

(19)



**Евразийское
патентное
ведомство**

(21) **202292535** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.01.12

(22) Дата подачи заявки
2021.03.05

(51) Int. Cl. *C12N 5/10* (2006.01)
C12N 15/09 (2010.01)
C12N 15/63 (2006.01)
C12N 15/85 (2006.01)
C12N 15/90 (2006.01)

(54) СПОСОБЫ И КОМПОЗИЦИИ, ОБЕСПЕЧИВАЮЩИЕ ПОДАВЛЕНИЕ ЗАЩИТЫ ХОЗЯИНА ДЛЯ МОДУЛИРОВАНИЯ ГЕНОМА

(31) 62/985,750; 63/035,653; 63/147,529

(32) 2020.03.05; 2020.06.05; 2021.02.09

(33) US

(86) PCT/US2021/021213

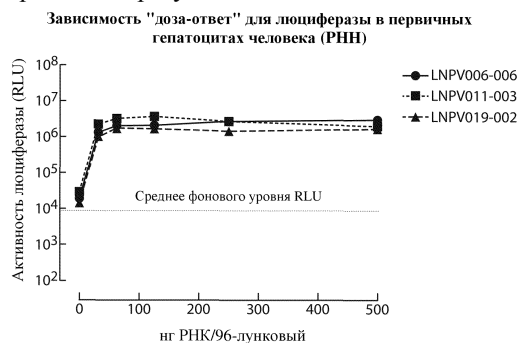
(87) WO 2021/178898 2021.09.10

(71) Заявитель:
**ФЛЭГШИП ПАЙОНИРИНГ
ИННОВЕЙШНЗ VI, ЭлЭлСи (US)**

(72) Изобретатель:
**Ситорик Роберт Джеймс, Рубенс Якоб
Розенблум, Котта-Рамузино Чечилия
Джованна Сильвия, Саломон Уильям
Эдвард, Ван Цзы Цзюнь (US)**

(74) Представитель:
Медведев В.Н. (RU)

(57) В настоящем изобретении предусмотрены, например, композиции и способы модулирования ответа хозяина на систему Gene Writer. В некоторых вариантах осуществления модуляция ответа хозяина приводит в результате к повышению уровня интеграции гетерологичной последовательности нуклеиновой кислоты, представляющей интерес, в геном-мишень. В некоторых вариантах осуществления модуляция ответа хозяина приводит в результате к повышению стабильности, например сохранению вставки или ее экспрессии. В некоторых вариантах осуществления модуляция ответа хозяина приводит в результате к снижению цитотоксичности.



202292535

A1

A1

202292535

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-575580ЕА/042

СПОСОБЫ И КОМПОЗИЦИИ, ОБЕСПЕЧИВАЮЩИЕ ПОДАВЛЕНИЕ ЗАЩИТЫ ХОЗЯИНА ДЛЯ МОДУЛИРОВАНИЯ ГЕНОМА

РОДСТВЕННЫЕ ЗАЯВКИ

Настоящая заявка заявляет приоритет по заявке на патент США под регистрационным номером 62/985750, поданной 5 марта 2020 г., по заявке на патент США под регистрационным номером 63/035653, поданной 5 июня 2020 г., и по заявке на патент США под регистрационным номером 63/147529, поданной 9 февраля 2021 г., каждая из которых включена в данный документ посредством ссылки во всей своей полноте.

ПРЕДПОСЫЛКИ ИЗОБРЕТЕНИЯ

В последние годы методики интеграции генов в геном продвинулись вперед, однако для определенных вариантов применения эффективность интеграции генов по-прежнему остается слишком низкой. В данной области техники существует потребность в улучшенных композициях и способах повышения эффективности интеграции генов.

КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

В настоящем изобретении предусмотрен, например, способ модификации молекулы ДНК-мишени в клетке-хозяине млекопитающего, при этом способ включает:

а) приведение клетки-хозяина в контакт (например, непосредственно или опосредованно, например, посредством обеспечения доступа к клетке, например, посредством системного введения) с системой модификации гена; и

б) приведение клетки-хозяина в контакт (например, непосредственно или опосредованно, например, посредством обеспечения доступа к клетке, например, посредством системного введения) со средством, которое способствует активности системы модификации гена (например, модулятором ответа хозяина или эпигенетическим модификатором),

где система модификации гена содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта. В настоящем изобретении также предусмотрен, например, способ модификации молекулы ДНК-мишени в клетке-хозяине млекопитающего, при этом способ включает приведение клетки-хозяина в контакт (например, непосредственно или опосредованно, например, посредством обеспечения доступа к клетке, например, посредством системного введения) со следующим:

I) системой модификации гена и необязательно средством доставки для системы модификации гена, где система модификации гена содержит:

а) полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и

б) нуклеиновую кислоту-матрицу, при этом нуклеиновая кислота-матрица

содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

II) средством, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор), связанным с компонентом системы модификации гена или средством доставки.

Например, средство, которое способствует активности системы модификации гена, может быть ковалентно связано с компонентом системы модификации гена, например, слито с компонентом системы модификации гена, например, полипептидом Gene Writer или нуклеиновой кислотой, кодирующей полипептид Gene Writer, например, нуклеиновой кислотой-матрицей Gene Writer (например, РНК- или ДНК-матрицей) или нуклеиновой кислотой, кодирующей матрицу Gene Writer (например, ДНК, кодирующей РНК-матрицу), дополнительной нуклеиновой кислотой системы Gene Writing (например, gRNA) или средством доставки системы модификации гена, например, AAV или наночастицей (например, LNP). В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, встроено в средство доставки или составлено совместно с ним.

В настоящем изобретении также предусмотрен набор, содержащий

a) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

b) средство, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор).

В настоящем изобретении также предусмотрен набор, содержащий

систему модификации гена, содержащую полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, при этом нуклеиновая кислота-матрица содержит: I) систему модификации гена и необязательно средство доставки для системы модификации гена, где система модификации гена содержит:

a) полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и

b) нуклеиновую кислоту-матрицу, при этом нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

II) средство, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор), связанное с компонентом системы модификации гена или средством доставки.

Например, средство, которое способствует активности системы модификации гена, может быть ковалентно связано с компонентом системы модификации гена, например, слито с компонентом системы модификации гена, например, полипептидом Gene Writer или

нуклеиновой кислотой, кодирующей полипептид Gene Writer, нуклеиновой кислотой-матрицей Gene Writer (например, РНК- или ДНК-матрицей) или нуклеиновой кислотой, кодирующей матрицу Gene Writer (например, ДНК, кодирующей РНК-матрицу), дополнительной нуклеиновой кислотой системы Gene Writing (например, gRNA) или средством доставки системы модификации гена, например, AAV или наночастицей (например, LNP). В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, встроено в средство доставки или составлено совместно с ним.

В настоящем изобретении также предусмотрена композиция, содержащая

- а) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и
- б) средство, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор).

В настоящем изобретении также предусмотрена композиция, содержащая

систему модификации гена, содержащую полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, при этом нуклеиновая кислота-матрица содержит: I) систему модификации гена и необязательно средство доставки для системы модификации гена, где система модификации гена содержит:

- а) полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и
- б) нуклеиновую кислоту-матрицу, при этом нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

II) средство, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор), связанное с компонентом системы модификации гена или средством доставки.

В некоторых вариантах осуществления эпигенетический модификатор предусматривает ингибитор HDAC или ингибитор гистонметилтрансферазы, например, описанные в данном документе.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, предусматривает антитело, полипептид (например, доминантно-отрицательную мутантную форму полипептида в пути ответа хозяина), фермент (например, эндопептидазу, например, Ig-расщепляющую эндопептидазу, например, IdeS), малую молекулу или нуклеиновую кислоту (например, молекулу для RNAi). В некоторых вариантах осуществления фермент представляет собой фермент дикого типа или его функциональный фрагмент или вариант. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена,

предусматривает нуклеиновую кислоту, которая ковалентно связана с полипептидом GeneWriter или нуклеиновой кислотой-матрицей. Например, нуклеиновая кислота может кодировать белок, например, терапевтический белок, который способствует активности системы модификации гена. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой малую молекулу. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой домен полипептида.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена (например, ингибитор ответа хозяина), предусматривает белок или домен, которые ингибируют процесс у хозяина. В некоторых вариантах осуществления средство ингибирует или секвестрирует белок хозяина (например, фермент хозяина) или комплекс хозяина. В некоторых вариантах осуществления белок хозяина (или комплекс, содержащий белок хозяина) ингибирует систему модификации гена. В некоторых вариантах осуществления фермент хозяина (или комплекс, содержащий фермент хозяина) ингибирует систему модификации гена. Например, белок хозяина может представлять собой фермент репарации ДНК, который ингибирует систему модификации гена. В некоторых вариантах осуществления белок хозяина участвует в репарации, направляемой гомологией (HDR), например, белок, описанный в данном документе.

В некоторых вариантах осуществления белок хозяина, который ингибируется или секвестрируется, представляет собой белок, который ингибирует требуемый результат редактирования системы модификации гена. В некоторых вариантах осуществления ингибирование системы модификации гена означает ингибирование модификации гена на одной или нескольких стадиях во время процесса Gene Writing, необязательно включающее следующее: (i) связывание ДНК-мишени, (ii) расщепление однонитевой ДНК-мишени, (iii) ассоциацию матрицы Gene Writing с ДНК-мишенью, например, отжиг матрицы, (iv) праймированную мишенью полимеризацию ДНК на матрице Gene Writing, (v) второй ник нити, противоположной ДНК-мишени, (vi) синтез второй нити ДНК с использованием вновь полимеризованной ДНК из (iv) в качестве матрицы полимеризации или необязательно синтез второй нити с использованием дополнительной матрицы Gene Writing, (vii) осуществление активности флэп-эксонуклеазы в отношении ДНК-мишени и/или (viii) лигирование вновь синтезированной ДНК со свободным 5'-концом в геноме мишени. В некоторых вариантах осуществления средство слито с полипептидом Gene Writer.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, предусматривает белок или домен, которые стимулируют процесс у хозяина. В некоторых вариантах осуществления средство активирует или рекрутирует белок хозяина (например, фермент хозяина) или комплекс хозяина. В некоторых вариантах осуществления фермент хозяина представляет собой (или комплекс содержит) фермент репарации ДНК, который способствует активности системы модификации гена, например, ДНК-полимеразу или ДНК-лигазу. В некоторых вариантах

осуществления средство слито с полипептидом Gene Writer.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, содержит белок или домен, которые связываются с белком клетки хозяина. В некоторых вариантах осуществления связывание белка клетки-хозяина с компонентом системы модификации гена функционирует с обеспечивает рекрутирования активности этого белка хозяина (или комплекса, содержащего белок хозяин) к сайту мишени. В некоторых вариантах осуществления белок клетки-хозяина содержит 5'-экзонуклеазу, например, EXO1. В некоторых вариантах осуществления белок клетки-хозяина содержит специфическую в отношении структуры эндонуклеазу, например, FEN1. В некоторых вариантах осуществления средство слито с полипептидом Gene Writer. В некоторых вариантах осуществления полипептид Gene Writer содержит домен Cas, например, домен Cas9 никазы или каталитически неактивный домен Cas9. В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит в направлении от 5'- к 3'-концу следующее: (1) спейсер gRNA; (2) остов gRNA; (3) последовательность гетерологичного объекта и (4) 3'-концевой домен гомологии.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, предусматривает белок или домен, которые заменяют или дополняют белок, комплекс или путь хозяина. В некоторых вариантах осуществления средство содержит 5'-экзонуклеазу, например, EXO1 или ее активный фрагмент или вариант. В некоторых вариантах осуществления средство (например, EXO1) предусматривает последовательность в соответствии с NCBI:NP_006018.4 или UniProt: Q9UQ84, каждая из которых включена в данный документ посредством ссылки, или последовательность, на по меньшей мере 70%, 80%, 85%, 90%, 95%, 98% или 99% идентичную ей. В некоторых вариантах осуществления средство содержит специфическую в отношении структуры эндонуклеазу, например, FEN1 или ее активный фрагмент или вариант. В некоторых вариантах осуществления средство (например, FEN1) предусматривает последовательность в соответствии с NCBI:NP_004102.1 или UniProt: P39748, каждая из которых включена в данный документ посредством ссылки, или последовательность, на по меньшей мере 70%, 80%, 85%, 90%, 95%, 98% или 99% идентичную ей. В некоторых вариантах осуществления полипептид Gene Writer содержит домен Cas, например, каталитически неактивный домен Cas. В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит в направлении от 5'- к 3'-концу следующее: (1) спейсер gRNA; (2) остов gRNA; (3) последовательность гетерологичного объекта и (4) 3'-концевой домен гомологии.

В некоторых вариантах осуществления средство слито со средством доставки или компонентом средства доставки, например, AAV, например, капсидом AAV. В некоторых вариантах осуществления средство способствует снижению иммунного ответа хозяина. В некоторых вариантах осуществления средство предусматривает протеазу, например, экзопептидазу или эндопептидазу, которая расщепляет компонент иммунного ответа хозяина, например, иммуноглобулин или цитокин. В некоторых вариантах осуществления

средство предусматривает эндопептидазу, расщепляющую антитело хозяина, например, антитело, которое связывает средство доставки, например, антитело, которое нейтрализует или ингибирует средство доставки, например, антитело, которое нейтрализует или ингибирует AAV. В некоторых вариантах осуществления эндопептидаза представляет собой Ig-расщепляющую эндопептидазу, например, IdeS. В некоторых вариантах осуществления IdeS расщепляет IgG ниже шарнирной области. Способы предупреждения иммунного ответа, вызванного введением средства генной терапии, или для лечения пациента с иммунитетом к вирусному капсиду в анамнезе с использованием IdeS и других ферментных полипептидов, разрушающих иммуноглобулин G, описаны в Leborgne et al. Nat Med 26:1096-1101 (2020) и в PCT/EP2019/069280.

В некоторых вариантах осуществления белок IdeS, используемый с системой, представляет собой бактериальную эндопептидазу IgG или бактериальную цистеиновую эндопептидазу семейства IdeS/Mac. В некоторых вариантах осуществления белок IdeS, используемый с системой, представляет собой эндопептидазу IgG или цистеиновую эндопептидазу семейства IdeS/Mac из *Streptococcus pyogenes* или *Streptococcus equi*. В некоторых вариантах осуществления Ig-расщепляющая эндопептидаза (например, IdeS) содержит последовательность в соответствии с WP_012678049.1 или WP_002992557.1 или последовательность, на по меньшей мере 70%, 80%, 85%, 90%, 95%, 98% или 99% идентичную таковой. В некоторых вариантах осуществления IdeS может представлять собой модифицированный вариант, например, IdeS с последовательностью, соответствующей SEQ ID No: 3-18, 23 или 48 из PCT/EP2019/069280, которая включена в данный документ во всей своей полноте, включая последовательности IdeS, соответствующие SEQ ID No: 18, 23 и 48. В некоторых вариантах осуществления Ig-расщепляющая эндопептидаза может представлять собой IdeZ. В некоторых вариантах осуществления Ig-расщепляющая эндопептидаза (например, IdeZ) содержит последовательность в соответствии с WP_014622780.1 или последовательность, на по меньшей мере 70%, 80%, 85%, 90%, 95%, 98% или 99% идентичную ей. Другие протеазы, которые можно использовать в настоящем изобретении, включают, например, без ограничения ферменты IgdE из *S. suis*, *S. porcinus* и *S. equi*. В некоторых вариантах осуществления протеаза может представлять собой IdeMC или ее гомолог. Другие эндопептидазы, которые можно использовать в настоящем изобретении, включают, например, без ограничения IdeZ с N-концевым метионином и сигнальным пептидом и без них, и гибридные белки IdeS/IdeZ, описанные в документе WO 2016/128559, который включен в данный документ посредством ссылки во всей своей полноте. Другие протеазы, которые можно использовать в настоящем изобретении, включают, например, без ограничения протеазы, описанные в Jordan et al. (N Engl. J Med. 377;5, 2017), Lannergard and Guss (FEMS Microbiol Lett., 262(2006); 230-235) и Hulting et al. (FEMS Microbiol Lett., 298(2009), 44-50). В некоторых вариантах осуществления средство способствует иммунотолерантности.

В некоторых вариантах осуществления средство может представлять собой

иммуносупрессивное средство. В некоторых вариантах осуществления средство может подавлять поглощение макрофагами, например, представляет собой CD47, или его фрагмент или вариант, или средство, которое способствует экспрессии CD47 в клетке-мишени. В некоторых вариантах осуществления средство может представлять собой растворимый иммуносупрессивный цитокин, например, IL-10, или его фрагмент или вариант, или средство, которое может способствовать экспрессии растворимого иммуносупрессивного цитокина, например, IL-10 или его фрагмента или варианта, в клетке-мишени. В некоторых вариантах осуществления средство может представлять собой растворимый иммуносупрессивный белок или его фрагмент или вариант, или средство, которое может способствовать экспрессии растворимого иммуносупрессивного белка или его фрагмента или варианта в клетке-мишени. В некоторых вариантах осуществления растворимый иммуносупрессивный белок может представлять собой PD-1, PD-L1, CTLA4 или BTLA или их фрагмент или вариант.

В некоторых вариантах осуществления средство может представлять собой толерогенный белок, например, агонист ILT-2 или ILT-4, например, HLA-E или HLA-G, или любой другой эндогенный агонист ILT-2 или ILT-4, или его функциональный фрагмент или вариант. В некоторых вариантах осуществления средство может способствовать экспрессии толерогенного белка, например, агониста ILT-2 или ILT-4, например, HLA-E или HLA-G, или любого другого эндогенного агониста ILT-2 или ILT-4, или его функционального фрагмента или варианта. В некоторых вариантах осуществления средство может предусматривать белок, который подавляет активность системы комплемента, например, снижает активность регуляторного белка системы комплемента, например, белка, который связывает фактор, ускоряющий распад (DAF, CD55), например, белок 1, подобный фактору H (FH) (FHL-1), например, C4b-связывающий белок (C4BP), например, рецептор комплемента 1 (CD35), например, мембранный кофакторный белок (MCP, CD46), например, профектин (CD59). В некоторых вариантах осуществления средство может способствовать экспрессии белка, который подавляет активность системы комплемента, например, регуляторных белков системы комплемента, например, белков, которые связывают фактор, ускоряющий распад (DAF, CD55), например, белка 1, подобного фактору H (FH) (FHL-1), например, C4b-связывающего белка (C4BP), например, рецептора комплемента 1 (CD35), например, мембранного кофакторного белка (MCP, CD46), например, профектина (CD59).

В некоторых вариантах осуществления средство может предусматривать белок, который ингибирует фермент, представляющий собой CD/C5-конвертазу классического или альтернативного пути системы комплемента, например, белок, который регулирует сборку MAC. В некоторых вариантах осуществления средство может способствовать экспрессии белка, который ингибирует ферменты, представляющие собой CD/C5-конвертазу классического или альтернативного пути системы комплемента, например, белка, который регулирует сборку MAC. В некоторых вариантах осуществления средство может предусматривать антиген гистосовместимости, например, HLA-E или HLA-G. В

некоторых вариантах осуществления средство может способствовать экспрессии антигена гистосовместимости, например, HLA-E или HLA-G. В некоторых вариантах осуществления средство способствует гликозилированию, например, содержит сиаловую кислоту, которая осуществляет, например, подавление активации NK-клеток. В некоторых вариантах осуществления средство может способствовать образованию профиля поверхностного гликозилирования, например, содержит сиаловую кислоту, которая осуществляет, например, подавление активации NK-клеток.

В некоторых вариантах осуществления средство может представлять собой терапевтическое средство, нацеленное на систему комплемента, например, регуляторный белок системы комплемента, например, ингибитор системы комплемента, например, белок, который связывается с компонентом системы комплемента, например, C1-ингибитор или его вариант или фрагмент. В некоторых вариантах осуществления средство может представлять собой растворимый регулятор. В некоторых вариантах осуществления средство может представлять собой мембраносвязанный регулятор, например, DAF/CD55, MCP/CD46 или CD59. В некоторых вариантах осуществления средство представляет собой малую молекулу, белок, слитый белок, антитело или конъюгат антитело-лекарственное средство. В некоторых таких случаях терапевтическое средство, нацеленное на компонент системы комплемента, описано в работах Ricklin et al. *Nat Biotechnol* 25(11): 1265-1275 (2007) и Schaubert-Plewa et al *Gene Ther* 12(3): 238-45 (2005), обе из которых включены в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления средство может представлять собой средство, которое обеспечивает снижение уровня иммуноактивирующего средства. В некоторых вариантах осуществления средство подавляет экспрессию МНС класса I или МНС класса II. В некоторых вариантах осуществления средство подавляет экспрессию одного или нескольких костимулирующих белков. В некоторых вариантах осуществления костимулирующие белки включают без ограничения следующее: LAG3, ICOS-L, ICOS, Oх40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3 или B7-H4. В некоторых вариантах осуществления средство, которое обеспечивает снижение уровня иммуноактивирующего средства, предусматривает малую молекулу или ингибирующую РНК.

В некоторых вариантах осуществления средство по сути не вызывает иммуногенного ответа иммунной системы, например, врожденной иммунной системы. В некоторых вариантах осуществления иммуногенный ответ системы врожденного иммунитета включает ответ, осуществляемый клетками врожденного иммунитета, в том числе без ограничения NK-клетками, макрофагами, нейтрофилами, базофилами, эозинофилами, дендритными клетками, тучными клетками или гамма/дельта Т-клетками.

В некоторых вариантах осуществления средство по сути не вызывает иммуногенного ответа иммунной системы, например, адаптивной иммунной системы. В некоторых вариантах осуществления иммуногенный ответ адаптивной иммунной системы включает иммуногенный ответ адаптивной иммунной клетки, в том числе без ограничения

изменение, например, повышение количества или активности Т-лимфоцитов (например, CD4 Т-клеток, CD8 Т-клеток или гамма-дельта Т-клеток) или В-лимфоцитов.

В некоторых вариантах осуществления средство способствует установлению иммунотолерантности в отношении средства доставки, например, вирусного капсида, например, капсида AAV. В некоторых вариантах осуществления средство способствует установлению иммунотолерантности в отношении компонента системы модификации гена, например, полипептида Gene Writer или нуклеиновой кислоты, кодирующей полипептид Gene Writer, нуклеиновой кислоты-матрицы Gene Writer (например, РНК- или ДНК-матрицы) или нуклеиновой кислоты, кодирующей матрицу Gene Writer (например, ДНК, кодирующей РНК-матрицу), дополнительной нуклеиновой кислоты системы Gene Writer (например, gRNA) или средства доставки системы модификации гена, например, AAV или наночастицы. В некоторых вариантах осуществления средство способствует установлению иммунотолерантности в отношении одного или нескольких продуктов, экспрессируемых из генома после осуществления активности системы модификации гена, например, терапевтического белка, например, терапевтического белка, экспрессируемого из кодирующей последовательности, интегрированной в геном, или варианта белка хозяина, созданного с помощью направленной модификации эндогенной кодирующей последовательности.

В некоторых вариантах осуществления приведение клетки-хозяина в контакт с полипептидом Gene Writer и средством, которое способствует активности системы модификации гена, приводит в результате к повышению уровней последовательности гетерологичного объекта в геноме клетки-хозяина по сравнению с таковыми в аналогичной в ином отношении клетке, которую не приводили в контакт со средством, которое способствует активности системы модификации гена, например, где количество копий последовательности гетерологичного объекта в геноме популяции клеток-хозяев является на по меньшей мере 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% или 100% большим или является в по меньшей мере 2 раза, 5 раз или 10 раз большим, чем количество копий последовательности гетерологичного объекта в геноме аналогичных в ином отношении клеток, которые были приведены в контакт с системой модификации гена, но не со средством, которое способствует активности системы модификации гена.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

На фигуре 1 А и В описан анализ на основе активности люциферазы для первичных клеток. LNP, составленные в соответствии с примером 9, анализировали в отношении доставки полезной нагрузки к первичным гепатоцитам человека (А) и мыши (В) в соответствии с примером 10. В результате анализа с использованием люциферазы было выявлено дозозависимую активность люциферазы в лизатах клеток, что указывает на успешную доставку РНК в клетки и экспрессию люциферазы светлячка из полезной нагрузки мРНК.

На фигуре 2 показана LNP-опосредованная доставка полезной нагрузки РНК в печень мыши. LNP, содержащие мРНК люциферазы светлячка, составляли и доставляли

мышам iv, а образцы печени собирали и анализировали в отношении активности люциферазы через 6, 24 и 48 часов после введения. Активность репортерного гена в различных составах соответствовала ранжированию LIPIDV005>LIPIDV004>LIPIDV003. Экспрессия РНК была временной, и уровни ферментов возвращались к уровню, близкому к фоновому уровню в среде-носителе, через 48 часов после введения.

ПОДРОБНОЕ ОПИСАНИЕ

Определения

Используемый в данном документе термин “средство, которое способствует активности системы модификации гена” относится к средству (например, соединению, совокупности соединений, нуклеиновой кислоте, полипептиду или комплексу), которое способствует требуемому изменению нуклеиновой кислоты-мишени (например, вставке последовательности гетерологичного объекта в сайт-мишень в нуклеиновой кислоте-мишени) в присутствии системы модификации гена. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой модулятор ответа хозяина или эпигенетический модификатор. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, оказывает действие на сайт-мишень, эндогенный белок или эндогенную РНК.

Используемый в данном документе термин “антитело” относится к молекуле, которая специфически связывается с конкретным антигеном или является иммунологически реакционно-способной в отношении его, и включает по меньшей мере переменный домен тяжелой цепи и обычно включает по меньшей мере переменные домены тяжелой цепи и легкой цепи иммуноглобулина. Антитела и их антигенсвязывающие фрагменты, варианты или производные включают без ограничения следующее: поликлональные, моноклональные, полиспецифические, человеческие, гуманизированные, приматизированные или химерные антитела, гетероконъюгированные антитела (например, би-, три- и квадроспецифические антитела, диатела, триатела и тетратела), однодоменные антитела (sdAb), эпитоп-связывающие фрагменты, например, Fab, Fab' и F(ab')₂, Fd, Fv, одноцепочечные Fv (scFv), rlgG, одноцепочечные антитела, связанные дисульфидной связью Fv (sdFv), фрагменты, содержащие либо домен VL, либо домен VH, фрагменты, продуцируемые библиотекой экспрессии Fab, и антиидиотипические (специфические в отношении Id) антитела. Молекулы антител могут относиться к любому типу (например, IgG, IgE, IgM, IgD, IgA и IgY), классу (например, IgG1, IgG2, IgG3, IgG4, IgA1 и IgA2) или подклассу молекулы иммуноглобулина. Более того, если не указано иное, термин “моноклональное антитело” (mAb) подразумевает включение как интактных молекул, так и фрагментов антител (таких, как, например, фрагменты Fab и F(ab')₂), которые способны специфически связываться с белком-мишенью. Фрагменты Fab и F(ab')₂ не содержат Fc-фрагмента интактного антитела. Термин “ингибирующее антитело” относится к антителам, которые способны связываться с антигеном-мишенью и ингибировать или снижать его функцию и/или ослаблять один или

несколько путей передачи сигнала, опосредованных антигеном. Например, ингибирующие антитела могут связываться с лиганд-связывающим доменом рецептора или внеклеточными областями трансмембранного белка и блокировать их. Молекулы ингибирующих антител, которые проникают в клетку, могут блокировать функцию антигена, представляющего собой фермент, или антигена, представляющего собой сигнальную молекулу. Ингибирующие антитела ингибируют или снижают функцию антигена и/или ослабляют один или несколько путей передачи сигнала, опосредованных антигеном, на по меньшей мере 10% (например, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% или больше). Термин “агонистическое антитело” относится к антителам, которые способны связываться с антигеном-мишенью и обеспечивать повышение его активности или функции, например, обеспечивать повышение или активацию одного или нескольких путей передачи сигнала, опосредованных антигеном. Например, агонистическое антитело может связываться с внеклеточной областью трансмембранного белка и оказывать на нее агонистическое действие. Молекулы агонистических антител, которые проникают в клетку, могут обеспечивать повышение функции антигена, представляющего собой фермент, или антигена, представляющего собой сигнальную молекулу. Агонистические антитела обеспечивают активацию или повышение функции антигена и/или одного или нескольких путей передачи сигнала, опосредованных антигеном, на по меньшей мере 10% (например, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% или больше).

Термин “антигенсвязывающий фрагмент”, используемый в данном документе, относится к одному или нескольким фрагментам иммуноглобулина, которые сохраняют способность специфически связываться с антигеном-мишенью. Антигенсвязывающую функцию иммуноглобулина могут выполнять фрагменты полноразмерного антитела. Фрагменты антител могут представлять собой Fab, F(ab')₂, scFv, SMIP, диатело, триатело, аффитело, нанотело, аптамер или доменное антитело. Примеры связывающих фрагментов, охватываемых термином “антигенсвязывающий фрагмент” антитела, включают без ограничения следующее: (i) Fab-фрагмент - одновалентный фрагмент, состоящий из доменов VL, VH, CL и CH1; (ii) F(ab')₂-фрагмент - двухвалентный фрагмент, содержащий два фрагмента Fab, связанных дисульфидным мостиком в шарнирной области; (iii) Fd-фрагмент, состоящий из доменов VH и CH1; (iv) Fv-фрагмент, состоящий из доменов VL и VH одного плеча антитела, (v) dAb (Ward et al., Nature 341:544-546, 1989), содержащий домены VH и VL; (vi) фрагмент dAb, состоящий из домена VH; (vii) dAb, который состоит из домена VH или VL; (viii) выделенная определяющая комплементарность область (CDR); и (ix) комбинацию двух или более выделенных CDR, которые могут быть необязательно соединены синтетическим линкером. Помимо этого, несмотря на то, что два домена фрагмента Fv, VL и VH, кодируются отдельными генами, они могут быть соединены с использованием способов на основе рекомбинации с помощью линкера, который позволяет получать их в виде одной белковой цепи, в которой области VL и VH попарно объединяются с образованием одновалентных молекул (известных как одноцепочечный Fv

(scFv)). Эти фрагменты антител могут быть получены с помощью общепринятых методик, известных специалистам в данной области техники, и фрагменты могут подвергаться скринингу в отношении их применимости таким же образом, как и интактные антитела. Антигенсвязывающие фрагменты могут быть получены с помощью методик рекомбинантной ДНК, ферментативного или химического расщепления интактных иммуноглобулинов или в определенных случаях с помощью процедур химического синтеза пептидов, известных в данной области техники.

Полипептид “Gene Writer”, используемый в данном документе, относится к полипептиду, способному интегрировать последовательность нуклеиновой кислоты (например, последовательность, представленную на матричной нуклеиновой кислоте) в молекулу ДНК-мишени (например, в клетке-хозяине млекопитающего, такую как молекула геномной ДНК в клетке-хозяине). В некоторых вариантах осуществления полипептид Gene Writer способен интегрировать последовательность по сути без использования аппарата хозяина. В некоторых вариантах осуществления полипептид Gene Writer интегрирует последовательность в произвольное положение в геноме, а в некоторых вариантах осуществления полипептид Gene Writer интегрирует последовательность в конкретный сайт-мишень. В некоторых вариантах осуществления полипептид Gene Writer содержит один или несколько доменов, которые в совокупности обеспечивают следующее: 1) связывание нуклеиновой кислоты-матрицы, 2) связывание молекулы ДНК-мишени и 3) обеспечивают интеграцию по меньшей мере части нуклеиновой кислоты-матрицы в ДНК-мишень. Полипептиды Gene Writer предусматривают как встречающиеся в природе полипептиды, такие как РНК-ретротранспозазы, ДНК-рекомбиназы (например, тирозиновые рекомбиназы, сериновые рекомбиназы и т. д.), так и ДНК-транспозазы, а также сконструированные варианты изложенного выше, например, содержащие одну или несколько аминокислотных замен по сравнению с последовательностью, встречающейся в естественных условиях. Полипептиды Gene Writer также предусматривают гетерологичные конструкции, например, такие, в которых один или несколько доменов, перечисленных выше, являются взаимно гетерологичными, вне зависимости от того, получены ли они с помощью слияния (или другого варианта конъюгирования) гетерологичных доменов, являющихся в ином случае доменами дикого типа, либо вариантов слияния модифицированных доменов, например, с помощью замены или слияния гетерологичного субдомена или другого замещенного домена. Иллюстративный полипептид Gene Writer и системы, содержащие его, а также способы его применения, которые можно использовать в способах, предусмотренных в данном документе, описаны, например, в РСТ/US19/48607, поданной 28 августа 2019 г.; 62/876165, поданной 19 июля 2019 г.; 62/939525, поданной 22 ноября 2019 г.; и 62/967934, поданной 30 января 2020 г., каждая из которых включена в данный документ посредством ссылки.

