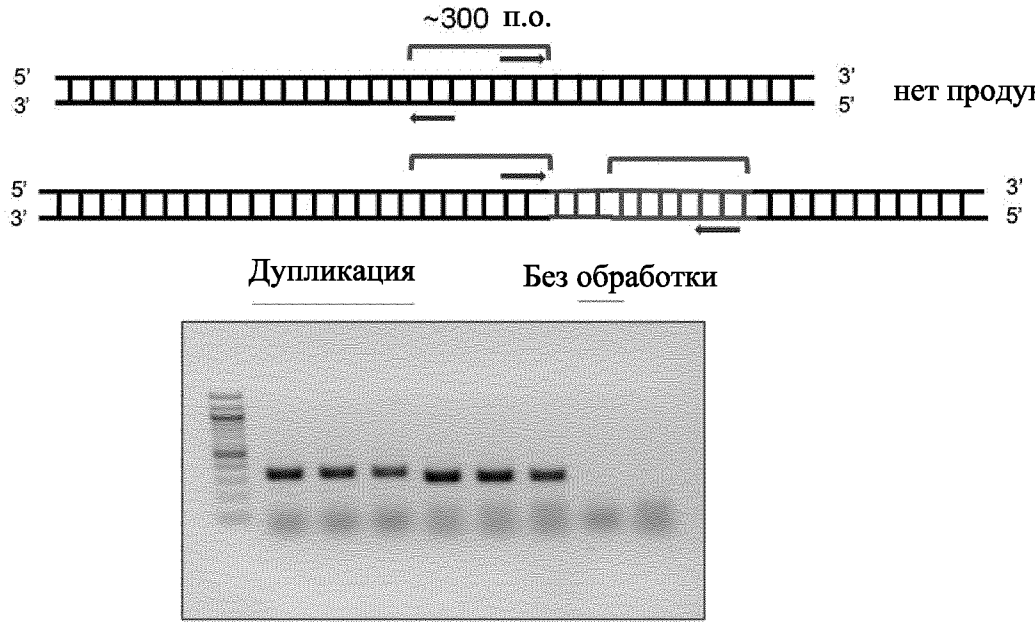
Фиг. 99

Частичная дупликация AAVS1 (~300 п.о.)

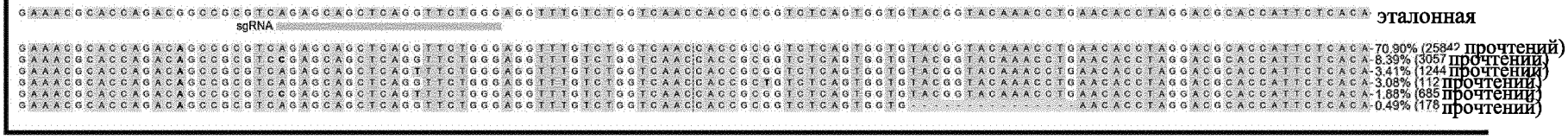


Фиг. 100A

Частичная дупликация AAVS1 (около 300 п.о.)

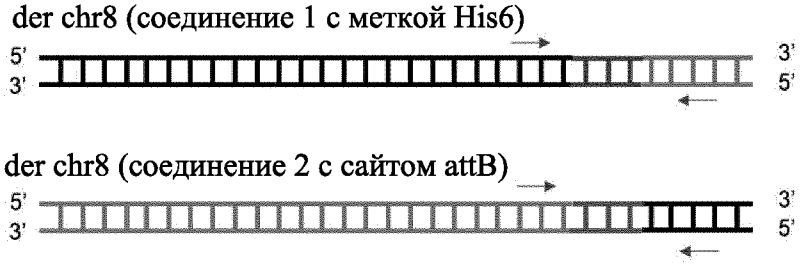


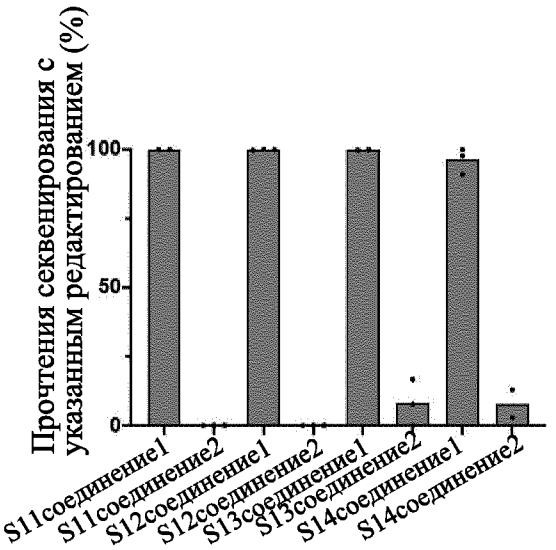
Прочтения, выровненные с ожидаемыми аллелями) - около 94% всех выровненных прочтений, около 6% инделей



(SEQ ID NOs: 4337,4337-4341)

Фиг. 100B





HTS ампликонов ПЦР (pegRNA набор № 2 rep 1 в качестве примера)

Прочтения, выровненные с ожидаемым соединение 1 (содержит метку His) - 99,8% всех выровненных прочтений, 0,2% инделей

(SEQ ID NO: 4349)

CAAGGACGCGACTCTCCCGACGCGGGGAGGCTATTCTGCCCATAACACCATCATCACCATCATTTAGAGACGCAAACACAGCCACCACCCAAGTGATCAC- эталонная

sgRNA

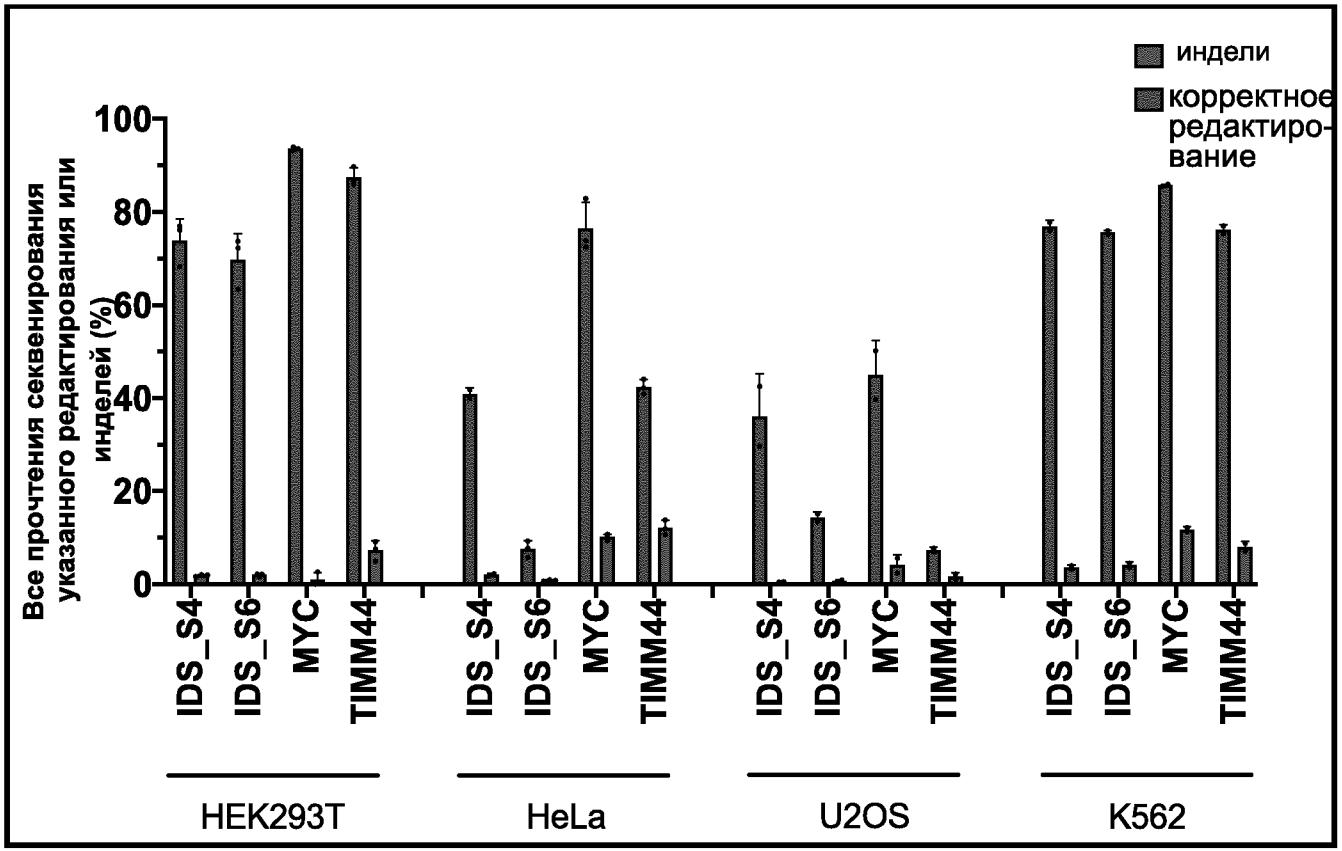
CAAGGACGCGACTCTCCCGACGCGGGGAGGCTATTCTGCCCATAACACCATCATCACCATCATTTAGAGACGCAAACACAGCCACCACCCAAGTGATCAC- 97.39% (48480 прочтений)

(SEQ ID NO: 4349)

Выделенные замены

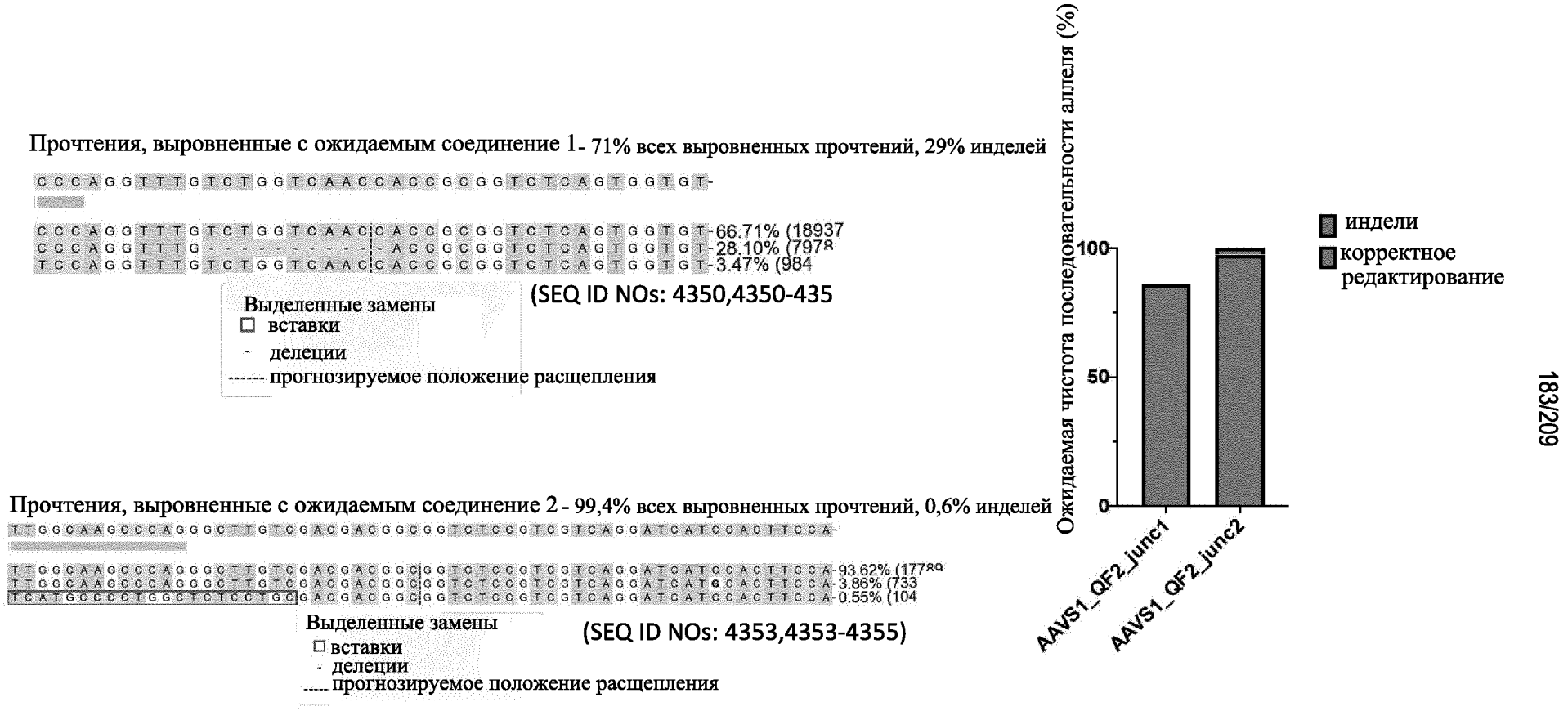
- вставки
- делеции
- прогнозируемое положение расщепления