Термин “модулятор ответа хозяина”, используемый в данном документе, относится к средству, которое модифицирует системные (например, адаптивные, врожденные или адаптивные и врожденные иммунные ответы), внутриклеточные (ответ на повреждение и

репарацию ДНК, клеточный врожденный иммунитет) или системные и внутриклеточные ответы на полипептид Gene Writer, нуклеиновую кислоту, кодирующую полипептид Gene Writer, или активность полипептида Gene Writer. В некоторых вариантах осуществления средство предусматривает соединение, совокупность соединений, нуклеиновую кислоту, полипептид или комплекс. Иллюстративные средства включают малые молекулы и большие молекулы, такие как биологический препарат, например, нуклеиновая кислота или полипептид, а также комбинации больших и малых молекул, такие как конъюгат антитело-лекарственное средство. В определенных вариантах осуществления модулятор ответа хозяина ингибирует (уменьшает, репрессирует или блокирует; например, на по меньшей мере 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99% или больше по сравнению с контролем, например, в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500 или 1000 раз) ответ хозяина, в то время как в других вариантах осуществления модулятор ответа хозяина обеспечивает повышение (стимулирует или способствует; например, на по меньшей мере 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99% или больше по сравнению с контролем, например, в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500 или 1000 раз) уровня ответа хозяина. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой ингибитор ответа хозяина. Таким образом, в некоторых таких вариантах осуществления ответ хозяина ингибирует активность Gene Writer, а ингибитор ответа хозяина обеспечивает снижение уровня ответа хозяина, что таким образом способствует активности Gene Writer. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой стимулятор ответа хозяина. Таким образом, в некоторых таких вариантах осуществления ответ хозяина способствует активности Gene Writer, а стимулятор ответа хозяина обеспечивает повышение ответа хозяина, что таким образом способствует активности Gene Writer. Иллюстративные классы модуляторов ответа хозяина включают антитела (в том числе конъюгаты антител), модуляторы, представляющие собой нуклеиновые кислоты (например, ингибирующие РНК, в том числе конъюгаты, содержащие такие молекулы), системы на основе CRISPR, другие модуляторы, содержащие полипептиды (в том числе доминантно-негативные полипептиды и конъюгаты, содержащие их), низкомолекулярные модуляторы и комбинации изложенного выше. Модулятор ответа хозяина может представлять собой компонент, эндогенный для клетки, или компонент, чужеродный для клетки, например, компонент, который в ином случае не находился бы в клетке. В некоторых вариантах осуществления модулятор ответа хозяина предусматривает природный компонент клетки-хозяина, например, нуклеиновую кислоту или белок, или нуклеиновую кислоту, кодирующую белок, клетки-хозяина. В некоторых вариантах осуществления модулятор ответа хозяина не предусматривает природного компонента клетки-хозяина, например, не предусматривает нуклеиновой кислоты, белка или нуклеиновой кислоты, кодирующей белок, встречающихся в клетке-хозяине в естественных условиях. В некоторых вариантах осуществления модулятор ответа хозяина содержит компонент, не встречающийся в клетке в естественных условиях, например, содержит нуклеиновую кислоту, белок или малую

молекулу, не встречающиеся в клетке-хозяине в естественных условиях.

Используемый в данном документе термин “эпигенетический модификатор” относится к средству (например, соединению, совокупности соединений, нуклеиновой кислоте, полипептиду или комплексу), которое изменяет эпигенетическое состояние нуклеиновой кислоты. В некоторых вариантах осуществления эпигенетический модификатор обеспечивает повышение или снижение уровня метилирования ДНК. В некоторых вариантах осуществления эпигенетический модификатор обеспечивает повышение или снижение уровня ковалентной модификации гистона. В некоторых вариантах осуществления эпигенетический модификатор обеспечивает повышение или снижение количества гистонов в области нуклеиновой кислоты. В некоторых вариантах осуществления эпигенетический модификатор обеспечивает изменение положения гистонов в области нуклеиновой кислоты.

Геномная инженерия предусматривает огромный терапевтический потенциал, в том числе возможность окончательного излечения генетических заболеваний. Существующие способы геномной инженерии, однако, ограничены среди прочего ограниченной способностью существующих систем обеспечивать эффективную интеграцию последовательностей, таких как последовательности, состоящие из нескольких оснований, в ДНК, что обусловлено, по меньшей мере частично, использованием эндогенного аппарата хозяина для осуществления редактирования. Более того, даже определенные автономные (т. е. без использования эндогенного аппарата хозяина) системы конструирования генома, например, основанные на мобильных генетических элементах, могут ингибироваться с помощью путей ответа хозяина, например, путей, ингибирующих активность мобильных генетических элементов. Соответственно, существует потребность в улучшенных способах конструирования генома, которые учитывают как потребность в улучшенных системах конструирования генома, так и потребность в уменьшении воздействия путей ответа хозяина, которые в противном случае ограничивают эффективность этих систем.

В настоящем изобретении предусмотрены среди прочего способы конструирования генома, в которых используются улучшенные системы конструирования генома и которые обеспечивают ингибирование путей ответа хозяина, которые ингибируют эти системы. Настоящее изобретение основано, по меньшей мере частично, на наблюдении заявителя, что определенные пути защиты хозяина могут обеспечивать ингибирование способов модификации генома, например, с помощью ингибирования систем, которые в противном случае способны обеспечивать автономное (т. е. без использования эндогенного аппарата хозяина) модифицирование молекулы ДНК в клетке млекопитающего, такой как геном клетки. В некоторых вариантах осуществления модуляция ответа хозяина приводит в результате к повышению стабильности, например, сохранению вставки или ее экспрессии. В некоторых вариантах осуществления модуляция ответа хозяина приводит в результате к снижению цитотоксичности.

Общие сведения по видам ответа хозяина

В некоторых вариантах осуществления система модификации гена, описанная в

данном документе, индуцирует ответ хозяина. В некоторых вариантах осуществления ответ хозяина предусматривает повышение уровня эндогенного белка, снижение уровня эндогенного белка, повышение активности эндогенного белка, снижение активности эндогенного белка, повышение уровня эндогенной РНК или снижение уровня эндогенной РНК.

В некоторых вариантах осуществления средство (например, модулятор ответа хозяина или эпигенетическое модифицирующее средство), которое способствует активности системы модификации гена, не слито с компонентом системы модификации гена. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, слито с компонентом системы модификации гена. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, ковалентно связано с компонентом системы модификации гена. В некоторых вариантах осуществления средство ковалентно связано или слито с полипептидом Gene Writer или с нуклеиновой кислотой, кодирующей полипептид Gene Writer. В некоторых вариантах осуществления средство ковалентно связано или слито с нуклеиновой кислотой-матрицей (например, РНК, ДНК или ДНК, кодирующей РНК-матрицу). В некоторых вариантах осуществления средство ковалентно связано или слито с gRNA. В некоторых вариантах осуществления средство представляет собой нуклеиновую кислоту, например, РНК, например, ингибирующую РНК, малую молекулу, большую молекулу, например, биологический препарат, например, полипептид, например, антитело (в том числе конъюгаты антитело-лекарственное средство) или фермент или их функциональный фрагмент, например, домен. В некоторых вариантах осуществления средство модулирует, например, ингибирует или стимулирует, процесс у хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина ингибирует ответ хозяина (например, нежелательный ответ хозяина) на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или 90% по сравнению с таковым в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или системой модификации гена, не слитой с модулятором ответа хозяина. В некоторых вариантах осуществления модулятор ответа хозяина ингибирует ответ хозяина до уровня, характерного для подобной в ином отношении клетки, которую не приводили в контакт с системой модификации гена.

В некоторых вариантах осуществления модулятор ответа хозяина ингибирует процесс у хозяина (например, ингибирует или секвестрирует фермент репарации ДНК хозяина, который может препятствовать Gene Writing) на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или 90% по сравнению с таковым в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или с системой модификации гена, не слитой с модулятором ответа хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина обеспечивает снижение иммунного ответа хозяина (например, модулятор предусматривает фермент, например, эндопептидазу, например, Ig-расщепляющую эндопептидазу, например, IdeS,

которая разрушает антитела хозяина, в том числе нейтрализующие антитела к AAV, слитые с компонентом средства доставки, например, AAV, например, капсидом AAV или, например, молекулой, которая способствует установлению иммунотолерантности). В некоторых вариантах осуществления иммунный ответ хозяина снижается на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или 90% по сравнению с таковым в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, не слитой с модулятором ответа хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина обеспечивает повышение уровня ответа хозяина (например, требуемого ответа хозяина) на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или 90% по сравнению с таковым в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или системой модификации гена, не слитой с модулятором ответа хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина стимулирует процесс у хозяина (например, активирует или рекрутирует белок или комплекс хозяина, например, фермент репарации ДНК хозяина, который стимулирует Gene Writing) на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или 90% по сравнению с таковым в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или с системой модификации гена, не слитой с модулятором ответа хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина обеспечивает повышение уровня молекулы хозяина, например, нуклеиновой кислоты, белка или нуклеиновой кислоты, кодирующей белок, посредством обеспечения получения дополнительных копий этой молекулы, например, большего количества копий нуклеиновой кислоты, белка или нуклеиновой кислоты, кодирующей белок. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой белок, эндогенный для клетки, и приводит в результате к повышению уровней этого белка в клетке. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой нуклеиновую кислоту, эндогенную для клетки, и приводит в результате к повышению уровней этой нуклеиновой кислоты в клетке. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой нуклеиновую кислоту, кодирующую белок, эндогенный для клетки. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой молекулу РНК, например, мРНК, которая кодирует эндогенный белок клетки и приводит в результате к его сверхэкспрессии, например, к уровням экспрессии, которые являются на по меньшей мере 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или больше или в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500 или 1000 раз большими по сравнению с таковыми в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или с системой модификации гена, не слитой с модулятором ответа хозяина. В некоторых вариантах осуществления модулятор ответа хозяина представляет собой молекулу ДНК, например,

эписомальную ДНК, которая кодирует эндогенный белок клетки и приводит в результате к его сверхэкспрессии, например, к уровням экспрессии, которые являются на по меньшей мере 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или больше или в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500 или 1000 раз большими по сравнению с таковыми в подобной в ином отношении клетке, приводимой в контакт с системой модификации гена, но не с модулятором ответа хозяина, или с системой модификации гена, не слитой с модулятором ответа хозяина.

В некоторых вариантах осуществления модулятор ответа хозяина, например, усилитель или ингибитор ответа хозяина, представляет собой фермент. В некоторых вариантах осуществления фермент слит с компонентом средства доставки, например, AAV. В некоторых вариантах осуществления фермент слит с капсидом AAV. В некоторых вариантах осуществления фермент представляет собой эндопептидазу, например, Ig-расщепляющую эндопептидазу. В некоторых вариантах осуществления фермент представляет собой IdeS, которая разрушает антитела хозяина, в том числе нейтрализующие антитела к AAV.

В некоторых вариантах осуществления модулятор ответа хозяина, например, усилитель ответа хозяина, представляет собой белок или его функциональный фрагмент, например, домен. В некоторых вариантах осуществления белок или домен стимулируют процесс у хозяина, например, активируют или рекрутируют белок или комплекс хозяина. В некоторых вариантах осуществления белок или домен стимулируют Gene Writing, например, посредством замены или дополнения белка, комплекса или пути у хозяина. В некоторых вариантах осуществления белок представляет собой фермент репарации ДНК хозяина, который стимулирует Gene Writing. В некоторых вариантах осуществления белок или домен стимулируют запись последовательности в транс-положении. В некоторых вариантах осуществления белок или домен стимулируют запись последовательности в цис-положении. В некоторых вариантах осуществления домен представляет собой домен, который рекрутирует 5'-экзонуклеазу хозяина, например, EXO1, для записи последовательности в цис-положении. В некоторых вариантах осуществления домен представляет собой домен, который рекрутирует структурно-специфическую эндонуклеазу, например, FEN1, для записи последовательности в цис-положении.

В некоторых вариантах осуществления модулятор ответа хозяина, например, ингибитор ответа хозяина, представляет собой белок или его функциональный фрагмент, например, домен. В некоторых вариантах осуществления белок или домен ингибируют процесс у хозяина, например, ингибируют или секвестрируют фермент репарации ДНК хозяина, который может препятствовать Gene Writing.

В некоторых вариантах осуществления ингибитор ответа хозяина ингибирует или секвестрирует белок хозяина (например, фермент хозяина) или комплекс хозяина. В некоторых вариантах осуществления белок хозяина участвует в репарации, направляемой гомологией (HDR). В некоторых вариантах осуществления белок хозяина, участвующий в HDR, выбран из PARP1, PARP2, MRE11, RAD50, NBS1, BARD1, BRCA2, BRCA1, RTS,

RECQ5, RPA3, PP4, PALB2, DSS1, RAD51, BACH1, FANCI, Topbp1, TOPO III, FEN1, MUS81, EME1, SLX1, SLX4, RECQ1, WRN, CtIP, EXO1, DNA2, комплекса MRN), белка, участвующего в комплементирующей группе анемии Фанкони (FANC) (например, FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCI, FANCL, FANCM, FANCN, FANCO, FANCP, FANCO, FANCR, FANCS, FANCT), препятствовании HDR (например, FBH1, RECQ5, BLM, FANCI, PARI, RECQ1, WRN, RTEL, RAP80, miR-155, miR-545, miR-107, miR-1255, miR-148, miR-193), одонитевом отжиге (SSA) (например, RPA, RPA1, RPA2, RPA3, RAD52, XPF, ERCC1), каноническом негомологичном соединении концов (C-NHEJ) (например, DNA-PK, DNA-PKcs, 53BP1, XRCC4, LIG4, XLF, АРТЕМИДА, APLF, PNK, Rif1, PTIP, ДНК-полимераза, Ku70, Ku80), альтернативном негомологичном соединении концов (Alt-NHEJ) (PARP1, PARP2, CtIP, LIG3, MRE11, Rad50, Nbs1, XPF, ERCC1, LIG1, ДНК-полимераза θ , комплекс MRN, XRCC1), репарации ошибочно спаренных оснований (MMR) (например, EXO1, MSH2, MSH3, MSH6, MLH1, PMS2, MLH3, ДНК-полимераза дельта, RPA, RFC, LIG1), эксцизионной репарации нуклеотидов (NER) (например, XPF, XPG, ERCC1, TTDA, UVSSA, USP7, CETN2, RAD23B, UV-DDB, подкомплекс САК, RPA, PCNA), эксцизионной репарации оснований (BER) (например, APE1, полимераза β , полимераза δ , полимераза ϵ , XRCC1, LIG3, FEN-1, PCNA, RECQL4, WRN, MYH, PNKP, APTX), репарации одонитевого разрыва (SSBR) (например, PARP1, PARP2, PARG, XRCC1, ДНК-полимераза β , ДНК-полимераза δ , ДНК-полимераза ϵ , PCNA, LIG1, PNK, PNKP, APE1, APTX, APLF, TDP1, LIG3, FEN1, CtIP, MRN, ERCC1), модификации хроматина (например, Ezh2, HDAC класса I, HDAC класса II, KDM4A/JMJD2A, FACT), клеточном цикле (например, CDK1, CDC7, ATM, ATR), синтезе ДНК на поврежденной матрице (TLS) (например, UBC13 или RAD18), клеточном метаболизме (например, mTOR), гибели клетки (например, p53), или разделении РНК:ДНК/R-петли (например, SETX, RNH1 или RNH2), или ответе с задействованием интерферона типа I (например, каспаза-1, IFN α , IFN β , NF- κ B, TNF- α).

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, модулирует путь, приведенный в таблице 0 в столбце под названием “Путь”. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, модулирует уровень или активность белка, приведенного в таблице 0 в столбце под названием “Белок”. В некоторых вариантах осуществления средство стимулирует или ингибирует путь или белок, приведенные в таблице 0. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой белок или его фрагмент, приведенные в таблице 0. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, предусматривает композицию, приведенную в таблице 0 в столбце под названием “Название молекулы”, например, композицию, описанную в столбце под названием “Цитирование”. В некоторых вариантах осуществления средство представляет собой ингибитор, и средство предусматривает нуклеиновую кислоту, например, ингибирующую РНК, например, siRNA. В некоторых

вариантах осуществления средство предусматривает малую молекулу, белок, слитый белок, антитело, полипептид (например, доминантно-отрицательную мутантную форму полипептида в пути ответа хозяина), фермент (например, эндопептидазу, например, Ig-расщепляющую эндопептидазу, например, IdeS). В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, предусматривает нуклеиновую кислоту, которая ковалентно связана с полипептидом GeneWriter или нуклеиновой кислотой-матрицей. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой малую молекулу. В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, представляет собой домен полипептида.

Таблица 0. Пути хозяина и мишени для модуляции

Категория	Путь	Белок	Тип молекулы	Название молекулы	Список литературы (включено в данный документ посредством ссылки во всей своей полноте)
Обнаружение ДНК	cGAS-STING	cGAS	Ингибитор	PF-06928215	Hall et al. (PLoS One. 12(9): e0184843, 2017)
Обнаружение ДНК	cGAS-STING	cGAS	Ингибитор	RU.365	Vincent et al. (Nat Commun. 29;8(1):750, 2017)
Обнаружение ДНК	cGAS-STING	cGAS	Ингибитор	RU.521	Vincent et al. (Nat Commun. 23; 8(1):1827, 2017)
Обнаружение ДНК	cGAS-STING	cGAS	Ингибитор	Сурамин	Wang et al. (Future Med Chem. 1;10(11):1301-1317, 2018)
Обнаружение ДНК	cGAS-STING	cGAS	Ингибитор	G150	Lama et al. (Nat Commun. 21;10(1):2261, 2019)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	C-176	Haag et al. (Nature. 2018 Jul; 559(7713):269-273)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	C-178	Haag et al. (Nature. 2018 Jul; 559(7713):269-273)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	H151	Haag et al. (Nature. 2018 Jul; 559(7713):269-273)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	Соединение 18	Siu et al. (ACS Med Chem Lett. 2018 Dec 6;10(1):92-97)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	Астин С	Li et al. (Cell Rep. 2018 Dec

					18;25(12):3405-3421.e7)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	Положительный результат скрининга 1	Siu et al. (ACS Med Chem Lett. 2018 Dec 6;10(1):92-97)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	Соединение 13	Siu et al. (ACS Med Chem Lett. 2018 Dec 6;10(1):92-97)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	siRNA-1 (SEQ ID NO: 6)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	siRNA-2 (SEQ ID NO: 7)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	siRNA-3 (SEQ ID NO: 8)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	siRNA-4 (SEQ ID NO: 9)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	E1A (hAd5)	Lau L, (Science. 2015 Oct 30;350(6260):568-71)
Обнаружение ДНК	cGAS-STING	STING	Ингибитор	E7 (HPV18)	Lau et al. (Science. 2015 Oct 30;350(6260):568-71)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	BX795	Clark et al. (J Biol Chem. 2009 May 22;284(21):14136-46.)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	Тозасертиб	Richters et al. (ACS Chem Biol. 2015 Jan 16;10(1):289-98)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	Тозасертиб-15a	Richters et al. (ACS Chem Biol. 2015 Jan 16;10(1):289-98)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	20b	McIver et al. (Bioorg Med Chem Lett. 2012 Dec 1;22(23):7169-73)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	Азабензимидазол,	Wang et al. (Bioorg Med Chem Lett.

				попадание 1a	2012 Mar 1;22(5):2063-9)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	CYT387	Pardanani et al. (Leukemia. 2009 Aug; 23(8):1441-5)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	Domainex	Hasan et al. (Pharmacol Res. 2016 Sep;111:336-342)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	Amgen, соединение II	Ou et al., (Mol Cell. 2011 Feb 18;41(4):458-70)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	MRT67307	Clark et al. (Biochem J. 2011 Feb 15;434(1):93-104)
Обнаружение ДНК	cGAS-STING	TBK1	Ингибитор	AZ13102909	Vu et al. (Mol Cancer Res. 2014 Oct;12(10):1509-19)
Обнаружение ДНК	cGAS-STING	IRF3	Ингибитор	siRNA-1 (SEQ ID NO: 2)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	IRF3	Ингибитор	siRNA-2 (SEQ ID NO: 3)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	IRF3	Ингибитор	siRNA-3 (SEQ ID NO: 4)	WO2018201144A1
Обнаружение ДНК	cGAS-STING	IRF3	Ингибитор	siRNA-4 (SEQ ID NO: 5)	WO2018201144A1
Обнаружение ДНК	SIDSP	DNA-PK	Ингибитор	Nu-7441	Leahy et al. (Bioorg Med Chem Lett. 2004 Dec 20;14(24):6083-7)
Обнаружение ДНК	SIDSP	DNA-PK	Ингибитор	hAd5, E1A	Burleigh et al. (Sci Immunol. 2020 Jan 24;5(43):eaba4219)
Обнаружение ДНК	SIDSP	DNA-PK	Ингибитор	HSV-1, ICP0	Burleigh et al. (Sci Immunol. 2020 Jan 24;5(43):eaba4219)
Обнаружение РНК	IFN1	RIG-I	Ингибитор	Протеаза HIV-1	Solis et al. (J Virol. 2011 Feb;

					85(3):1224-36)
Обнаружение РНК	IFN1	MDA5			
Обнаружение РНК	IFN1	Комплекс IKK	Ингибитор	BAY11	Awe et al. (Stem Cell Res Ther. 2013 Feb 6;4(1):15)
Обнаружение РНК	IFN1	TRIF	Ингибитор	Pepinh-TRIF	Toshchakov et al. (J Immunol. 2005 Jul 1;175(1):494-500)
Обнаружение РНК	IFN1	MyD88	Ингибитор	Pepinh-MYD	Loiarro et al. (J Biol Chem. 2005 Apr 22; 280(16):15809-14)
Обнаружение РНК	IFN1	IFN	Ингибитор	B18R вируса осповакцины	Kim et al. (PLoS One. 2017 Dec 7;12(12):e0189308)
		Созревание эндосом	Ингибитор	Хлорохин	
			Ингибитор	TSA	
		Созревание эндосом		Бафиломицин A1	
		RNASEH2			
		ADAR1			Orecchini et al. (RNA Biol. 2017 Nov 2;14(11):1485-1491)
Противовирусное средство	Интерферон типа I	IFN- α			
Противовирусное средство	Интерферон типа I	IFN- β			

Противовирусное средство	Интерферон типа II	IFN- γ			
HDR	Белок, препятствующий HDR	FBH1	Агонист		
HDR	Белок, препятствующий HDR	RECQ5	Агонист		
HDR	Белок, препятствующий HDR	BLM	Агонист		
HDR	Белок, препятствующий HDR	FANCI	Агонист		
HDR	Белок, препятствующий HDR	PARI	Агонист		
HDR	Белок, препятствующий HDR	RECQ1	Агонист		
HDR	Белок, препятствующий HDR	WRN	Агонист		

HDR	Белок, препятствующий HDR	RTEL	Агонист		
HDR	Белок, препятствующий HDR	Rap80	Агонист		
HDR	Белок, препятствующий HDR	miR-155	Агонист		
HDR	Белок, препятствующий HDR	miR-545	Агонист		
HDR	Белок, препятствующий HDR	miR-107	Агонист		
HDR	Белок, препятствующий HDR	miR-1255	Агонист		
HDR	Белок, препятствующий HDR	miR-148	Агонист		
HDR	Белок, препятствующий	miR-193	Агонист		

	HDR				
		p53	Ингибитор		
		p53	Ингибитор	Доминантно-отрицательная мРНК	Schiroli et al. (Cell Stem Cell. 2019 Apr 4;24(4):551-565.e8)
		BRCA1	Ингибитор	Доминантно-отрицательная мРНК	
Репарация ДНК		BRCA1	Ингибитор	siBRCA1	Mita et al. (Nat Struct Mol Biol. 2020 Feb;27(2):179-191)
Репарация ДНК		BRCA2	Ингибитор	siBRCA2	Mita et al. (Nat Struct Mol Biol. 2020 Feb;27(2):179-191)
Репарация ДНК		FANCD	Ингибитор	siFANCD2	Mita et al. (Nat Struct Mol Biol. 2020 Feb;27(2):179-191)
		MORC2	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
		TASOR	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
		MPP8	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
		SETX	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
		MOV10	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)

		SAFB	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
		RAD51	Ингибитор		Liu et al. (Nature. 2018 Jan 11;553(7687):228-232)
Модификатор хроматина	Chr	Ezh2		GSK343	Zeidler M. et al. Neoplasia. Nov 2005; Chou D. et al PNAS 2010; Puppe et al, Breast cancer research 2009
Модификатор хроматина	Chr	Ezh2		EPZ-6438	Zeidler M. et al. Neoplasia. Nov 2005; Chou D. et al PNAS 2010; Puppe et al, Breast cancer research 2009
Модификатор хроматина	Chr	Ezh2		GSK2816126	Zeidler M. et al. Neoplasia. Nov 2005; Chou D. et al PNAS 2010; Puppe et al, Breast cancer research 2009
Модификатор хроматина	Chr	Ezh2		SureCN6120847	Zeidler M. et al. Neoplasia. Nov 2005; Chou D. et al PNAS 2010; Puppe et al, Breast cancer research 2009
Модификатор хроматина	Chr	Ezh2		EPZ005687	Zeidler M. et al. Neoplasia. Nov 2005; Chou D. et al PNAS 2010; Puppe et al, Breast cancer research

					2009
Модификатор хроматина	Chr	HDAC класса I/класса II		Трихостатин А (TSA)	Tang et al, Nat Struct Mol Biol. 2013
Модификатор хроматина	Chr	HDAC класса I/класса II		Бутират натрия (NaB)	Tang et al, Nat Struct Mol Biol. 2013
Модификатор хроматина	Chr	KDM4A/ JMJD2A			Pfister et al, Cell Reports 2014
Модификатор хроматина	Chr	FACT		CBL0137	
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		Олапариб, AZD2281, KU-0059436	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		Инипариб, BSI-201	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		BMN 673	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		Рукапариб (AG014699, PF-01367338)	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		Велипариб, ABT-888	Caldecott Nature reviews Genetic, 2008

Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		CEP 9722	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		INO-1001	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		MK 4827	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		BGB-290	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		E701,GPI21016	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		MP-124	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		LT-673	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		NMS-P118	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		XAV939	Caldecott Nature reviews Genetic, 2008
Репарация однонитевого разрыва	SSBR	PARP1/PA RP2		3-Аминобензамид	Caldecott Nature reviews Genetic, 2008
Репарация ошибочно спаренных оснований	MMR	EXO1			

Репарация ошибочно спаренных оснований	MMR	MSH2			Guo-Min Li 2008 Cell Research, 2013 Cancer Research
Репарация ошибочно спаренных оснований	MMR	MSH3			Guo-Min Li 2008 Cell Research, 2013 Cancer Research
Репарация ошибочно спаренных оснований	MMR	MSH6			Guo-Min Li 2008 Cell Research, 2013 Cancer Research
Репарация ошибочно спаренных оснований	MMR	MLH1			Guo-Min Li 2008 Cell Research, 2013 Cancer Research
Репарация ошибочно спаренных оснований	MMR	PMS2			Guo-Min Li 2008 Cell Research, 2013 Cancer Research
Репарация ошибочно спаренных оснований	MMR	MLH3			
Репарация ошибочно спаренных оснований	MMR	ДНК-полимераза дельта			
Репарация ошибочно спаренных оснований	MMR	RPA			
Репарация ошибочно спаренных оснований	MMR	HMGB1			
Репарация ошибочно спаренных оснований	MMR	RFC			
Репарация ошибочно спаренных оснований	MMR	ДНК-			

спаренных оснований		лигаза I			
Эксцизионная репарация нуклеотидов	NER	XPA-G			
Эксцизионная репарация нуклеотидов	NER	POLH			
Эксцизионная репарация нуклеотидов	NER	XPF			
Эксцизионная репарация нуклеотидов	NER	ERCC1			
Эксцизионная репарация нуклеотидов	NER	XPA-G			
Эксцизионная репарация нуклеотидов	NER	LIG1			
Эксцизионная репарация нуклеотидов	NER	CSA			
Эксцизионная	NER	CSB			

репарация нуклеотидов					
Эксцизионная репарация нуклеотидов	NER	XPA			
Эксцизионная репарация нуклеотидов	NER	XPB			
Эксцизионная репарация нуклеотидов	NER	XPC			
Эксцизионная репарация нуклеотидов	NER	XPB			
Эксцизионная репарация нуклеотидов	NER	XPD			
Эксцизионная репарация нуклеотидов	NER	XPF		NSC 130813	
Эксцизионная репарация нуклеотидов	NER	XPG			
Эксцизионная репарация нуклеотидов	NER	ERCC1			

Эксцизионная репарация нуклеотидов	NER	TTDA			
Эксцизионная репарация нуклеотидов	NER	UVSSA			
Эксцизионная репарация нуклеотидов	NER	USP7			
Эксцизионная репарация нуклеотидов	NER	CETN2			
Эксцизионная репарация нуклеотидов	NER	RAD23B			
Эксцизионная репарация нуклеотидов	NER	UV-DDB			
Эксцизионная репарация нуклеотидов	NER	Подкомпле кс CAK			
Эксцизионная репарация	NER	RPA			

нуклеотидов					
Эксцизионная репарация нуклеотидов	NER	PCNA			
Эксцизионная репарация оснований	BER	APE1			
Эксцизионная репарация оснований	BER	Полимераз а бета			
Эксцизионная репарация оснований	BER	Полимераз а дельта			
Эксцизионная репарация оснований	BER	Полимераз а эпсилон			
Эксцизионная репарация оснований	BER	XRCC1			
Эксцизионная репарация оснований	BER	Лигаза III			
Эксцизионная репарация оснований	BER	FEN-1			
Эксцизионная репарация оснований	BER	PCNA			
Эксцизионная репарация оснований	BER	RECQL4			

Эксцизионная репарация оснований	BER	WRN			
Эксцизионная репарация оснований	BER	MYH			
Эксцизионная репарация оснований	BER	PNKP			
Эксцизионная репарация оснований	BER	APTX			
Однонитевой отжиг	SSA	RPA			
Однонитевой отжиг	SSA	RPA1			
Однонитевой отжиг	SSA	RPA2			
Однонитевой отжиг	SSA	RPA3			
Однонитевой отжиг	SSA	RAD52		AID 651668	Ciccia et al, Mol Cell 2010
Однонитевой отжиг	SSA	RAD52		AICAR	Ciccia et al, Mol Cell 2010
Однонитевой отжиг	SSA	XPF		NSC 130813	Ciccia et al, Mol Cell 2010
Однонитевой отжиг	SSA	ERCC1		NSC 130813	Ciccia et al, Mol Cell 2010
Негомологичное соединение концов	C-NHEJ	53BP1			Stephanie Panier and Simon J. Boulton, Nature Review 2014
Негомологичное соединение концов	C-NHEJ	XRCC4			
Негомологичное соединение концов	C-NHEJ	LIG4		SCR7	

Синтез ДНК на поврежденной матрице	TLS	Ubc13			
Синтез ДНК на поврежденной матрице	TLS	Rad18			
Клеточный метаболизм		mTOR		Рапамицин	
		APOBEC			
Негомологичное соединение концов	NHEJ	53 BP1			Stephanie Panier and Simon J. Boulton, Nature Review 2014
Негомологичное соединение концов	NHEJ	Rif1			Di Virgilio M. et al Science 2013; Zimmermann M et al Science 2013
Негомологичное соединение концов	NHEJ	PTIP			Zimmermann M and De langhe, Trends in cell Biology, 2014
Негомологичное соединение концов	NHEJ	KU 70-80			Betermier et al, PLoS Genetics 2014
Негомологичное соединение концов	NHEJ	DNApk			Betermier et al, PLoS Genetics 2014
Негомологичное соединение концов	NHEJ	Lig4			
Негомологичное	NHEJ	XLf			

соединение концов					
Негомологичное соединение концов	NHEJ	Артемида			Betermier et al, PLoS Genetics 2014
Альтернативный путь NHEJ	Alt-NHEJ	Лигаза I			
Альтернативный путь NHEJ	Alt-NHEJ	Лигаза III			
Альтернативный путь NHEJ	Alt-NHEJ	Полимераз а тэта			Koole et al, Nature com 2014; Chan et al Plos Genetics 2010
Препятствование репарации, направляемой гомологией	Препятствовани е HDR	Fbh1			Chapman et al, Molecular Cell 2012
Препятствование репарации, направляемой гомологией	Препятствовани е HDR	RTEL			Chapman et al, Molecular Cell 2012
Препятствование репарации, направляемой гомологией	Препятствовани е HDR	PARI			Chapman et al, Molecular Cell 2012
Препятствование репарации,	Препятствовани е HDR	Rap80			Hu Y. et al, Genes and Dev 2011, Adamson et al Nature Cell Bio 2012

направляемой гомологией					
Препятствование репарации, направляемой гомологией	Препятствовани е HDR	miRNA			Gasparini et al PNAS 2014; Dimitrov et al, Genes&Dev 2014; Huang et al, Mol Cancer res 2013; Neijenhuis et al, DNA repair 2013, Choi et al, eLIFE 2014
	SSR	PARP			Caldecott Nature reviews Genetic, 2008
	SSR	XRCC1			Caldecott Nature reviews Genetic, 2008
Репарация, направляемая гомологией	HDR	BRCA1			Lorraine Symington, Cold Spring Harb Perspect Biol 2014
Репарация, направляемая гомологией	HDR	CtIP			Lorraine Symington, Cold Spring Harb Perspect Biol 2014
Репарация, направляемая гомологией	HDR	EXO1			Lorraine Symington, Cold Spring Harb Perspect Biol 2014
Репарация, направляемая гомологией	HDR	DNA2			Lorraine Symington, Cold Spring Harb Perspect Biol 2014

Репарация, направляемая гомологией	HDR	Комплекс MRN			Lorraine Symington, Cold Spring Harb Perspect Biol 2014
Репарация, направляемая гомологией	HDR	BRCA2			Wolf-Dietrich Heyer Annu.Rev.Genet.2010
Модификатор хроматина	Chr	HDAC класса I			Tang et al, Nat Struct Mol Biol. 2013
Модификатор хроматина	Chr	HDAC класса II			Tang et al, Nat Struct Mol Biol. 2013
Клеточный цикл	Цикл	CDK1		RO-3306	
Клеточный цикл	Цикл	CDK1		AZD 5438	
Клеточный цикл	Цикл	CDC7		XL413	
Синтез ДНК на поврежденной матрице	TLS	Ubc13			
Синтез ДНК на поврежденной матрице	TLS	Rad18			
	R-петля	SEtX			
	R-петля	RNh1 и 2			
Комплементирующая	FANC	FANCM			

группа анемии Фанкони					
Комплементирующая группа анемии Фанкони	FANC	FANCI			
Комплементирующая группа анемии Фанкони	FANC	FANCD1			
Комплементирующая группа анемии Фанкони	FANC	FANCD2			
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		Олапариб, AZD2281, KU-0059436	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		Инипариб, BSI-201	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		BMN 673	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		Рукапариб (AG014699, PF-01367338)	

Репарация, направляемая гомологией	HDR	PARP1/PA RP2		Велипариб, АБТ-888	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		СЕР 9722	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		INO-1001	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		МК 4827	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		BGB-290	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		E701,GPI21016	
Репарация, направляемая гомологией	HDR	PARP1/PA RP2		MP-124	
Репарация, направляемая	HDR	PARP1/PA RP2		LT-673	

гомологией					
Репарация, направляемая гомологией	HDR	PARP1/PARP2		NMS-P118	
Репарация, направляемая гомологией	HDR	PARP1/PARP2		XAV939	
Репарация, направляемая гомологией	HDR	PARP1/PARP2		3-Аминобензамид	
Репарация, направляемая гомологией	HDR	MRE11			
Репарация, направляемая гомологией	HDR	RAD50			
Репарация, направляемая гомологией	HDR	NBS1			
Репарация, направляемая гомологией	HDR	BARD1			
Репарация,	HDR	BRCA2		siBRCA2	Wolf-Dietrich Heyer

направляемая гомологией					Annu.Rev.Genet.2010
Репарация, направляемая гомологией	HDR	RTS			
Репарация, направляемая гомологией	HDR	RECQ5			
Репарация, направляемая гомологией	HDR	RPA3			
Репарация, направляемая гомологией	HDR	PP4			
Репарация, направляемая гомологией	HDR	PALB2			
Репарация, направляемая гомологией	HDR	DSS1			
Репарация, направляемая гомологией	HDR	RAD51		B02	Wolf-Dietrich Heyer Annu.Rev.Genet.2010

Репарация, направляемая гомологией	HDR	BACH1			
Репарация, направляемая гомологией	HDR	FANCI			
Репарация, направляемая гомологией	HDR	Topbp1			
Репарация, направляемая гомологией	HDR	TOPO III			
Репарация, направляемая гомологией	HDR	FEN1			
Репарация, направляемая гомологией	HDR	MUS81			
Репарация, направляемая гомологией	HDR	EME1			
Репарация, направляемая	HDR	SLX1			

гомологией					
Репарация, направляемая гомологией	HDR	SLX4			
Репарация, направляемая гомологией	HDR	RECQ1			
Репарация, направляемая гомологией	HDR	WRN			
Комплементирующая группа анемии Фанкони	FANC	FANC A			
Комплементирующая группа анемии Фанкони	FANC	FANC B			
Комплементирующая группа анемии Фанкони	FANC	FANC C			
Комплементирующая группа анемии Фанкони	FANC	FANC D1			
Комплементирующая	FANC	FANC D2		siFANCD2	

группа Фанкони анемии					
Комплементирующая группа Фанкони анемии	FANC	FANC E			
Комплементирующая группа Фанкони анемии	FANC	FANC F			
Комплементирующая группа Фанкони анемии	FANC	FANC G			
Комплементирующая группа Фанкони анемии	FANC	FANC I			
Комплементирующая группа Фанкони анемии	FANC	FANC J			
Комплементирующая группа Фанкони анемии	FANC	FANC L			
Комплементирующая группа Фанкони анемии	FANC	FANC M			

Комплементирующая группа анемии Фанкони	FANC	FANC N			
Комплементирующая группа анемии Фанкони	FANC	FANC O			
Комплементирующая группа анемии Фанкони	FANC	FANC P			
Комплементирующая группа анемии Фанкони	FANC	FANC Q			
Комплементирующая группа анемии Фанкони	FANC	FANC R			
Комплементирующая группа анемии Фанкони	FANC	FANC S			
Комплементирующая группа анемии Фанкони	FANC	FANC T			
Препятствование репарации,	Белок, препятствующий	miR-155		miR-155-5p	

направляемой гомологией	HDR				
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-155		miR-155-3p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-545		miR-545-5p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-545		miR-545-3p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-107		miR-107	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-1255		miR-1255-A	
Препятствование	Белок,	miR-1255		miR-1255-B1	

репарации, направляемой гомологией	препятствующий HDR				
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-1255		miR-1255-B2	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-148		miR-148-5p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-148		miR-148-3p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-193		miR-193-5p	
Препятствование репарации, направляемой гомологией	Белок, препятствующий HDR	miR-193		miR-193-3p	

Негомологичное соединение концов	C-NHEJ	DNA Pk		NU7441	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		CC115	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		NK314	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		Вортманнин	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		LY294002	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		NU 7026	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		IC86621	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		IC87102	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		IC87361	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		OK1035	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		SU11752	Betermier et al, plos genetics 2014

Негомологичное соединение концов	C-NHEJ	DNA Pk		IC486241	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA Pk		Ванилин	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	DNA-PKcs			
Негомологичное соединение концов	C-NHEJ	53BP1			Stephanie Panier and Simon J. Boulton, Nature Review 2014
Негомологичное соединение концов	C-NHEJ	XLF			
Негомологичное соединение концов	C-NHEJ	АРТЕМИД А		siRNA	Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	APLF			
Негомологичное соединение концов	C-NHEJ	PNK			
Негомологичное соединение концов	C-NHEJ	Rif1			Di Virgilio M. et al Science 2013; Zimmermann M et al Science 2013
Негомологичное соединение концов	C-NHEJ	PTIP			Zimmermann M and De langhe, Trends in cell Biology, 2014
Негомологичное соединение концов	C-NHEJ	ДНК- полимераза			

Негомологичное соединение концов	C-NHEJ	Ku70			Betermier et al, plos genetics 2014
Негомологичное соединение концов	C-NHEJ	Ku80			Betermier et al, plos genetics 2014
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		Олапариб, AZD2281, KU-0059436	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		Инипариб, BSI-201	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		BMN 673	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		Рукапариб (AG014699, PF-01367338)	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		Велипариб, ABT-888	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		CEP 9722	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		INO-1001	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		МК 4827	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PA RP2		BGB-290	

Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		E701,GPI21016	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		MP-124	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		LT-673	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		NMS-P118	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		XAV939	
Альтернативный путь NHEJ	Alt-NHEJ	PARP1/PARP2		3-Аминобензамид	
Альтернативный путь NHEJ	Alt-NHEJ	CtIP			
Альтернативный путь NHEJ	Alt-NHEJ	LIG-3			
Альтернативный путь NHEJ	Alt-NHEJ	MRE11			
Альтернативный путь NHEJ	Alt-NHEJ	Rad50			
Альтернативный путь NHEJ	Alt-NHEJ	Nbs1			

Альтернативный путь NHEJ	Alt-NHEJ	CtIP			
Альтернативный путь NHEJ	Alt-NHEJ	XPF		NSC 130813	
Альтернативный путь NHEJ	Alt-NHEJ	ERCC1		NSC 130813	
Альтернативный путь NHEJ	Alt-NHEJ	Лигаза 1			
Альтернативный путь NHEJ	Alt-NHEJ	ДНК-полимераза тэта			Koole et al, Nature com 2014; Chan et al Plos Genetics 2010
Альтернативный путь NHEJ	Alt-NHEJ	Комплекс MRN			
Альтернативный путь NHEJ	Alt-NHEJ	XRCC1			
Репарация одностороннего разрыва	SSBR	PARG			
Репарация одностороннего разрыва	SSBR	XRCC1			Caldecott Nature reviews Genetic, 2008
Репарация одностороннего разрыва	SSBR	ДНК-полимераза бета			

Репарация однонитевого разрыва	SSBR	ДНК- полимераза дельта			
Репарация однонитевого разрыва	SSBR	ДНК- полимераза эпсилон			
Репарация однонитевого разрыва	SSBR	PCNA			
Репарация однонитевого разрыва	SSBR	LIG1			
Репарация однонитевого разрыва	SSBR	PNK			
Репарация однонитевого разрыва	SSBR	PNKP			
Репарация однонитевого разрыва	SSBR	APE1			
Репарация однонитевого разрыва	SSBR	APTX			
Репарация однонитевого разрыва	SSBR	APLF			
Репарация однонитевого разрыва	SSBR	TDP1			

Репарация однонитевого разрыва	SSBR	LIG3			
Репарация однонитевого разрыва	SSBR	FEN1			
Репарация однонитевого разрыва	SSBR	CtIP		MLN4924	
Репарация однонитевого разрыва	SSBR	MRN			
Репарация однонитевого разрыва	SSBR	ERCC1		NSC 130813	
Клеточный цикл	Цикл	ATM			
Ответ клеточной РНК	Клетка	MDA5			
		LIG4		SCR7	
		ATM			
		ATR			
Ингибиторы p53					
Ингибиторы ATR					
Репарация ДНК		p53, ATM, ATR			
Клеточный цикл		p53, ATM, ATR			
Гибель клетки		p53			

Обнаружение ДНК	TLR9	TLR9	Ингибитор	ODN A151	Kath et al. (BioRxiv. 2021.02.14.431017)
Обнаружение ДНК	cGAS	cGAS	Ингибитор	ODN A151	Kath et al. (BioRxiv. 2021.02.14.431017)
Обнаружение ДНК	AIM2	AIM2	Ингибитор	ODN A151	Kath et al. (BioRxiv. 2021.02.14.431017)
Обнаружение ДНК	cGAS	cGAS	Ингибитор	Ru.521	Kath et al. (BioRxiv. 2021.02.14.431017)
Обнаружение ДНК	STING		Ингибитор	H151	Kath et al. (BioRxiv. 2021.02.14.431017)
Обнаружение ДНК	STING	Tbx1	Ингибитор	BX795	Kath et al. (BioRxiv. 2021.02.14.431017)

В некоторых вариантах осуществления способы, описанные в данном документе, включают модулирование, например, активацию или подавление, одного или нескольких из следующего: ADAR1, AICDA, AIM2, ALKBH1, APE, APOBEC1, APOBEC3, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3F, APOBEC3G, APOBEC3H, ASZ1, ATG5, ATM, BECN1, BRCA1, BRCA2, BST2, C3, C3ORF26, C3ORF37, CALCOCO2, CBX1, CBX3, CDK9, CHAF1, CMPK1, CORO1B, CSDA, DCLRE1C, DDX17, DDX21, DDX39A, DDX4, DDX5, DDX58, DDX6, DGCR8, DHX9, DICER1, DNMT1, DNMT3A, DNMT3B, DNMT3L, DROSHA, EHMT2, ELAVL1, ERAL1, ERCC1, ERCC2, ERCC4, EXD1, EZH2, FAM120A, FAM98A, FANCA, FANCB, FANCM, FASTKD2, FCGR1B, FKBP4, FKBP6, GTSF1, H1FX, HAX1, HECTD1, HENMT1, HEXIM1, HIST1H1C, HIST1H2BO, HNRNPA1, HNRNPA2B1, HNRNPAB, HNRNPC, HNRNPL, HNRNPU, HSP90AA1, HSP90AB1, HSPA1A, HSPA8, IGF2BP1, IGF2BP2, IGF2BP3, ILF2, ILF3, IPO7, ISG20, KDM1A, KIAA0430, KPNA2, KPTN, LARP1, LARP7, LIG4, Ligase IV, MAEL, MATR3, MAVS, MDA5, MECP2, MEPCE, MIR128-1, MORC1, MOV10, MOV10L1, MRE11A, MRPL28, MTNR1A, MX2, NAP1L1, NAP1L4, NCF4, NCL, NOP56, NPM1, NUSAP1, PABPC1, PABPC4, PABPC4L, PALB2, PARP1, PCBP2, PCNA, PIWIL1, PIWIL2, PIWIL4, PLD6, PRKDC, PURA, PURB, RAD50, RAD54L, RALY, RBMX, RCL1, RDH8, RIG-I, RIOK1, RNaseH1, RNaseH2, RNaseH2A, RNaseH2B, RNaseH2C, RNase L, RNASEL, RPRD2, RPS27A, SAMHD1, SERBP1, SETDB1, SF3B3, SIRT6, SNRNP70, SNUPN, SQSTM1, SRP14, SRSF1, SRSF10, SRSF6, SSB, STAU1, STAU2, STK17A, SUV39H, SYNCRIP, TBX1, TDRD1, TDRD12, TDRD5, TDRD9, TDRKH, TEX19, TIMM13, TIMM8B, TLR3, TLR9, TOMM40, TOP1, TRA2A, TRA2B, TREX1, TRIM28, TRIM5 α , TROVE2, TUBB, TUBB2C, UBE2T, UHRF1, UNG, UQCRH, XRCC2, XRCC4, XRCC6, XRN1, YARS2, YBX1, YME1L1, ZAP, ZC3HAV1, ZCCHC3, и ZFR.

Ингибирующие РНК

В некоторых вариантах осуществления модулятор ответа хозяина, например, ингибитор ответа хозяина, предусматривает молекулу нуклеиновой кислоты, например, молекулу РНК. В некоторых вариантах осуществления модулятор ответа хозяина, например, ингибитор ответа хозяина, представляет собой ингибирующую молекулу РНК. В некоторых вариантах осуществления молекула ингибирующей РНК обеспечивает снижение уровня (например, уровня белка или уровня мРНК) фактора, кодируемого геном, описанным в данном документе, т. е. который опосредует ответ хозяина.

Определенные РНК могут ингибировать экспрессию генов посредством биологического процесса РНК-интерференции (RNAi). В некоторых вариантах осуществления молекулы RNAi предусматривают РНК или РНК-подобные структуры, обычно содержащие 15-50 пар оснований (например, приблизительно 18-25 пар оснований) и имеющие последовательность азотистых оснований, идентичную (комплементарную) или почти идентичную (по сути комплементарную) кодирующей последовательности в экспрессируемом гене-мишени в клетке. Молекулы для RNAi предусматривают без ограничения короткие интерферирующие РНК (siRNA), двухнитевые РНК (dsRNA),

микроРНК (miRNA), короткие шпилечные РНК (shRNA), меродуплексы и субстраты дайсера (патенты США №№ 8084599, 8349809 и 8513207). В некоторых вариантах осуществления в данном документе средство представляет собой молекулу для RNAi, которая ингибирует экспрессию гена, участвующего в ответе хозяина.

В некоторых вариантах осуществления молекулы для RNAi содержат последовательность, по сути комплементарную или полностью комплементарную всему гену-мишени или его фрагменту. Молекулы для RNAi могут быть комплементарны последовательностям на границе между интронами и экзонами для предотвращения созревания новообразованных ядерных РНК-транскриптов специфических генов до мРНК для транскрипции. Молекулы для RNAi, комплементарные конкретным генам, могут гибридизоваться с мРНК для этого гена и, например, предотвращать его трансляцию. Антисмысловая молекула может представлять собой ДНК, РНК или их производное или гибрид. Примеры таких производных молекул включают без ограничения пептидную нуклеиновую кислоту (PNA) и молекулы на основе фосфотиоатных связей, такие как дезоксирибонуклеиновый гуанидин (DNG) или рибонуклеиновый гуанидин (RNG).

Молекулы для RNAi могут быть доставлены в клетку в виде "готовой к применению" РНК, синтезированной *in vitro*, или в виде антисмыслового гена, трансфицируемого в клетки, в результате транскрипции которого образуются молекулы для RNAi. Гибридизация с мРНК в некоторых вариантах осуществления приводит в результате к разрушению гибридизированной молекулы под действием РНКазы и/или ингибированию образования трансляционных комплексов. И то, и другое может приводить в результате к невозможности продуцировать продукт исходного гена.

Длина молекулы для RNAi, которая гибридизуется с представляющим интерес транскриптом, может составлять примерно 10 нуклеотидов, от приблизительно 15 до 30 нуклеотидов или приблизительно 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или больше нуклеотидов. Степень идентичности антисмысловой последовательности с транскриптом-мишенью может составлять по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95%.

Молекулы для RNAi также могут содержать выступающий конец (например, могут содержать два выступающих конца) - обычно неспаренные выступающие нуклеотиды, которые непосредственно не задействованы в двойной спиральной структуре, в норме образуемой парами последовательностями пары смысловой и антисмысловой нити. Молекулы для RNAi могут содержать 3'- и/или 5'-выступающие концы, каждый из которых независимо состоит из приблизительно 1-5 оснований (например, 2 оснований) на каждой из смысловых нитей и антисмысловых нитей. Смысловая и антисмысловая нити молекулы для RNAi могут содержать одинаковое или разное количество нуклеотидных оснований. Антисмысловая и смысловая нити могут образовывать дуплекс, где только 5'-конец имеет тупой конец, только 3'-конец имеет тупой конец, как 5'-, так и 3'-концы заканчиваются тупым концом, или ни 5'-, ни 3'-конец не заканчиваются тупым концом. В другом варианте осуществления один или несколько нуклеотидов в выступающих концах содержат

тиофосфат, фосфотиоат, дезоксинуклеотидный инвертированный (связанный в направлении от 3' до 3') нуклеотид или представляют собой модифицированный рибонуклеотид или дезоксирибонуклеотид.

Молекулы малой интерферирующей РНК (siRNA) обычно содержат нуклеотидную последовательность, которая является идентичной участку мРНК-мишени длиной от приблизительно 15 до приблизительно 25 смежных нуклеотидов. В некоторых вариантах осуществления последовательность siRNA начинается с динуклеотида AA, характеризуется содержанием GC приблизительно 30-70% (приблизительно 30-60%, приблизительно 40-60% или приблизительно 45%-55%) и не характеризуется высоким процентным значением идентичности по отношению к любой нуклеотидной последовательности, отличной от мишени в геноме млекопитающего, в который предусмотрено ее введение, например, согласно определению с помощью стандартного поиска BLAST.

siRNA и shRNA характеризуются сходством с промежуточными соединениями в пути процессинга генов эндогенных микроРНК (miRNA) (Bartel, Cell 116:281-297, 2004). В некоторых вариантах осуществления siRNA могут функционировать в качестве miRNA и наоборот (Zeng et al., Mol Cell 9:1327-1333, 2002; Doench et al., Genes Dev 17:438-442, 2003). МикроРНК, такие как siRNA, используют RISC для подавления генов-мишеней, однако, в отличие от siRNA, большинство miRNA животных не расщепляют мРНК. Вместо этого miRNA уменьшают выход белка посредством подавления трансляции или удаления поли(A) и разрушения мРНК (Wu et al., Proc Natl Acad Sci USA 103:4034-4039, 2006). Известные сайты связывания miRNA расположены в пределах 3'-UTR мРНК; по-видимому, miRNA нацеливаются на сайты с практически абсолютной комплементарностью по отношению к нуклеотидам 2-8 с 5'-конца miRNA (Rajewsky, Nat Genet 38 Suppl:S8-13, 2006; Lim et al., Nature 433:769-773, 2005). Эта область известна как затравочная область. В некоторых вариантах осуществления экзогенные siRNA обеспечивают подавление мРНК, в составе которых имеется затравочная область, комплементарная siRNA (Birmingham et al., Nat. Methods 3:199-204, 2006). Несколько сайтов-мишеней в 3'-UTR обеспечивают более сильное подавление в некоторых вариантах осуществления (Doench et al., Genes Dev 17:438-442, 2003).

микроРНК

miRNA и другие малые интерферирующие нуклеиновые кислоты обычно регулируют экспрессию генов посредством расщепления/разрушения транскрипта РНК-мишени или репрессии трансляции матричной РНК-мишени (мРНК). miRNA могут в некоторых случаях экспрессироваться нативно, обычно в виде конечных нетранслируемых РНК-продуктов длиной 19-25. miRNA обычно проявляют свою активность посредством специфичных в отношении последовательности взаимодействий с 3'-нетранслируемыми областями (UTR) мРНК-мишеней. Эти эндогенно экспрессируемые miRNA могут образовывать шпилечные предшественники, которые впоследствии процессируются в miRNA-дуплекс, а затем в зрелую одонитевую молекулу miRNA. Эта зрелая miRNA обычно направляет многобелковый комплекс, miRISC, который идентифицирует области-

мишени 3' UTR мРНК-мишеней на основе их комплементарности зрелой miRNA. Применимые трансгенные продукты могут включать, например, miRNA или сайты связывания miRNA, которые регулируют экспрессию связанного полипептида. Неограничивающий перечень генов miRNA; продукты этих генов и их гомологов являются применимыми в качестве трансгенов или в качестве мишеней для малых интерферирующих нуклеиновых кислот (например, miRNA-губок, антисмысловых олигонуклеотидов), например, в способах, таких как способы, приведенные в US10300146, 22:25-25:48, включенном посредством ссылки. В некоторых вариантах осуществления один или несколько сайтов связывания для одной или нескольких вышеуказанных miRNA включены в трансген, например, трансген, доставляемый вектором на основе gAAV, например, для ингибирования экспрессии трансгена в одной или нескольких тканях животного, несущего трансген. В некоторых вариантах осуществления сайт связывания может быть выбран таким образом, чтобы обеспечивать контроль экспрессии трансгена тканеспецифическим образом. Например, сайты связывания для специфичного в отношении печени miR-122 могут быть включены в трансген для ингибирования экспрессии этого трансгена в печени. Дополнительные иллюстративные последовательности miRNA описаны, например, в патенте США № 10300146 (включенном в данный документ посредством ссылки во всей своей полноте). Однако для специфичного в отношении печени Gene Writing можно использовать сверхэкспрессию miR-122 вместо использования сайтов связывания для осуществления miR-122-специфичного разрушения. Эта miRNA положительно ассоциирована с дифференцировкой и созреванием печени, а также с усилением экспрессии генов, специфичных для печени. Таким образом, в некоторых вариантах осуществления кодирующая последовательность miR-122 может быть добавлена к компоненту системы Gene Writing для усиления направленной на печень терапии.

Ингибитор miR или ингибитор miRNA обычно представляет собой средство, которое блокирует экспрессию и/или процессинг miRNA. Примеры таких средств включают без ограничения антагонисты микроРНК, специфичные в отношении микроРНК антисмысловые последовательности, микроРНК-губки и специфичные к микроРНК олигонуклеотиды (двухнитевые, шпилечные, короткие олигонуклеотиды), которые ингибируют взаимодействие miRNA с комплексом Дроша. Ингибиторы микроРНК, например, miRNA-губки, могут экспрессироваться в клетках из трансгенов (например, как описано в работе Ebert, M. S. Nature Methods, Epub Aug. 12, 2007; включенной в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления микроРНК-губки или другие ингибиторы miR используют с AAV. МикроРНК-губки обычно специфически ингибируют miRNA посредством комплементарной гептамерной затравочной последовательности. В некоторых вариантах осуществления все семейство miRNA может быть подвергнуто сайленсингу с помощью одной последовательности губки. Другие способы сайленсинга функции miRNA (подавление мишеней miRNA) в клетках будут очевидны специалисту средней квалификации в данной области техники.

В некоторых вариантах осуществления miRNA, описанная в данном документе, содержит последовательность, приведенную в таблице 4 публикации согласно РСТ № WO2020014209, включенной в данный документ посредством ссылки. Также в данный документ посредством ссылки включен перечень иллюстративных последовательностей miRNA из WO2020014209.

В некоторых вариантах осуществления преимущественно подвергать сайленсингу один или несколько компонентов системы Gene Writing (например, мРНК, кодирующую полипептид Gene Writer, РНК-матрицу Gene Writer или последовательность гетерологичного объекта, экспрессируемую из генома после успешного осуществления способа Gene Writing) в части клеток. В некоторых вариантах осуществления преимущественно ограничивать экспрессию компонента системы Gene Writing выбранными типами клеток в ткани, представляющей интерес.

Например, известно, что в заданной ткани, например, в печени, макрофаги и иммунные клетки, например, клетки Купфера в печени, могут участвовать в поглощении средства доставки одного или нескольких компонентов системы Gene Writing. В некоторых вариантах осуществления по меньшей мере один сайт связывания для по меньшей мере одной miRNA, высоко экспрессируемой в макрофагах и иммунных клетках, например, клетках Купфера, включен в по меньшей мере один компонент системы Gene Writing, например, нуклеиновую кислоту, кодирующую полипептид Gene Writing, или трансген. В некоторых вариантах осуществления miRNA, которая нацелена на один или несколько сайтов связывания, приведена в таблице, ссылка на которую размещена в данном документе, например, miR-142, например, зрелая miRNA hsa-miR-142-5p или hsa-miR-142-3p.

В некоторых вариантах осуществления может быть предпочтительно снижение уровней Gene Writer и/или активности Gene Writer в клетках, в которых экспрессия Gene Writer или сверхэкспрессия трансгена может оказывать токсический эффект. Например, было показано, что доставка кассеты сверхэкспрессии трансгена в нейроны ганглиев задних корешков может привести в результате к развитию токсичности генной терапии (см. работу Hordeaux et al. Sci Transl Med 12(569):eaba9188 (2020), включенную в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления по меньшей мере один сайт связывания miRNA может быть включен в компонент системы Gene Writing, представляющий собой нуклеиновую кислоту, для снижения экспрессии данного компонента системы в нейроне, например, в нейроне ганглиев дорсальных корешков. В некоторых вариантах осуществления по меньшей мере один сайт связывания miRNA, включенный в компонент системы Gene Writing, представляющий собой нуклеиновую кислоту, для снижения экспрессии данного компонента системы в нейроне, представляет собой сайт связывания miR-182, например, зрелой miRNA hsa-miR-182-5p или hsa-miR-182-3p. В некоторых вариантах осуществления по меньшей мере один сайт связывания miRNA, включенный в компонент системы Gene Writing, представляющий собой нуклеиновую кислоту, для снижения экспрессии данного компонента системы в

нейроне, представляет собой сайт связывания miR-183, например, зрелой miRNA hsa-miR-183-5p или hsa-miR-183-3p. В некоторых вариантах осуществления для усиления рестрикции экспрессии одного или нескольких компонентов системы Gene Writing в отношении ткани или типа клеток, представляющих интерес, можно использовать комбинации сайтов связывания miRNA.

В приведенной ниже таблице A5 предусмотрены иллюстративные miRNA и соответствующие экспрессирующие клетки, например, miRNA, для которых в некоторых вариантах осуществления можно включать сайты связывания (комплементарные последовательности) в трансген или нуклеиновую кислоту полипептида, например, для снижения экспрессии в этой клетке, не являющейся мишенью.

Таблица A5. Иллюстративные miRNA из клеток и тканей, не являющихся мишенями

Тип клеток, подвергнутых сайленсингу	Название miRNA	Зрелая miRNA	Последовательность miRNA	SEQ ID NO
Клетки Купфера	miR-142	hsa-miR-142-5p	cauaaaguagaaagcacuacu	459
Клетки Купфера	miR-142	hsa-miR-142-3p	uguaguguuuccuacuuuauugga	460
Нейроны ганглиев задних корешков	miR-182	hsa-miR-182-5p	uuuggcaaugguagaacucacacu	461
Нейроны ганглиев задних корешков	miR-182	hsa-miR-182-3p	ugguucuagacuugccaacua	462
Нейроны ганглиев задних корешков	miR-183	hsa-miR-183-5p	uauggcacugguagaauucacu	463
Нейроны ганглиев задних корешков	miR-183	hsa-miR-183-3p	gugaauuaccgaagggccauaa	464
Гепатоциты	miR-122	hsa-miR-122-5p	uggagugugacaauagguguuuug	465
Гепатоциты	miR-122	hsa-miR-122-3p	aacgccauuaucacacuaaaaua	466

Молекулы для RNAi быстро конструируют и получают с помощью технологий, известных в данной области техники. Кроме того, существуют вычислительные средства, которые повышают вероятность нахождения эффективных и специфических мотивов последовательности (Pei et al. 2006, Reynolds et al. 2004, Khvorova et al. 2003, Schwarz et al. 2003, Ui-Tei et al. 2004, Heale et al. 2005, Chalk et al. 2004, Amarzguioui et al. 2004).

Молекула для RNAi модулирует экспрессию РНК, кодируемой геном. Поскольку несколько генов могут характеризоваться некоторой степенью гомологии

последовательностей друг с другом, в некоторых вариантах осуществления молекула для RNAi может быть сконструирована с обеспечением нацеливания на класс генов с достаточной гомологией последовательностей. В некоторых вариантах осуществления молекула для RNAi может содержать последовательность, которая характеризуется комплементарностью с последовательностями, которые являются общими для различных генов-мишеней или являются уникальными для специфического гена-мишени. В некоторых вариантах осуществления молекула для RNAi может быть сконструирована с обеспечением нацеливания на консервативные области последовательности РНК, которая является гомологичной для нескольких генов, тем самым обеспечивая нацеливание на несколько генов в семействе генов (например, различные изоформы генов, сплайс-варианты, мутантные гены и т. д.). В некоторых вариантах осуществления молекула для RNAi может быть сконструирована с обеспечением нацеливания на последовательность, которая является уникальной для конкретной последовательности РНК одного гена.

В некоторых вариантах осуществления молекулу для RNAi связывают с полимером для доставки, например, с помощью физиологически неустойчивой связи или линкера. Физиологически неустойчивый линкер выбирают так, чтобы он подвергался химической трансформации (например, расщеплению), находясь в определенных физиологических условиях (например, дисульфидная связь расщепляется в восстанавливающей среде цитоплазмы клетки). Высвобождение молекулы из полимера путем расщепления физиологически неустойчивой связи обеспечивает взаимодействие молекулы с подходящими клеточными компонентами для обеспечения ее активности.