Фиг. 101

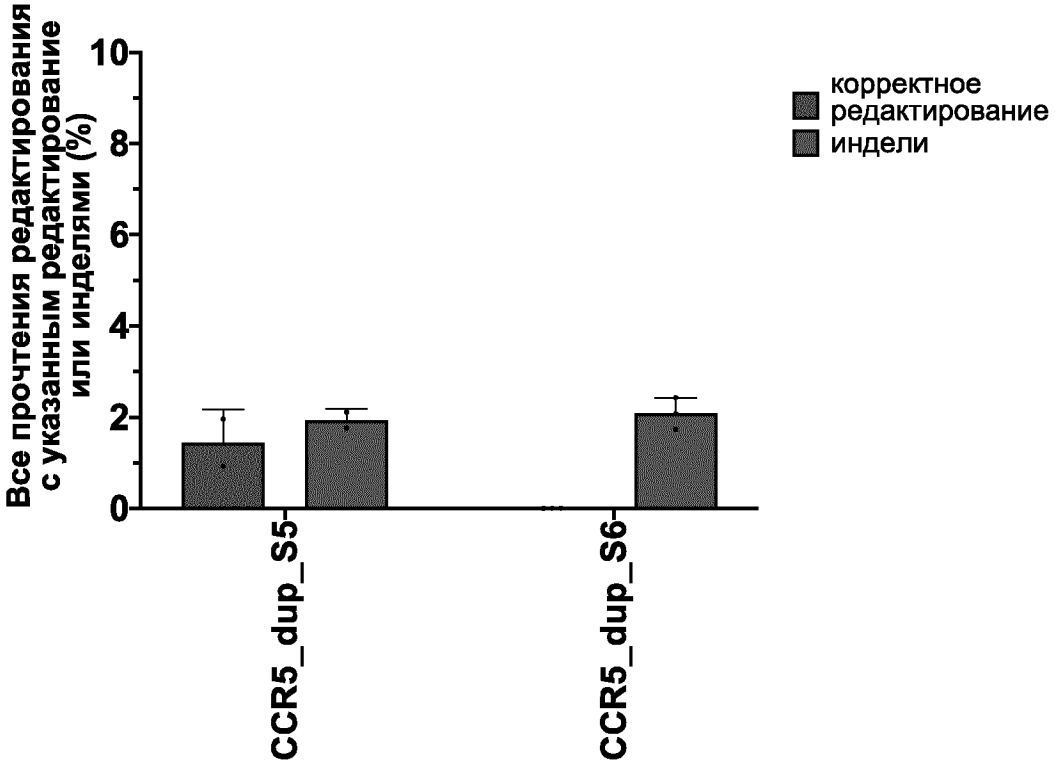


Разрез двойной pegRNA:
IDS S4: 116nt
IDS S6: 75nt
MYC: 263nt
TIMM44: 157nt

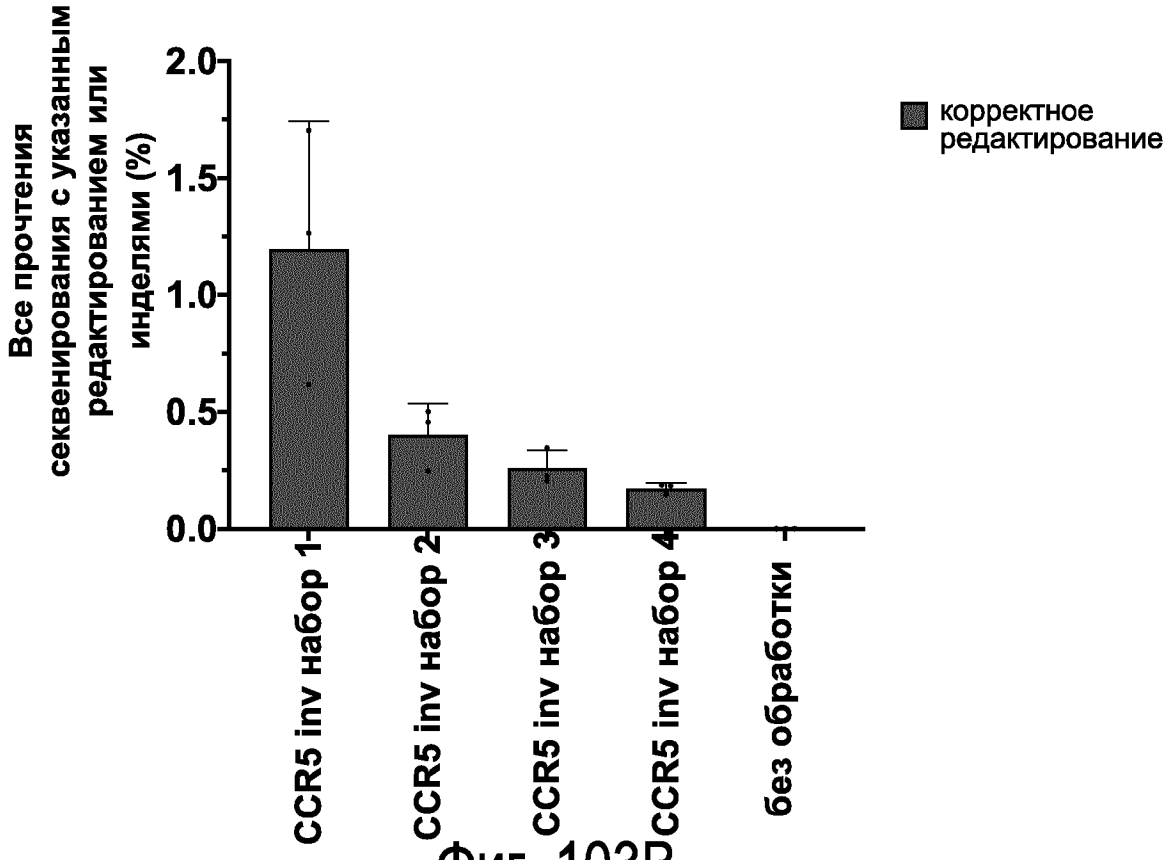
Фиг. 102A


Прочтения, выровненные с ожидаемым соединение 1- 71% всех выровненных прочтений, 29% инделей
CCCAGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGT-
CCCAGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGT-66.71% (18937
CCCAGGTTTG----------ACCGCGGTCTCAGTGGTGT-28.10% (7978
TCCAGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGT-3.47% (984
Выделенные замены
вставки
делеции
прогнозируемое положение расщепления
(SEQ ID NOs: 4350,4350-435
Прочтения, выровненные с ожидаемым соединение 2 - 99,4% всех выровненных прочтений, 0,6% инделей
TTGGCAAGCCCAGGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCACTTCCA-
TTGGCAAGCCCAGGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCACTTCCA-93.62% (17789
TTGGCAAGCCCAGGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATGCACTTCCA-3.86% (733
TCATGCCCCTGGCTCTCCTGCGACGACGGCGGTCTCCGTCGTCAGGATCATCCACTTCCA-0.55% (104
Выделенные замены
вставки
делеции
прогнозируемое положение расщепления
(SEQ ID NOs: 4353,4353-4355)
Ожидаемая чистота последовательности аллеля (%)
100
50
0
AAVS1_QF2_junc1
AAVS1_QF2_junc2
индели
корректное редактирование

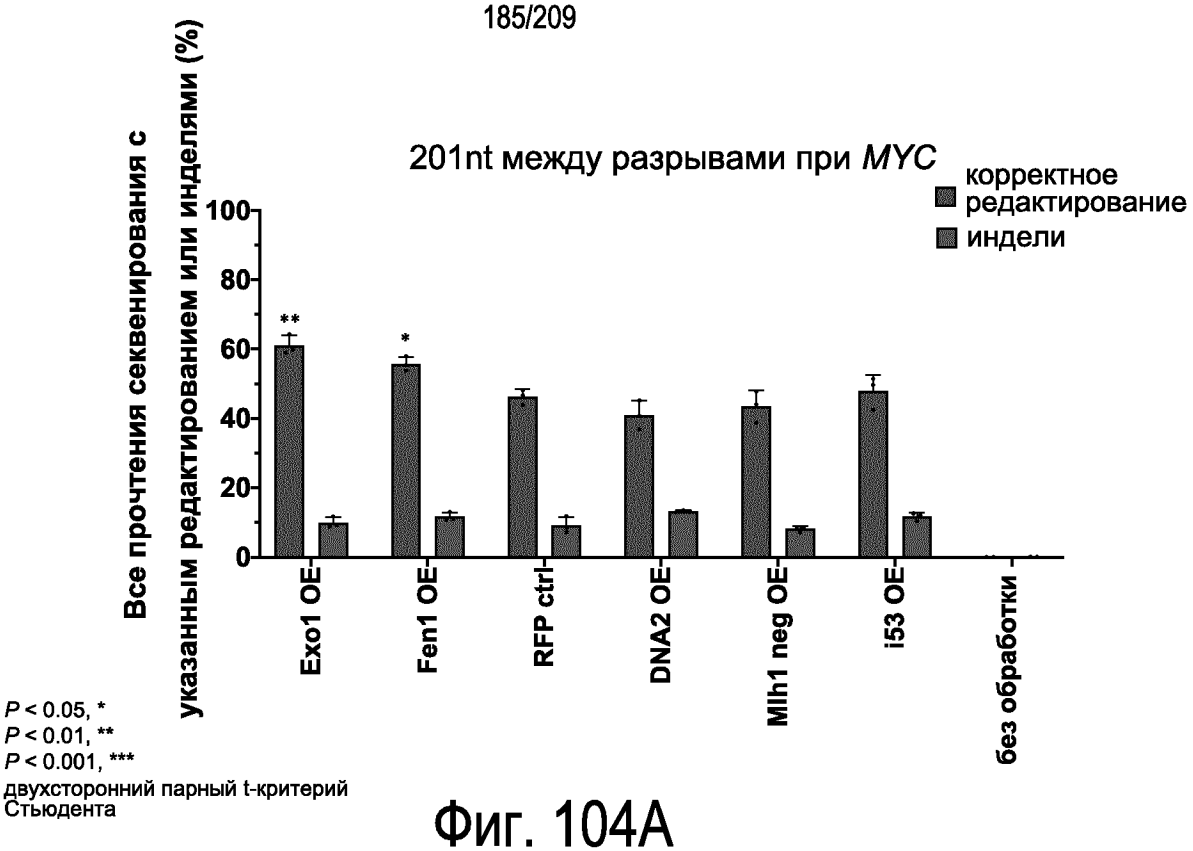
Фиг. 102B



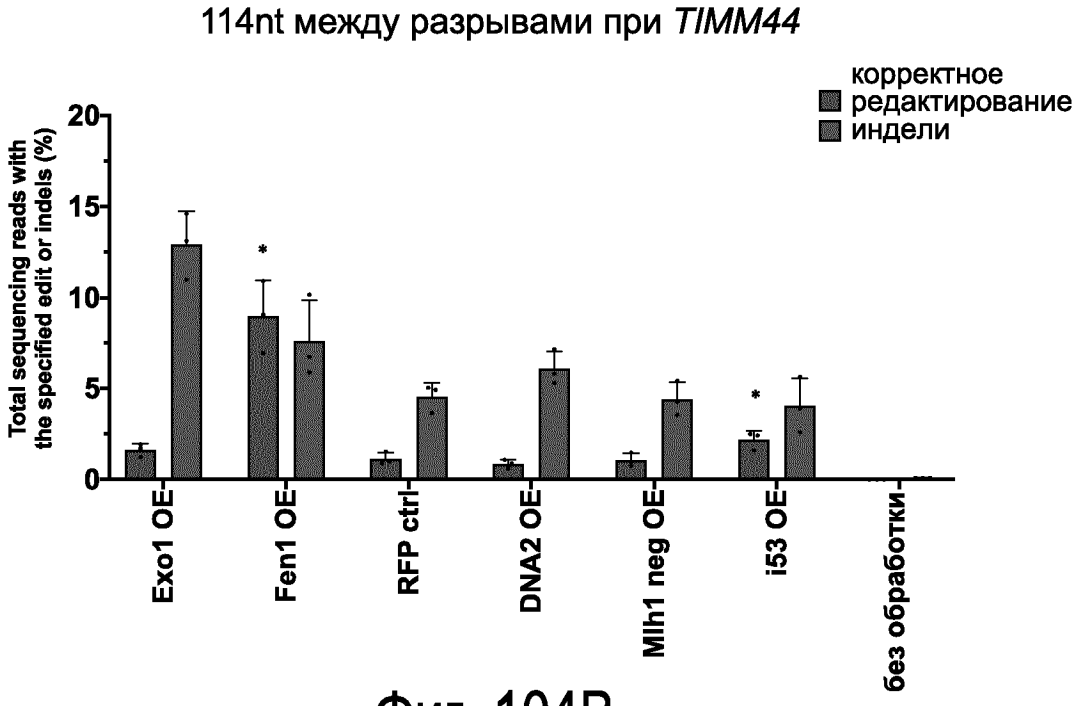
Фиг. 103A



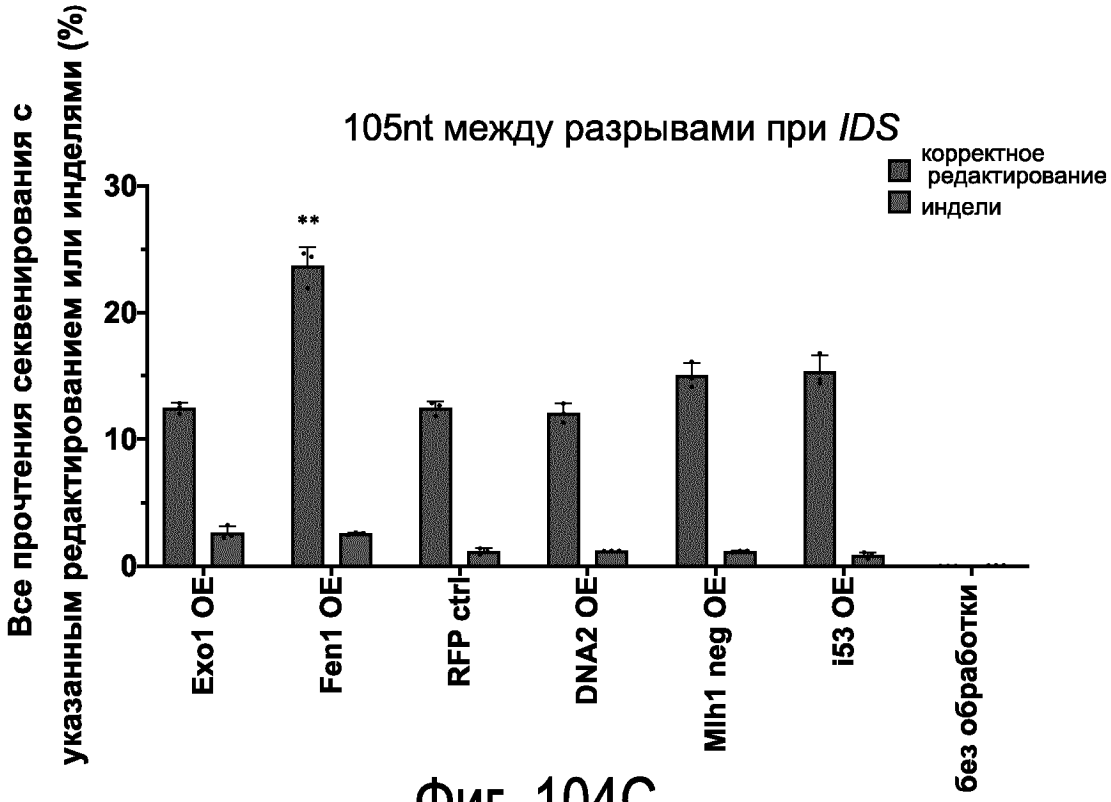
Фиг. 103B



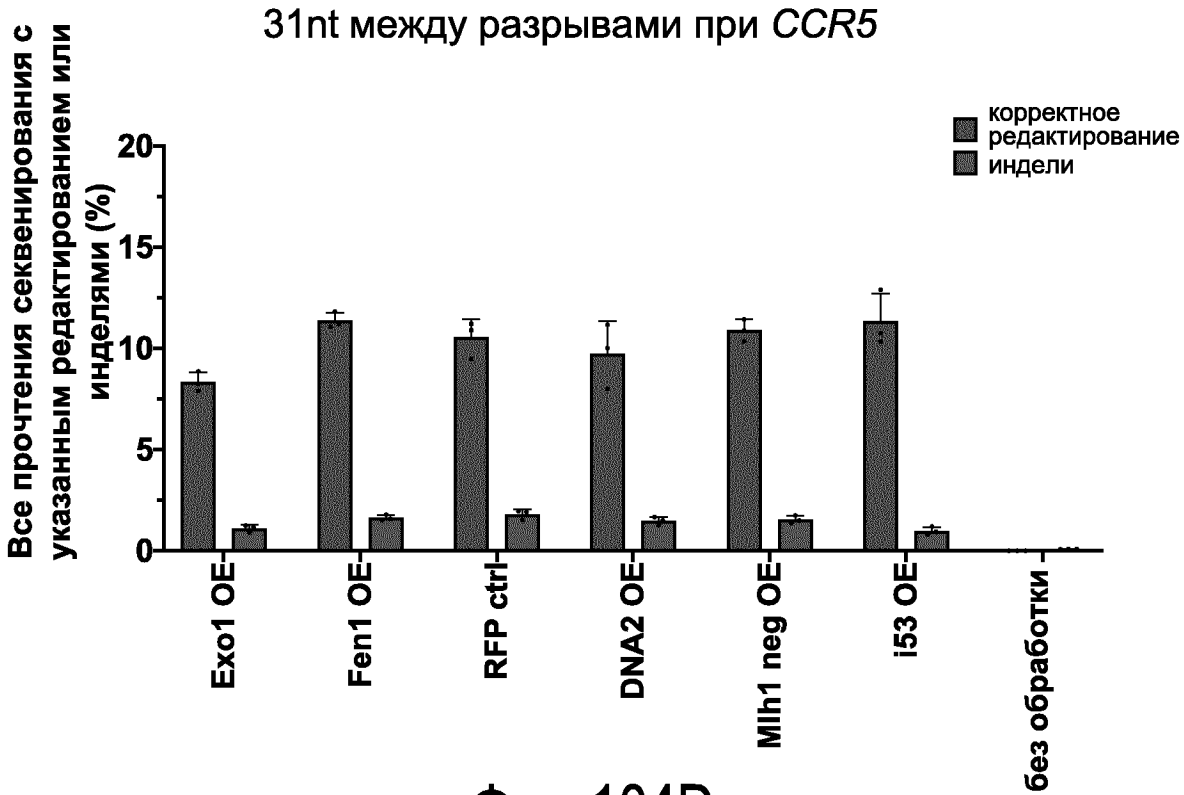
Фиг. 104A



Фиг. 104B

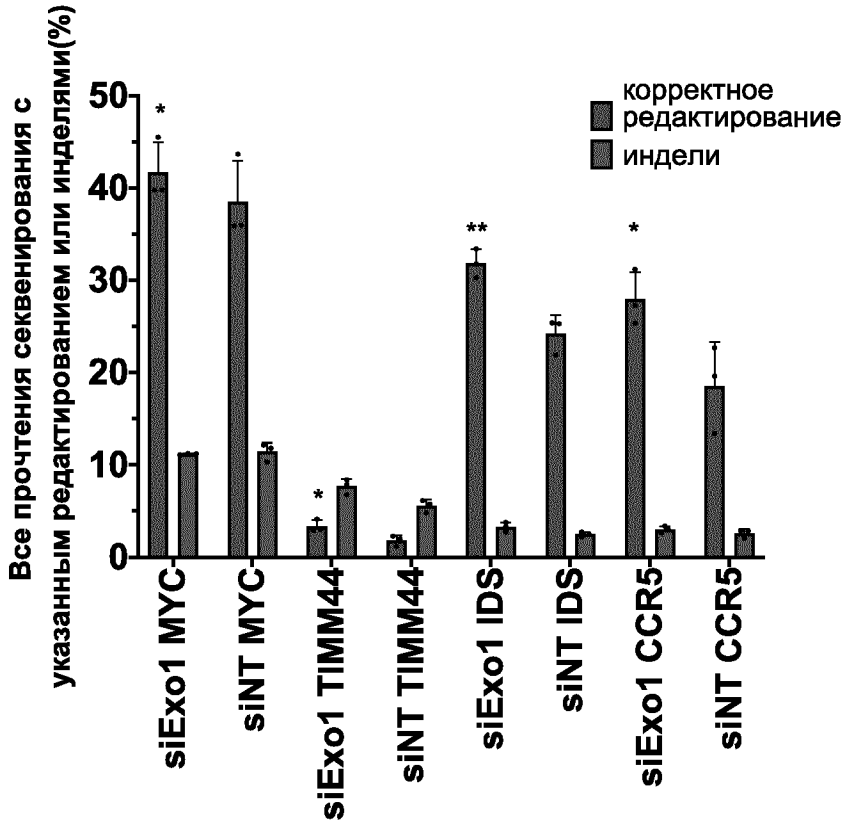


Фиг. 104C



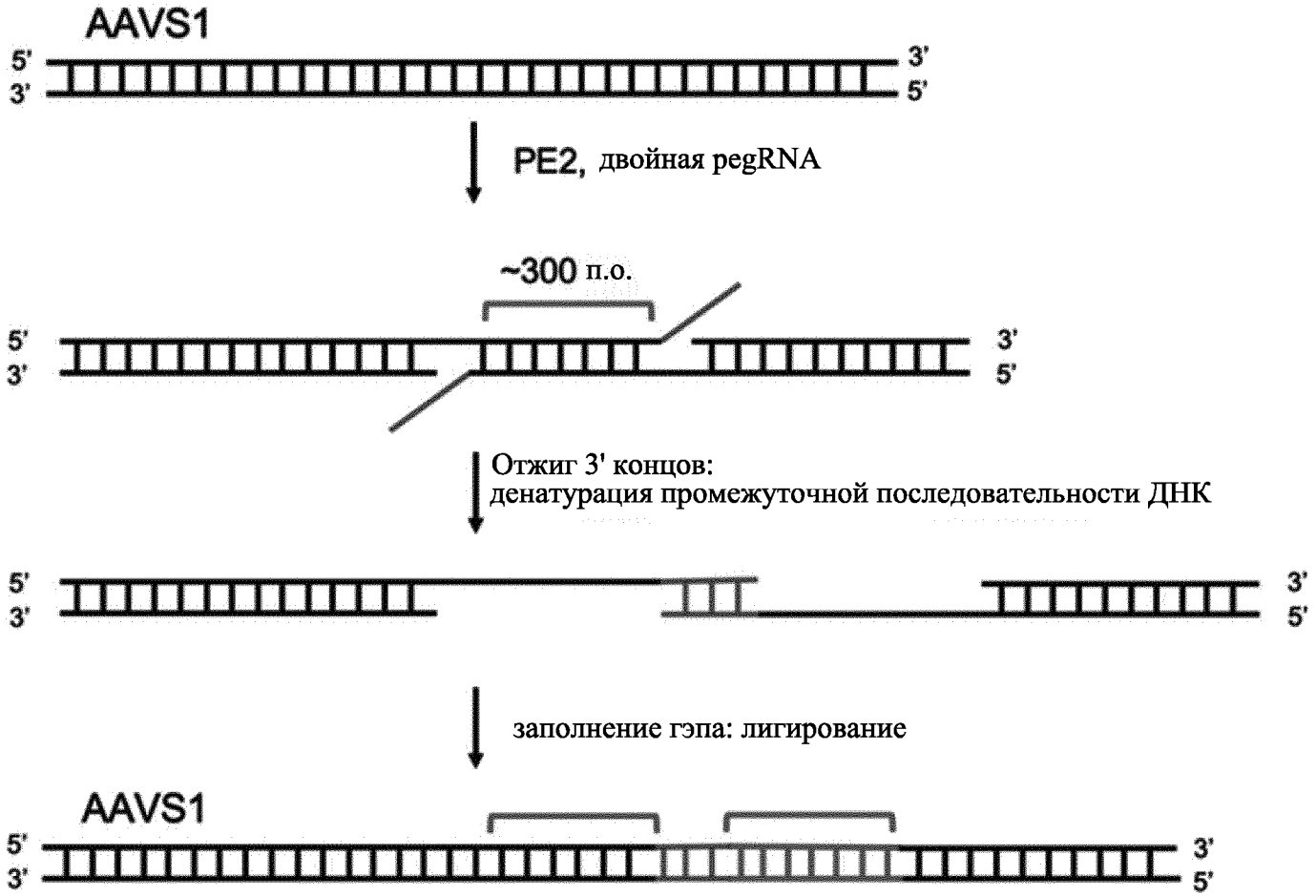
Фиг. 104D

Эксперимент по нокдауну siExo1



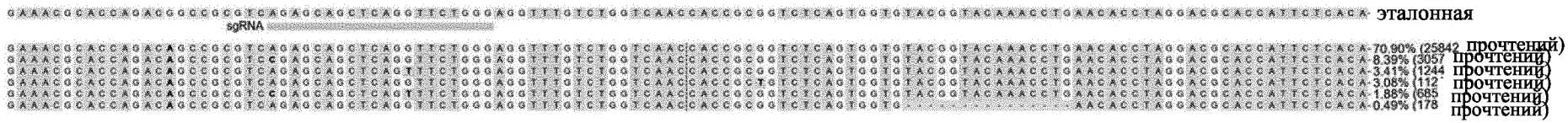
P < 0.05, *
P < 0.01, **
P < 0.001, ***

Фиг. 104E


AAVS1
5'
3'
3'
5'
PE2, двойная pegRNA
~300 п.о.
5'
3'
3'
5'
Отжиг 3' концов:
денатурация промежуточной последовательности ДНК
5'
3'
3'
5'
заполнение гэпа: лигирование
AAVS1
5'
3'
3'
5'

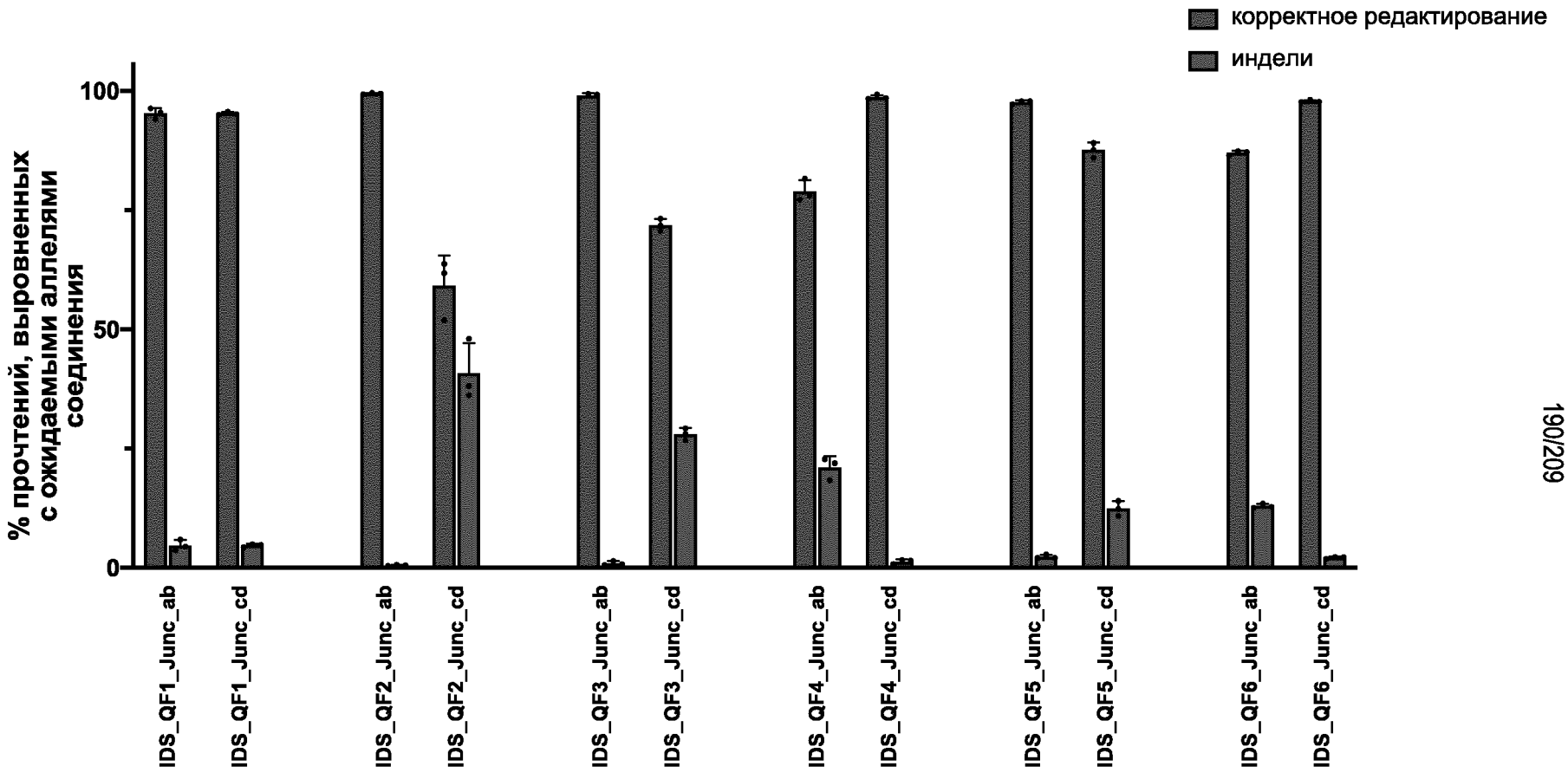
Фиг. 105A

Прочтения, выровненные с ожидаемыми аллелями (дупликация) - около 94% всех выровненных прочтений, около 6% инделей