Конъюгат молекула для RNAi-полимер может быть образован путем образования ковалентной связи молекулы с полимером. Полимер полимеризуют или модифицируют так, чтобы он содержал реакционноспособную группу А. Молекулу для RNAi также полимеризуют или модифицируют так, чтобы она содержала реакционноспособную группу В. Реакционноспособные группы А и В выбирают таким образом, чтобы их можно было связывать посредством обратимой ковалентной связи с использованием способов, известных в данной области техники.

Конъюгацию молекулы для RNAi с полимером можно выполнять в присутствии избытка полимера. Поскольку молекула для RNAi и полимер могут иметь противоположные заряды во время конъюгации, наличие избытка полимера может приводить к снижению или устранению агрегации конъюгата. В качестве альтернативы можно применять избыток полимера-носителя, такого как поликатион. Избыток полимера может быть удален из конъюгированного полимера перед введением конъюгата животному или в культуру клеток. В качестве альтернативы избыток полимера может быть введен совместно с конъюгатом животному или в культуру клеток.

Например, молекула ингибирующей РНК предусматривает короткую интерферирующую РНК, короткую шпилечную РНК и/или микроРНК, которые нацелены на экспрессию гена в случае гена, участвующего в ответе хозяина. siRNA представляет собой двухнитевую молекулу РНК, длина которой, как правило, составляет приблизительно

19-25 пар оснований. shRNA представляет собой молекулу РНК, содержащую шпилечный изгиб, которая обеспечивает снижение экспрессии гена-мишени, например, посредством RNAi. shRNA могут доставляться в клетки в виде плазмид, например, вирусных или бактериальных векторов, например, с помощью трансфекции, электропорации или трансдукции). МикроРНК представляет собой некодирующую молекулу РНК, длина которой, как правило, составляет приблизительно 22 нуклеотида. miRNA, как правило, связываются с сайтами-мишенями на молекулах мРНК и приводят к сайленсингу мРНК, например, вызывая расщепление мРНК, дестабилизацию мРНК или подавление трансляции мРНК. В вариантах осуществления молекула ингибирующей РНК обеспечивает снижение уровня и/или активности отрицательного регулятора функции или положительного регулятора функции. В других вариантах осуществления молекула ингибирующей РНК обеспечивает снижение уровня и/или активности ингибитора положительного регулятора функции.

Молекула ингибирующей РНК может быть модифицированной, например, содержать модифицированные нуклеотиды, например, 2'-фтор, 2'-О-метил, 2'-дезоксидезокси, незамкнутую нуклеиновую кислоту, 2'-гидроксид, фосфотиоат, 2'-тиоуридин, 4'-тиоуридин, 2'-дезоксидеоксиуридин. Без ограничения какой-либо теорией, полагают, что определенная модификация может обеспечивать повышение устойчивости к нуклеазам и/или стабильности в сыворотке крови или снижение иммуногенности.

В некоторых вариантах осуществления молекула ингибирующей РНК обеспечивает снижение уровня и/или активности или функции фактора, кодируемого геном, участвующим в ответе хозяина. В вариантах осуществления молекула ингибирующей РНК ингибирует экспрессию фактора, кодируемого геном, участвующим в ответе хозяина. В других вариантах осуществления молекула ингибирующей РНК обеспечивает повышение уровня разрушения фактора, кодируемого геном, участвующим в ответе хозяина, и/или снижение времени полужизни фактора, кодируемого геном, участвующим в ответе хозяина. Молекула ингибирующей РНК может быть химически синтезирована или транскрибирована *in vitro*.

Получение и применение ингибирующих терапевтических средств на основе некодирующей РНК, такой как рибозимы, РНКазы Р, siRNA и miRNA, дополнительно описано в Sioud, RNA Therapeutics: Function, Design, and Delivery (Methods in Molecular Biology). Humana Press 2010; and Kaczmarek et al. 2017. Advances in the delivery of RNA therapeutics: from concept to clinical reality. Genome Medicine 9:60.

CRISPR

Систему на основе CRISPR можно использовать для ингибирования экспрессии гена, участвующего в ответе хозяина, например, для инактивации гена, участвующего в ответе хозяина, описанного в данном документе, или для снижения или ингибирования экспрессии гена, участвующего в ответе хозяина (например, посредством генетического или эпигенетического редактирования). В определенных вариантах осуществления система на основе CRISPR в качестве ингибитора содержит отрицательный эффектор и одну или

несколько направляющих РНК, которые нацелены на ген, участвующий в ответе хозяина.

В системах на основе CRISPR используются РНК-направляемые нуклеазы, называемые CRISPR-ассоциированными эндонуклеазами или эндонуклеазами "Cas" (например, Cas9 или Cpf1), для расщепления ДНК. В типичной системе CRISPR/Cas эндонуклеаза направляется на нуклеотидную последовательность-мишень (например, сайт в геноме, который подлежит редактированию последовательности) с помощью специфичных в отношении последовательности некодирующих "направляющих РНК", которые нацеливаются на одно- или двухнитевые последовательности ДНК. Были идентифицированы три класса (I-III) систем на основе CRISPR. В системах на основе CRISPR класса II используется одна эндонуклеаза Cas (вместо нескольких белков Cas). Одна система на основе CRISPR класса II содержит эндонуклеазу Cas типа II, такую как Cas9, CRISPR-РНК ("crRNA") и транс-активирующую crRNA ("tracrRNA"). crRNA содержит "направляющую РНК", как правило, последовательность РНК длиной приблизительно 20 нуклеотидов, которая соответствует последовательности ДНК-мишени. tracrRNA также содержит область, которая связывается с crRNA с образованием частично двухнитевой структуры, которая расщепляется РНКазой III, что приводит к образованию гибрида crRNA/tracrRNA. Затем гибрид crRNA/tracrRNA направляет эндонуклеазу Cas9 для распознавания и расщепления последовательности ДНК-мишени. Последовательность ДНК-мишени обычно должна находиться вблизи "мотива, прилегающего к протоспейсеру" ("PAM"), который является специфическим для рассматриваемой эндонуклеазы Cas; однако, последовательности PAM встречаются по всему рассматриваемому геному. Эндонуклеазы CRISPR, идентифицированные у различных видов прокариот, имеют уникальные требования к последовательности PAM; примеры последовательностей PAM включают 5'-NGG (*Streptococcus pyogenes*), 5'-NNAGAA (CRISPR1 *Streptococcus thermophilus*), 5'-NGGNG (CRISPR3 *Streptococcus thermophilus*) и 5'-NNNGATT (*Neisseria meningitidis*). Некоторые эндонуклеазы, например, эндонуклеазы Cas9, ассоциированы с G-богатыми сайтами PAM, например, с 5'-NGG, и осуществляют расщепление ДНК-мишени с образованием "тупых" концов в местоположении на 3 нуклеотида выше сайта PAM (в направлении 5'-конца от него). Другая система на основе CRISPR класса II содержит эндонуклеазу Cpf1 типа V, которая характеризуется меньшим размером, чем Cas9; примеры включают AsCpf1 (из *Acidaminococcus* sp.) и LbCpf1 (из *Lachnospiraceae* sp.). Cpf1-ассоциированные комплексы CRISPR подвергаются процессингу в зрелые crRNA без потребности в tracrRNA; другими словами, в случае системы с использованием Cpf1 для расщепления последовательности ДНК-мишени может использоваться только нуклеаза Cpf1 и crRNA. Эндонуклеазы Cpf1, как правило, ассоциированы с T-богатыми сайтами PAM, например, с 5'-TTN. Cpf1 также может распознавать PAM-мотив 5'-CTA. Cpf1 расщепляет ДНК-мишень посредством введения смещенного или ступенчатого двухнитевого разрыва с образованием выступающего 5'-конца длиной 4 или 5 нуклеотидов, например, посредством расщепления ДНК-мишени с образованием смещенного или ступенчатого разреза, охватывающего 5 нуклеотидов, расположенного на 18 нуклеотидов

ниже сайта PAM (в направлении 3'-конца от него) на кодирующей нити и на 23 нуклеотида ниже сайта PAM на комплементарной нити; при этом выступающий конец длиной 5 нуклеотидов, который образуется в результате такого смещенного расщепления, обеспечивает возможность более точного редактирования генома за счет вставки ДНК посредством гомологичной рекомбинации по сравнению со вставкой в ДНК при расщеплении с образованием тупых концов. См., например, Zetsche et al. (2015) *Cell*, 163:759-771.

Для целей редактирования генов комплексы CRISPR могут быть сконструированы таким образом, чтобы они содержали одну или несколько последовательностей направляющих РНК, соответствующих необходимой последовательности ДНК-мишени; см., например, Cong et al. (2013) *Science*, 339:819-823; Ran et al. (2013) *Nature Protocols*, 8:2281-2308. Чтобы произошло расщепление ДНК Cas9, как правило, используется последовательность gRNA из по меньшей мере приблизительно 16 или 17 нуклеотидов; при этом для достижения поддающегося обнаружению уровня расщепления ДНК в случае Cpf1, как правило, используется последовательность gRNA из по меньшей мере приблизительно 16 нуклеотидов. На практике последовательности направляющей РНК обычно конструируют таким образом, чтобы их длина составляла 17-24 нуклеотида (например, 19, 20 или 21 нуклеотид) и они были комплементарны целевой последовательности гена или нуклеиновой кислоты. Специализированные генераторы и алгоритмы для создания gRNA являются коммерчески доступными для применения в конструировании эффективных направляющих РНК. Редактирования гена также достигали с использованием химерной “одиночной направляющей РНК” (“sgRNA”), одиночной молекулы РНК, которая содержит как область tracrRNA (например, которая связывает нуклеазу), так и по меньшей мере одну область crRNA (например, которая направляет нуклеазу к последовательности, предназначенной для редактирования). sgRNA, как правило, представляют собой сконструированные молекулы, которые имитируют встречающийся в природе комплекс crRNA-tracrRNA. Также было продемонстрировано, что химически модифицированные sgRNA являются эффективными в редактировании генома; см., например, Hendel et al. (2015) *Nature Biotechnol.* 985-991.

В то время как Cas9 дикого типа, как правило, приводит к образованию двухнитевых разрывов (DSB) в конкретных последовательностях ДНК, которые служат мишенью для gRNA, доступен целый ряд эндонуклеаз CRISPR с модифицированными функциональными свойствами, например, “никазный” вариант Cas9 обеспечивает образование одонитевого разрыва; каталитически неактивная Cas9 (“dCas9”) препятствует транскрипции путем создания стерического препятствия, и, как правило, не разрезает ДНК-мишень, или не разрезает ее на уровне, который, является поддающимся обнаружению. dCas9 может быть дополнительно слита с эффектором для обеспечения репрессии (CRISPRi) или активации (CRISPRa) экспрессии гена-мишени. Например, Cas9 может быть слита с транскрипционным репрессором (например, доменом KRAB) или транскрипционным активатором (например, слиянием dCas9-VP64). Каталитически неактивная Cas9 (dCas9),

слитая с нуклеазой FokI ("dCas9-FokI"), может использоваться для образования DSB в последовательностях-мишенях, гомологичных двум gRNA. См., например, данные по многочисленным плазмидам, кодирующие CRISPR/Cas9, которые раскрыты в депозитории Addgene (Addgene, 75 Sidney St., Suite 550A, Кембридж, Массачусетс, 02139; addgene.org/crispr/) и находятся в открытом доступе. "Двойная никаза" Cas9, которая вводит два отдельных двухнитевых разрыва, каждый из которых направляется отдельной направляющей РНК, описана как осуществляющая более точное редактирование генома в Ran et al. (2013) Cell, 154:1380-1389. В одном варианте осуществления ингибитор, раскрытый в данном документе, предусматривает систему на основе CRISPRi для снижения экспрессии гена, участвующего в ответе хозяина.

Технология CRISPR для редактирования генов эукариот раскрыта в публикациях заявок на патент США 2016/0138008A1 и US2015/0344912A1, а также в патентах США 8697359, 8771945, 8945839, 8999641, 8993233, 8895308, 8865406, 8889418, 8871445, 8889356, 8932814, 8795965 и 8906616. Эндонуклеаза Cpf1 и соответствующие направляющие РНК и сайты РАМ раскрыты в публикации заявки на патент США 2016/0208243 A1. Технология CRISPR для создания дисфункции mtDNA в митохондриальном геноме посредством системы CRISPR/Cas9 раскрыта в Jo, A., et al., BioMed Res. Int'l, vol 2015, ID статьи 305716, 10 страниц, <http://dx.doi.org/10.1155/2015/305716>. Совместная доставка Cas9 и sgRNA с помощью наночастиц раскрыта в Mout, R., et al., ACS Nano, Jan 31, 2017, ID статьи doi: 10.1021/acsnano.6b07600.

В некоторых вариантах осуществления композиция, содержащая gRNA и нацеленную нуклеазу, например, Cas9, например, Cas9 дикого типа, никазу Cas9 (например, Cas9 D10A), неактивную форму Cas9 (dCas9), eSpCas9, Cpf1, C2C1 или C2C3 или нуклеиновую кислоту, кодирующую такую нуклеазу, используется для модулирования экспрессии генов. Выбор нуклеазы и gRNA определяется тем, представляет ли собой мутация-мишень делецию, замену или добавление нуклеотидов, например, делецию, замену или добавление нуклеотидов в последовательность-мишень. Продукты слияния, в которых каталитически неактивная эндонуклеаза, например, неактивная форма Cas9 (dCas9, например, D10A; H840A), связана со всем или частью (например, биологически активной частью) (одного или нескольких) эффекторных доменов, образуют химерные белки, которые могут быть связаны с полипептидом для направления композиции к специфическим сайтам ДНК с помощью одной или нескольких последовательностей РНК (sgRNA) для модулирования активности и/или экспрессии одной или нескольких последовательностей нуклеиновых кислот-мишеней (например, для метилирования или деметилирования последовательности ДНК).

Обнаружение РНК

В некоторых вариантах осуществления модулятор ответа хозяина ингибирует один или несколько белков, участвующих в обнаружении РНК и ответе на нее, например, TLR3, TLR4, TLR7, TLR8, MyD88, TRIF, IKK, NF- κ B, IRF3, IRF7, IFN- α , IFN- β , TNF α , IL-6, IL-

12, JAK-1, TYK-2, STAT1, STAT2, IRF-9, PKR, OAS, ADAR, RIG-I, MDA5, LGP2, MAVS, NLRP3, NOD2 или каспазу 1 или любую их комбинацию.

Без ограничения какой-либо теорией, в некоторых вариантах осуществления активация TLR4 блокирует трансляцию мРНК без снижения поглощения LNP клетками. Ингибирование TLR4 или эффекторной протеинкиназы R, находящейся дальше в его пути передачи сигнала, может обеспечить улучшение экспрессии мРНК, доставляемой в клетки в свободном виде или в LNP (Lokugamage et al. Adv Materials 2019). В некоторых вариантах осуществления ингибитор TLR4 или эффектора, находящегося дальше в пути передачи сигнала, например, протеинкиназы R, используется для повышения эффективности системы Gene Writing. В некоторых вариантах осуществления модулятор ответа хозяина, который представляет собой ингибитор одного или нескольких белков, участвующих в обнаружении РНК и ответе (например, TLR4), обеспечивает повышение экспрессии полипептида GeneWriter из мРНК, например, обеспечивает повышение уровней белка Gene Writer до уровней, являющихся на по меньшей мере 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200% или 300% выше по сравнению с таковым в подобной в ином отношении клетке, не приводимой в контакт с модулятором ответа хозяина.

Эпигенетические модификаторы

В некоторых вариантах осуществления средство, которое способствует активности полипептида Gene Writer (например, способствует вставке последовательности гетерологичного объекта с помощью полипептида Gene Writer), представляет собой эпигенетический модификатор. Без ограничения какой-либо теорией, в некоторых вариантах осуществления структура хроматина в сайте вставки влияет на эффективность вставки, например, открытый хроматин может быть более подходящим для вставки, чем гетерохроматин. Соответственно, активность Gene Writer можно повысить путем совместного введения эпигенетического модификатора.

В некоторых вариантах осуществления эпигенетический модификатор действует специфически на сайт-мишень. В некоторых вариантах осуществления эпигенетический модификатор действует на совокупность сайтов в геноме (например, на геном в целом), причем один из совокупности сайтов представляет собой сайт-мишень. Эпигенетический модификатор может предусматривать, например, модифицирующий хроматин фермент (или нуклеиновую кислоту, кодирующую его), ингибитор модифицирующего эндогенный хроматин фермента (например, ингибитор, представляющий собой нуклеиновую кислоту) или малую молекулу (например, низкомолекулярный ингибитор модифицирующего эндогенный хроматин фермента).

В некоторых вариантах осуществления эпигенетический модификатор, который способствует транспозиции, представляет собой ингибитор HDAC или ингибитор гистонметилтрансферазы. Эти ингибиторы действуют на деацетилазы гистонов и метилтрансферазы гистонов соответственно, блокируя их активность и обеспечивая расслабление хроматина, что может улучшить доступность ДНК-мишени для систем Gene Writing. В некоторых вариантах осуществления ингибиторы HDAC, ингибиторы

метилтрансфераз гистонов или комбинация первого и второго могут предоставляться вместе с системой Gene Writing для повышения эффективности интеграции. Ингибиторы HDAC и ингибитор метилтрансферазы гистонов описаны в документе WO 2020077357A1, который включен в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления ингибитор HDAC представляет собой ингибитор всех видов HDAC, ингибитор HDAC класса I, ингибитор HDAC класса II или ингибитор HDAC класса I и II. Неограничивающие примеры ингибиторов всех видов HDAC включают трихостатин А (TSA), вориноостат, CAY10433 (нацелен на класс I и II) или фенилбутират натрия (нацелен на класс I и IIa). Неограничивающие примеры ингибиторов HDAC класса I (нацеливающих на HDAC 1, 2, 3 или 8) включают MS-275, CAY10398 или энтиноостат. Неограничивающие примеры ингибиторов HDAC класса II (нацеливающих на HDAC 4, 5, 6, 7, 9 или 10) включают MC-1568, скриптайд или CAY10603. Вальпроевая кислота (VPA) может ингибировать несколько деацетилаз гистонов, относящихся как к классу I, так и к классу II.

Ингибитор метилтрансферазы гистонов может представлять собой селективный ингибитор метилтрансфераз гистонов G9a/GLP, который метилирует лизин 9 гистона 3 (H3K9). Неограничивающие примеры ингибиторов G9a/GLP включают BIX01294, UNC0642, A-366, UNC0224, UNC0631, UNC0646, BRD4770 или UNC0631. Неограничивающие примеры лизинметилтрансфераз гистонов включают хетоцин, EPZ005687, EPZ6438, GSK126, GSK343, EII, UNC199, EPZ004777, EPZ5676, LLY-507, AZ505 или A-893. Ингибитор метилтрансферазы гистонов может представлять собой 2-циклогексил-N-(1-изопропилпиперидин-4-ил)-6-метокси-7-(3-(пирролидин-1-ил)пропокси)хиназолин-4-амин (ETNC0638), BIX01294, ETNC0642, A-366, UNC0224, UNC0631, UNC0646, BRD4770, UNC0631, хетоцин, EPZ005687, EPZ6438, GSK126, GSK343, EII, UNC199, EPZ004777, EPZ5676, LLY-507, AZ505 или A-893. В некоторых вариантах осуществления ингибитор метилтрансферазы гистонов представляет собой UNC0638.

В некоторых вариантах осуществления эпигенетический модификатор содержит нацеливающий фрагмент, который направляет его к сайту-мишени. В некоторых вариантах осуществления нацеливающий фрагмент предусматривает ДНК-связывающий домен, например, домен цинкового пальца, домен эффектора TAL или каталитически неактивный белок Cas.

Полипептиды Gene Writer

Полипептид Gene Writer, как правило, представляет собой по сути автономную белковую систему, способную интегрировать нуклеиновую кислоту-матрицу в молекулу ДНК-мишени (например, в клетке-хозяине млекопитающего, как, например, в молекулу геномной ДНК в клетке-хозяине), по сути без использования аппарата хозяина.

Gene Writer, подходящие для использования в композициях и способах, описанных в данном документе, включают, например, ретротранспозазы, ДНК-транспозазы и рекомбиназы (например, сериновые рекомбиназы и тирозиновые рекомбиназы).

Иллюстративные полипептиды Gene Writer, и системы, их содержащие, и способы их применения описаны, например, в PCT/US19/48607, поданной 28 августа 2019 г.; 62/876165, поданной 19 июля 2019 г.; 62/939525, поданной 22 ноября 2019 г.; и 62/967934, поданной 30 января 2020 г., каждая из которых включена в данный документ посредством ссылки, включая последовательности аминокислот и нуклеиновых кислот, приведенные в перечисленных заявках.

Например, таблица 3 документа PCT/US19/48607 включена в данный документ посредством ссылки во всей своей полноте. В некоторых вариантах осуществления полипептид Gene Writer содержит аминокислотную последовательность из столбца 8 таблицы 3 из заявки PCT/US19/48607, или последовательность любого домена из нее (например, ДНК-связывающего домена, РНК-связывающего домена, домена эндонуклеазы или домена RT), или последовательность, на по меньшей мере 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% идентичную таковой. В некоторых вариантах осуществления РНК-матрица содержит последовательность из таблицы 3 из заявки PCT/US19/48607 (например, одну или обе из 5'-нетранслируемой области из столбца 6 и 3'-нетранслируемой области из столбца 7) или последовательность, на по меньшей мере 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% идентичную ей.

Иллюстративные полипептиды GeneWriter и последовательности доменов RT также описаны, например, в предварительной заявке на патент США № 63/035627, поданной 5 июня 2020 г., например, в таблице 1, таблице 3, таблице 30 и таблице 31; полная заявка включена в данный документ посредством ссылки, включая указанные последовательности и таблицы. Соответственно, полипептид GeneWriter, описанный в данном документе, может содержать аминокислотную последовательность в соответствии с любой из таблиц, упомянутых в этом абзаце, или последовательность домена из нее (например, домена RT, ДНК-связывающего домена, РНК-связывающего домена или домена эндонуклеазы), или функциональный фрагмент или вариант любого из представленного выше, или аминокислотную последовательность, на по меньшей мере 70%, 80%, 85%, 90%, 95% или 99% идентичную таковой.

В некоторых вариантах осуществления полипептид Gene Writer содержит один или несколько доменов, которые в совокупности обеспечивают следующее: 1) связывание нуклеиновой кислоты-матрицы, 2) связывание молекулы ДНК-мишени и 3) обеспечивают интеграцию по меньшей мере части нуклеиновой кислоты-матрицы в ДНК-мишень. В некоторых вариантах осуществления полипептид Gene Writer представляет собой встречающийся в природе полипептид. В некоторых вариантах осуществления полипептид Gene Writer представляет собой сконструированный полипептид, например, содержащий одну или несколько аминокислотных замен во встречающейся в природе последовательности. В некоторых вариантах осуществления полипептиды Gene Writer содержат два или более домена, которые являются взаимно гетерологичными, что достигается, например, с помощью слияния (или другого варианта конъюгирования) гетерологичных доменов, являющихся в ином случае доменами дикого типа, а также

вариантов слияния модифицированных доменов, например, с помощью замены или слияния гетерологичного субдомена или другого замещенного домена. Например, в некоторых вариантах осуществления справедливо одно или несколько из следующего: домен RT является гетерологичным относительно DBD; DBD является гетерологичным относительно домена эндонуклеазы; или домен RT является гетерологичным относительно домена эндонуклеазы.

В некоторых вариантах осуществления система Gene Writer способна производить в сайте-мишени замену, состоящую из по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90 или 100 или более нуклеотидов. В некоторых вариантах осуществления замена представляет собой мутацию типа транзиции. В некоторых вариантах осуществления замена представляет собой мутацию типа трансверсии. В некоторых вариантах осуществления при замене происходит превращение аденина в тимин, аденина в гуанин, аденина в цитозин, гуанина в тимин, гуанина в цитозин, гуанина в аденин, тимина в цитозин, тимина в аденин, тимина в гуанин, цитозина в аденин, цитозина в гуанин или цитозина в тимин.

В некоторых вариантах осуществления вставка, делеция, замена или их комбинация обеспечивают повышение или снижение экспрессии (например, транскрипции или трансляции) гена. В некоторых вариантах осуществления вставка, делеция, замена или их комбинация обеспечивают повышение или снижение экспрессии (например, транскрипции или трансляции) гена посредством изменения, добавления или делеции последовательностей в промоторе или энхансере, например, последовательностей, которые связывают факторы транскрипции. В некоторых вариантах осуществления вставка, делеция, замена или их комбинация обеспечивают изменение трансляции гена (например, изменение аминокислотной последовательности), вставку или делецию стартового или стоп-кодона, изменение или восстановление рамки трансляции гена. В некоторых вариантах осуществления вставка, делеция, замена или их комбинация изменяют сплайсинг гена, например, посредством вставки, делеции или изменения акцепторного или донорного сайта сплайсинга. В некоторых вариантах осуществления вставка, делеция, замена или их комбинация изменяют период полужизни транскрипта или белка. В некоторых вариантах осуществления вставка, делеция, замена или их комбинация изменяют локализацию белка в клетке (например, обеспечивают перемещение из цитоплазмы в митохондрии, из цитоплазмы во внеклеточное пространство (например, добавляют метку секреции)). В некоторых вариантах осуществления вставка, делеция, замена или их комбинация изменяют (например, улучшают) фолдинг белка (например, для предупреждения накопления ошибочно свернутых белков). В некоторых вариантах осуществления вставка, делеция, замена или их комбинация изменяют, повышают, снижают активность гена, например, белка, кодируемого данным геном.

Перенацеливание (например, полипептида или молекулы нуклеиновой кислоты Gene Writer или системы, описанной в данном документе) обычно предусматривает: (i) направление полипептида на связывание и расщепление в сайте-мишени и/или (ii)

конструирование РНК-матрицы, комплементарной последовательности-мишени. В некоторых вариантах осуществления РНК-матрица характеризуется комплементарностью в отношении последовательности-мишени, находящейся в направлении 5'-конца от ника в первой нити, например, таким образом, что 3'-конец РНК-матрицы отжигается, и 5'-конец сайта-мишени служит в качестве праймера, например, для TPRT. В некоторых вариантах осуществления домен эндонуклеазы полипептида и 5'-конец РНК-матрицы также модифицированы, как описано.

В некоторых вариантах осуществления полипептид Gene Writer содержит модификацию ДНК-связывающего домена, например, по сравнению с полипептидом дикого типа. В некоторых вариантах осуществления ДНК-связывающий домен предусматривает добавление, делецию, замену или модификацию аминокислотной последовательности исходного ДНК-связывающего домена. В некоторых вариантах осуществления ДНК-связывающий домен модифицирован таким образом, чтобы он содержал гетерологичный функциональный домен, который специфически связывается с последовательностью нуклеиновой кислоты-мишени (например, ДНК), представляющей интерес. В некоторых вариантах осуществления функциональный домен заменяет по меньшей мере часть (например, полностью) предшествующего ДНК-связывающего домена полипептида. В некоторых вариантах осуществления функциональный домен содержит цинковый палец (например, цинковый палец, который специфически связывается с последовательностью нуклеиновой кислоты-мишени (например, ДНК), представляющей интерес. В некоторых вариантах осуществления функциональный домен содержит домен Cas (например, домен Cas, который специфически связывается с последовательностью нуклеиновой кислоты-мишени (например, ДНК), представляющей интерес. В вариантах осуществления домен Cas предусматривает Cas9 или его мутантную форму или вариант (например, как описано в данном документе). В вариантах осуществления домен Cas ассоциирован с направляющей РНК (gRNA), например, как описано в данном документе. В вариантах осуществления домен Cas направляется к последовательности нуклеиновой кислоты-мишени (например, ДНК), представляющей интерес, посредством gRNA. В вариантах осуществления домен Cas кодируется той же молекулой нуклеиновой кислоты (например, РНК), что и gRNA. В вариантах осуществления домен Cas кодируется в молекуле нуклеиновой кислоты (например, РНК), отличной от gRNA.

В некоторых вариантах осуществления полипептид Gene Writer содержит модификацию домена эндонуклеазы, например, по сравнению с полипептидом дикого типа. В некоторых вариантах осуществления домен эндонуклеазы предусматривает добавление, делецию, замену или модификацию аминокислотной последовательности исходного домена эндонуклеазы. В некоторых вариантах осуществления домен эндонуклеазы модифицирован таким образом, чтобы он содержал гетерологичный функциональный домен, который специфически связывается с последовательностью нуклеиновой кислоты-мишени (например, ДНК), представляющей интерес, и/или индуцирует ее расщепление эндонуклеазой. В некоторых вариантах осуществления домен эндонуклеазы содержит

цинковый палец. В некоторых вариантах осуществления домен эндонуклеазы предусматривает домен Cas (например, Cas9 или его мутантную форму или вариант). В вариантах осуществления домен эндонуклеазы, содержащий домен Cas, ассоциирован с направляющей РНК (gRNA), например, как описано в данном документе. В некоторых вариантах осуществления домен эндонуклеазы модифицирован таким образом, чтобы он содержал функциональный домен, который не нацелен на последовательность конкретной нуклеиновой кислоты-мишени (например, ДНК). В вариантах осуществления домен эндонуклеазы предусматривает домен FokI.

В некоторых вариантах осуществления домен обратной транскриптазы (RT) демонстрирует повышенную строгость инициации праймируемой мишенью обратной транскрипции (TPRT), например, по сравнению с эндогенным доменом RT. В некоторых вариантах осуществления домен RT иницирует TPRT, если 3 нуклеотида в сайте-мишени непосредственно перед ником в первой нити, например, при праймировании РНК-матрицы посредством геномной ДНК, характеризуются по меньшей мере 66% или 100% комплементарностью с 3 гомологичным нуклеотидам в РНК-матрице. В некоторых вариантах осуществления домен RT иницирует TPRT при наличии несовпадения менее 5 нуклеотидов (например, несовпадения менее 1, 2, 3, 4 или 5 нуклеотидов) между гомологичной РНК-матрицей и ДНК-мишенью, праймирующей обратную транскрипцию. В некоторых вариантах осуществления домен RT модифицирован таким образом, что строгость в отношении несовпадений при праймировании реакции TPRT увеличивается, например, где домен RT не допускает каких-либо несовпадений или допускает меньшее количество несовпадений в области праймирования по сравнению с доменом RT дикого типа (например, немодифицированным). В некоторых вариантах осуществления домен RT содержит домен RT HIV-1. В вариантах осуществления домен RT HIV-1 иницирует более низкие уровни синтеза даже при несовпадении трех нуклеотидов по сравнению с альтернативным доменом RT (например, как описано в работе Jamburuthugoda and Eickbush J Mol Biol 407(5):661-672 (2011), включенной в данный документ посредством ссылки во всей своей полноте).

В некоторых вариантах осуществления полипептид Gene Writing содержит домен эндонуклеазы, содержащий нуклеазу Cas9, например, Cas9 H840A. В вариантах осуществления Cas9 H840A характеризуется следующей аминокислотной последовательностью:

<i>Нуклеаза</i>	<i>Cas9</i>	<i>(H840A):</i>
DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA		
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLAESFLVEEDKKHERHPIFGN		
IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV		
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLI		
ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAIL		
LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG		
YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI		

LRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV
 DKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII
 KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTIAHLFDDKVMKQLKRRRYTGW
 GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS
 LHEHIANLAGSPAIIKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR
 ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD
 VDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQR
 KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
 VITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDY
 K VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
 WDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV
 KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP
 EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENII
 HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQLGGD
 (SEQ ID NO: 1).

В некоторых вариантах осуществления полипептид Gene Writing содержит домен RT из обратной транскриптазы ретровируса, например, RT M-MLV дикого типа, например, содержащий следующую последовательность:

M-MLV

(*WT*):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKATSTPVS
 IKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVN
 KRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPENGI
 SGQLTWTRLPPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQ
 GTRALLQTLGNLGYRASAKKAQICQKQVKYLYGILLKEGQRWLTEARKETVMGQPTPK
 TPRQLREFLG TAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTA
 PALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLR
 MVAIAI AVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLD TD
 RVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSS
 LLQEGQRKAGAAVTTETEVIAKALPAGTSAQRAELIALTQALKMAEGKKLNVTDSR
 YAFATAH IHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQKGHSAEA
 RGNRMADQAARKAAITETPDTSTLLI (SEQ ID NO: 2).