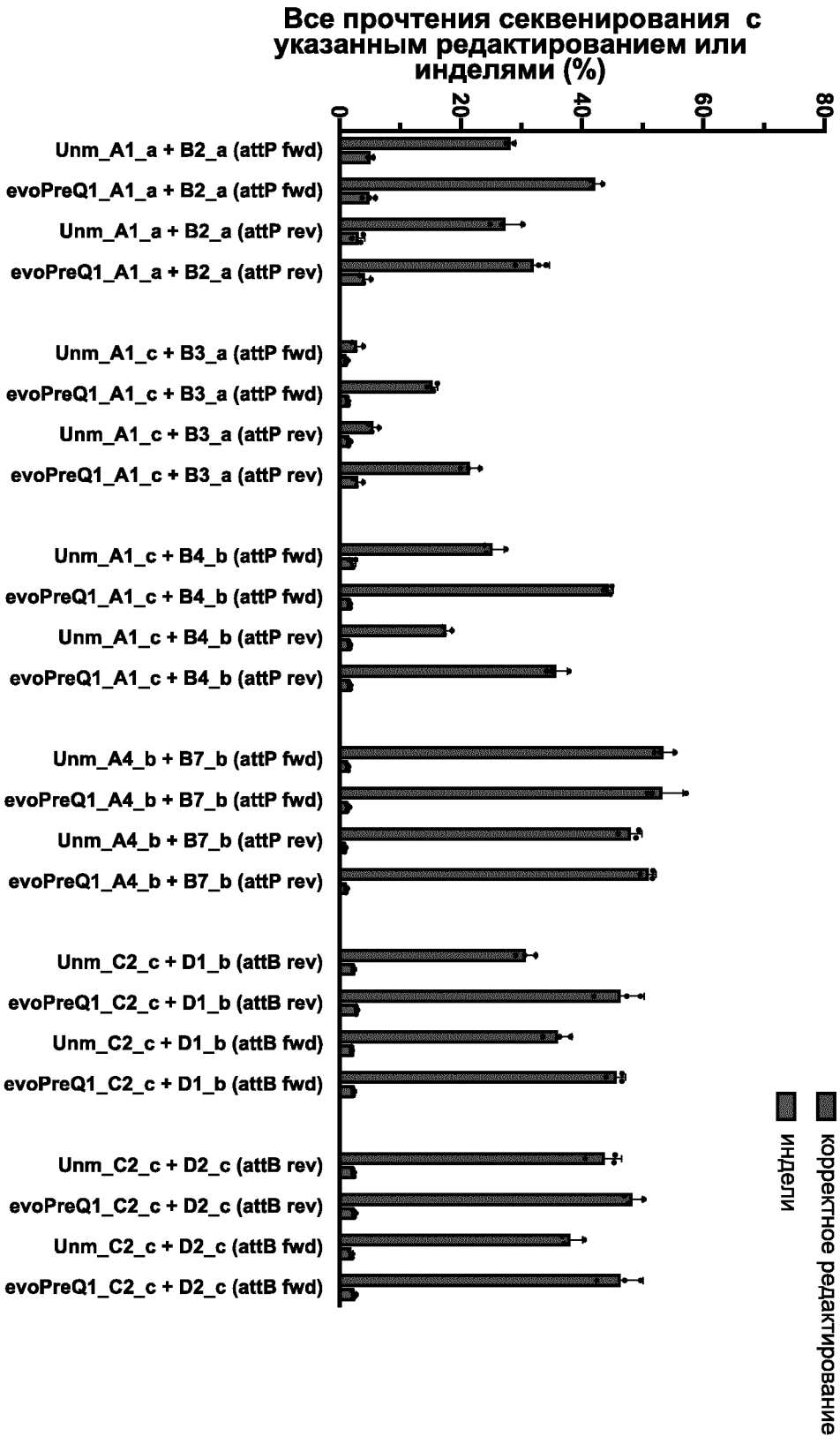


(SEQ ID NOs: 4342-4348)

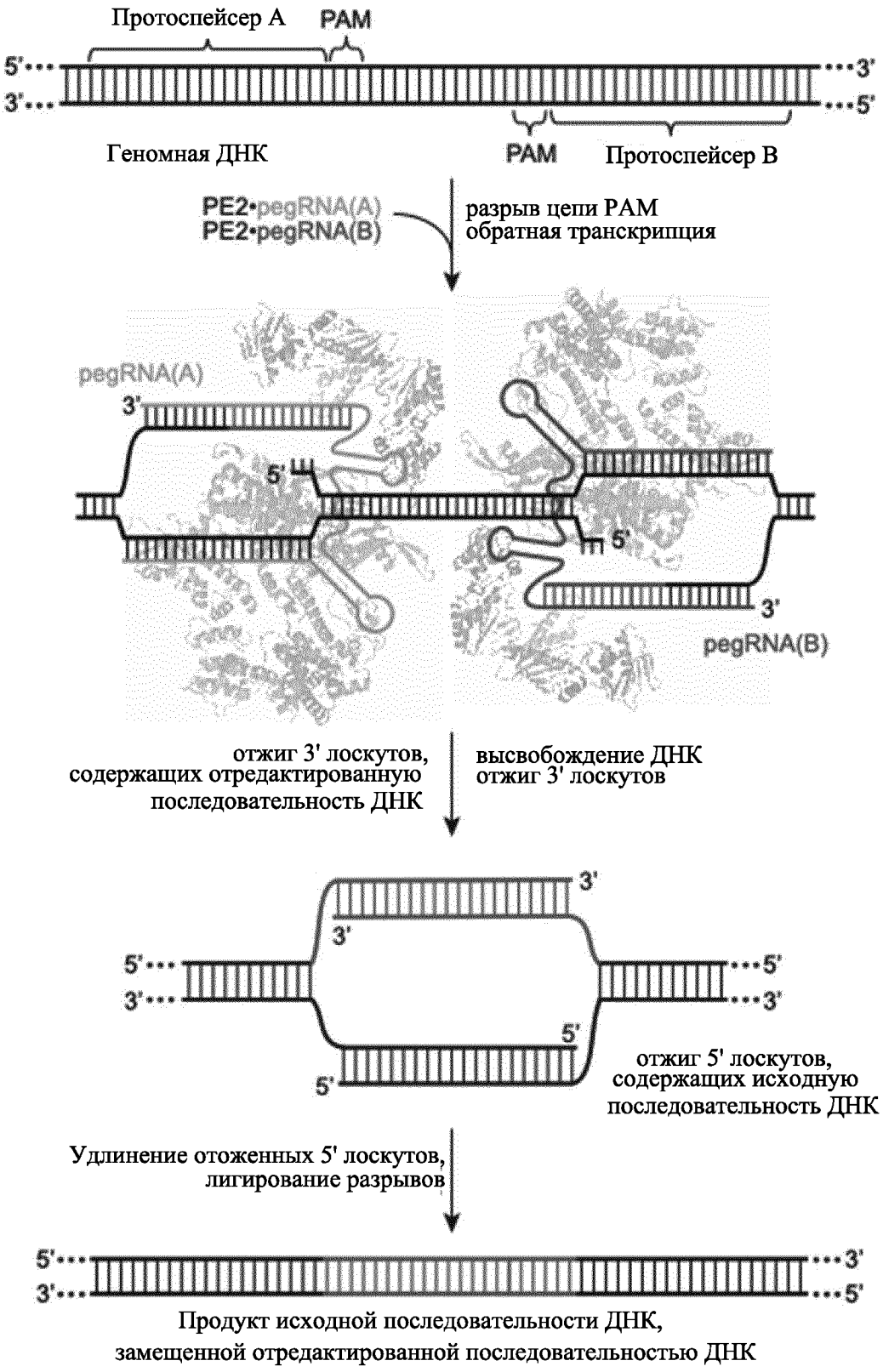
Фиг. 105B


корректное редактирование
индели
% прочтений, выровненных с ожидаемыми аллелями соединения
100
50
0
IDS_QF1_Junc_ab
IDS_QF1_Junc_cd
IDS_QF2_Junc_ab
IDS_QF2_Junc_cd
IDS_QF3_Junc_ab
IDS_QF3_Junc_cd
IDS_QF4_Junc_ab
IDS_QF4_Junc_cd
IDS_QF5_Junc_ab
IDS_QF5_Junc_cd
IDS_QF6_Junc_ab
IDS_QF6_Junc_cd

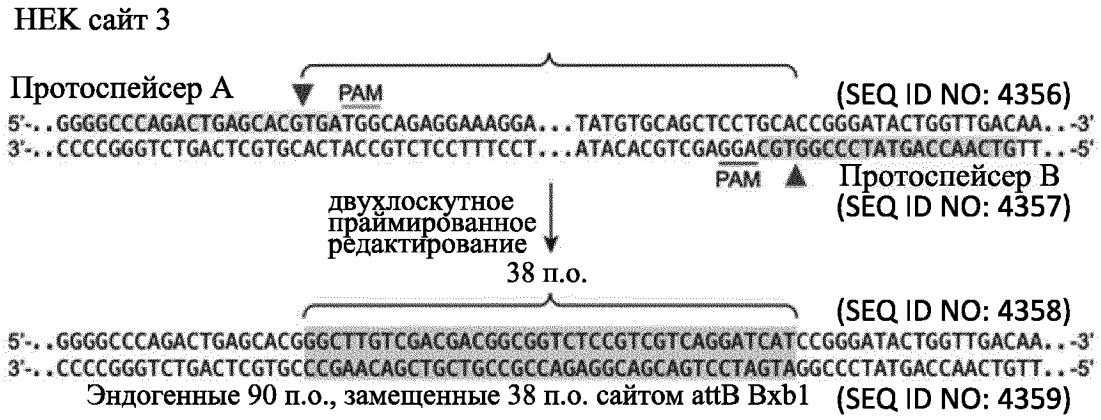
Фиг. 106


Все прочтения секвенирования с указанным редактированием или инделями (%)
0
20
40
60
80
Unm_A1_a + B2_a (attP fwd)
evoPreQ1_A1_a + B2_a (attP fwd)
Unm_A1_a + B2_a (attP rev)
evoPreQ1_A1_a + B2_a (attP rev)
Unm_A1_c + B3_a (attP fwd)
evoPreQ1_A1_c + B3_a (attP fwd)
Unm_A1_c + B3_a (attP rev)
evoPreQ1_A1_c + B3_a (attP rev)
Unm_A1_c + B4_b (attP fwd)
evoPreQ1_A1_c + B4_b (attP fwd)
Unm_A1_c + B4_b (attP rev)
evoPreQ1_A1_c + B4_b (attP rev)
Unm_A4_b + B7_b (attP fwd)
evoPreQ1_A4_b + B7_b (attP fwd)
Unm_A4_b + B7_b (attP rev)
evoPreQ1_A4_b + B7_b (attP rev)
Unm_C2_c + D1_b (attB rev)
evoPreQ1_C2_c + D1_b (attB rev)
Unm_C2_c + D1_b (attB fwd)
evoPreQ1_C2_c + D1_b (attB fwd)
Unm_C2_c + D2_c (attB rev)
evoPreQ1_C2_c + D2_c (attB rev)
Unm_C2_c + D2_c (attB fwd)
evoPreQ1_C2_c + D2_c (attB fwd)
инделы
корректное редактирование

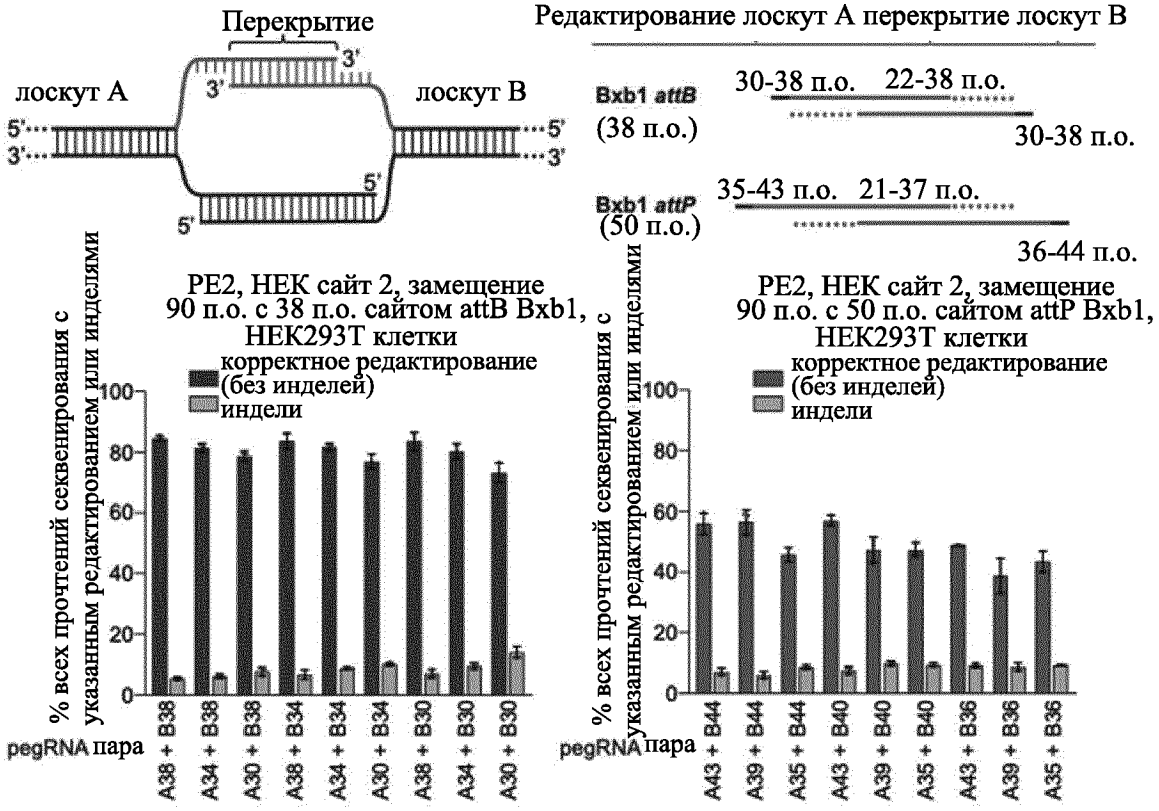
Фиг. 107


Протоспейсер A
PAM
5'
3'
3'
5'
Геномная ДНК
PAM
Протоспейсер B
PE2•pegRNA(A)
PE2•pegRNA(B)
разрыв цепи PAM
обратная транскрипция
pegRNA(A)
3'
5'
5'
3'
pegRNA(B)
отжиг 3' лоскутов,
содержащих отредактированную
последовательность ДНК
высвобождение ДНК
отжиг 3' лоскутов
3'
3'
5'
3'
5'
5'
3'
отжиг 5' лоскутов,
содержащих исходную
последовательность ДНК
Удлинение отоженных 5' лоскутов,
лигирование разрывов
5'
3'
3'
5'
Продукт исходной последовательности ДНК,
замещенной отредактированной последовательностью ДНК

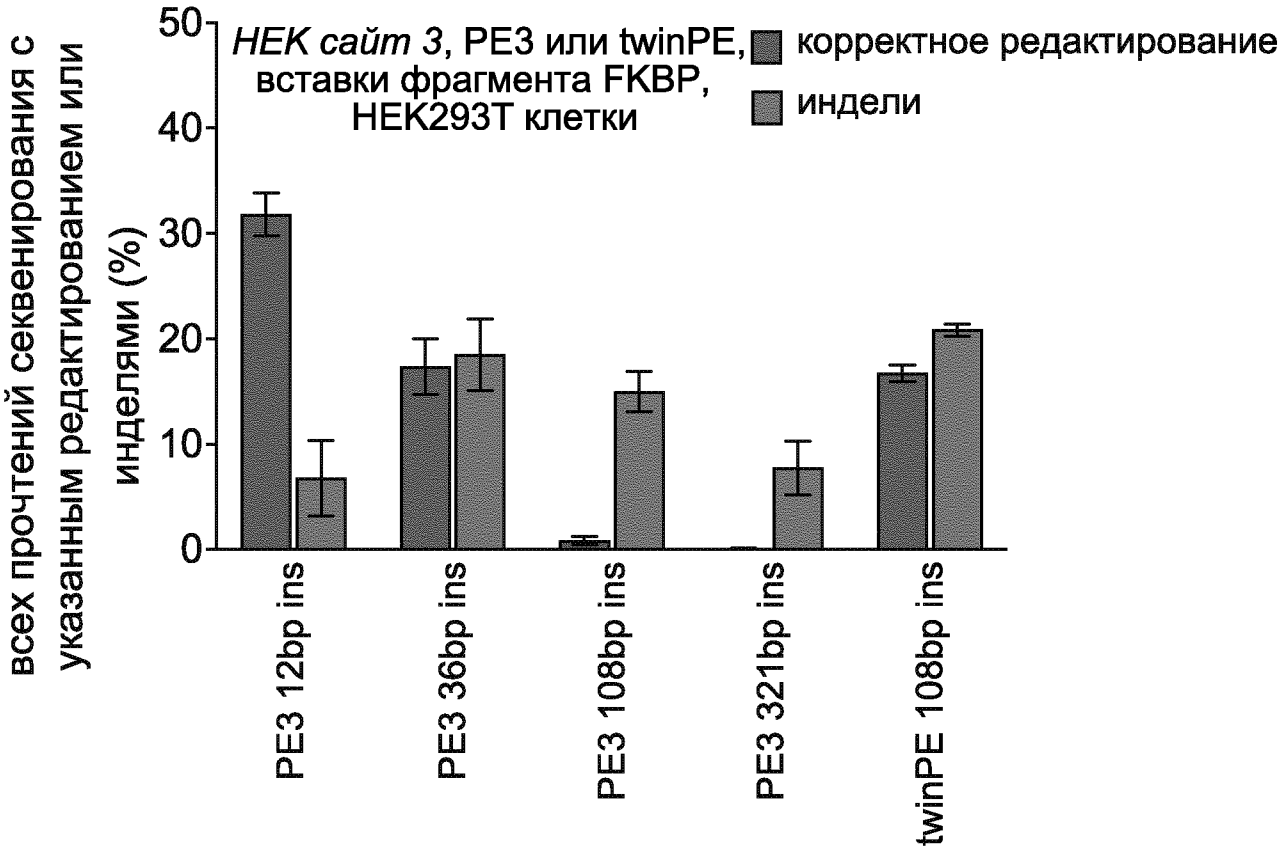
Фиг. 108A



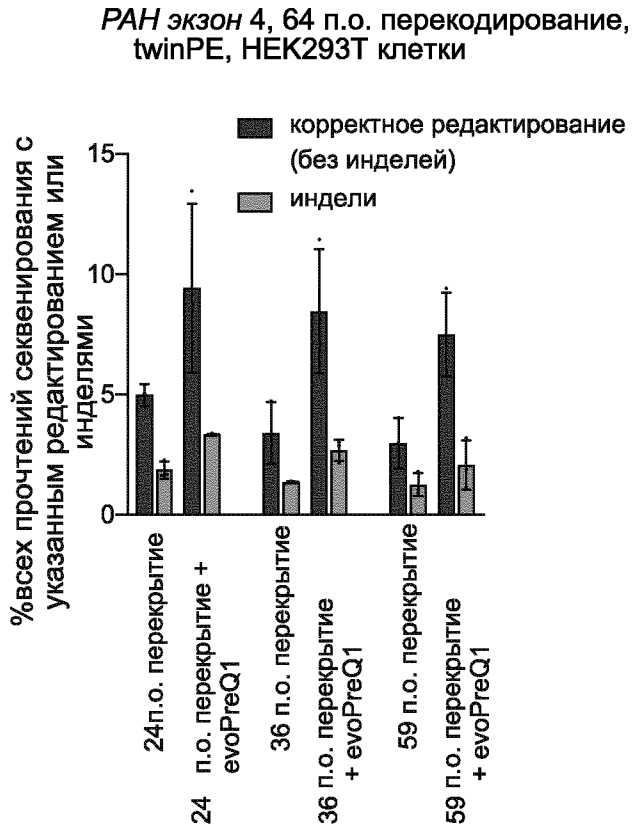
Фиг. 108B

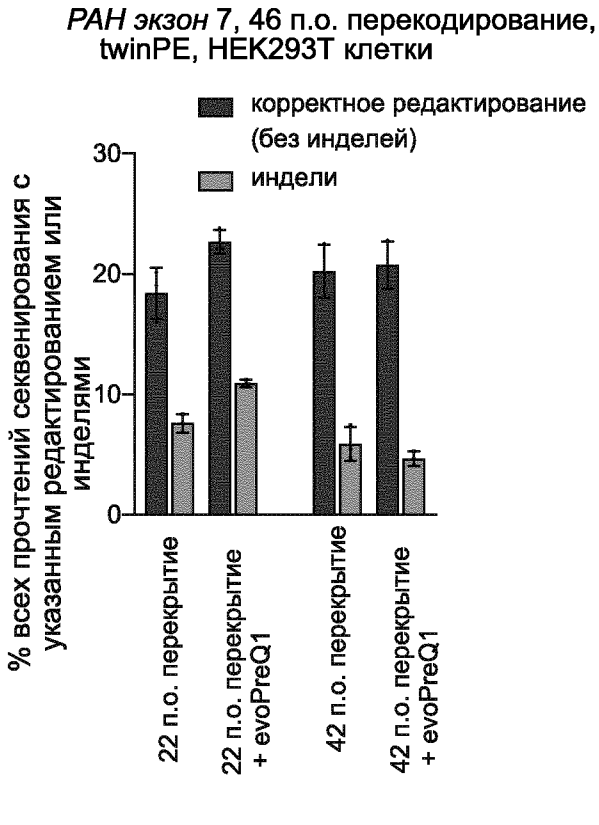


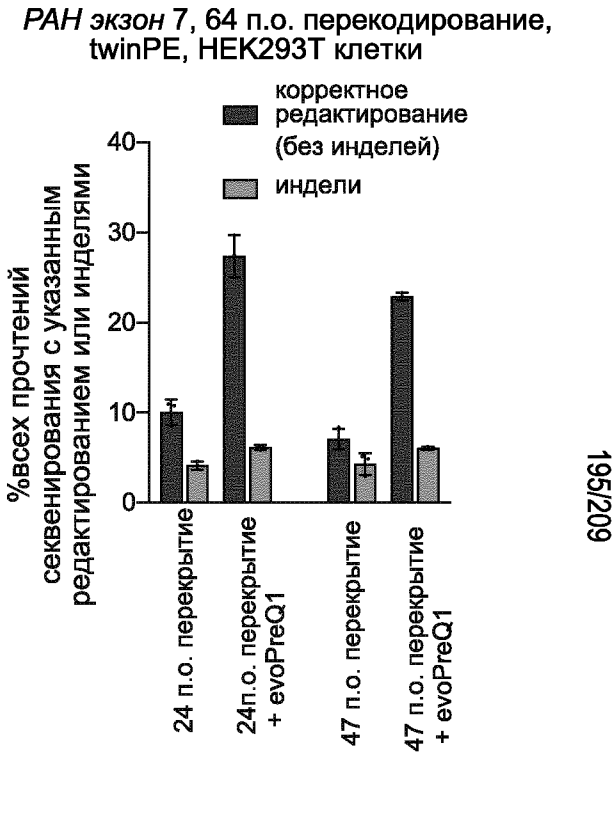
Фиг. 108C


всех прочтений секвенирования с указанным редактированием или инделями (%)
50
40
30
20
10
0
HEK сайт 3, PE3 или twinPE, вставки фрагмента FKBP, HEK293T клетки
корректное редактирование
индели
PE3 12bp ins
PE3 36bp ins
PE3 108bp ins
PE3 321bp ins
twinPE 108bp ins

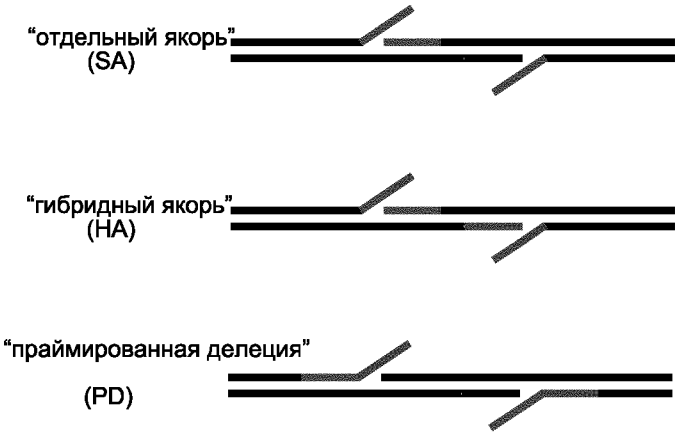
Фиг. 109A


PAH экзон 4, 64 п.о. перекодирование, twinPE, HEK293T клетки
корректное редактирование (без инделей)
индели
%всех прочтений секвенирования с указанным редактированием или инделями
0
5
10
15
24п.о. перекрытие
24 п.о. перекрытие + evoPreQ1
36 п.о. перекрытие
36 п.о. перекрытие + evoPreQ1
59 п.о. перекрытие
59 п.о. перекрытие + evoPreQ1

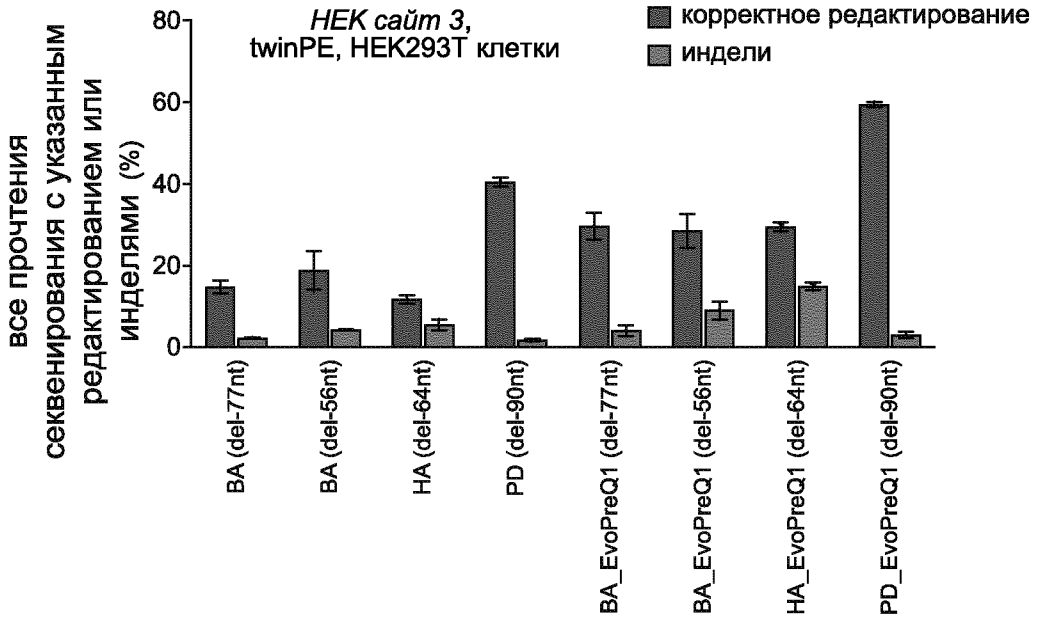

PAH экзон 7, 46 п.о. перекодирование, twinPE, HEK293T клетки
корректное редактирование (без инделей)
индели
% всех прочтений секвенирования с указанным редактированием или инделями
0
10
20
30
22 п.о. перекрытие
22 п.о. перекрытие + evoPreQ1
42 п.о. перекрытие
42 п.о. перекрытие + evoPreQ1


PAH экзон 7, 64 п.о. перекодирование, twinPE, HEK293T клетки
корректное редактирование (без инделей)
индели
%всех прочтений секвенирования с указанным редактированием или инделями
0
10
20
30
40
24 п.о. перекрытие
24п.о. перекрытие + evoPreQ1
47 п.о. перекрытие
47 п.о. перекрытие + evoPreQ1