В некоторых вариантах осуществления полипептид Gene Writing содержит домен RT из обратной транскриптазы ретровируса, например, RT M-MLV, например, содержащий следующую последовательность:

TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKAT
 STPVS IKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDL
 REVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRD
 PEMGISGQLTWTRLPPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSEL

DCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLG YLLKEGQRWLTEARKETVMG
QPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGT LFNWGPDQQKAYQEIKQ
ALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWP
PCLRMVAIAI AVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALL
LDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT
DGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLVY
TDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQKGH
SAEARGNRMADQAARKAAITETPDTSTLL (SEQ ID NO: 3)

В некоторых вариантах осуществления полипептид Gene Writing содержит домен RT из обратной транскриптазы ретровируса, содержащий последовательность аминокислот 659-1329 NP_057933. В вариантах осуществления полипептид Gene Writing дополнительно содержит одну дополнительную аминокислоту на N-конце последовательности аминокислот 659-1329 из NP_057933, например, как показано ниже:

TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKAT
STPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQD
LREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAF
EWRDPENGISGQLTWTRL PQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDL
LLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLG YLLKEGQRWL T
EARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGT LFNWGPD
QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK
KLDPVAAGWPPCLRMVAIAI AVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSN
ARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP
LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALK
MAEGKKLVYTDDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLS
IIHCPGHQKGHSAEARGNRMADQAARKAA (SEQ ID NO: 4)

Коровая последовательность RT (жирный шрифт), аннотированная, как указано выше.

RNAseH (шрифт с подчеркиванием), как указано выше.

В вариантах осуществления полипептид Gene Writing дополнительно содержит одну дополнительную аминокислоту на C-конце последовательности аминокислот 659-1329 из NP_057933. В вариантах осуществления полипептид Gene Writing содержит домен РНКазы H1 (например, аминокислоты 1178-1318 из NP_057933).

В некоторых вариантах осуществления домен ретровирусной обратной транскриптазы, например, RT М-MLV, может содержать одну или несколько мутаций относительно последовательности дикого типа, которые могут обеспечивать улучшение характеристик RT, например, термостабильности, процессивности и/или связывания с матрицей. В некоторых вариантах осуществления домен RT М-MLV содержит, по сравнению с приведенной выше последовательностью М-MLV (WT), одну или несколько мутаций, например, выбранных из D200N, L603W, T330P, T306K, W313F, D524G, E562Q, D583N, P51L, S67R, E67K, T197A, H204R, E302K, F309N, L435G, N454K, H594Q, D653N,

R110S, K103L, например, комбинацию мутаций, таких как D200N, L603W и T330P, необязательно дополнительно включая T306K и W313F. В некоторых вариантах осуществления RT M-MLV, используемая в данном документе, содержит мутации D200N, L603W, T330P, T306K и W313F. В вариантах осуществления мутантная RT M-MLV содержит следующую аминокислотную последовательность:

M-MLV

(PE2):

TLNIEDEYRLHETSKPEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKATSTPVS
IKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLVPKKPGTNDYRPVQDLREVN
KRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPENGI
SGQLTWTRLPPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLALLAATSELDCCQ
GTRALLQTLGNLGYRASAKKAQICQKQVKYLGILLKEGQRWLTEARKETVMGQPTPK
TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDPQKAYQEIKQALLTA
PALGLPDLTKPFELFVDEKQGYAKGVLTQKLGWRRPVAYLSKKLDPVAAGWPPCLR
MVAIAVLT KDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTD
RVQFGPVVALNPATLLPLPEEGLQHNCILDILAEAHGTRPDLTDQPLPDADHTWYTDGSS
LLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVTDSR
YAFATAIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQKGHSAEA
RGNRMADQAARKAAITETPDTSTLLI (SEQ ID NO: 5).

В некоторых вариантах осуществления полипептид Gene Writer может содержать линкер, например, пептидный линкер, например, линкер, описанный в таблице 1. В некоторых вариантах осуществления полипептид Gene Writer содержит гибкий линкер между эндонуклеазой и доменом RT, например, линкер, содержащий аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESSGGSSGGSS. В некоторых вариантах осуществления домен RT полипептида Gene Writer может быть расположен с С-концевой стороны от домена эндонуклеазы. В некоторых вариантах осуществления домен RT полипептида Gene Writer может быть расположен с N-концевой стороны от домена эндонуклеазы.

Таблица 1. Иллюстративные линкерные последовательности

Аминокислотная последовательность	SEQ ID NO
GGG	101
GGSSGGG	102
GGSSGGSSGGG	103
GGSSGGSSGGSSGGG	104
GGSSGGSSGGSSGGSSGGG	105
GGSSGGSSGGSSGGSSGGSSGGG	106
GGGGG	107
GGGGSSGGGGG	108

GGGGS GGGGS GGGGS	109
GGGGS GGGGS GGGGS GGGGS	110
GGGGS GGGGS GGGGS GGGGS GGGGS	111
GGGGS GGGGS GGGGS GGGGS GGGGS GGGGS	112
GGG	113
GGGG	114
GGGGG	115
GGGGGG	116
GGGGGGG	117
GGGGGGGG	118
GSS	119
GSSGSS	120
GSSGSSGSS	121
GSSGSSGSSGSS	122
GSSGSSGSSGSSGSS	123
GSSGSSGSSGSSGSSGSS	124
EAAAK	125
EAAAKEAAAK	126
EAAAKEAAAKEAAAK	127
EAAAKEAAAKEAAAKEAAAK	128
EAAAKEAAAKEAAAKEAAAKEAAAK	129
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	130
PAP	131
PAPAP	132
PAPAPAP	133
PAPAPAPAP	134
PAPAPAPAPAP	135
PAPAPAPAPAPAP	136
GGSGGG	137
GGGGGS	138
GGSGSS	139
GSSGGS	140
GGSEAAAK	141
EAAAKGGS	142

GGSPAP	143
PAPGGS	144
GGGGSS	145
GSSGGG	146
GGGEAAK	147
EAAAKGGG	148
GGGPAP	149
PAPGGG	150
GSSEAAK	151
EAAAKGSS	152
GSSPAP	153
PAPGSS	154
EAAAKPAP	155
PAPEAAK	156
GGSGGGSS	157
GGSGSSGGG	158
GGGGGSGSS	159
GGGGSSGGS	160
GSSGGSGGG	161
GSSGGGGGS	162
GGSGGGEAAK	163
GGSEAAKGGG	164
GGGGGSEAAK	165
GGGEAAKGGS	166
EAAAKGGSGGG	167
EAAAKGGGGGS	168
GGSGGGPAP	169
GGSPAPGGG	170
GGGGGSPAP	171
GGGPAPGGS	172
PAPGGSGGG	173
PAPGGGGGS	174
GGSGSSEAAK	175
GGSEAAKGSS	176

GSSGGSEAAAK	177
GSSEAAAKGGS	178
EAAAKGGSGSS	179
EAAAKGSSGGS	180
GGSGSSPAP	181
GGSPAPGSS	182
GSSGGSPAP	183
GSSPAPGGS	184
PAPGGSGSS	185
PAPGSSGGS	186
GGSEAAAKPAP	187
GGSPAPEAAAK	188
EAAAKGGSPAP	189
EAAAKPAPGGS	190
PAPGGSEAAAK	191
PAPEAAAKGGS	192
GGGSSEAAAK	193
GGGEAAAKGSS	194
GSSGGGEAAAK	195
GSSEAAAKGGG	196
EAAAKGGGGSS	197
EAAAKGSSGGG	198
GGGGSSPAP	199
GGGPAPGSS	200
GSSGGGPAP	201
GSSPAPGGG	202
PAPGGGGSS	203
PAPGSSGGG	204
GGGEAAAKPAP	205
GGGPAPEAAAK	206
EAAAKGGGPAP	207
EAAAKPAPGGG	208
PAPGGGEAAAK	209
PAPEAAAKGGG	210

GSSEAAAKPAP	211
GSSPAPEAAAK	212
EAAAKGSSPAP	213
EAAAKPAPGSS	214
PAPGSSEAAAK	215
PAPEAAAKGSS	216
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	217
GGGGSEAAAKGGGGS	218
EAAAKGGGGSEAAAK	219
SGSETPGTSESATPES	220
GSAGSAAGSGEF	221
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	222

В некоторых вариантах осуществления полипептид Gene Writer содержит последовательность dCas9, содержащую мутацию D10A и/или H840A, например, следующую последовательность:

SMDKKYSIGLAIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFD
SGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
NGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
HLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIIKQILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ
KGQKNSRERMKRIIEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLES
EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
GETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEKPNIDFLEA
KGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
KLKGSPEQNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIR
EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQL

GGD (SEQ ID NO: 7)

В некоторых вариантах осуществления полипептид Gene Writer ковалентно связан или слит со средством, которое способствует активности системы модификации гена (например, модулятором ответа хозяина или эпигенетическим модификатором). В некоторых вариантах осуществления модулятор ответа хозяина, например, энхансер или ингибитор ответа хозяина, представляет собой белок или его функциональный фрагмент, например, домен.

В некоторых вариантах осуществления белок или домен, слитые с полипептидом Gene Writing, стимулируют процесс у хозяина, например, активируют или рекрутируют белок или комплекс хозяина. В некоторых вариантах осуществления белок или домен стимулируют Gene Writing, например, посредством замены или дополнения белка, комплекса или пути у хозяина. В некоторых вариантах осуществления белок представляет собой фермент репарации ДНК хозяина, который стимулирует Gene Writing. В некоторых вариантах осуществления белок или домен стимулируют запись последовательности в транс-положении. В некоторых вариантах осуществления белок или домен стимулируют запись последовательности в цис-положении. В некоторых вариантах осуществления домен представляет собой домен, который рекрутирует 5'-экзонуклеазу хозяина, например, EXO1, для записи последовательности в цис-положении. В некоторых вариантах осуществления домен представляет собой домен, который рекрутирует структурно-специфическую эндонуклеазу, например, FEN1, для записи последовательности в цис-положении. В некоторых вариантах осуществления белок или домен слиты с полипептидом Gene Writing, при этом белок или домен ингибируют процесс у хозяина, например, ингибируют или секвестрируют фермент репарации ДНК хозяина, который может препятствовать Gene Writing.

В некоторых вариантах осуществления нуклеиновая кислота-матрица, описанная в данном документе, например, РНК-матрица, ковалентно связана или слита со средством, которое способствует активности системы модификации гена (например, модулятором ответа хозяина или эпигенетическим модификатором).

В некоторых вариантах осуществления молекула РНК-матрицы для применения в системе содержит в направлении от 5'-к 3'-концу следующее: (1) спейсер gRNA; (2) остов gRNA; (3) последовательность гетерологичного объекта и (4) 3'-концевой домен гомологии. В некоторых вариантах осуществления

(1) присутствует спейсер Cas9 из ~18-22 нуклеотидов, например, 20 нуклеотидов;

(2) присутствует остов gRNA, содержащий одну или несколько шпилечных петель, например, 1, 2 или 3 петли, для обеспечения ассоциации матрицы с доменом нуклеазы Cas9; в некоторых вариантах осуществления остов gRNA несет последовательность, в направлении от 5'-к 3'-концу, GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGGACCGAGTCGGTCC (SEQ ID NO: 8).;

(3) в некоторых вариантах осуществления длина последовательности

гетерологичного объекта составляет, например, 7-74, например, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70 или 70-80 нуклеотидов или 80-90 нуклеотидов; в некоторых вариантах осуществления первое (расположенное ближе всего к 5'-концу) основание последовательности не представляет собой С;

(4) в некоторых вариантах осуществления 3'-концевой домен гомологии, который связывает праймирующую последовательность-мишень после того, как происходит никирование, составляет, например, 3-20 нуклеотидов, например, 7-15 нуклеотидов, например, 12-14 нуклеотидов. В некоторых вариантах осуществления 3'-концевой домен гомологии характеризуется содержанием GC, составляющим 40-60%.

Вторая gRNA, ассоциированная с системой, может способствовать обеспечению полной интеграции. В некоторых вариантах осуществления вторая gRNA может быть нацелена на местоположение, которое находится на расстоянии 0-200 нуклеотидов от ника первой нити, например, на расстоянии 0-50, 50-100, 100-200 нуклеотидов от ника первой нити. В некоторых вариантах осуществления вторая gRNA может связывать свою последовательность-мишень только после осуществления редактирования, например, gRNA связывает последовательность, присутствующую в последовательности гетерологичного объекта, но не в исходной последовательности-мишени.

В некоторых вариантах осуществления система Gene Writing, описанная в данном документе, используется для осуществления редактирования в клетках HEK293, K562, U2OS или HeLa. В некоторых вариантах осуществления система Gene Writing применяется для осуществления редактирования в первичных клетках, например, первичных кортикальных нейронах мышей E18.5.

В некоторых вариантах осуществления домен обратной транскриптазы или RT (например, описанный в данном документе) содержит последовательность RT MoMLV или ее вариант. В вариантах осуществления последовательность RT MoMLV содержит одну или несколько мутаций, выбранных из D200N, L603W, T330P, T306K, W313F, D524G, E562Q, D583N, P51L, S67R, E67K, T197A, H204R, E302K, F309N, L435G, N454K, H594Q, D653N, R110S и K103L. В вариантах осуществления последовательность RT MoMLV содержит комбинацию мутаций, таких как D200N, L603W и T330P, необязательно дополнительно включая T306K и/или W313F.

В некоторых вариантах осуществления домен эндонуклеазы (например, как описано в данном документе) содержит nCAS9, например, содержащую мутацию H840A.

В некоторых вариантах осуществления длина последовательности гетерологичного объекта (например, системы, описанной в данном документе) составляет приблизительно 1-50, 50-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000 или больше нуклеотидов.

В некоторых вариантах осуществления домены RT и эндонуклеазы соединены гибким линкером, например, содержащим аминокислотную последовательность SGGSSGGSSGSETPGTSESATPESGGSSGGSS (SEQ ID NO: 6).

В некоторых вариантах осуществления домен эндонуклеазы является N-концевым

относительно домена RT. В некоторых вариантах осуществления домен эндонуклеазы является С-концевым относительно домена RT.

В некоторых вариантах осуществления система осуществляет включение последовательности гетерологичного объекта в сайт-мишень посредством TPRT, например, как описано в данном документе.

Gene Writer, содержащие последовательности локализации

В определенных вариантах осуществления РНК системы редактирования генов GeneWriter™ дополнительно содержит последовательность внутриклеточной локализации, например, последовательность ядерной локализации. Последовательность ядерной локализации может представлять собой последовательность РНК, способствующую импорту РНК в ядро. В определенных вариантах осуществления сигнал ядерной локализации расположен в РНК-матрице. В определенных вариантах осуществления полипептид ретротранспозаза кодируется первой РНК, а РНК-матрица представляет собой вторую, отдельную РНК, и сигнал ядерной локализации расположен в РНК-матрице, а не в РНК, кодирующей полипептид ретротранспозазу. Без ограничения какой-либо теорией, в некоторых вариантах осуществления РНК, кодирующая ретротранспозазу, нацелена главным образом на цитоплазму для способствования ее трансляции, тогда как РНК-матрица нацелена главным образом на ядро для способствования ее ретротранспозиции в геном. В некоторых вариантах осуществления сигнал ядерной локализации расположен на 3'-конце, 5'-конце или во внутренней области РНК-матрицы. В некоторых вариантах осуществления сигнал ядерной локализации расположен в направлении 3'-конца от гетерологичной последовательности (например, расположен непосредственно за гетерологичной последовательностью в направлении 3'-конца) или расположен в направлении 5'-конца от гетерологичной последовательности (например, расположен непосредственно за гетерологичной последовательностью в направлении 5'-конца). В некоторых вариантах осуществления сигнал ядерной локализации находится за пределами 5'-UTR или за пределами 3'-UTR РНК-матрицы. В некоторых вариантах осуществления сигнал ядерной локализации находится между 5'-UTR и 3'-UTR, при этом сигнал ядерной локализации необязательно не транскрибируется вместе с транскриптом (например, сигнал ядерной локализации находится в антисмысловой ориентации или располагается ниже сигнала терминации транскрипции или сигнала полиаденилирования). В некоторых вариантах осуществления последовательность ядерной локализации находится внутри интрона. В некоторых вариантах осуществления совокупность одинаковых или разных сигналов ядерной локализации находятся в РНК, например, в РНК-матрице. В некоторых вариантах осуществления сигнал ядерной локализации имеет длину, составляющую менее 5, 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900 или 1000 п. о. Можно использовать разнообразные последовательности ядерной локализации РНК. Например, в работе Lubelsky and Ulitsky, *Nature* 555 (107-111), 2018 описаны последовательности РНК, которые управляют локализацией РНК в ядре. В некоторых вариантах осуществления сигнал ядерной локализации представляет собой сигнал ядерной

локализации РНК, полученный из SINE (SIRLOIN). В некоторых вариантах осуществления сигнал ядерной локализации связывается с белком, характеризующимся повышенным содержанием в ядре. В некоторых вариантах осуществления сигнал ядерной локализации связывается с белком HNRNPK. В некоторых вариантах осуществления сигнал ядерной локализации богат пиримидиновыми остатками, например, представляет собой С/Т-богатую, С/У-богатую, С-богатую, Т-богатую или У-богатую область. В некоторых вариантах осуществления сигнал ядерной локализации получен из длинной некодирующей РНК. В некоторых вариантах осуществления сигнал ядерной локализации получен из длинной некодирующей РНК MALAT1 или представляет собой М-область MALAT1 длиной 600 нуклеотидов (описанную в работе Miyagawa et al., *RNA* 18, (738-751), 2012). В некоторых вариантах осуществления сигнал ядерной локализации получен из длинной некодирующей РНК BORG или представляет собой мотив AGCCC (описанный в работе Zhang et al., *Molecular and Cellular Biology* 34, 2318-2329 (2014)). В некоторых вариантах осуществления последовательность ядерной локализации описана в работе Shukla et al., *The EMBO Journal* e98452 (2018). В некоторых вариантах осуществления сигнал ядерной локализации получен из не содержащего LTR ретротранспозона, ретротранспозона, содержащего LTR, ретровируса или эндогенного ретровируса.

В некоторых вариантах осуществления полипептид, описанный в данном документе, содержит одну или несколько (например, 2, 3, 4, 5) последовательностей, нацеленных на ядро, например, последовательность ядерной локализации (NLS). В некоторых вариантах осуществления NLS представляет собой двухкомпонентную NLS. В некоторых вариантах осуществления NLS обеспечивает импорт белка, содержащего NLS, в ядро клетки. В некоторых вариантах осуществления NLS слита с N-концом полипептида Gene Writer, описанного в данном документе. В некоторых вариантах осуществления NLS слита с С-концом полипептида Gene Writer. В некоторых вариантах осуществления NLS слита с N-концом или С-концом домена Cas. В некоторых вариантах осуществления линкерная последовательность расположена между NLS и соседним доменом полипептида Gene Writer.

В некоторых вариантах осуществления NLS содержит аминокислотную последовательность MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 9), PKKRKVEGADKRTADGSEFESPKKKRKV (SEQ ID NO: 10), RKSGKIAAIWKRPRKPKKKRKV (SEQ ID NO: 11) KRTADGSEFESPKKKRKV (SEQ ID NO: 12), KKTELQTTNAENKTKKL (SEQ ID NO: 13) или KRGINDRNFWRGNGRKTR (SEQ ID NO: 14), KRPAATKKAGQAKKKK (SEQ ID NO: 15) или ее функциональный фрагмент или вариант. Иллюстративные последовательности NLS также описаны в документе PCT/EP2000/011690, содержание которого включено в данный документ посредством ссылки для раскрытия иллюстративных последовательностей ядерной локализации. В некоторых вариантах осуществления NLS содержит аминокислотную последовательность, раскрытую в таблице 2. NLS из этой таблицы можно использовать с одной или несколькими копиями в полипептиде в одном или нескольких местоположениях

полипептида, например, 1, 2, 3 или больше копий NLS в N-концевом домене, между пептидными доменами, в С-концевом домене или в комбинации местоположений, для обеспечения улучшения в отношении субклеточной локализации в ядре. В одном полипептиде можно использовать несколько уникальных последовательностей. Последовательности могут быть в естественных условиях однокомпонентными или двухкомпонентными, например, иметь один или два отрезка из основных аминокислот, или могут использоваться в качестве химерных двухкомпонентных последовательностей. Ссылки на последовательности соответствуют номерам доступа UniProt, за исключением тех случаев, когда они указаны как SeqNLS для последовательностей, полученных с использованием алгоритма предсказания субклеточной локализации (Lin et al BMC Bioinform 13:157 (2012), включенная в данный документ посредством ссылки во всей своей полноте).

Таблица 2. Иллюстративные сигналы ядерной локализации для применения в системах Gene Writing

Последовательность	Ссылки на последовательность	SEQ ID No.
AHFKISGEKRPSTDPGKK AKNPKKKKKKKDP	Q76IQ7	223
AHRAKKMSKTHA	P21827	224
ASPEYVNLPIGNG	SeqNLS	225
CTKRPRW	O88622, Q86W56, Q9QYM2, O02776	226
DKAKRVSRNKSEKKRR	O15516, Q5RAK8, Q91YB2, Q91YB0, Q8QGQ6, O08785, Q9WVS9, Q6YGGZ4	227
EELRLKEELLKGIYA	Q9QY16, Q9UHL0, Q2TBP1, Q9QY15	228
EEQLRRRKNSRLNNTG	G5EFF5	229
EVLKVIRTGKRKKKAWK RMVTKVC	SeqNLS	230
HHHHHHHHHHHHHHQPH	Q63934, G3V7L5, Q12837	231
HKKKHPDASVNFSEFSK	P10103, Q4R844, P12682, B0CM99, A9RA84, Q6YKA4, P09429, P63159, Q08IE6, P63158, Q9YH06, B1MTB0	232
HKRTKK	Q2R2D5	233
IINGRKLKLLKSRRRSSQT SNNSTTSRRS	SeqNLS	234
KAEQERRK	Q8LH59	235
KEKRKRREELFIEQKKRK	SeqNLS	236
KKGKDEWFSGKKP	P30999	237

KKGPSVQKRKKT	Q6ZN17	238
KKKTVINDLLHYKKEK	SeqNLS, P32354	239
KKNGGKGKKNKPSAKIKK	SeqNLS	240
KKPKWDDFKKKKK	Q15397, Q8BKS9, Q562C7	241
KKRKKD	SeqNLS, Q91Z62, Q1A730, Q969P5, Q2KHT6, Q9CPU7	242
KKRRKRRRK	SeqNLS	243
KKRRRRARK	Q9UMS6, D4A702, Q91YE8	244
KKSKRGR	Q9UBS0	245
KKSRKRG	B4FG96	246
KKSTALSRELKGIMRRR	SeqNLS, P32354	247
KKSYQDPEIIAHSRPRK	Q9U7C9	248
KKTGKNRKLKSKRVKTR	Q9Z301, O54943, Q8K3T2	249
KKVSIAGQSGKLWRWKR	Q6YUL8	250
KKYENVVIKRSRKRGRPRK	SeqNLS	251
KNKKRK	SeqNLS	252
KPKKKR	SeqNLS	253
KRAMKDDSHGNSTSPKR RK	Q0E671	254
KRANSNLVAAYEKAKKK	P23508	255
KRASEDTTSGSPPKSSA GPKR	Q9BZZ5, Q5R644	256
KRFKRRWMVRKMKTKK	SeqNLS	257
KRGLNSSFETSPKKVK	Q8IV63	258
KRGNSSIGPNDLSKRKQR KK	SeqNLS	259
KRIHSVLSQSQIDPSKKV KRAK	SeqNLS	260
KRKGKLKNKGSKRKK	O15381	261
KRRRRRRREKRKR	Q96GM8	262
KRSNDRTYSPEEEKQRRRA	Q91ZF2	263
KRTVATNGDASGAHRAK KMSK	SeqNLS	264

KRVYNKGEDEQEHLPKG KKR	SeqNLS	265
KSGKAPRRRAVSMDNSN K	Q9WVH4, O43524	266
KVNFLDMSLDDIIYKELE	Q9P127	267
KVQHRIAKKTTRRRR	Q9DXE6	268
LSPSLSPL	Q9Y261, P32182, P35583	269
MDSLLMNRRKFLYQFKN VRWAKGRRETYLC	Q9GZX7	270
MPQNEYIELHRKRYGYRL DYHEKKRKKESTREHER SKKAKKMIGLKAKLYHK	SeqNLS	271
MVQLRPRASR	SeqNLS	272
NNKLLAKRRKGGASPKD DPMDDIK	Q965G5	273
NYKRPMDGTYGPPAKRH EGE	O14497, A2BH40	274
PDTKRAKLDSETTMVK KK	SeqNLS	275
PEKRTKI	SeqNLS	276
PGGRGKKK	Q719N1, Q9UBP0, A2VDN5	277
PGKMDKGEHRQERRDRP Y	Q01844, Q61545	278
PKKGDKYDKTD	Q45FA5	279
PKKKSRL	O35914, Q01954	280
PKKNKPE	Q22663	281
PKKRAKV	P04295, P89438	282
PKPKKLKVE	P55263, P55262, P55264, Q64640	283
PKRGRGR	Q9FYS5, Q43386	284
PKRRLVDDA	P0C797	285
PKRRRTY	SeqNLS	286
PLFKRR	A8X6H4, Q9TXJ0	287
PLRKAKR	Q86WB0, Q5R8V9	288
PPAKRKCIF	Q6AZ28, O75928, Q8C5D8	289

PPARRRRL	Q8NAG6	290
PPKKKRKV	Q3L6L5, P03070, P14999, P03071	291
PPNKRMKVKH	Q8BN78	292
PPRIYPQLPSAPT	P0C799	293
PQRSPFPKSSVKR	SeqNLS	294
PRPRKVPR	P0C799	295
PRRRVQRKR	SeqNLS, Q5R448, Q5TAQ9	296
PRRVRLK	Q58DJ0, P56477, Q13568	297
PSRKRPR	Q62315, Q5F363, Q92833	298
PSSKKRKV	SeqNLS	299
PTKKRVK	P07664	300
QRPGPYDRP	SeqNLS	301
RGKGGKGLGKGGAKRHR K	SeqNLS	302
RKAGKGGGGHKTTKKRS AKDEKVP	B4FG96	303
RKIKLKRAK	A1L3G9	304
RKIKRKRAK	B9X187	305
RKKEAPGPREELESRGR	O35126, P54258, Q5IS70, P54259	306
RKKRKGK	SeqNLS, Q29243, Q62165, Q28685, O18738, Q9TSZ6, Q14118	307
RKKRRQRRR	P04326, P69697, P69698, P05907, P20879, P04613, P19553, P0C1J9, P20893, P12506, P04612, Q73370, P0C1K0, P05906, P35965, P04609, P04610, P04614, P04608, P05905	308
RKKSIPLSIKNLKRKHKR KKNKTR	Q9C0C9	309
RKLVKPKNTKMKTCLRT NPY	Q14190	310
RKRLILSDKGQLDWKK	SeqNLS, Q91Z62, Q1A730, Q2KHT6, Q9CPU7	311
RKRLKSK	Q13309	312
RKRRVRDNM	Q8QPH4, Q809M7, A8C8X1, Q2VNC5, Q38SQ0, O89749, Q6DNQ9, Q809L9,	313

	Q0A429, Q20NV3, P16509, P16505, Q6DNQ5, P16506, Q6XT06, P26118, Q2ICQ2, Q2RCG8, Q0A2D0, Q0A2H9, Q9IQ46, Q809M3, Q6J847, Q6J856, B4URE4, A4GCM7, Q0A440, P26120, P16511,	
RRRSPKDKKEKDLDGAG KRRKT	Q7RTP6	314
RKRTPRVGQTGENDMN KRRRK	O94851	315
RLPVRRRRRR	P04499, P12541, P03269, P48313, P03270	316
RLRFRKPKSK	P69469	317
RQQRKR	Q14980	318
RRDLNSSFETSPKKVK	Q8K3G5	319
RRDRAKLR	Q9SLB8	320
RRGDGRRR	Q80WE1, Q5R9B4, Q06787, P35922	321
RRGRKRKA EKQ	Q812D1, Q5XXA9, Q99JF8, Q8MJG1, Q66T72, O75475	322
RRKKRR	Q0VD86, Q58DS6, Q5R6G2, Q9ERI5, Q6AYK2, Q6NYC1	323
RRKRSKSEDMDSVESKRR R	Q7TT18	324
RRKRSR	Q99PU7, D3ZHS6, Q92560, A2VDM8	325
RRPKGKTLQKRKPK	Q6ZN17	326
RRRGFERFGPDNMGRKR K	Q63014, Q9DBR0	327
RRRGKNKVAAQNCRK	SeqNLS	328
RRRKRR	Q5FVH8, Q6MZT1, Q08DH5, Q8BQP9	329
RRRQKQKGGASRRR	SeqNLS	330
RRRREGPRARRR	P08313, P10231	331
RRTIRLKL VYDKCDRSCK IQKKNRNKCQYCRFHKC LSVGMSHNAIRFGRMPRS EKAKLKAE	SeqNLS	332

RRVPQRKEVSRCKCRK	Q5RJN4, Q32L09, Q8CAK3, Q9NUL5	333
RVGGRRQAVECIEDLLNE PGQPLDLCKRPRP	P03255	334
RVVKLRIAP	P52639, Q8JMN0	335
RVVRRR	P70278	336
SKRKTKISRKTR	Q5RAY1, O00443	337
SYVKTVPNRTRTYIKL	P21935	338
TGKNEAKKRKIA	P52739, Q8K3J5, Q5RAU9	339
TLSPASSPSSVSCPVIPAST DESPGSALNI	SeqNLS	340
VSKKQRTGKKIH	P52739, Q8K3J5, Q5RAU9	341
SPKKKRKVE		342
KRTAD GSEFE SPKKKRKVE		343
PAAKRVKLD		344
PKKKRKV		345
MDSLLMNRRKFLYQFKN VRWAKGRRETYLC		346
SPKKKRKVEAS		347
MAPKKKRKVGHRGVP		348

В некоторых вариантах осуществления NLS представляет собой двухкомпонентную NLS. Двухкомпонентная NLS обычно содержит два кластера основных аминокислот, разделенных спейсерной последовательностью (длина которой может составлять, например, приблизительно 10 аминокислот). В однокомпонентной NLS обычно отсутствует спейсер. Примером двухкомпонентной NLS является NLS нуклеоплазмина, содержащая последовательность KR[PAATKKAGQA]KKKK (SEQ ID NO: 15), где спейсер заключен в скобки. Другая иллюстративная двухкомпонентная NLS содержит последовательность PKKKRKVEGADKRTADGSEFESPKKKRKV (SEQ ID NO: 16). Иллюстративные NLS описаны в международной заявке WO2020051561, которая включена в данный документ посредством ссылки во всей своей полноте, в том числе в отношении последовательностей ядерной локализации.