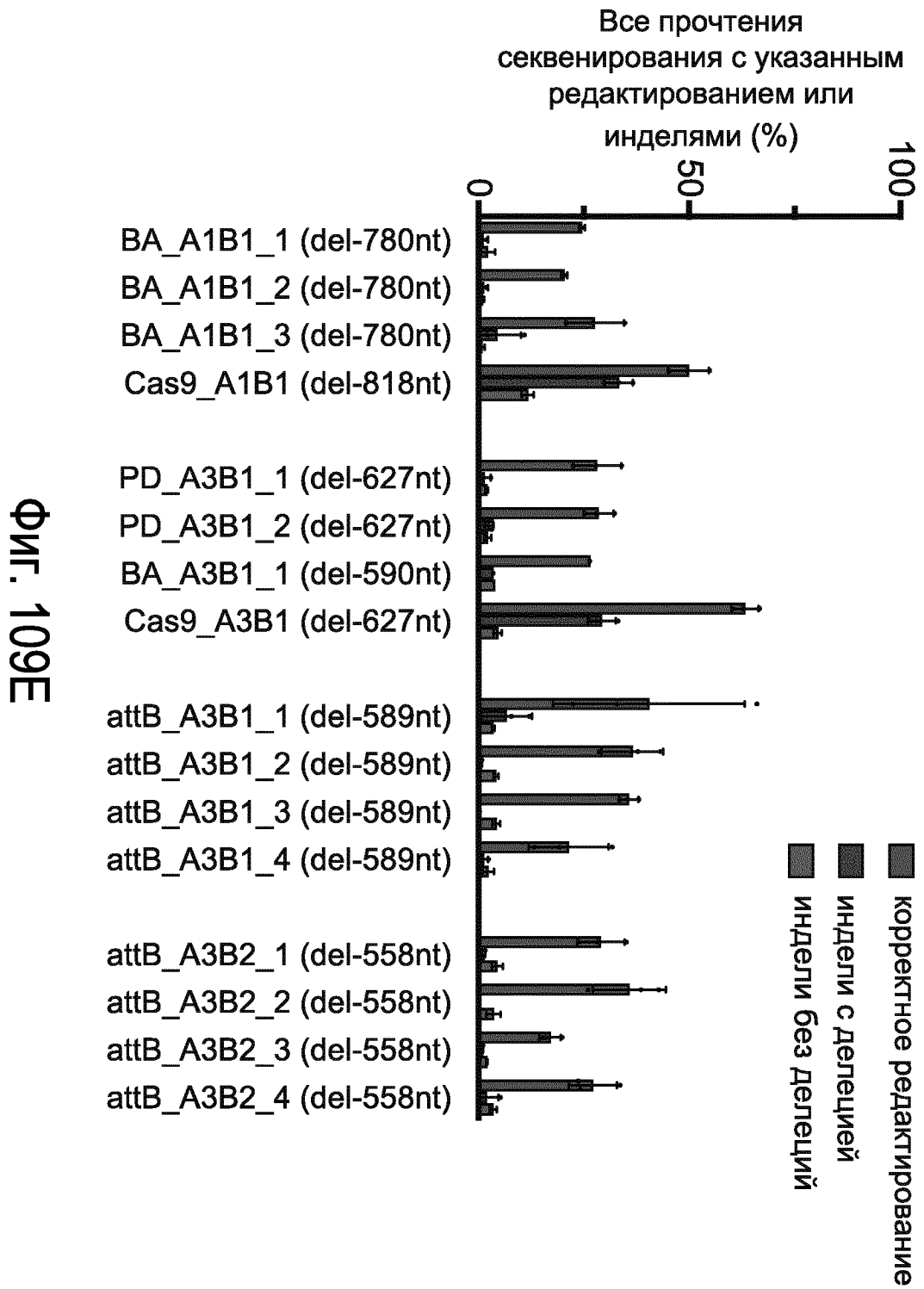
Фиг. 109B



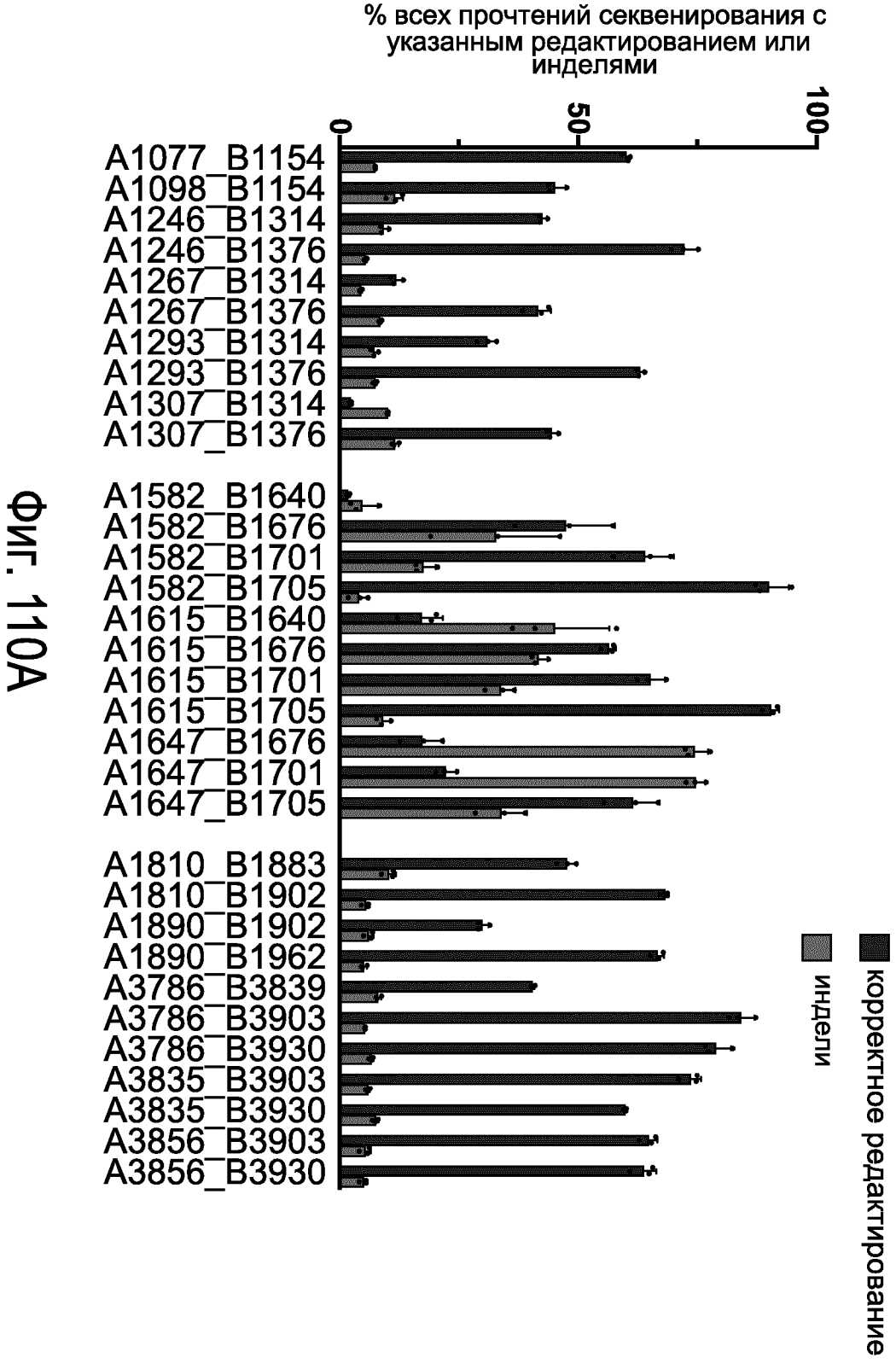
Фиг. 109C



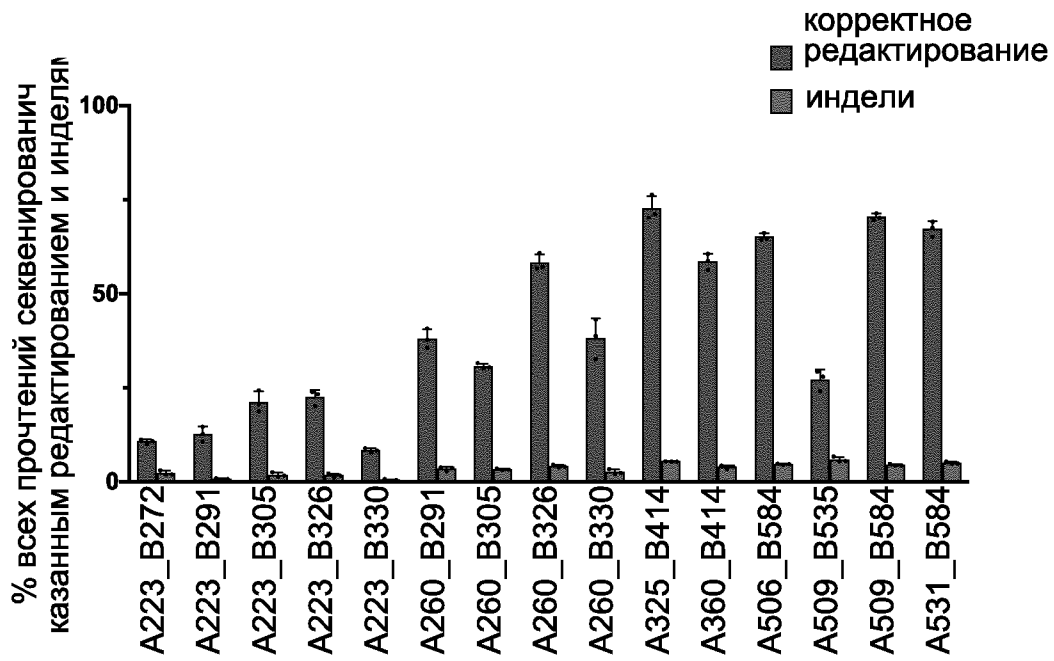
Фиг. 109D


Все прочтения секвенирования с указанным редактированием или инделями (%)
0
50
100
BA_A1B1_1 (del-780nt)
BA_A1B1_2 (del-780nt)
BA_A1B1_3 (del-780nt)
Cas9_A1B1 (del-818nt)
PD_A3B1_1 (del-627nt)
PD_A3B1_2 (del-627nt)
BA_A3B1_1 (del-590nt)
Cas9_A3B1 (del-627nt)
attB_A3B1_1 (del-589nt)
attB_A3B1_2 (del-589nt)
attB_A3B1_3 (del-589nt)
attB_A3B1_4 (del-589nt)
attB_A3B2_1 (del-558nt)
attB_A3B2_2 (del-558nt)
attB_A3B2_3 (del-558nt)
attB_A3B2_4 (del-558nt)
корректное редактирование
индели с делецией
индели без делеций

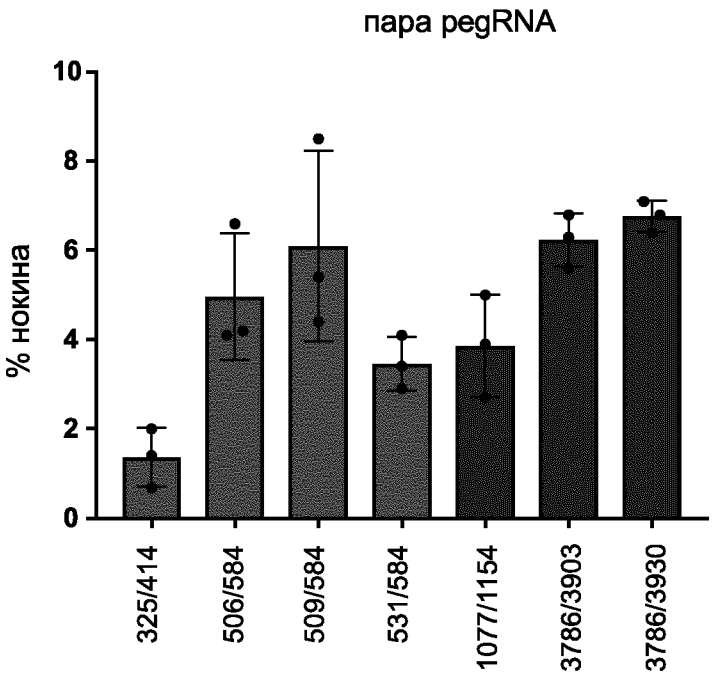
Фиг. 109E


% всех прочтений секвенирования с указанным редактированием или инделями
0
50
100
A1077_B1154
A1098_B1154
A1246_B1314
A1246_B1376
A1267_B1314
A1267_B1376
A1293_B1314
A1293_B1376
A1307_B1314
A1307_B1376
A1582_B1640
A1582_B1676
A1582_B1701
A1582_B1705
A1615_B1640
A1615_B1676
A1615_B1701
A1615_B1705
A1647_B1676
A1647_B1701
A1647_B1705
A1810_B1883
A1810_B1902
A1890_B1902
A1890_B1962
A3786_B3839
A3786_B3903
A3786_B3930
A3835_B3903
A3835_B3930
A3856_B3903
A3856_B3930
индели
корректное редактирование