В определенных вариантах осуществления полипептид системы редактирования генов GeneWriter™ дополнительно содержит последовательность внутриклеточной локализации, например, последовательность ядерной локализации и/или последовательность ядрышковой локализации. Последовательность ядерной локализации и/или последовательность ядрышковой локализации может представлять собой аминокислотные последовательности, которые способствуют импорту белка в ядро и/или

ядрышко, где он может способствовать интеграции гетерологичной последовательности в геном. В определенных вариантах осуществления полипептид системы редактирования генов Gene Writer™ (например, ретротранспозаза) дополнительно содержит последовательность ядрышковой локализации. В определенных вариантах осуществления полипептид ретротранспозаза кодируется первой РНК, а РНК-матрица представляет собой вторую, отдельную РНК, и сигнал ядрышковой локализации кодируется РНК, кодирующей полипептид ретротранспозазу, а не РНК-матрицей. В некоторых вариантах осуществления сигнал ядрышковой локализации расположен на N-конце, C-конце или во внутренней области полипептида. В некоторых вариантах осуществления используется совокупность одинаковых или разных сигналов ядрышковой локализации. В некоторых вариантах осуществления длина сигнала ядерной локализации составляет менее 5, 10, 25, 50, 75 или 100 аминокислот. Могут использоваться разнообразные сигналы ядрышковой локализации полипептида. Например, в работе Yang et al., *Journal of Biomedical Science* 22, 33 (2015), описан сигнал ядерной локализации, который также функционирует в качестве сигнала ядрышковой локализации. В некоторых вариантах осуществления сигнал ядрышковой локализации может также представлять собой сигнал ядерной локализации. В некоторых вариантах осуществления сигнал ядрышковой локализации может перекрываться с сигналом ядерной локализации. В некоторых вариантах осуществления сигнал ядрышковой локализации может содержать отрезок из основных остатков. В некоторых вариантах осуществления сигнал ядрышковой локализации может быть богат остатками аргинина и лизина. В некоторых вариантах осуществления сигнал ядрышковой локализации может быть получен из белка, характеризующегося повышенным содержанием в ядрышке. В некоторых вариантах осуществления сигнал ядрышковой локализации может быть получен из белка, характеризующегося повышенным содержанием в локусах рибосомной РНК. В некоторых вариантах осуществления сигнал ядрышковой локализации может быть получен из белка, который связывается с рРНК. В некоторых вариантах осуществления сигнал ядрышковой локализации может быть получен из MSP58. В некоторых вариантах осуществления сигнал ядрышковой локализации может представлять собой однокомпонентный мотив. В некоторых вариантах осуществления сигнал ядрышковой локализации может представлять собой двухкомпонентный мотив. В некоторых вариантах осуществления сигнал ядрышковой локализации может состоять из нескольких однокомпонентных или двухкомпонентных мотивов. В некоторых вариантах осуществления сигнал ядрышковой локализации может состоять из смеси однокомпонентных и двухкомпонентных мотивов. В некоторых вариантах осуществления сигнал ядрышковой локализации может представлять собой два двухкомпонентных мотива. В некоторых вариантах осуществления мотив ядрышковой локализации может представлять собой KRASSQALGTIPKRRSSSRFIKRKK (SEQ ID NO: 17). В некоторых вариантах осуществления сигнал ядрышковой локализации может быть получен из киназы, индуцирующей ядерный фактор κВ. В некоторых вариантах осуществления сигнал ядрышковой локализации может представлять собой мотив RKKRKKK (SEQ ID NO: 18)

(описанный в работе Birbach et al., Journal of Cell Science, 117 (3615-3624), 2004).

GeneWriter, содержащие домены Cas

В некоторых вариантах осуществления GeneWriter, описанная в данном документе, содержит домен Cas. В некоторых вариантах осуществления домен Cas может направлять GeneWriter к сайту-мишени, определенному посредством gRNA, за счет чего осуществляется запись последовательности в “цис-положении”. В некоторых вариантах осуществления транспозаза слита с доменом Cas. В некоторых вариантах осуществления домен Cas используется для замены эндогенного домена транспозазы, например, для замены домена эндонуклеазы или ДНК-связывающего домена. В некоторых вариантах осуществления домен эндонуклеазы предусматривает домен CRISPR/Cas (также называемый в данном настоящем документе CRISPR-ассоциированным белком). В некоторых вариантах осуществления ДНК-связывающий домен предусматривает домен CRISPR/Cas. В некоторых вариантах осуществления домен CRISPR/Cas предусматривает белок, участвующий в системе коротких палиндромных повторов, регулярно расположенных группами (CRISPR), например, белок Cas, и необязательно связывает направляющую РНК, например, одиночную направляющую РНК (sgRNA). Дополнительное описание систем на основе CRISPR можно найти, например, в разделе данного документа под заглавием “CRISPR”.

Различные CRISPR-ассоциированные (Cas) гены или белки могут использоваться в технологиях, предусмотренных настоящим изобретением, и выбор белка Cas будет зависеть от конкретных условий способа. Конкретные примеры белков Cas включают системы класса II, в том числе Cas1, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cpf1, C2C1 или C2C3. В некоторых вариантах осуществления белок Cas, например, белок Cas9, может быть получен из любого из разнообразных видов прокариот. В некоторых вариантах осуществления конкретный белок Cas, например, конкретный белок Cas9, выбран для распознавания конкретной последовательности мотива, прилегающего к протоспейсеру (PAM). В некоторых вариантах осуществления ДНК-связывающий домен или домен эндонуклеазы содержит последовательность, нацеливающуюся на полипептид, такой как белок Cas, например, Cas9. В некоторых вариантах осуществления белок Cas, например, белок Cas9, может быть получен из бактерий или архей или синтезирован с применением известных способов. В определенных вариантах осуществления белок Cas может быть получен из грамположительных бактерий или грамотрицательных бактерий. В некоторых вариантах осуществления белок Cas может быть получен из бактерий рода *Streptococcus* (например, *S. pyogenes* или *S. thermophilus*), *Francisella* (например, *F. novicida*), *Staphylococcus* (например, *S. aureus*), *Acidaminococcus* (например, *Acidaminococcus* sp. BV3L6), *Neisseria* (например, *N. meningitidis*), *Cryptococcus*, *Corynebacterium*, *Haemophilus*, *Eubacterium*, *Pasteurella*, *Prevotella*, *Veillonella* или *Marinobacter*. В некоторых вариантах осуществления Gene Writer может содержать белок Cas, приведенный в таблице 3А или таблице 4, или его функциональный фрагмент или последовательность, на по меньшей мере 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% идентичную таковым.

Таблица 3А. Белки CRISPR/Cas, виды и мутации

Вариант	Исходный хозяин	Последовательность белка	Мутация никазы	SEQ ID No.
Nme2Cas 9	Neisseria meningitidis	MAAFKPNPINYILGLDIGIASVGWA MVEIDEEENPIRLIDLGVRFERA EV PKTGDSLAMARRLARSVRRLTRRA HRLLRARRLLKREGVLQAADF DEN GLIKSLPNTPWQLRAAALDRKLTPL EWSAVLLHLIKHRGYLSQRKNEGET ADKELGALLKGVANNAHALQTGDF RTPAELALNKFEEKESGHIRNQRGDY SHTFSRKDLQAEILLFEKQKEFGNP HVSGGLKEGIETLLMTQRPALSGDA VQKMLGHCTFEPAEPKAAKNTYTA ERFIWLTCLNNLRILEQGSERPLTDT ERATLMDEPYRKSKLTYAQARKLL GLEDTAFFKGLRYGKDNAEASTLME MKAYHAISRALEKEGLKDKKSPLNL SSELQDEIGTAFSLFKTDEDITGRLK DRVQPEILEALLKHISFDK FVQISLKA LRRIVPLMEQGKRYDEACAEIYGDH YGKKNTEEKIYLPPIPADEIRNPVVL RALSQARKVINGVVRRYGSPARIHIE TAREVGKSFKDRKEIEKRQEENRKD REKAAAKFREYFPNFVGEPKSKDIL KLRLYEQQH GKCLYSGKEINL VRLN EKG YVEIDHALPFSRTWDDSFNNKV LVLGSENQNKGNQTPYEYFNGKDN SREWQEFKARVETSRFPRSKKQRILL QKFDEDGFKECNLNDTRYVNRFLCQ FVADHILLTGKGKRRVFASNGQITN LLRGFWGLRKVRAENDRH HALDAV VVACSTVAMQQKITRFVRYKEMNA FDGKTIDKETGKVLHQKTHFPQPWE FFAQEVMIRVFGKPDGKPEFEEADTP	N611A	349

		EKLRTLLEAEKLSSRPEAVHEYVTPLF VSRAPNRKMSGAHKDTLRSKRFB KHNEKISVKRVWLTEIKLADLENMV NYKNGREIELYEALKARLEAYGGNA KQAFDPKDNPFYKKGGQLVKAVRV EKTQESGVLLNKKNAYTIADNGDM VRVDVFCKVDKKGKNQYFIVPIYA WQVAENILPDIDCKGYRIDDSYTFCF SLHKYDLIAFQKDEKSKVEFAYYIN CDSSNGRFYLA WHDKGSKEQQFRIS TQNLVLIQKYQVNELGKEIRPCRLK KRPPVR		
PpnCas9	Pasteurella pneumotropic a	MQNNPLNYILGLDLGASIGWAVVEI DEESSPIRLIDVGVRTFERAEVAKTG ESLALSRRLARSSRRLIKRRRAERLKK AKRLLKAEKILHSIDEKLPINWQLR VKGLKEKLERQEWAAVLLHLSKHR GYLSQRKNEGKSDNKELGALLSGIA SNHQMLQSSEYRTPAEIAVKKFQVE EGHIRNQRGSYTHTF SRLDLLAEME LLFQRQAELGNSYTSTTLLENLTALL MWQKPALAGDAILKMLGKCTFEPS EYKAAKNSYSAERFVWLTKLNNLRI LENGTERALNDNERFALLEQPYEKS KLTYAQVRAMLALSDNAIFKGVRY LGEDKKTVESKTTLIEMKFYHQIRKT LGS AELKKEWNE LKGN S DLLDEIGT AFSLYKTDDDICRYLEGKLPERV LN ALLENLNFDKFIQLSLKALHQILPLM LQGQRYDEAVSAIYGDHYGKKSTET TRLLPTIPADEIRNPVVLRTL TQARK VINAVVRLYGSPARIHIETAREVGKS YQDRKKLEKQQEDNRKQRESAVKK FKEMFPHFVGEPKGKDILKMRLYEL QQA KCLYSGKSLELHRLLEKGYVEV	N605A	350

		DHALPFSRTWDDSFNNKVLVLANE NQNKGNLTPYEWLDGKNNSERWQ HFVVRVQTSGFSAKKQRILNHKLD EKGFIERNLNDTRYVARFLCNFIADN MLLVGKGKRNVFASNGQITALLRHR WGLQKVREQNDRHHALDAVVVAC STVAMQQKITRFVRYNEGNVFSGER IDRETGEIPLHFPSPWAFFKENVEIRI FSENPKLELENRLPDYPQYNHEWVQ PLFVSRMPTRKMTGQGHMETVKSA KRLNEGLSVLKVPLTQLKLSDLERM VNRDREIALYESLKARLEQFGNDPA KAFAEPFYKKGGALVKAVRLEQTQ KSGVLVRDGNGVADNASMVRVDVF TKGGKYFLVPIYTWQVAKGILPNRA ATQGKDENDWDIMDEMATFQFSLC QNDLIKLVTKKKTIFGYFNGLNRATS NINIKEHDLDKSKGKLGILEVGVKL AISLEKYQVDELGKNIRPCRPTKRQH VR		
SauCas9	Staphylococcus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTSYVVK EAKQLLKVQKAYHQLDQSFIDTYID LLETRRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEY YEFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS	N580A	351

		NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVDLSQQK EIP TTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEII RTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAI PLEDLLNPNFNYEVDHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKGYKHHAE DALIIANADFIFKEWK KLDKAKKVMENQMFE EKQAESMPE IETE QEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRELINDTLYSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHPQTYQKLKLIME QYGDEKNPLYKYEEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKPYRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQAEFIASFYNN DLIKINGEL YRVIGVNNDLLNRIEVN MIDITYREYLENMNDKRPPRIIKTAS KTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG		
SauCas9- KKH	Staphylococ- cus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTS DYVK EAKQLLKVQKAYHQLDQSFIDTYID	N580A	352

		LLETRRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEY YEKFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVDLSQQK EIP TTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEII RTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAI PLEDLLNPNFYEV DHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKGYKHHAEDALIIANADFIFKEWK KLDKAKKVMENQMFEKQAESMPE IETE QEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRKLINDTL YSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKYEEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKP YRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQAEFIASFYKN DLIKINGEL YRVIGVNNDLLNRIEVN MIDITYREYLENMNDKRPPHIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG		
SauriCas9	Staphylococc	MQENQQKQNYILGLDIGITSVGYGLI	N588A	353

us auricularis	DSKTREVIDAGVRLFPEADSENNNSN RRSKRGARRLKRRRIHRLNRVKDLL ADYQMIDLNNVPKSTDPTYTIRVKGL REPLTKEEFAIALLHIAKRRGLHNISV SMGDEEQDNELSTKQQLQKNAQQL QDKYVCELQLERLTNINKVRGEKNR FKTEDFVKEVKQLCETQRQYHNIDD QFIQQYIDL VSTRREYFEGPGNGSPY GWDGDLLKWYEKLMGRCTYFPEEL RSVKYAYSADLFNALNDLNNLVVT RDDNPKLEYEYKYHIIENVFKQKKN PTLKQIAKEIGVQDYDIRGYRITKSG KPQFTSFKLYHDLKNIFEQAKYLED VEMLDEIAKILTIYQDEISIKKALDQL PELLTESEKSQIAQLTGYTGTHRLSL KCIHIVIDELWESPENQMEIFTRLNL KPKKVEMSEIDSIPTTLVDEFILSPVV KRAFIQSIKVINAVINRFGLPEDIIIE AREKNSKDRRKFINKLQKQNEATRK KIEQLLAKYGNTNAKYMIEKIKLHD MQEGKCLYSLEAIPLEDLLSNPTHYE VDHIIPRSVSFDNSLNNKVLVKQSEN SKKGNRTPYQYLSSNESKISYNQFK QHILNLSKAKDRISKKKRDMLEER DINKFEVQKEFINRNLVDTRYATREL SNLLKTYFSTHDYAVKVKTINGGFT NHLRKVWDFKKHRNHGYKHHAED ALVIANADFLFKTHKALRRTDKILEQ PGLEVNDTTVKVDTEEKYQELFETP KQVKNIKQFRDFKYSHRVDKKPNR QLINDTLYSTREIDGETYVVQTLKDL YAKDNEKVKKLFTERPQKILMYQH DPKTFEKLMTILNQYAEAKNPLAAY YEDKGEYVTKYAKKGNGPAIHKIKY IDKKLGSYLDVSNKYPETQNKLVKL		
----------------	---	--	--

		SLKSFRFDIYKCEQGYKMVSIGYLD VLKKDNYYYIPKDKYEA EKQKKKIK ESDLFVG SFYYNDLIMYEDELFRVIG VNSDINN LVELNMVDITYKDFCEVN NVTGEKRIKKTIGKRVV LIEKYTTDI LGNLYKTPLPKKPQLIFKRGEL		
SauriCas9 -KKH	Staphylococcus auricularis	MQENQQKQNYILGLDIGITSVGYGLI DSKTREVIDAGVRLFPEADSENNSN RRSKRGARRLKRRRIHRLNRVKDLL ADYQMIDLNNVPKSTDPTYTIRVKGL REPLTKEEFAIALLHIAKRRGLHNISV SMGDEEQDNELSTKQQLQKNAQQL QDKYVCELQLERLTNINKVRGEKNR FKTEDFVKEVKQLCETQRQYHNIDD QFIQQYIDL VSTRREYFEGPGNGSPY GWDGDLLK WYEKLMGRCTYFPEEL RSVKYAYSADLFNALNDLNNLVVT RDDNPKLEYEYKYHIIENVFKQKKN PTLKQIAKEIGVQDYDIRGYRITKSG KPQFTSFKLYHDLKNIFEQAKYLED VEMLDEIAKILTIYQDEISIKKALDQL PELLTESEKSQIAQLTGYTGTHRLSL KCIHIVIDELWESPENQMEIFTRLNL KPKKVEMSEIDSIPTTLVDEFILSPVV KRAFIQSIKVINAVINRFGLPEDIIEL AREKNSKDRRK FINKLQKQNEATRK KIEQLLAKYGNTNAKYMIEKIKLHD MQEGKCLYSLEAIPLEDLLSNPTHYE VDHIIPRSVSFDNSLNNKVLVKQSEN SKKGNRTPYQYLSSNESKISYNQFK QHILNLSKAKDRISKKKRDMLEER DINKFEVQKEFINRNLVDTRYATREL SNLLKTYFSTHDYAVKVKTINGGFT NHLRKVWDFKKHRNHGYKHHAED ALVIANADFLFKTHKALRRTDKILEQ	N588A	354

		PGLEVNDTTVKVDTEEKYQELFETP KQVKNIKQFRDFKYSHRVDKKPNR KLINDTLYSTREIDGETYVVQTLKDL YAKDNEKVKKLFTERPQKILMYQH DPKTFEKLMTILNQYAEAKNPLAAY YEDKGEYVTKYAKKGNGPAIHKIKY IDKKLGSYLDVSNKYPETQNKLVKL SLKSFRFDIYKCEQGYKMVSIGYLD VLKKDNYYYIPKDKYEA EKQKKKIK ESDLFVGSFYKNDLIMYEDELFRVIG VNSDINNVELNMVDITYKDFCEVN NVTGEKHIKKTIGKRVVLEKYTTDI LGNLYKTPLPKKPQLIFKRGEL		
ScaCas9- Sc++	Streptococcus canis	MEKKYSIGLDIGTNSVGWAVITDDY KVPSKKFKVLGNTNRKSIKKNLMGA LLFDSGETAEATRLKRTARRRYTRR KNRIRYLQEIFANEMAKLDDSFQRL EESFLVEEDKKNERHPIFGNLADEVA YHRNYPTIYHLRKKLADSPEKADLR LIYLALAHIIKFRGHFLIEGKLNAENS DVAKLFIYQLIQTYNQLFEESPLDEIE VDAKGILSARLSKSKRLEKLI VFPN EKKNGLFGNIIALALGLTPNFKSNFD LTEDAKLQLSKDTYDDDLDELLGQI GDQYADLFSAAKNLSDAILLSDILRS NSEVTKAPLSASMVKRYDEHHQDL ALLKTLVRQQFPEKYAEIFKDDTKN GYAGYVGADKKLRKRSGKLATEEE FYKFIKPILEKMDGAEELLAKLNRD DLLRKQRTFDNGSIPHQIHLKELHAI LRRQEEFY PFLKENREKIEKILTFRIP YYVGPLARGNSRFAWLTRKSEEAIT PWNFEEVVDKGASAQSFIERMTNFD EQLPNKKVLPKHSLLYEYFTVYNEL TKVKYVTERMRKPEFLSGEQKKAIV	N872A	355

		DLLFKTNRKVTVKQLKEDYFKKIEC FDSVEIIGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFE DREMIEERLKTYAHLFDDKVMKQL KRRHYTGWGRLSRKMINGIRDKQS GKTILDFLKSDGFSNRNFMQLIHDDS LTFKEEIEKAQVSGQGDSLHEQIADL AGSPAIKKGILQTVKIVDELVKVMG HKPENIVIAMARENQTTTKGLQQSR ERKKRIEEGIKELESQILKENPVENTQ LQNEKLYLYYLQNGRDMYVDQELD INRLSDYDVDHIVPQSFIKDDSIDNK VLTRSVENRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKA ERGGLSEADKAGFIKRQLVETRQITK HVARILDSRMNTKRDKNDKPIREVK VITLKSCLVSDFRKDFQLYKVRDINN YHHAHDAYLNAVVGITALIKKYPKL ESEFVYGDYKVYDVRKMIKSEQEI GKATAKRFFYSNIMNFFKTEVKLAN GEIRKRPLIETNGETGEVVWNKEKD FATVRKVLAMPQVNIVKKTEVQTG GFSKESILSKRESAKLIPRKKGWDTR KYGGFGSPTVAYSILVVAKEVGKA KKLKSVKVLVGITIMEKGSYEKDPKG FLEAKGYKDIKKELIFKLPKYSLFEL ENGRRRMLASAKELQKANELVLPQ HLVRLLYYTQNISATTGSNNLGYIEQ HREEFKEIFEKIIDFSEKYILKNKVNS NLKSSFDEQFAVSDSILLSNSFVSLLK YTSFGASGGFTFLDLVDKQGRLRYQ TVTEVLDTLIYQSITGLYETRTDLS QLGGD		
SpyCas9	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA	N863A	356

LLFDSGETAEATRLKRTARRRYTRR
KNRICYLQEIFSNEMAKVDDSFHRL
EESFLVEEDKKHERHPIFGNIVDEVA
YHEKYPTIYHLRKKLVDSTDKADLR
LIYLALAHMIKFRGHFLIEGDLNPDN
SDVDKLFQILVQTYNQLFEENPINAS
GVDAKAILSARLSKSRLENLIAQLP
GEKKNGLFGNLIASLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLNLLA
QIGDQYADLFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKQRTFDNG
SIPHQIHLGELHAILRRQEDFYFPLKD
NREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENE
DILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSG
QGDSLHEHIANLAGSPAIAKKGILQTV
KVVDLVKVMGRHKPENIVIAMARE
NQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYL
QNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKKMKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNT

		KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVREINNYHHAHDAYLNA VVGITALIKKYPKLESEFVYGDYKVY DVRKMIKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPPKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVILADANLDKVL SAYNKH DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDTLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- NG	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF	N863A	357

AWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENE
DILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSG
QGDSLHEHIANLAGSPAIIKKGILQTV
KVVDELVKVMGRHKPENIVIAMARE
NQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYL
QNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKKMKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNT
KYDENDKLIREVKVITLKSCLVSDFR
KDFQFYKVBREINNYHHAHDAYLNA
VVGTAIIKKYPKLESEFVYGDYKVY
DVRKMIAKSEQEIGKATAKYFFYSNI
MNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSPQVNI
VKKTEVQTGGFSKESIRPKRNSDKLI
ARKKDWDPKKYGGFVSPTVAYSVL
VVAKVEKGKSKKLKSVKELLGITIM
ERSSFENPIDFLEAKGYKEVKKDLII
KLPKYSLELENGRKRMLASARFLQ
KGNELALPSKYVNFLYLASHYEKLG
GSPEDNEQKQLFVEQHKHYLDEIEQ
ISEFSKRVILADANLDKVL SAYNKHR
DKPIREQAENIIHLFTLTNLGAPRAFK
YFDTTIDRKVYRSTKEVL DATLIHQ

		ITGLYETRIDLSQLGGD		
SpyCas9-SpRY	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAERTRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIAKKGILQTV KVVDLVKVMGRHKPENIVIAMARE NQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGK	N863A	358

		SDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVREINNYHHAHDAYLNA VVGITALIKKYPKLESEFVYGDYKVY DVRKMIKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESIRPKRNSDKLI ARKKDWDPKKYGGFLWPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLENGRKRMLASAKQLQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVILADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTRLGAPRAFK YFDTTIDPKQYRSTKEVL DATLIHQ ITGLYETRIDLSQLGGD		
St1Cas9	Streptococcus thermophilus	MSDLVLGLDIGISVGVGILNKVTGE IIHKNSRIFPAAQAENNLVRRITNRQG RRLARRKKHRRVRLNRLFEEGLIT DFTKISINLNPYQLRVKGLTDELSNE ELFIALKNMVKHRGISYLDASDDG NSSVG DY AQIVKENSQLETKTPGQI QLERYQTYGQLRGDFTVEKDGGKH RLINVFPTSAYRSEALRILQTQQEFNP QITDEFINRYLEILTGKRKYHGPNG EKSRTDYGRYRTSGETLDNIFGILIG KCTFYPDEFRAAKASYTAQEFNLLN DLNNLTVPTETKKLSKEQKNQIINYV KNEKAMGPAKLKFKYIAKLLSCDVA DIKGYRIDKSGKAEIHTFEAYRKMK TLETLDIEQMDRETLDKLAYVLTNL	N622A	359

		TEREGIQEALEHEFADGSFSQKQVDE LVQFRKANSSIFGKGWHNFSVKLM MELIPELYETSEEQMTILTRLGKQKT TSSSNKTKYIDEKLLTEEIYNPVVAK SVRQAIKIVNAAIKEYGDFDNIVIEM ARETNEDDEKKAIQKIQKANKDEKD AAMLKAANQYNGKAELPHSVFHGH KQLATKIRLWHQQGERCLYTGKTISI HDLINNSNQFEVDHILPLSITFDDSLA NKVLVYATANQEKGQRTPYQALDS MDDAWSFRELKAFVRESKTLNKK KEYLLTEEDISKFDVRKKFIERNLVD TRYASRVVLNALQEHFRAHKIDTKV SVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTL VSYSEDQLLDIETGELISDDEYKESV FKAPYQHFVDTLKSKEFEDSILFSYQ VDSKFNRKISDATIYATRQAKVGKD KADETYVLGKIKDIYTQDGYDAFM KIYKKDKSKFLMYRHDPQTFEKVIE PILENYPNKQINEKGKEVPCNPFLKY KEEHGYIRKYSKKGNGPEIKSLKYY DSKLGNHIDITPKDSNNKVVLQSVSP WRADVYFNKTTGKYEILGLKYADL QFEKGTGTYKISQEKYNDIKKKEGV DSDSEFKFTLYKNDLLLVKDTETKE QQLFRFLSRTMPKQKHVELKPYDK QKFEGGEALIKVLGNVANSQGCKK GLGKSNISIYKVRTDVLGNQHIIKNE GDKPKLDF		
BlatCas9	Brevibacillus laterosporus	MAYTMGIDVGIASCGWAIVDLERQR IIDIGVRTFEKAENPKNGEALAVPRR EARSSRRRLRRKKHRIERLKHMFVR NGLAVDIQHLEQTLRSQNEIDVWQL RVDGLDRMLTQKEWLRVLIHLAQR	N607A	360

RGFQSNRKTGSSSEDGQVLVNV TEN
DRLMEEKDYRTVAEMMVKDEKFS D
HKRNKNGNYHGVVSRSSLLVEIHT L
FETQRQHHNSLASKDFELEYVNIWS
AQRPVATKDQIEKMIGTCTFLPKEK
RAPKASWHFQYFMLLQTINHIRITNV
QGTRSLNKEEIEQVVNMALTKSKVS
YHDTRKILDLSEEYQFVGLDYGKED
EKKKVESKETIIKLDDYHKL N KIFNE
VELAKGETWEADDYDTVAYALTFF
KDDEDIRDYLNKYKDSKNRLVKN
LANKEYTNELIGKVSTLSFRKVGHL S
LKALRKIIPFLEQGMTYDKACQAAG
FDFQGISKKKRSVVL PVIDQISNPVV
NRALTQTRKVINALIKKYGSPETIHIE
TARELSKTFDERKNITKDYKENRDK
NEHAKKHLSELGIINPTGLDIVKYKL
WCEQQGRCMYSNQPISFERLKESGY
TEVDHIIPYSRSMNDSYNNRVLVMT
RENREKGNQTPFEYMGNDTQRWYE
FEQRVTTNPQIKKEKRQNLLLKGFT
NRRELEMLERNLNDTRYITKYL SHFI
STNLEFSPSDKKKKVVNTSGRITSHL
RSRWGLEKNRGQNDLHHAMDAIVI
AVTSDSFIQQVTNYYKRKERRELNG
DDKFPLPWKFFREEVIARLSPNPKEQ
IEALPNHFYSEDELADLQPIFVSRMP
KRSITGEAHQAQFRRVVGKTKEGKN
ITAKKTALVDISYDKNGDFNMYGRE
TDPATYEAIKERYLEFGGNVKKAFS
TDLHKPKKDGTKGPLIKSVRIMENK
TLVHPVNKGKGVVYNSSIVRTDVFQ
RKEKYLLPVYVTDVTKGKLPNKVI
VAKKGYHDWIEVDDSFTFLFSLYPN
DLIFIRQNPKKKISLKKRIESHSISDSK

		EVQEIHAYYKGVDSSTAIEFIIHDG SYYAKGVGVQNLDCEKYQVDILG NYFKVKGEKRLELETSDSNHKGKD VNSIKSTSR		
cCas9-v16	Staphylococcus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTSFYVK EAKQLLKVQKAYHQLDQSFIDTYID LLETRRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEY YEKFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVVDLSQQK EIP TTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEII RTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAI PLEDLLNPNFYEV DHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKG YKHH AEDALIIANADFIKWK KLDKAKKVMENQMFEEKQAESMPE IETE QEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRKLINDTL YSTRKDD	N580A	361

		KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKYEEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKPYRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQAEFIASFYKN DLIKINGELYRVIGVNSDKNNLIEVN MIDITYREYLENMNDKRPPHIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG		
cCas9-v17	Staphylococcus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTSDYVK EAKQLLKVQKAYHQLDQSFIDTYID LLETRRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEY YEFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVDLSQQK EIPTTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEIIRTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAI PLEDLLNNPFNYEVDHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLAKGKGRIS	N580A	362

		KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKGYPKHAEDALIANADFIKWK KLDKAKKVMENQMFEEKQAESMPE IETEQEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRKLINDTL YSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKYEEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKL SLKPYRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQAEFIASFYKN DLIKINGEL YRVIGVNNSTRNIVELN MIDITYREYLENMNDKRPPHIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQ IHKKG		
cCas9-v21	Staphylococcus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTSDYVK EAKQLLKVQKAYHQLDQSFIDTYID LLETRRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN ADLYNALNDLNNLVITRDENEKLEY YEKFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVDLSQQK	N580A	363

		EIP TTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEII RTTGKE NAKY LIEKIKLHDMQEGKCLYSLEAI PLEDLLNPNFNYEVDHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLA KGKGRIS KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKG YKHHAE DALIIANADFI FKEWK KLDKAKKVMENQMFE EKQAESMPE IETE QEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRKLINDTL YSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKY YEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKPYRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQAEFIASFYKN DLIKINGEL YRVIGVNSDDRNI IELN MIDITYREYLENMNDKRPPHIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG		
cCas9-v42	Staphylococcus aureus	MKRNYILGLDIGITSVGYGIIDYETR DVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRRHRIQRVKKLLFDYNL LTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEED TGNELSTKEQISRNSKALEEKYVAEL QLERLKKDGEVRGSINRFKTS DYVK EAKQLLKVQKAYHQLDQSFIDTYID LLET RRTYYEGPGEGSPFGWKDIKE WYEMLMGHCTYFPEELRSVKYAYN	N580A	364

		ADLYNALNDLNNLVITRDENEKLEY YEFQIIENVFKQKKKPTLKQIAKEIL VNEEDIKGYRVTSTGKPEFTNLKVY HDIKDITARKEIENAELLDQIAKILTI YQSSEDIQEELTNLNSLTQEEIEQIS NLKGYTGTHNLSLKAINLILDELWH TNDNQIAIFNRLKLVKKVVDLSQQK EIPTTLVDDFILSPVVKRSFIQSIKVIN AIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEII RTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAI PLEDLLNPNFYEV DHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLS SSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRN LVDTRYATRGLMNLLRSYFRVNNL DVKVKSINGGFTSFLRRKWKFKKER NKGKYKHHAE DALIIANADFIFKEWK KLDKAKKVMENQMFE EKQAESMPE IETE QEYKEIFITPHQIKHIKDFKDYK YSHRVDKKPNRKLINDTL YSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLI NKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKY YEETGNYLTKYS KKDNGPVIKKIKYYGNKLNAHLDIT DDYPNSRNKVVKLSLKP YRFDVYLD NGVYKFVTVKNLDVIKKENYYEVN SKCYEEAKKLKKISNQA EFIA SFYKN DLIKINGEL YRVIGVNNNRLNKIELN MIDITYREYLENMNDKRPPHIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQ IIKKG		
CdiCas9	Corynebacterium diphtheriae	MKYHVGIDVGTF SVGLAAIEVDDAG MPIKTLSLVSHIHDSGLDPDEIKSAV TRLASSGIARRTRRLYRRKRRRLQQ	H573A (альтернативная)	365

LDKFIQRQGWPVIELEDYSDPLYPW
 KVRaelAASYIADEKERGEKLSVAL
 RHIArHRGWRNPYAKVSSLYLPDGP
 SDAFKAIREEIKRASGQVPETATVG
 QMVTLCeLGTLKLRGEGGVLSARLQ
 QSDYAREIQEICRMQEIGQELYRKIID
 VVFAAESPKGSASSRVGKDPLQPGK
 NRALKASDAFQRYRIAALIGNLRVR
 VDGEKRILSVEEKNLVFDHLVNLTP
 KKEPEWVTIAEILGIDRGQLIGTATM
 TDDGERAGARPPTHDTNRSIVNSRIA
 PLVDWWKTASALEQHAMVKALSN
 AEVDDFDSPEGAKVQAFFADLDDD
 VHAKLDSLHLPVGRAAYSedTLVRL
 TRRMLSDGVDLYTARLQEFGIEPSW
 TPPTPRIGEPVGNPAVDRVLKTVSR
 WLESATKTWGAPERVIIIEHVREGFV
 TEKRAREMDGDMRRRAARNAKLfq
 EMQEKLNvQGKPSRADLWRYQSVQ
 RQNCQCAYCGSPITFSNSEMDHIVPR
 AGQGStNTRENlVAVChRCNqSKG
 NTPFAIWAKNTSIEGVSVKEAVERTR
 HWVTDtGMRSTDFKKfTKAVVERF
 QRATMDEEIDARSMESVAWMANEL
 RSRVAQHfASHGTTVRVYRGSLTAE
 ARRASGISGKLKFFDGVGKSRLDRR
 HHAI DA AVIAFTSDYVAETLAVRSN
 LKQSQAHRQEAPQWREFTGKDAEH
 RAAWRVWCQKMEKLSALLTEDLRD
 DRVVVMSNVRLRLGNGSAHKETIG
 KLSKVKLSSQLSVSDIDKASSEALW
 CALTREPGFDPKEGLPANPERHIRVN
 GTHVYAGDNIGLFPVSAGSIALRGG
 YAE LGSSFH HARVYKITSGKKPAFA
 MLRVYTIDLLPYRNQDLFSVELKPQ