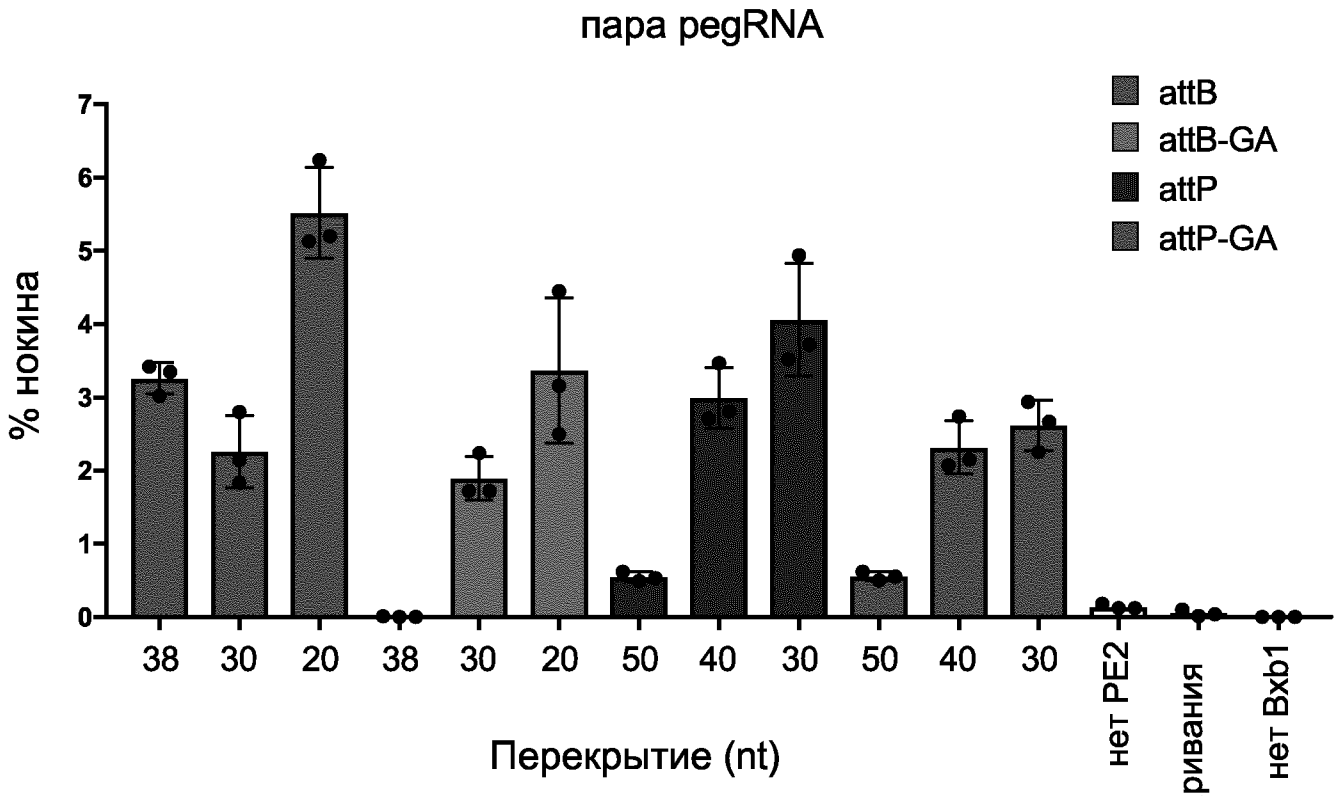
Фиг. 110A



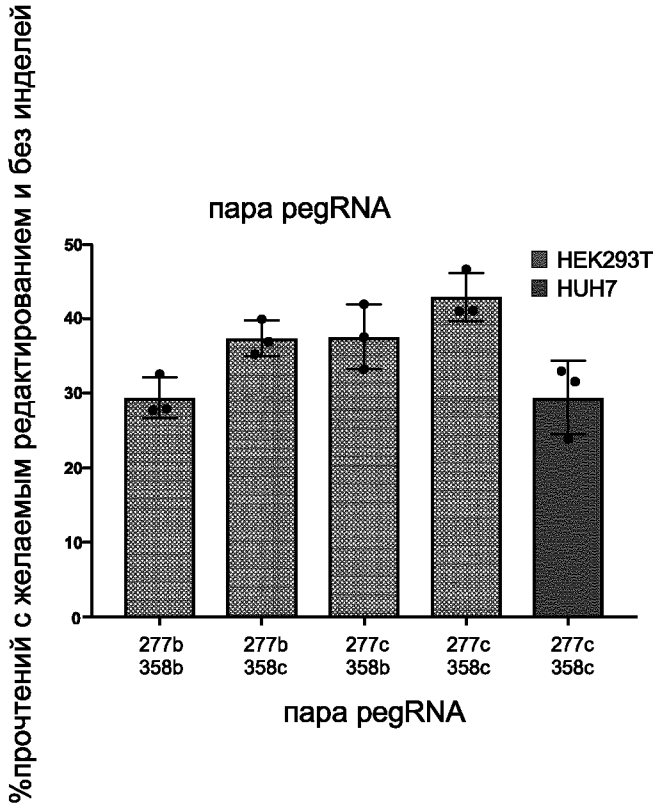
Фиг. 110B



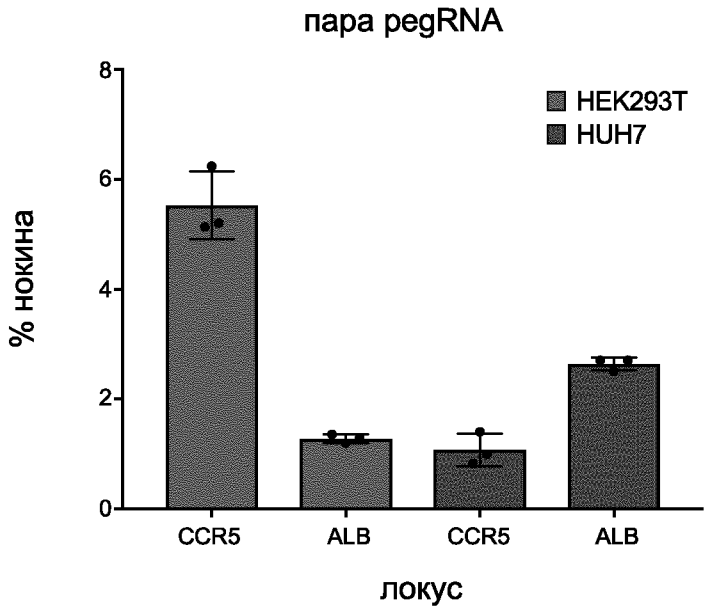
Фиг. 110C


пара pegRNA
attB
attB-GA
attP
attP-GA
% нокина
0
1
2
3
4
5
6
7
38
30
20
38
30
20
50
40
30
50
40
30
нет PE2
ошибка спаривания
нет Bxb1
Перекрытие (nt)

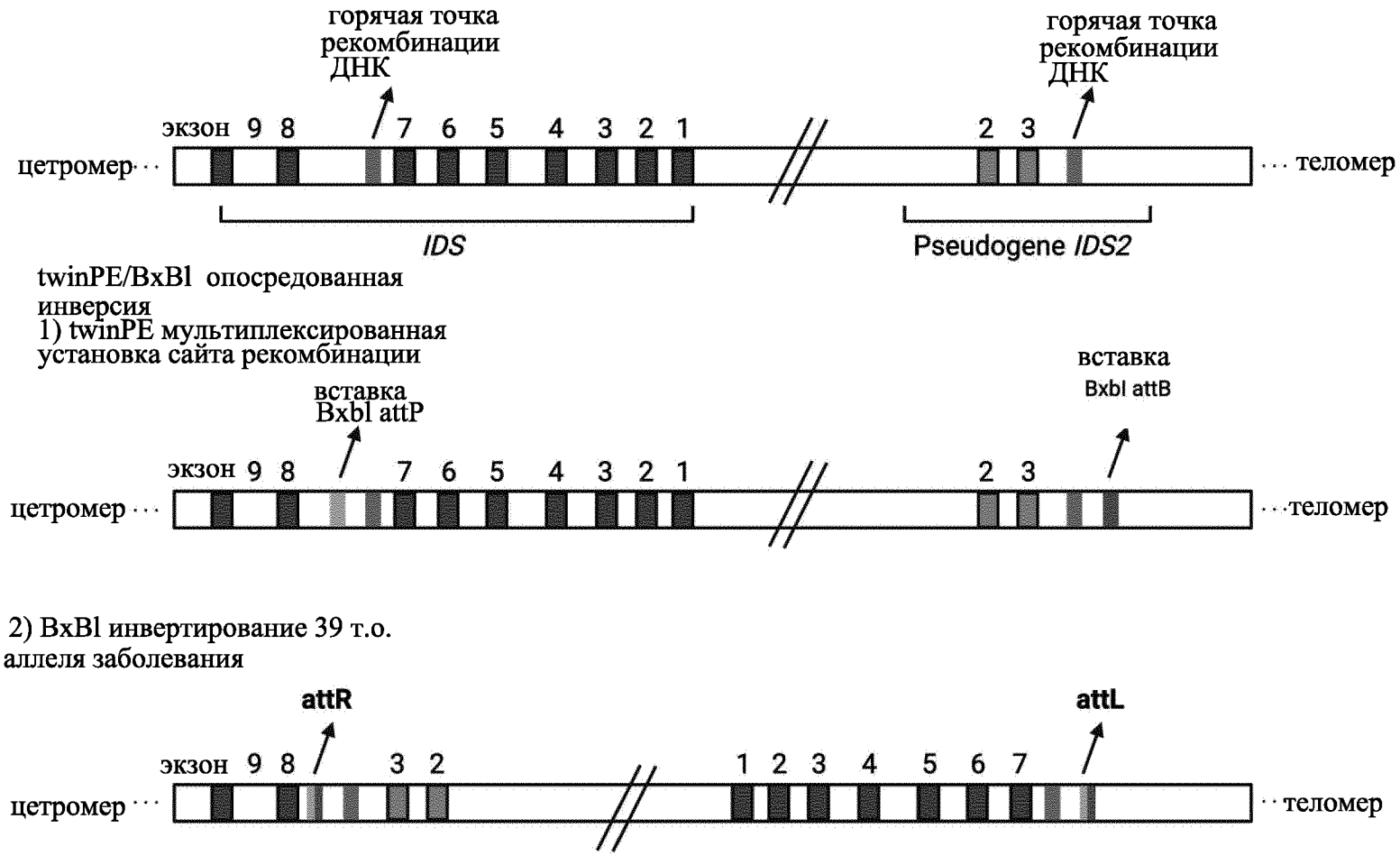
Фиг. 110D



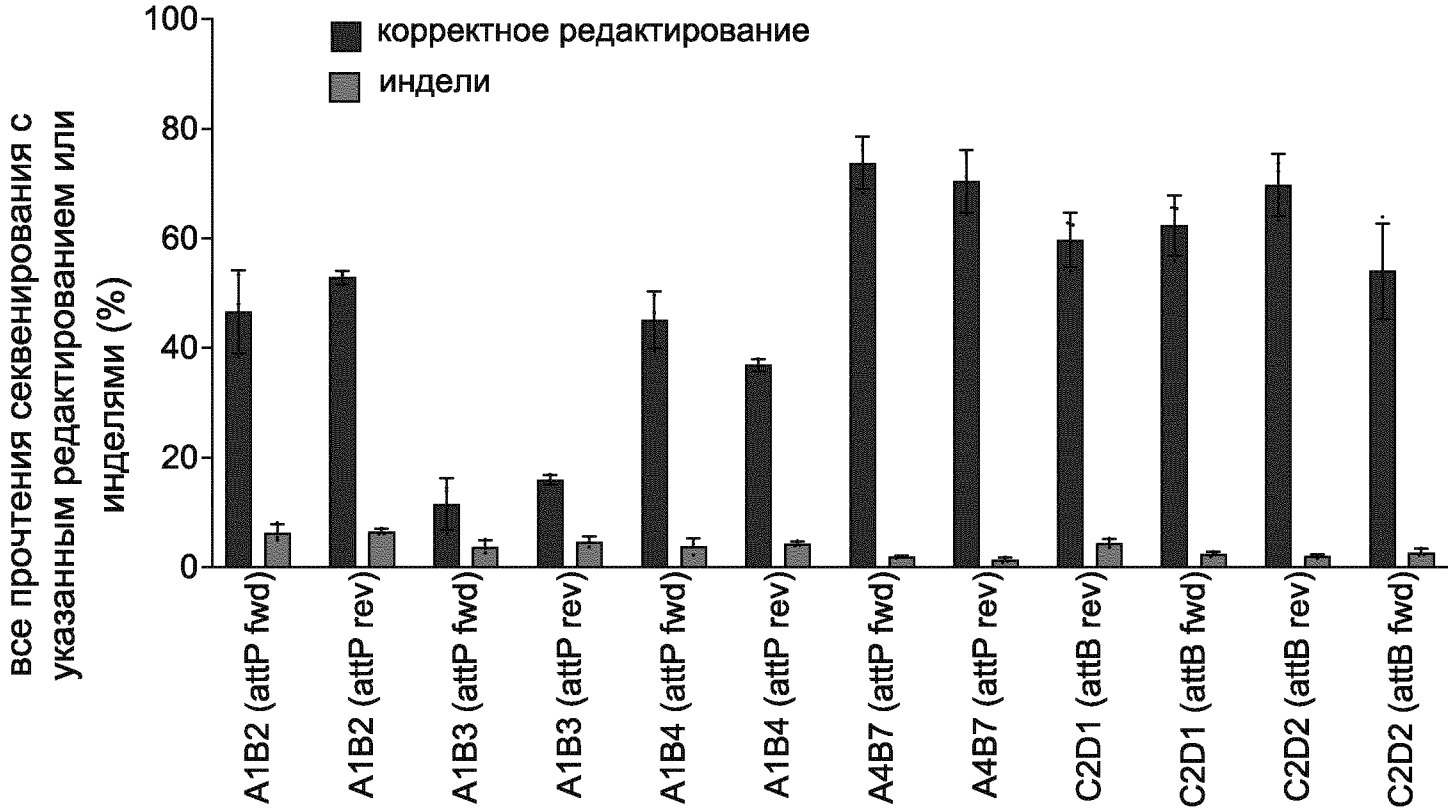
Фиг. 110E



Фиг. 110F

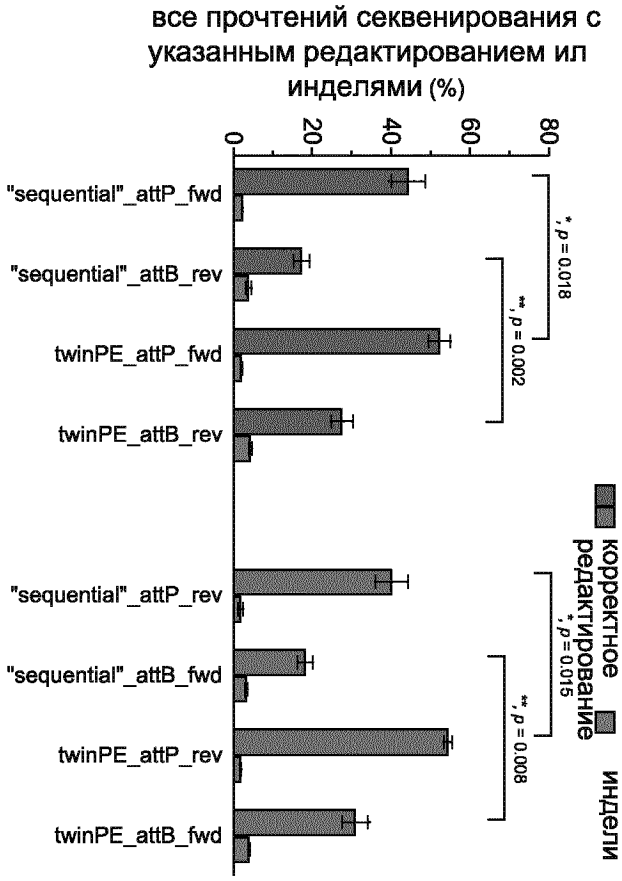

горячая точка рекомбинации ДНК
горячая точка рекомбинации ДНК
экзон 9 8 7 6 5 4 3 2 1
2 3
цетромер
теломер
IDS
Pseudogene IDS2
twinPE/BxBl опосредованная инверсия
1) twinPE мультиплексированная установка сайта рекомбинации
вставка BxbI attP
вставка BxbI attB
экзон 9 8 7 6 5 4 3 2 1
2 3
цетромер
теломер
2) BxBl инвертирование 39 т.о. аллеля заболевания
attR
attL
экзон 9 8 3 2
1 2 3 4 5 6 7
цетромер
теломер

Фиг. 111A

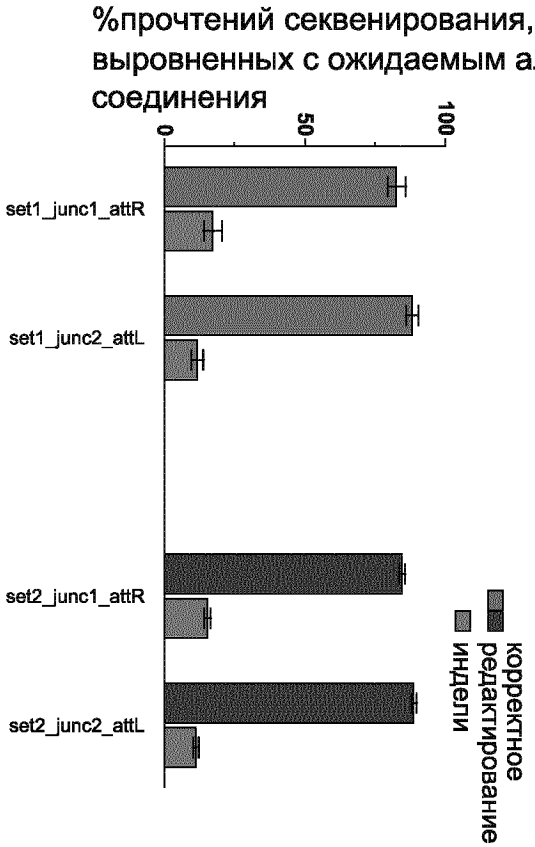

корректное редактирование
индели
все прочтения секвенирования с указанным редактированием или инделями (%)
100
80
60
40
20
0
A1B2 (attP fwd)
A1B2 (attP rev)
A1B3 (attP fwd)
A1B3 (attP rev)
A1B4 (attP fwd)
A1B4 (attP rev)
A4B7 (attP fwd)
A4B7 (attP rev)
C2D1 (attB rev)
C2D1 (attB fwd)
C2D2 (attB rev)
C2D2 (attB fwd)

Фиг. 111B

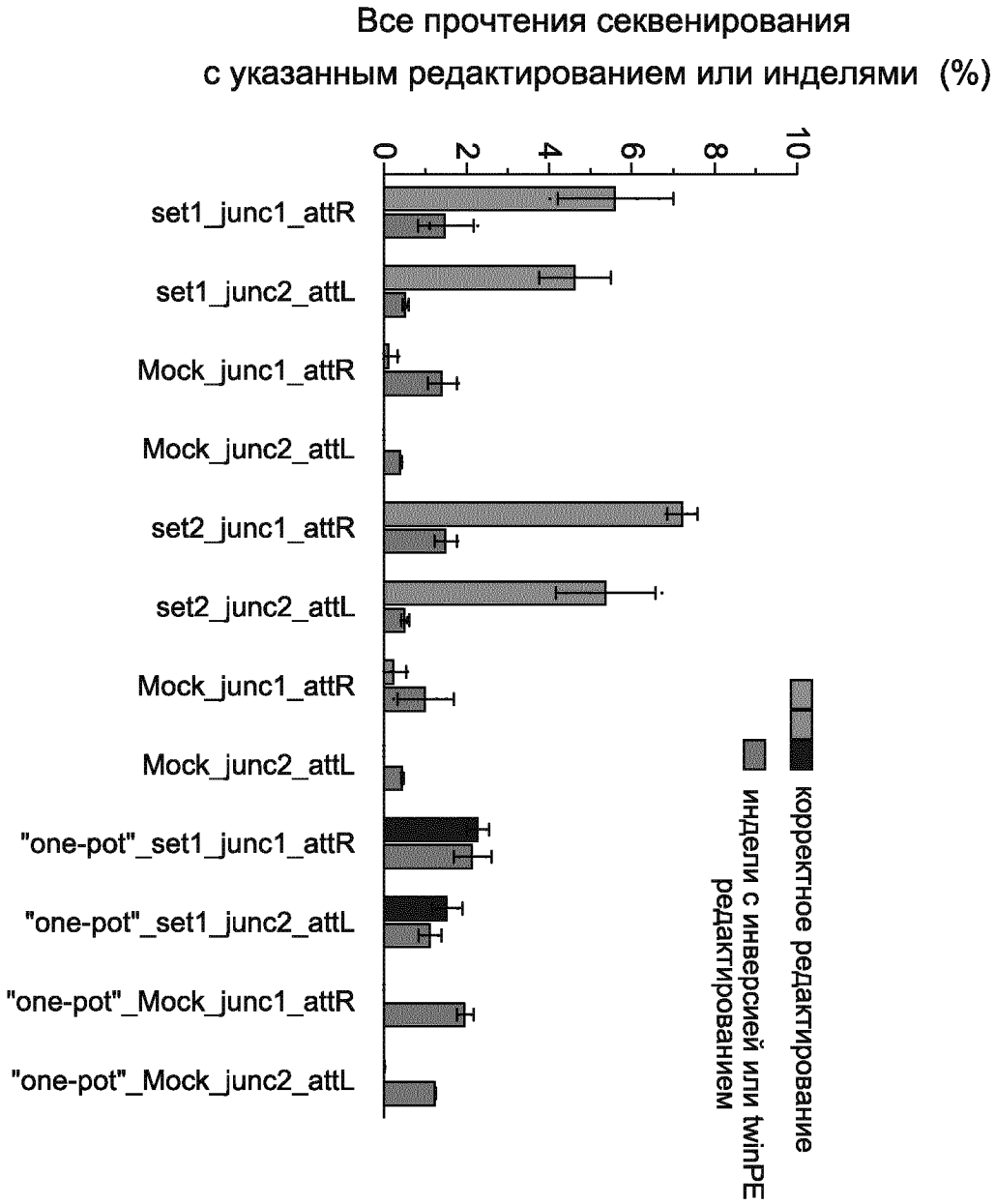
Фиг. 111C



Фиг. 111D

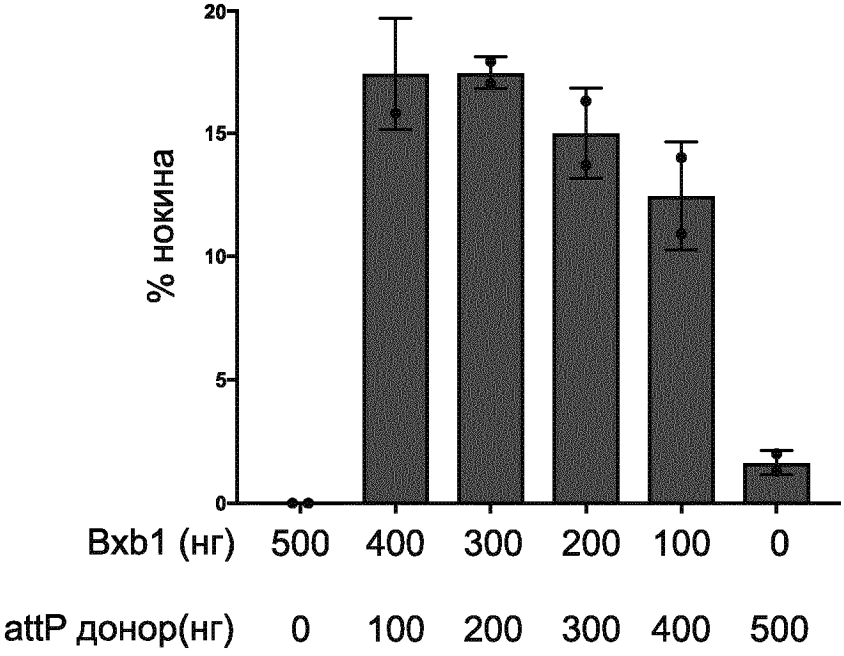


Фиг. 111E

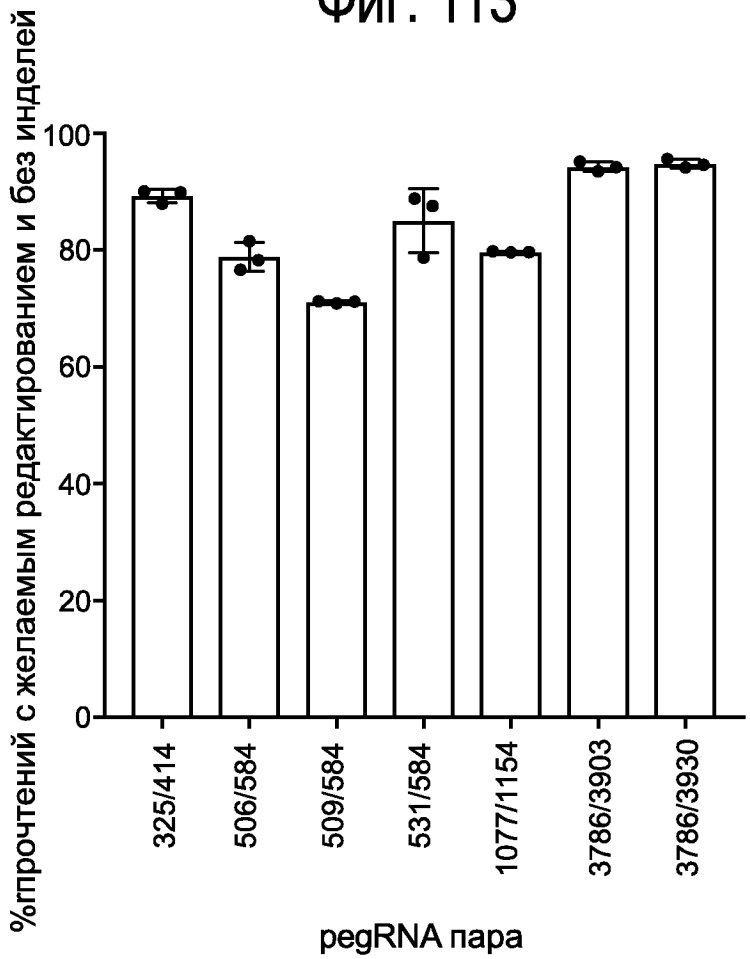

Все прочтения секвенирования
с указанным редактированием или инделями (%)
0
2
4
6
8
10
set1_junc1_attR
set1_junc2_attL
Mock_junc1_attR
Mock_junc2_attL
set2_junc1_attR
set2_junc2_attL
Mock_junc1_attR
Mock_junc2_attL
"one-pot"_set1_junc1_attR
"one-pot"_set1_junc2_attL
"one-pot"_Mock_junc1_attR
"one-pot"_Mock_junc2_attL
корректное редактирование
индели с инверсией или twinPE редактированием

Все прочтения редактирования с указанным редактированием или инделями (%)
10
5
0
корректное редактирование
индели
экзон10_1
экзон10_2
экзон10_3
экзон10_4
экзон10_5
экзон10_6
экзон10_7
экзон10_8
экзон10_9
экзон11.1_1
экзон11.1_2
экзон11.1_3
экзон11.1_4
экзон11.1_5
экзон11.1_6
экзон11.1_7
экзон11.1_8
экзон11.1_9
экзон11.2_1
экзон11.2_2
экзон11.2_3
экзон11.2_4
экзон11.2_5
экзон11.2_6
экзон11.2_7
экзон11.2_8
экзон11.2_9
экзон12.1_1
экзон12.1_2
экзон12.1_3
экзон12.1_4
экзон12.1_5
экзон12.1_6
экзон12.1_7
экзон12.1_8
экзон12.1_9
экзон12.2_1
экзон12.2_2
экзон12.2_3
экзон12.2_4
экзон12.2_5
экзон12.2_6
экзон12.2_7
экзон12.2_8
экзон12.2_9

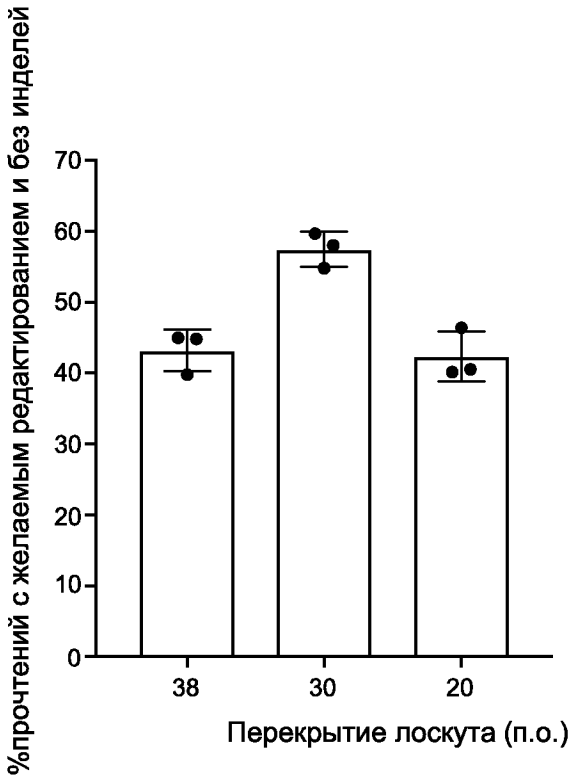
Фиг. 112



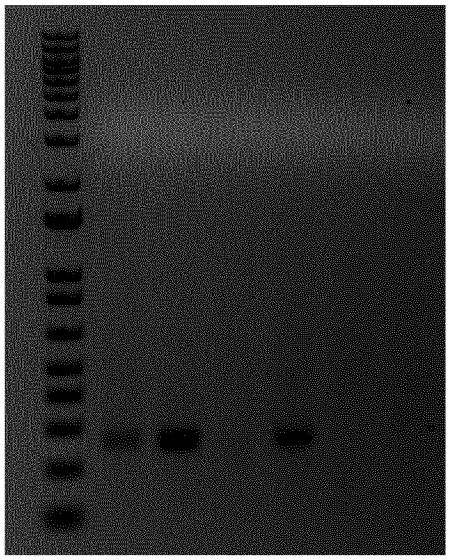
Фиг. 113



Фиг. 114

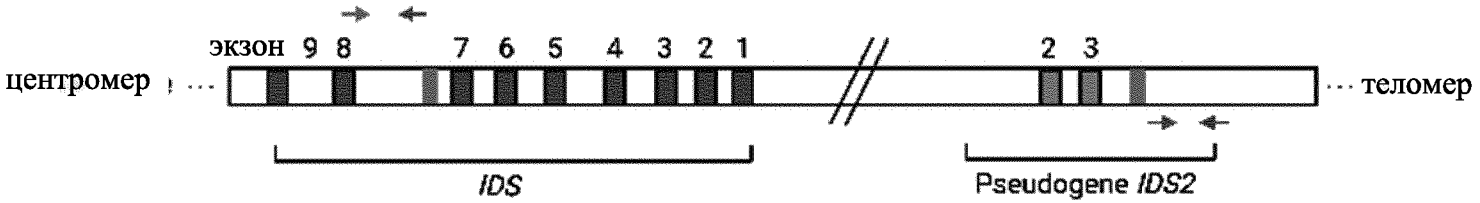


Фиг. 115A



1: лэддер
2: 38-п.о. перекрытие
3: 30-п.о. перекрытие
4: 20-п.о. перекрытие
5: нет PE2
6: нет Bxb1
7: ошибка спаривания

Фиг. 115B



← Обратный праймер (синий), который может связываться с идентичной последовательностью как в неинвертированных, так и в инвертированных аллелях и, таким образом, амплифицировать оба редактирования с одной и той же парой праймеров

→ ← Пара праймеров 1 (зеленый и синий) захватывает неотредактированные аллели (403 п.о.), twinPE редактирование (337 п.о.), и инвертированные аллели (326 п.о.) при соединении 1

→ ← Пара праймеров 2 (красный и синий) захватывает неотредактированные аллели (346 п.о.), twinPE редактирование (326 п.о.), и инвертированные аллели (320 п.о.) при соединении 2

Фиг. 116