		TMSMRQAEKKLRDALATGNAEYLG WLVDDEL VVDTSKIATDQVKAVE AELGTIRRWRVDGFFSPSKLRLRPLQ MSKEGIKKESAPELSKIIDRPGWLPA VNKLFS DGNVTVVRRDSLGRVRLES TAHLPVTWKVQ		
CjeCas9	Campylobacter jejuni	MARILAFDIGISSIGWAFSENDELKD CGVRIFTKVENPKTGESLALPRRLAR SARKRLARRKARLNHLKHLIANEFK LNYEDYQSFDESLAKAYKGSISPYE LRFRALNELLSKQDFARVILHIAKRR GYDDIKNSDDKEKGAILKAIKQNEE KLANYQSVGEYLYKEYFQKFENS KEFTNVRNKKESYERCIAQSFLKDEL KLIFKKQREFGFSFSKKFEEEVLSVA FYKRALKDFSHLVGNCSFFTDEKRA PKNSPLAFMFVALTRIINLLNNLKNT EGILYTKDDL NALLNEVLKNGTLTY KQTKKLLGLSDDYEFKGEKGTYFIE FKKYKEFIKALGEHNL SQDDLNEIA KDITLIKDEIKLKKALAKYDLNQNQI DSL SKLEFKDHLNISFKALKLVTPLM LEGKKYDEACNELNLKVAINEDKK DFLPAFNETYYKDEV TNPVVLRAIK EYRKVLNALLKKYGKVHKINIELAR EVGKNHSQRAKIEKEQNENYKAKK DAELECEKLGLKINSKNILKLRLFKE QKEFCAYS GEKIKISDLQDEKMLEID HIYPYSRSFDDSYMKNVLVFTKQNQ EKLNQTPFEAFGNDS AKWQKIEVLA KNLPTKKQKRILDKNYKDKEQKNF KDRNLNDTRYIARLV LNYTKDYLDF LPLSDDENTKLNDTQKGSKVHVEAK SGMLTSALRHTWGFS AKDRNNHLH HAIDAVIIAYANNSIVKAFSDFKKEQ	N582A	366

		ESNSAELYAKKISELDYKNKRKFFEP FSGFRQKVLDKIDEIFVSKPERKKPS GALHEETFRKEEEFYQSYGGKEGVL KALELGKIRKVNGKIVKNGDMFRV DIFKHKKTNKFYAVPIYTMDFALKV LPNKAVARSKKGEIKDWILMDENYE FCFSLYKDSLILIQTKDMQEPEFVYY NAFTSSTVSLIVSKHDNKFETLSKNQ KILFKNANEKEVIAKSIGIQNLKVFE KYIVSALGEVTKAEFRQREDFKK		
GeoCas9	Geobacillus stearothermo philus	MRYKIGLDIGITSVGWAVMNLDIPRI EDLGVRIFDRAENPQTGESLALPRRL ARSARRRLRRRKHRLERIRRLVIREG ILTKEELDKLFEEKHEIDVWQLRVEA LDRKLNDELARVLLHLAKRRGFKS NRKSERSNKENSTMLKHIEENRAILS SYRTVGEMIVKDPKFALHKRNKGEN YTNTIARDDLEREIRLIFSKQREFGN MSCTEEFENEYITIWASQRPVASKDD IEKKVGFCFEPKEKRAPKATYTFQS FIAWEHINKLRLISPSGARGLTDEER RLLYEQAFQKNKITYHDIRTLLHLPD DTYFKGIVYDRGESRKQENIRFLEL DAYHQIRKAVDKVYGGKSSSFLPI DFDTFGYALTLFKDDADIHSYLRNE YEQNGKRMPNLANKVYDNELIEELL NLSFTKFGHLSLKALRSILPYMEQGE VYSSACERAGYTFTGPKKKQKTMLL PNIPPIANPVVMRALTQARKVVNAII KKYGSPVSIHIELARDLSQTFDERRK TKKEQDENRKKNETAIRQLMEYGLT LNPTGHDIVKFKLWSEQNGRCAYSL QPIEIERLLEPGYVEVDHVIPYSRSLD DSYTNKVLVLTRENREKGNRIPA EY LGVGTERWQQFETFVL TNKQFSKKK	N605A	367

		RDRLLRLHYDENEETEFKNRNLNDT RYISRFFANFIREHLKFAESDDKQKV YTVNGRVT AHLRSRWEFNKNREES DLHHA VDAVIVACTTPSDIAKVTAF YQRREQNKELAKKTEPHFPQPWPHF ADELRARLSKHPKESIKALNLGNYD DQKLESLQPVFVSRMPKRSVTGAAH QETLRRYVGIDERSGKIQT VVKTKLS EIKLDASGHFPMYGKESDPRTYEAIR QRLLEHNNDPKKAFQEPLYKPKKNG EPGPVIRTVKIIDTKNQVIPLNDGKT VAYNSNIVRVDVFEEKDGKYYCVPV YTMDIMKGILPNKAIEPNKPYSEWK EMTEDYTFRFSLYPNDLIRIELPREKT VKTAAGEEINVKDV FVYYKTIDSAN GGLELISHDHRFSLRGVGSRTLKRFE KYQVDVLGNIYKVRGEKRVGLASS AHSKPGKTIRPLQSTRD		
iSpyMacC as9	Streptococcus spp.	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINAS GVDAKAILSARLSKSRKLENLIAQLP GEKKNGLFGNLIASLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLKREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFY PFLKD	N863A	368

NREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEEVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPKHS
 LLYEYFTVYNELTKVKYVTEGMRK
 PAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRF
 NASLGTYHDLLKIIKDKDFLDNEENE
 DILEDIVLTLTLFEDREMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSR
 KLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSG
 QGDSLHEHIANLAGSPAIIKKGILQTV
 KVVDEL VKVMGRHKPENIVIAMARE
 NQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYL
 QNGRDMYVDQELDINRLSDYVDH
 IVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNT
 KYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVBREINNYHHAHDAYLNA
 VVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNI
 MNFFKTEITLANGEIRKRPLIETNGET
 GEIVWDKGRDFATVRKVLSPQVNI
 VKKTEIQTVGQNGGLFDDNPKSPLE
 VTPSKLVPLKKELNPKKYGGYQKPT
 TAYPVLLITDTKQLIPISVMNKKQFE
 QNPVKFLRDRGYQQVGKNDFIKLPK
 YTLVDIGDGIKRLWASSKEIHKGNQ
 LVVSKKSQILLYHAHHLSDLSNDY
 LQNHNQFQDVLFNEIISFSKKCKLGK
 EHIQKIENVYSNKKNSASIEELAESFI
 KLLGFTQLGATSPFNFLGVKLNQKQ

		YKGKKDYILPCTEGTLIRQSITGLYE TRVDLSKIGEDSGGSGGSKRTADGS EFES		
NmeCas9	Neisseria meningitidis	MAAFKPNSINYILGLDIGIASVGWA MVEIDEEENPIRLIDLGVRFERA PKTGDSLAMARRLARSVRRLTRRA HRLLRTRLLKREGVLQAANFDENG LIKSLPNTPWQLRAAALDRKLTPL WSAVLLHLIKHRGYLSQRKNEGETA DKELGALLKGVAGNAHALQTGDFR TPAELALNKFEKESGHIRNQSDYSH TFSRKDLQAEILLFEKQKEFGNPHV SGGLKEGIETLLMTQRPALSGDAVQ KMLGHCTFEPAPKAANKNTYTAERF IWLTKLNNLRILEQGSRPLDTERA TLMDEPYRKSKLTYAQARKLLGLED TAFKGLRYGKDNAEASTLMEMKA YHAISRALEKEGLKDKKSPLNLSPEL QDEIGTAFSLFKTDEDITGRLKDRIQP EILEALLKHISFDKFVQISLKALRRIV PLMEQGKRYDEACAEIYGDHYGKK NTEEKIYLPPIPADEIRNPVVLRLALSQ ARKVINGVVRRYGSPARIHIETAREV GKSFKDRKEIEKRQEENRKDREKAA AKFREYFPNFVGEPSKDILKLRLYE QQHGKCLYSGKEINLGRLEKGYVE IDHALPFSRTWDDSFNNKVLVLGSE NQNKGNQTPYEYFNGKDNSREWQE FKARVETSRFPRSKKQRILLQKFDED GFKERNLNDTRYVNRFLCQFVADR MRLTGKGKKRVFASNGQITNLLRGF WGLRKVRAENDRHHALDAVVVAC STVAMQQKITRFVRYKEMNAFDGK TIDKETGEVLHQKTHFPQPWEFFAQ EVMIRVFGKPDGKPEFEEADTLEKL	N611A	369

		RTLLAEKLSSRPEAVHEYVTPLFVSR APNRKMSGQGHMETVKSARKLDEG VSVLRVPLTQLKLDLEKMVNRERE PKLYEALKARLEAHKDDPAKAFAP FYKYDKAGNRTQQVKAVRVEQVQ KTGVWVRNHNGIADNATMVRVDV FEKGDKYYLVPIYSWQVAKGILPDR AVVQKGDEEDWQLIDDSFNFKFSLH PNDLVEVITKKARMFGYFASCHRG GNINIRIHDLDHKIGKNGILEGIGV KTSFQKYQIDELGKEIRPCRLKKR PPVR		
ScaCas9	Streptococcus canis	MEKKYSIGLDIGTNSVGWAVITDDY KVPSKKFKVLGNTNRKSIKKNLMGA LLFDSGETAEATRLKRTARRRYTRR KNRIRYLQEIFANEMAKLDDSFQRL EESFLVEEDKKNERHPIFGNLADEVA YHRNYPTIYHLRKKLADSPEKADLR LIYLALAHIIKFRGHFLIEGKLNAENS DVAKLFYQLIQTYNQLFEESPLDEIE VDAKGILSARLSKSKRLEKLIAVFPN EKKNGLFGNIIALALGLTPNFKSNFD LTEDAKLQLSKDTYDDDLDELLGQI GDQYADLFSAAKNLSDAILLSDILRS NSEVTKAPLSASMVKRYDEHHQDL ALLKTLVRQQFPEKYAEIFKDDTKN GYAGYVGIGIKHRKRTTKLATQEEF YKFIKPILEKMDGAEELLAKLNRDD LLRKQRTFDNGSIPHQIHLKELHAIL RRQEEFYFPLKENREKIEKILTFRIPY YVGPLARGNSRFAWLTRKSEEAITP WNFEEVVDKGASAQSFIERMTNFDE QLPNKKVLPKHSLLEYFTVYNELT KVKYVTERMRKPEFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECF	N872A	370

		DSVEIIGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLK RRHYTGWGRLSRKMINGIRDKQSG KTILDFLKSDGFSNRNFMQLIHDDSL TFKEEIEKAQVSGQGDSLHEQIADLA GSPAIIKKGILQTVKIVDELVKVMGH KPENIVIEMARENQTTTKGLQQSRER KKRIEEGIKELESQILKENPVENTQL QNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFIKDDSIDNKV LTRSVENRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAER GGLSEADKAGFIKRQLVETRQITKH VARILDSRMNTKRDKNDKPIREVKV ITLKSCLVSDFRKDFQLYKVRDINNY HHAHDAYLNAVVGTAIIKKYPKLES EFVYGDYKVYDVRKMIKSEQEIGK ATAKRFFYSNIMNFFKTEVKLANGEI RKRPLIETNGETGEVVWNKEKDFAT VRKVLAMPQVNIVKKTEVQTGGFS KESILSKRESAKLIPRKKGWDTRKY GGFGSPTVAYSILVVAKVEKGKAKK LKS VKVLVGITIMEKGSYEKDPIGFL EAKGYKDIKKELIFKLPKYSLFELEN GRRRMLASATELQKANELVLPQHL VRLLYYTQNISATTGSNNLGYIEQHR EEFKEIFEKIIDFSEKYILKNKVNSNL KSSFDEQFAVSDSILLSNSFVSLLKYT SFGASGGFTFLDLVDKQGRLRYQTV TEVL DATLIYQSITGLYETRTDLSQL GGD		
ScaCas9- HiFi-Sc++	Streptococcus canis	MEKKYSIGLDIGTNSVGWAVITDDY KVPSKKFKVLGNTNRKSIKKNLMGA LLFDSGETAEATRLKRTARRRYTRR	N872A	371

KNRIRYLQEIFANEMAKLDDSFQRL
EESFLVEEDKKNERHPIFGNLADEVA
YHRNYPTIYHLRKKLADSPEKADLR
LIYLALAHIIKFRGHFLIEGKLNAENS
DVAKLFYQLIQTYNQLFEESPLDEIE
VDAKGILSARLSKSKRLEKLIIVFPN
EKKNGLFGNIIALALGLTPNFKSNFD
LTEDAKLQLSKDTYDDDLDELLGQI
GDQYADLFSAAKNLSDAILLSDILRS
NSEVTKAPLSASMVKRYDEHHQDL
ALLKTLVRQQFPEKYAEIFKDDTKN
GYAGYVGADKKLRKRSGKLATEEE
FYKFIKPILEKMDGAEELLAKLNRD
DLLRKQRTFDNGSIPHQIHLKELHAI
LRRQEEFYFPLKENREKIEKILTFRIP
YYVGPLARGNSRFAWLTRKSEEAIT
PWNFEEVVDKGASAQSFIERMTNFD
EQLPNKKVLPKHSLLYEYFTVYNEL
TKVKYVTERM RKPEFLSGEQKKAIV
DLLFKTNRKVTVKQLKEDYFKKIEC
FDSVEIIGVEDRFNASLGTYHDLLKII
KDKDFLDNEENEDILEDIVLTLTLFE
DREMIEERLKTYAHLFDDKVMKQL
KRRHYTGWGRLSRKMINGIRDKQS
GKTILDFLKSDGFSNANFMQLIHDDS
LTFKEEIEKAQVSGQGDSLHEQIADL
AGSPAIIKKGILQTVKIVDELVKVMG
HKPENIVIAMARENQTTTKGLQQSR
ERKKRIEEGIKELESQILKENPVENTQ
LQNEKLYLYYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFIKDDSIDNK
VLTRSVENRGKSDNVPSEEVVKKM
KNYWRQLLNAKLITQRKFDNLTKA
ERGGLSEADKAGFIKRQLVETRQITK
HVARILDSRMNTKRDKNDKPIREVK

		VITLKSCLVSDFRKDFQLYKVRDINN YHHAHDAYLNAVVGTAIIKKYPKL ESEFVYGDYKVYDVRKMIKSEQEI GKATAKRFFYSNIMNFFKTEVKLAN GEIRKRPLIETNGETGEVVWNKEKD FATVRKVLAMPQVNIVKKTEVQTG GFSKESILSKRESAKLIPRKKGWDTR KYGGFGSPTVAYSILVVAKEKGKA KKLKS VKVLVGITIMEKGSYEKDP FLEAKGYKDIKKELIFKLPKYSLFEL ENGRRRMLASAKELQKANELVLPQ HLVRLLYYTQNISATTGSNNLGYIEQ HREEFKEIFEKIIDFSEKYILKNKVNS NLKSSFDEQFAVSDSILLSNSFVSLK YTSFGASGGFTFLDLVKQGRRLRYQ TVTEVLDTLIYQSITGLYETRTDLS QLGGD		
SpyCas9- 3var- NRRH	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFILVQTYNQLFEENPINAS GVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSAILLSDIL RVNTEITKAPLSASMVKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG IIPHQIHLGELHAILRRQGDFYPFLKD NREKIEKILTFRIPYYVGPLARGNSRF	N863A	372

AWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENE
DILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRLRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSG
QGDSLHEHIANLAGSPAIKKGILQTV
KVVDDELVKVMGGHKPENIVIEMAR
ENQTTQKGQKNSRERMKRIEEGIKE
LGSQILKEHPVENTQLQNEKLYLYY
LQNGRDMYVDQELDINRLSDYDVD
HIVPQSFLKDDSIDNKVLTRSDKNRG
KSDNVPSEEVVKKMKNYWRQLLNA
KLITQRKFDNLTKAERGGLSELDKA
GFIKRQLVETRQITKHVAQILDSRMN
TKYDENDKLIREVKVITLKSCLVSDF
RKDFQFYKVREINNYHHAHDAYLN
AVVGTALIKKYPKLESEFVYGDYKV
YDVRKMIKSEQEIGKATAKYFFYS
NIMNFFKTEITLANGEIRKRPLIETNG
ETGEIVWDKGRDFATVRKVLSMPQ
VNIVKKTEVQTGGFSKESILPKGNSD
KLIARKKDWDPKKYGGFNSPTAAYS
VLVVAKVEKGKSKKLKSVKELLGIT
IMERSSEKNPIGFLEAKGYKEVKKD
LIIKLPKYSLELENGRKRMLASAGV
LHKGNELALPSKYVNFLYLASHYEK
LKGSPEDNEQKQLFVEQHKHYLDEII
EQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGVPA
AFKYFDTTIDKKRYTSTKEVLDATLI

		HQSITGLYETRIDLSQLGGD		
SpyCas9- 3var- NRTH	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMVKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG IIPHQIHLGELHAILRRQGDFYPFLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRLRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIAKKGILQTV KVVDELVKVMGGHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKE LGSQILKEHPVENTQLQNEKLYLYY LQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRG	N863A	373

		KSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKA GFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSCLVSDF RKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKV YDVRKMIKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSPQ VNIVKKTEVQTGGFSKESILPKGNSD KLIARKKDWDPKKYGGFNSPTVAYS VLVVAKVEKGKSKKLKSVKELLGIT IMERSSEKNPIGFLEAKGYKEVKKD LIKLPKYSLELENGRKRMLASASV LHKGNELALPSKYVNFLYLASHYEK LKGSSDNKQKQLFVEQHKHYLDEI IEQISEFSKRVILADANLDKVL SAYN KHRDKPIREQAENIIHLFTLTNLGAS AAFKYFDTTIGRKLYTSTKEVL DAT LIHQ SITGLYETRIDLSQLGGD		
SpyCas9- 3var- NRCH	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNE MAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIASLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLS DAILLSDIL RVNTEITKAPLSASMVKRYDEHHQD LTLLKALVRQQLP EKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM	N863A	374

DGTEELLVKLNREDLLRKQRTFDNG
IIPHQIHLGELHAILRRQGDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGAS
AQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRK
PAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENE
DILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRLRYTGWGRLSR
KLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSG
QGDSLHEHIANLAGSPAIIKKGILQTV
KVVDLVKVMGGHKPENIVIEMAR
ENQTTQKGQKNSRERMKRIEEGIKE
LGSQILKEHPVENTQLQNEKLYLYY
LQNGRDMYVDQELDINRLSDYDVD
HIVPQSFLKDDSIDNKVLTRSDKNRG
KSDNVPSEEVVKMKMKNYWRQLLNA
KLITQRKFDNLTKAERGGLSELDKA
GFIKRQLVETRQITKHVAQILDSRMN
TKYDENDKLIREVKVITLKSCLVSDF
RKDFQFYKVREINNYHHAHDAYLN
AVVGTAIIKKYPKLESEFVYG DYKV
YDVRKMIKSEQEIGKATAKYFFYS
NIMNFFKTEITLANGEIRKRPLIETNG
ETGEIVWDKGRDFATVRKVLSMPQ
VNIVKKTEVQTGGFSKESILPKGNSD
KLIARKKDWDPPKYGGFNSPTVAYS
VLVVAKVEKGKSKKLKSVKELLGIT
IMERSSEKNPIDFLEAKGYKEVKKD
LIIKLPKYSLFELENGRKRMLASAGV
LQKGNELALPSKYVNFLYLASHYEK
LKGSPEDNEQKQLFVEQHKHYLDEII

		EQISEFSKRVILADANLDKVLSAYNK HRDKPIREQAENIIHLFTLTNLGAPA AFKYFDTTINRKQYNTTKEVLDTLI RQSITGLYETRIDLSQLGGD		
SpyCas9- HF1	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIIKKGILQTV KVVDDELVKVMGRHKPENIVIAMARE NQTTQKGQKNSRERMKRIEEGIKEL	N863A	375

		GSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGK SDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVBREINNYHHAHDAYLNA VVGITALIKKYPKLESEFVYGDYKVY DVRKMIKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVILADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDTLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- QQR1	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIASLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILSDIL	N863A	376

RVNTEITKAPLSASMIKRYDEHHQD
 LTLKALVRQQQLPEKYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNG
 SIPHQIHLGELHAILRRQEDFYFPLKD
 NREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEEVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPKHS
 LLYEYFTVYNELTKVKYVTEGMRK
 PAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRF
 NASLGTYHDLLKIIKDKDFLDNEENE
 DILEDIVLTLTLFEDREMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSR
 KLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSG
 QGDSLHEHIANLAGSPAIKKGILQTV
 KVVDELVKVMGRHKPENIVIAMARE
 NQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYL
 QNGRDMYVDQELDINRLSDYVDH
 IVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNT
 KYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVBREINNYHHAHDAYLNA
 VVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNI
 MNFFKTEITLANGEIRKRPLIETNGET
 GEIVWDKGRDFATVRKVLSPQVNI
 VKKTEVQTGGFSKESILPKRNSDKLI
 ARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIM
 ERSSFENPIDFLEAKGYKEVKKDLII

		KLPKYSLFELNGRKRMLASARELQ KGNELALPSKYVNFLYLASHYEKLLK GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVILADAQLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTFKQKQYRSTKEVLDTLIH QSITGLYETRIDLSQLGGD		
SpyCas9- SpG	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQVLVQTYNQLFEENPINAS GVDAKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIASLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSDAILSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPKEYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSG	N863A	377

		QGDSLHEHIANLAGSPAIIKKGILQTV KVVDEL VKVMGRHKPENIVIAMARE NQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYVDH IVPQSFLKDDSIDNKVLTRSDKNRGK SDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVBREINNYHHAHDAYLNA VVGITALIKKYPKLESEFVYGDYKVY DVRKMIAKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFLWPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAKQLQ KGNELALPSKYVNFLYLASHYEKLG GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVILADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- VQR	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAAKILSARLSKSRRLLENLIAQLP	N863A	378

GEKKNGLFGNLIASLGLTPNFKSNF
 DLAEDAKLQLSKDTYDDDLNLLA
 QIGDQYADLFLAAKNLSDAILSDIL
 RVNTEITKAPLSASMIKRYDEHHQD
 LTLKALVRQQLPKEYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNG
 SIPHQIHLGELHAILRRQEDFYFPLKD
 NREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEEVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPKHS
 LLYEYFTVYNELTKVKYVTEGMRK
 PAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRF
 NASLGTYHDLLKIIKDKDFLDNEENE
 DILEDIVLTLTLFEDREMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSR
 KLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSG
 QGDSLHEHIANLAGSPAIKKGILQTV
 KVVDELVKVMGRHKPENIVIAMARE
 NQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYL
 QNGRDMYVDQELDINRLSDYDVDH
 IVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNT
 KYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVBREINNYHHAHDAYLNA
 VVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNI
 MNFFKTEITLANGEIRKRPLIETNGET
 GEIVWDKGRDFATVRKVLSMPQVNI
 VKKTEVQTGGFSKESILPKRNSDKLI

		ARKKDWDPKKYGGFVSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLLK GSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVLADANLDKVL SAYNKH DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKQYRSTKEVL DATLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- VRER	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAIL SARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLNLLA QIGDQYADLFLAAKNLSAILLSDIL RVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA	N863A	379

		HLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSG QGDSLHEHIANLAGSPAIAKKGILQTV KVVDELVKVMGRHKPENIVIAMARE NQTTQKGQKNSRERMKRIIEGIKEL GSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYVDH IVPQSFLKDDSIDNKVLTRSDKNRGK SDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVREINNYHHAHDAYLNA VVGTAIIKKYPKLESEFVYGDYKVY DVRKMIAKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFVSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLELENGRKRMLASARELQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIEQ ISEFSKRVLADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKEYRSTKEVLDTLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- xCas	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR	N863A	380

LIYLALAHMIKFRGHFLIEGDLNPDN
 SDVDKLFQVLVQTYNQLFEENPINAS
 GVDAILKARLSKSRLENLIAQLP
 GEKKNGLFGNLIASLGLTPNFKSNF
 DLAEDTKLQLSKDTYDDDLDNLLA
 QIGDQYADLFLAAKNLSDAILLSDIL
 RVNTEITKAPLSASMIKLYDEHHQD
 LTLLKALVRQQLEKYKEIFFDQSKN
 GYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNG
 IIPHQIHLGELHAILRRQEDFYFPLKD
 NREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEKVVDKGA
 SAQSFIERMTNFDKNLPNEKVLPHKS
 LLYEYFTVYNELTKVKYVTEGMRK
 PAFLSGDQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRF
 NASLGTYHDLLKIIKDKDFLDNEENE
 DILEDIVLTLTLFEDREMIEERLKTYA
 HLFDDKVMKQLKRRRYTGWGRLSR
 KLINGIRDKQSGKTILDFLKSDGFAN
 RNFIQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVK
 VVDELVKVMGRHKPENIVIAMAREN
 QTTQKGQKNSRERMKRIEELG
 SQILKEHPVENTQLQNEKLYLYYLQ
 NGRDMYVDQELDINRLSDYDVDHI
 VPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNT
 KYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNA
 VVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNI

		MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFENPIDFLEAKGYKEVKKDLII KLPKYSLELENGRKRMLASAGVLQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVLADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD		
SpyCas9- xCas-NG	Streptococcus pyogenes	MDKKYSIGLDIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINAS GVDAKAILSARLSKSRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNF DLAEDTKLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKLYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNG IIPHQIHLGELHAILRRQEDFYFPLKD NREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEKVVDKGA SAQSFIERMTNFDKNLPNEKVLPKHS LLYEYFTVYNELTKVKYVTEGMRK PAFLSGDQKKAIVDLLFKTNRKVTV	N863A	381

		KQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFAN RNFIQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIIKKGILQTVK VVDELVKVMGRHKPENIVIAMAREN QTTQKGQKNSRERMKRIEEGIKELG SQILKEHPVENTQLQNEKLYLYYLQ NGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKNRGK SDNVPSEEVVKKMKNYWRQLLNAK LITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVBREINNYHHAHDAYLNA VVGTAIIKKYPKLESEFVYGDYKVY DVRKMIKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSPQVNI VKKTEVQTGGFSKESIRPKRNSDKLI ARKKDWDPKKYGGFVSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIM ERSSFEKNPIDFLEAKGYKEVKKDLII KLPKYSLELENGRKRMLASARFLQ KGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIEEQ ISEFSKRVILADANLDKVL SAYNKHR DKPIREQAENIIHLFTLTNLGAPRAFK YFDTTIDRKVYRSTKEVL DATLIHQ ITGLYETRIDLSQLGGD		
St1Cas9- CNRZ106 6	Streptococcus thermophilus	MSDLVLGLDIGIGSVGVGILNKVTGE IIHKNSRIFPAAQAENNLVRRITNRQG RRLARRKKHRRVRLNRLFEEGLIT	N622A	382

DFTKISINLNPYQLRVKGLTDELSNE
ELFIALKNMVKHRGISYLLDDASDDG
NSSVGDYAQIVKENSQLETKTPGQI
QLERYQTYGQLRGDFTVEKDGGKKH
RLINVFPTSAYRSEALRILQTQQEFNP
QITDEFINRYLEILTGKRKYHGPNG
EKSRTDYGRYRTSGETLDNIFGILIG
KCTFYPDEFRAAKASYTAQEFNLLN
DLNNLTVPTETKKLSKEQKNQIINYV
KNEKAMGPAKLFKYIAKLLSCDVA
DIKGYRIDKSGKAEIHTFEAYRKMK
TLETLDIEQMDRETLDKLAYVLTNL
TEREGIQEALHEFADGSFSQKQVDE
LVQFRKANSSIFGKGWHNFSVKLM
MELIPELYETSEEQMTILTRLGKQKT
TSSSNKTKYIDEKLLTEEIYNPVVAK
SVRQAIKIVNAAIKEYGDFDNIVIEM
ARETNEDDEKKAIQKIQKANKDEKD
AAMLKAANQYNGKAELPHSVFHGH
KQLATKIRLWHQQGERCLYTGKTISI
HDLINNSNQFEVDHILPLSITFDDSLA
NKVLVYATANQEKGQRTPYQALDS
MDDAWSFRELKAFVRESKTLNKK
KEYLLTEEDISKFDVRKKFIERNLVD
TRYASRVVLNALQEHFRAHKIDTKV
SVVRGQFTSQLRRHWGIEKTRDTYH
HHAVDALIAASSQLNLWKKQKNTL
VSYSEEQLLDIETGELISDDEYKESVF
KAPYQHFDVTLKSKEFEDSILFSYQV
DSKFNRKISDATIYATRQAKVGKDK
KDETYVLGKIKDIYTQDGYDAFMKI
YKKDKSKFLMYRHDPQTFEKVIEPIL
ENYPNKQMNEKGKEVPCNPFLKYK
EEHGYIRKYSKKGNGPEIKSLKYD
SKLLGNPIDITPENSKNKVVLQSLKP

		WRTDVYFNKATGKYEILGLKYADL QFEKGTGTYKISQEKYNDIKKKEGV DSDSEFKFTLYKNDLLL VKDTETKE QQLFRFLSRTL PKQKHVELKPYDK QKFEGGEALIKVLGNVANGGQCIKG LAKSNISYKVRTDVLGNQHIIKNEG DKPKLDF		
St1Cas9- LMG1831	Streptococcus thermophilus	MSDLVLGLDIGIGSVGVGILNKVTGE IHHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEESSLIT DFTKISINLNPYQLRVKGLTDELSNE ELFIALKNMVKHRGISYLLDDASDDG NSSVG DY AQIVKENS KQLETKTPGQI QLERYQTYGQLRGDFTVEKD GK KH RLINVFPTSAYRSEALRILQTQQEFNP QITDEFINRYLEILTGKRKYHGPNG EKSRTDYGRYRTSGETLDNIFGILIG KCTFYPDEFRAAKASYTAQEFNLLN DLNNLTVPTETKKLSKEQKNQIINYV KNEKAMGPAKLFKYIAKLLSCDVA DIKGYRIDKSGKAEIHTFEAYRKMK TLETLDIEQMDRETLDKLA YVLT LN TEREQIEALEHEFADGSFSQKQVDE LVQFRKANSSIFGKGWHNFSVKLM MELIPELYETSEEQMTILTRLGKQKT TSSSNKTKYIDEKLLTEEIYNPVVAK SVRQAIKIVNAAIKEYGDFDNIVIEM ARETNEDDEKKA IQKI QKANKDEKD AAMLKAANQYNGKAELPHSVFHGH KQLATKIRLWHQQGERCLYTGKTISI HDLINNSNQFEVDHILPLSITFDDSLA NKVLVYATANQEKGQRTPYQALDS MDDAWSFRELKAFVRESKTL SNKK KEYLLTEEDISKFDVRKKFIERNLVD TRYASRVVLNALQEHFRAHKIDTKV	N622A	383

		SVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTL VSYSEEQLLDIETGELISDDEYKESVF KAPYQHFVDTLKSKEFEDSILFSYQV DSKFNRKISDATIYATRQAKVGKDK KDETYVLGKIKDIYTQDGYDAFMKI YKKDKSKFLMYRHDPQTFEKVIEPIL ENYPNKQMNEKGKEVPCNPFLKYK EEHGYIRKYSKKGNGPEIKSLKYYD SKLLGNPIDITPENSKNKVVLQSLKP WRTDVYFNKNTGKYEILGLKYADL QFEKKTGTYKISQEKYNGIMKEEGV DSDSEFKFTLYKNDLLL VKDTETKE QQLFRFLSRTMPNVKYYVELKPYSK DKFEKNESLIEILGSADKSGRCIKGL GKSNISIIYKVRTDVLGNQHIIKNEGD KPKLDF		
St1Cas9- MTH17C L396	Streptococcus thermophilus	MSDLVLGLDIGIGSVGVGILNKVTGE IIHKNSRIFPAAQAENNLVRRITNRQG RRLARRKKHRRVRLNRLFEEGLIT DFTKISINLNPYQLRVKGLTDELSNE ELFIALKNMVKHRGISYLLDDASDDG NSSVG DY AQIVKENS KQLETKTPGQI QLERYQTYGQLRGDFTVEKDGGKH RLINVFPTSAYRSEALRILQTQQEFNP QITDEFINRYLEILTGKRKYHGPNG EKSRTDYGRYRTSGETLDNIFGILIG KCTFYPDEFRAAKASYTAQEFNLLN DLNNLTVPTETKKLSKEQKNQIINYV KNEKAMGPAKLFKYIAKLLSCDVA DIKGYRIDKSGKAEIHTFEAYRKMK TLETLDIEQMDRETLDKLAYVLTNL TEREGIQEALHEFADGSFSQKQVDE LVQFRKANSSIFGKGWHNFSVKLM MELIPELYETSEEQMTILTRLGKQKT	N622A	384

		TSSSNKTKYIDEKLLTEEIYNPVVAK SVRQAIKIVNAAIKEYGDFDNIVIEM ARETNEDDEKKAIQKIQKANKDEKD AAMLKAANQYNGKAELPHSVFHGH KQLATKIRLWHQQGERCLYTGKTISI HDLINNSNQFEVDHILPLSITFDDSLA NKVLVYATANQEKGQRTPYQALDS MDDAWSFRELKAFVRESKTLNKK KEYLLTEEDISKFDVRKKFIERNLVD TRYASRVVLNALQEHFRAHKIDTKV SVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTL VSYSEDQLLDIETGELISDDEYKESV FKAPYQHFVDTLKSKEFEDSILFSYQ VDSKFNRKISDATIYATRQAKVGKD KADETYVLGKIKDIYTQDGYDAFM KIYKKDKSKFLMYRHDPQTFEKVIE PILENYPNKQINEKGKEVPCNPFLKY KEEHGYIRKYSKKGNGPEIKSLKYY DSKLGNHIDITPKDSNNKVVLQSLKP WRTDVYFNKNTGKYEILGLKYSDM QFEKGTGKYSISKEQYENIKVREGV DENSEFKFTLYKNDLLLLKDSenge QILLRFTSRNDTSKHYVELKPYNRQ KFEGSEYLIKSLGTVAKGGQCIKGLG KSNISIYKVRTDVLGNQHIIKNEGDK PKLDF		
St1Cas9-TH1477	Streptococcus thermophilus	MSDLVLGLDIGIGSVGVGILNKVTGE IIHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEEGLIT DFTKISINLNPYQLRVKGLTDELSNE ELFIALKNMVKHRGISYLLDDASDDG NSSVGDYAQIVKENSQLETKTPGQI QLERYQTYGQLRGDFTVEKDGGKH RLINVFPTSAYRSEALRILQTQQEFNP	N622A	385

QITDEFINRYLEILTGKRKYHGPNG
 EKSRTDYGRYRTSGETLDNIFGILIG
 KCTFYPDEFRAAKASYTAQEFNLLN
 DLNNLTVPTETKKLSKEQKNQIINYV
 KNEKAMGPAKLFKYIAKLLSCDVA
 DIKGYRIDKSGKAEIHTFEAYRKMK
 TLETLDIEQMDRETLDKLAYVLTN
 TEREQIEALEHEFADGSFSQKQVDE
 LVQFRKANSSIFGKGWHNFSVKLM
 MELIPELYETSEEQMTILTRLGKQKT
 TSSSNKTKYIDEKLLTEEIYNPVVAK
 SVRQAIKIVNAAIKEYGDFDNIVIEM
 ARETNEDDEKKAIQKIQKANKDEKD
 AAMLKAANQYNGKAELPHSVFHHG
 KQLATKIRLWHQQGERCLYTGKTISI
 HDLINNSNQFEVDHILPLSITFDDSLA
 NKVLVYATANQEKGQRTPYQALDS
 MDDAWSFRELKAFVRESKTLNKK
 KEYLLTEEDISKFDVRKKFIERNLVD
 TRYASRVVLNALQEHFRAHKIDTKV
 SVVRGQFTSQLRRHWGIEKTRDTYH
 HHAVDALIAASSQLNLWKKQKNTL
 VSYSEDQLLDIETGELISDDEYKESV
 FKAPYQHFVDTLKSKEFEDSILFSYQ
 VDSKFNRKISDATIYATRQAKVGKD
 KADETYVLGKIKDIYTQDGYDAFM
 KIYKKDKSKFLMYRHDPQTFEKVIE
 PILENYPNKQINEKGKEVPCNPFLKY
 KEEHGYIRKYSKKGNGPEIKSLKYY
 DSKLGNHIDITPKDSNNKVVLQSLKP
 WRTDVYFNKNTGKYEILGLKYSMD
 QFEKGTGKYSISKEQYENIKVREGV
 DENSEFKFTLYKNDLLLLKDSERGE
 QILLRFTSRNDTSKHVELKPYNRQ
 KFEGSEYLIKSLGTVVKGGRCIKGLG

		KSNISIIYKVRTDVLGNQHIIKNEGDK PKLDF		
--	--	--------------------------------------	--	--

В таблице 3В представлены параметры для определения необходимых компонентов для конструирования gRNA и/или РНК-матриц в применении вариантов Cas, перечисленных в таблице 3А, для Gene Writing. Ярус указывает на предпочтительные варианты Cas, если они доступны для использования в данном локусе. Место разреза указывает на валидированные или предсказанные требования для мотива, прилегающего к протоспейсеру (PAM), валидированное или предсказанное местоположение сайта разреза (относительно наиболее вышерасположенного основания сайта PAM). gRNA для данного фермента может быть собрана посредством соединения последовательностей crRNA, тетрапетли и tracrRNA и дополнительного добавления спейсера на 5'-конце с длиной в пределах от спейсера (min) до спейсера (max), которая соответствует протоспейсеру в сайте-мишени. Кроме того, предсказанное местоположение ника ssDNA в мишени является важным для конструирования 3'-концевой области РНК-матрицы, которую необходимо отжечь с последовательностью, расположенной непосредственно рядом с ником в направлении 5'-конца, чтобы инициировать праймируемую мишенью обратную транскрипцию.

Таблица 3В. Параметры для определения необходимых компонентов для конструирования gRNA и/или РНК-матриц для применения вариантов Cas, перечисленных в таблице 3А, для Gene Writing

Вариант	PAM	Разрез	Ярус	Спейсер (min)	Спейсер (max)	crRNA	Тетрапетля	tracrRNA
Nme2Cas9	NNNN CC	-3	1	22	24	GTTGT AGCTC CCTTT CTCAT TTCG (SEQ ID NO: 386)	GAAA	CGAAATGA GAACCGTT GCTACAAT AAGGCCGT CTGAAAAG ATGTGCCG CAACGCTC TGCCCCTT AAAGCTTC TGCTTTAA GGGGCATC GTTTA (SEQ ID NO: 387)
PpnCas9	NNNN		1	21	24	GTTGT	GAAA	GCGAAATG

	RTT					AGCTC CCTTT TTCAT TTCGC (SEQ ID NO: 388)		AAAAACGT TGTTACAA TAAGAGAT GAATTTCT CGCAAAGC TCTGCCTC TTGAAATT TCGGTTTC AAGAGGCA TCTTTTT (SEQ ID NO: 389)
SauCas9	NNGR R;NNG RRT	-3	1	21	23	GTTTT AGTA CTCTG (SEQ ID NO: 390)	GAAA	CAGAATCT ACTAAAAC AAGGCAAA ATGCCGTG TTTATCTC GTCAACTT GTTGGCGA GA (SEQ ID NO: 391)
SauCas9 -KKH	NNNR R;NNN RRT	-3	1	21	21	GTTTT AGTA CTCTG TAAT (SEQ ID NO: 392)	GAAA	ATTACAGA ATCTACTA AAACAAGG CAAAATGC CGTGTTTA TCTCGTCA ACTTGTTG GCGAGA (SEQ ID NO: 393)
SauriCas 9	NNGG	-3	1	21	21	GTTTT AGTA CTCTG (SEQ	GAAA	CAGAATCT ACTAAAAC AAGGCAAA ATGCCGTG

						ID NO: 394)		TTTATCTC GTCAACTT GTTGGCGA GATTTTT (SEQ ID NO: 395)
SauriCas 9-KKH (SEQ ID NO: 401)	NNRG	-3	1	21	21	GTTTT AGTA CTCTG (SEQ ID NO: 396)	GAAA	CAGAATCT ACTAAAAC AAGGCAAA ATGCCGTG TTTATCTC GTCAACTT GTTGGCGA GATTTTT (SEQ ID NO: 397)
ScaCas9 -Sc++	NNG	-3	1	20	20	GTTTT AGAG CTA (SEQ ID NO: 398)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 399)
SpyCas9	NGG	-3	1	20	20	GTTTT AGAG CTA (SEQ ID NO: 400)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 401)
SpyCas9	NG	-3	1	20	20	GTTTA	GAAA	CAGCATAG

-NG	(NGG= NGA= NGT>N GC)					AGAG CTATG CTG (SEQ ID NO: 402)		CAAGTTTA AATAAGGC TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG C (SEQ ID NO: 403)
SpyCas9 -SpRY	NRN> NYN	-3	1	20	20	GTTTT AGAG CTA (SEQ ID NO: 404)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 405)
St1Cas9	NNAG AAW> NNAG GAW= NNGG AAW	-3	1	20	20	GTCTT TGTAC TCTG (SEQ ID NO: 407)	GTAC	CAGAAGCT ACAAAGAT AAGGCTTC ATGCCGAA ATCAACAC CCTGTCAT TTTATGGC AGGGTGTT TT (SEQ ID NO: 406)
BlatCas9	NNNN CNAA >NNN NCND D>NN NNC	-3	1	19	23	GCTAT AGTTC CTTAC T (SEQ ID NO: 408)	GAAA	GGTAAGTT GCTATAGT AAGGGCAA CAGACCCG AGGCGTTG GGGATCGC CTAGCCCG

								TGTTTACG GGCTCTCC CCATATTC AAAATAAT GACAGACG AGCACCTT GGAGCATT TATCTCCG AGGTGCT (SEQ ID NO: 409)
cCas9- v16	NNVA CT;NN VATG M;NNV ATT;N NVGC T;NNV GTG;N NVGT T	-3	2	21	21	GUCU UAGU ACUC UG (SEQ ID NO: 410)	GAAA	CAGAAUCU ACUAAGAC AAGGCAAA AUGCCGUG UUUAUCUC GUCAACUU GUUGGCGA GAUUUUU UU (SEQ ID NO: 411)
cCas9- v17	NNVR RN	-3	2	21	21	GUCU UAGU ACUC UG (SEQ ID NO: 412)	GAAA	CAGAAUCU ACUAAGAC AAGGCAAA AUGCCGUG UUUAUCUC GUCAACUU GUUGGCGA GAUUUUU UU (SEQ ID NO: 413)
cCas9- v21	NNVA CT;NN VATG M;NNV	-3	2	21	21	GUCU UAGU ACUC UG	GAAA	CAGAAUCU ACUAAGAC AAGGCAAA AUGCCGUG

	ATT;N NVGC T;NNV GTG;N NVGT T					(SEQ ID NO: 414)		UUUAUCUC GUCAACUU GUUGGCGA GAUUUUU UU (SEQ ID NO: 415)
cCas9- v42	NNVR RN	-3	2	21	21	GUCU UAGU ACUC UG (SEQ ID NO: 416)	GAAA	CAGAAUCU ACUAAGAC AAGGCAAA AUGCCGUG UUUAUCUC GUCAACUU GUUGGCGA GAUUUUU UU (SEQ ID NO: 417)
CdiCas9	NNRH HHY;N NRAA AY		2	22	22	ACUG GGGU UCAG (SEQ ID NO: 418)	GAAA	CUGAACCU CAGUAAGC AUUGGCUC GUUUCCAA UGUUGAU UGCUCGCG CGGUGCUC CUUAUUUU UAAGGGCG CCGGC (SEQ ID NO: 419)
CjeCas9	NNNN RYAC	-3	2	21	23	GTTTT AGTCC CT (SEQ ID NO: 420)	GAAA	AGGGACTA AAATAAAG AGTTTGCG GGACTCTG CGGGGTTA CAATCCCC TAAAACCG

								CTTTTTT (SEQ ID NO: 421)
GeoCas9	NNNN CRAA		2	21	23	GUCA UAGU UCCCC UGA (SEQ ID NO: 421)	GAAA	UCAGGGUU ACUAUGAU AAGGGCUU UCUGCCUA AGGCAGAC UGACCCGC GGCGUUGG GGAUCGCC UGUCGCCC GCUUUUGG CGGGCAUU CCCCAUCC UU(SEQ ID NO: 422)
iSpyMac Cas9	NAAN	-3	2	19	21	GTTTT AGAG CTA (SEQ ID NO: 423)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 44)
NmeCas 9	NNNN GAYT; NNNN GYTT; NNNN GAYA; NNNN GTCT	-3	2	20	24	GTTGT AGCTC CCTTT CTCAT TTCG (SEQ ID NO: 425)	GAAA	CGAAATGA GAACCGTT GCTACAAT AAGGCCGT CTGAAAAG ATGTGCCG CAACGCTC TGCCCCTT AAAGCTTC

								TGCTTTAA GGGGCATC GTTTA (SEQ ID NO: 426)
ScaCas9	NNG	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 427)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 428)
ScaCas9 -HiFi- Sc++	NNG	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 429)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 430)
SpyCas9 -3var- NRRH	NRRH	-3	2	20	20	GTTTA AGAG CTATG CTG (SEQ ID NO: 431)	GAAA	CAGCATAG CAAGTTTA AATAAGGC TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG C (SEQ ID NO: 432)
SpyCas9 -3var- NRTH	NRTH	-3	2	20	20	GTTTA AGAG CTATG	GAAA	CAGCATAG CAAGTTTA AATAAGGC

						CTG (SEQ ID NO: 433)		TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG C (SEQ ID NO: 434)
SpyCas9 -3var- NRCH	NRCH	-3	2	20	20	GTTTA AGAG CTATG CTG (SEQ ID NO: 435)	GAAA	CAGCATAG CAAGTTTA AATAAGGC TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG C (SEQ ID NO: 436)
SpyCas9 -HF1	NGG	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 437)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 438)
SpyCas9 -QQR1	NAAG	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 439)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 440)

SpyCas9 -SpG	NGN	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 441)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 442)
SpyCas9 -VQR	NGAN	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 443)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 444)
SpyCas9 -VRER	NGCG	-3	2	20	20	GTTTT AGAG CTA (SEQ ID NO: 445)	GAAA	TAGCAAGT TAAAATAA GGCTAGTC CGTTATCA ACTTGAAA AAGTGGCA CCGAGTCG GTGC (SEQ ID NO: 446)
SpyCas9 -xCas	NG;GA A;GAT	-3	2	20	20	GTTTA AGAG CTATG CTG (SEQ ID NO: 447)	GAAA	CAGCATAG CAAGTTTA AATAAGGC TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG

								C (SEQ ID NO: 448)
SpyCas9 -xCas- NG	NG	-3	2	20	20	GTTTA AGAG CTATG CTG (SEQ ID NO: 449)	GAAA	CAGCATAG CAAGTTTA AATAAGGC TAGTCCGT TATCAACT TGAAAAAG TGGCACCG AGTCGGTG C (SEQ ID NO: 450)
St1Cas9- CNRZ10 66	NNAC AA	-3	2	20	20	GTCTT TGTAC TCTG (SEQ ID NO: 451)	GTAC	CAGAAGCT ACAAAGAT AAGGCTTC ATGCCGAA ATCAACAC CCTGTCAT TTTATGGC AGGGTGTT TT (SEQ ID NO: 452)
St1Cas9- LMG183 1	NNGC AA	-3	2	20	20	GTCTT TGTAC TCTG (SEQ ID NO: 453)	GTAC	CAGAAGCT ACAAAGAT AAGGCTTC ATGCCGAA ATCAACAC CCTGTCAT TTTATGGC AGGGTGTT TT (SEQ ID NO: 454)
St1Cas9- MTH17 CL396	NNAA AA	-3	2	20	20	GTCTT TGTAC TCTG	GTAC	CAGAAGCT ACAAAGAT AAGGCTTC

						(SEQ ID NO: 455)		ATGCCGAA ATCAACAC CCTGTCAT TTTATGGC AGGGTGTT TT (SEQ ID NO: 456)
St1Cas9- TH1477	NNGA AA	-3	2	20	20	GTCTT TGAC TCTG (SEQ ID NO: 457)	GTAC	CAGAAGCT ACAAAGAT AAGGCTTC ATGCCGAA ATCAACAC CCTGTCAT TTTATGGC AGGGTGTT TT (SEQ ID NO: 458)

В некоторых вариантах осуществления белок Cas требует, чтобы мотив, прилегающий к протоспейсеру (PAM), присутствовал в последовательности ДНК-мишени или примыкал к ней для того, чтобы белок Cas мог связываться и/или функционировать. В некоторых вариантах осуществления PAM представляет собой следующее или содержит, в направлении от 5'- к 3'- концу, следующее: NGG, YG, NNGRRT, NNNRRT, NGA, TYCV, TATV, NTTN или NNNGATT, где N означает любой нуклеотид, Y означает C или T, R означает A или G, а V означает A, C или G. В некоторых вариантах осуществления белок Cas представляет собой белок, приведенный в таблице 4. В некоторых вариантах осуществления белок Cas содержит одну или несколько мутаций, обеспечивающих изменение его PAM. В некоторых вариантах осуществления белок Cas содержит мутации E1369R, E1449H и R1556A или аналогичные замены аминокислот, соответствующих указанным положениям.. В некоторых вариантах осуществления белок Cas содержит мутации E782K, N968K и R1015H или аналогичные замены аминокислот, соответствующих указанным положениям.. В некоторых вариантах осуществления белок Cas содержит мутации D1135V, R1335Q и T1337R или аналогичные замены аминокислот, соответствующих указанным положениям.. В некоторых вариантах осуществления белок Cas содержит мутации S542R и K607R или аналогичные замены аминокислот, соответствующих указанным положениям.. В некоторых вариантах осуществления белок Cas содержит мутации S542R, K548V и N552R или аналогичные замены аминокислот, соответствующих указанным положениям..

Таблица 4. Белки CRISPR/Cas, виды и мутации

Название	Фермент	Вид	Количество а. к.	PAM	Мутации для обеспечения изменения распознавания PAM	Мутации для обеспечения каталитической инертности
<i>FnCas9</i>	Cas9	<i>Francisella novicida</i>	1629	5'-NGG-3'	Wt	D11A/H969A/N995A
<i>FnCas9 RHA</i>	Cas9	<i>Francisella novicida</i>	1629	5'-YG-3'	E1369R/E1449H/R1556A	D11A/H969A/N995A
<i>SaCas9</i>	Cas9	<i>Staphylococcus aureus</i>	1053	5'-NNGRRT-3'	Wt	D10A/H557A
<i>SaCas9 KKH</i>	Cas9	<i>Staphylococcus aureus</i>	1053	5'-NNNRRT-3'	E782K/N968K/R1015H	D10A/H557A
<i>SpCas9</i>	Cas9	<i>Streptococcus pyogenes</i>	1368	5'-NGG-3'	Wt	D10A/D839A/H840A/N863A
<i>SpCas9 VQR</i>	Cas9	<i>Streptococcus pyogenes</i>	1368	5'-NGA-3'	D1135V/R1335Q/T1337R	D10A/D839A/H840A/N863A
<i>AsCpf1 RR</i>	Cpf1	<i>Acidaminococcus</i> sp. BV3L6	1307	5'-TYCV-3'	S542R/K607R	E993A
<i>AsCpf1 RVR</i>	Cpf1	<i>Acidaminococcus</i> sp. BV3L6	1307	5'-TATV-3'	S542R/K548V/N552R	E993A
<i>FnCpf1</i>	Cpf1	<i>Francisella novicida</i>	1300	5'-NTTN-3'	Wt	D917A/E1006A/D1255A
<i>NmCas</i>	Cas9	<i>Neisseria</i>	1082	5'-	Wt	D16A/D587A/

9		<i>meningitidis</i>		NNNGA TT-3'		H588A/N611A
---	--	---------------------	--	----------------	--	-------------

В некоторых вариантах осуществления белок Cas является каталитически активным и разрезает одну или обе нити сайта ДНК-мишени. В некоторых вариантах осуществления за осуществлением разреза в сайте ДНК-мишени следует образование изменения, например, вставки или делеции, например, посредством механизма клеточной репарации.

В некоторых вариантах осуществления белок Cas модифицирован таким образом, чтобы в нем была деактивирована или частично деактивирована нуклеаза, например, Cas9 с недостаточностью нуклеазы. В то время как Cas9 дикого типа обеспечивает образование двухнитевых разрывов (DSB) в конкретных последовательностях ДНК, которые служат мишенью для gRNA, доступен целый ряд эндонуклеаз CRISPR с модифицированными функциональными свойствами, например: "никазная" версия Cas9, которая была частично деактивирована, образует только однонитевой разрыв; каталитически неактивная Cas9 ("dCas9") не разрезает ДНК-мишень. В некоторых вариантах осуществления связывание dCas9 с последовательностью ДНК может препятствовать транскрипции в этом сайте за счет создания стерического препятствия. В некоторых вариантах осуществления связывание dCas9 с якорной последовательностью может препятствовать (например, уменьшать или предупреждать) формированию и/или сохранению геномного комплекса (например, ASMC). В некоторых вариантах осуществления ДНК-связывающий домен содержит каталитически неактивную Cas9, например, dCas9. В данной области техники известно много каталитически неактивных белков Cas9. В некоторых вариантах осуществления dCas9 содержит мутации в каждом домене эндонуклеазы белка Cas, например, мутации D10A и H840A или N863A. В некоторых вариантах осуществления каталитически неактивный или частично неактивный домен CRISPR/Cas содержит белок Cas, содержащий одну или несколько мутаций, например, одну или несколько мутаций, приведенных в таблице 4. В некоторых вариантах осуществления белок Cas, описанный в конкретной строке таблицы 4, содержит одну, две, три или все мутации, приведенные в той же строке таблицы 4. В некоторых вариантах осуществления белок Cas, например, не описанный в таблице 4, содержит одну, две, три или все мутации, приведенные в строке таблицы 4, или соответствующую мутацию в соответствующем сайте в этом белке Cas.

В некоторых вариантах осуществления каталитически неактивный, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D11 (например, мутацию D11A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию H969 (например, мутацию H969A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию N995 (например, мутацию N995A) или аналогичную замену аминокислоты,

соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, содержит мутации в одном, двух или трех положениях из D11, H969 и N995 (например, мутации D11A, H969A и N995A) или аналогичные замены аминокислот, соответствующих указанным положениям.

В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D10 (например, мутацию D10A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию H557 (например, мутацию H557A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, содержит мутацию D10 (например, мутацию D10A) и мутацию H557 (например, H557A) или аналогичные замены аминокислот, соответствующих указанным положениям.

В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D839 (например, мутацию D839A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию H840 (например, мутацию H840A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию N863 (например, мутацию N863A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, содержит мутацию D10 (например, D10A), мутацию D839 (например, D839A), мутацию H840 (например, H840A) и мутацию N863 (например, N863A) или аналогичные замены аминокислот, соответствующих указанным положениям.

В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию E993 (например, мутацию E993A) или аналогичную замену аминокислоты, соответствующей указанному положению.

В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D917 (например, мутацию D917A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию E1006 (например, мутацию E1006A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный

белок Cas9 содержат мутацию D1255 (например, мутацию D1255A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, содержит мутацию D917 (например, D917A), мутацию E1006 (например, E1006A) и мутацию D1255 (например, D1255A) или аналогичные замены аминокислот, соответствующих указанным положениям.

В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D16 (например, мутацию D16A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию D587 (например, мутацию D587A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию H588 (например, мутацию H588A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, или частично деактивированный белок Cas9 содержат мутацию N611 (например, мутацию N611A) или аналогичную замену аминокислоты, соответствующей указанному положению. В некоторых вариантах осуществления каталитически неактивный белок Cas9, например, dCas9, содержит мутацию D16 (например, D16A), мутацию D587 (например, D587A), мутацию H588 (например, H588A) и мутацию N611 (например, N611A) или аналогичные замены аминокислот, соответствующих указанным положениям.

В некоторых вариантах осуществления ДНК-связывающий домен или домен эндонуклеазы могут содержать молекулу Cas, содержащую gRNA (например, нуклеиновую кислоту-матрицу, например, РНК-матрицу, содержащую gRNA) или связанную с ней (например, ковалентно).

В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают Cas9 *Streptococcus pyogenes* (SpCas9) или его функциональный фрагмент или вариант. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают модифицированный SpCas9. В вариантах осуществления модифицированный SpCas9 содержит модификацию, которая обеспечивает изменение специфичности мотива, прилегающего к протоспейсеру (PAM). В вариантах осуществления PAM характеризуется специфичностью в отношении последовательности нуклеиновой кислоты 5'-NGT-3'. В вариантах осуществления модифицированный SpCas9 содержит одну или несколько аминокислотных замен, например, в одном или нескольких из положений L1111, D1135, G1218, E1219, A1322, R1335, например, выбранных из L1111R, D1135V, G1218R, E1219F, A1322R, R1335V. В вариантах осуществления модифицированный SpCas9 содержит аминокислотную замену T1337R и одну или несколько дополнительных аминокислотных замен, например, выбранных из L1111,

D1135L, S1136R, G1218S, E1219V, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R1335Q, T1337, T1337L, T1337Q, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337H, T1337Q, и T1337M, или их соответствующих аминокислотных замен. В вариантах осуществления модифицированный SpCas9 содержит: (i) одну или несколько аминокислотных замен, выбранных из D1135L, S1136R, G1218S, E1219V, A1322R, R1335Q, и T1337; и (ii) одну или несколько аминокислотных замен, выбранных из L1111R, G1218R, E1219F, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, T1337L, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H, T1337Q, и T1337M, или их соответствующих аминокислотных замен.

В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен (например, описанный в данном документе) предусматривают домен Cas, например, домен Cas9. В вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают нуклеазоактивный домен Cas, Cas-никазный домен (nCas) или нуклеазонеактивный домен Cas (dCas). В вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают нуклеазоактивный домен Cas9, Cas9-никазный домен (nCas9) или нуклеазонеактивный домен Cas9 (dCas9). В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают домен Cas9 из Cas9 (например, dCas9 и nCas9), Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h или Cas12i. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают Cas9 (например, dCas9 и nCas9), Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h или Cas12i. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают Cas9 из *S. pyogenes* или *S. thermophilus* или ее функциональный фрагмент. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают последовательность Cas9, например, описанную в работе Chylinski, Rhun, and Charpentier (2013) RNA Biology 10:5, 726-737, включенной в данный документ посредством ссылки. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают HNH-нуклеазный субдомен и/или субдомен RuvC1 из Cas, например, Cas9, например, описанной в данном документе, или его вариант. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h или Cas12i. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают полипептид Cas (например, фермент) или его функциональный фрагмент. В вариантах осуществления полипептид Cas (например, фермент) выбран из Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5d, Cas5t, Cas5h, Cas5a, Cas6, Cas7, Cas8, Cas8a, Cas8b, Cas8c, Cas9 (e.g., Csn1 или Csx12), Cas10, Cas10d, Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, Cas12i, Csy1, Csy2, Csy3, Csy4, Cse1, Cse2, Cse3, Cse4, Cse5e, Csc1, Csc2, Csa5, Csn1, Csn2, Csm1, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10,

Csx16, CsaX, Csx3, Csx1, Csx1S, Csx11, Csf1, Csf2, CsO, Csf4, Csd1, Csd2, Cst1, Cst2, Csh1, Csh2, Csa1, Csa2, Csa3, Csa4, Csa5, эффекторных белков Cas типа II, эффекторных белков Cas типа V, эффекторных белков Cas типа VI, CARF, DinG, Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12b/C2c1, Cas12c/C2c3, SpCas9(K855A), eSpCas9(1.1), SpCas9-HF1, сверхточного варианта Cas9 (НураCas9), их гомологов, их модифицированных или сконструированных версий и/или их функциональных фрагментов. В вариантах осуществления Cas9 содержит одну или несколько замен, например, выбранных из H840A, D10A, P475A, W476A, N477A, D1125A, W1126A и D1127A. В вариантах осуществления Cas9 содержит одну или несколько мутаций в положениях, выбранных из D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 и/или A987, например, одну или несколько замен, выбранных из D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A и/или D986A. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают последовательность Cas (например, Cas9) из *Corynebacterium ulcerans*, *Corynebacterium diphtheria*, *Spiroplasma syphilidicola*, *Prevotella intermedia*, *Spiroplasma taiwanense*, *Streptococcus iniae*, *Belliella baltica*, *Psychroflexus torquis*, *Streptococcus thermophilus*, *Listeria innocua*, *Campylobacter jejuni*, *Neisseria meningitidis*, *Streptococcus pyogenes* или *Staphylococcus aureus* или ее фрагмент или вариант.

В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают домен Cpf1, например, содержащий одну или несколько замен, например, в положениях D917, E1006A, D1255 или любой их комбинации, например, выбранных из D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A и D917A/E1006A/D1255A.

В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен предусматривают spCas9, spCas9-VRQR (SEQ ID NO: 19), spCas9-VRER (SEQ ID NO: 20), xCas9 (sp), saCas9, saCas9-KKH, spCas9-MQKSER (SEQ ID NO: 21), spCas9-LRKIQK (SEQ ID NO: 22) или spCas9-LRVSQL (SEQ ID NO: 23).

В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен содержат аминокислотную последовательность, указанную в таблице 37 ниже, или аминокислотную последовательность, характеризующуюся по меньшей мере 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% идентичностью последовательности с ней. В некоторых вариантах осуществления домен эндонуклеазы или ДНК-связывающий домен содержат аминокислотную последовательность, которая характеризуется не более чем 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, или 50 различиями (например, мутациями) по сравнению с любой из аминокислотных последовательностей, описанных в данном документе.

Таблица 37. Каждая из эталонных последовательностей включена посредством ссылки во всей своей полноте.

Название	Аминокислотная последовательность или эталонная последовательность
Cas9 <i>Streptococcus pyogenes</i>	
Иллюстративный линкер	SGSETPGTSESATPES (SEQ ID NO: 24)
Иллюстративный линкерный мотив	(SGGS) _n (SEQ ID NO: 25)
Иллюстративный линкерный мотив	(GGGS) _n (SEQ ID NO: 26)
Иллюстративный линкерный мотив	(GGGGS) _n (SEQ ID NO: 27)
Иллюстративный линкерный мотив	(G) _n
Иллюстративный линкерный мотив	(EAAAK) _n (SEQ ID NO: 28)
Иллюстративный линкерный мотив	(GGS) _n
Иллюстративный линкерный мотив	(XP) _n
Cas9 из <i>Streptococcus pyogenes</i>	Эталонная последовательность в NCBI: NC_002737.2 и эталонная последовательность в Uniprot: Q99ZW2
Cas9 из <i>Corynebacterium ulcerans</i>	Эталон. послед. в NCBI: NC_015683.1, NC_017317.1
Cas9 из <i>Corynebacterium diphtheria</i>	Эталон. послед. в NCBI: NC_016782.1, NC_016786.1
Cas9 из <i>Spiroplasma syrophidicola</i>	Эталон. послед. в NCBI: NC_021284.1
Cas9 из <i>Prevotella intermedia</i>	Эталон. послед. в NCBI: NC_017861.1
Cas9 из <i>Spiroplasma taiwanense</i>	Эталон. послед. в NCBI: NC_021846.1
Cas9 из <i>Streptococcus iniae</i>	Эталон. послед. в NCBI: NC_021314.1
Cas9 из <i>Belliella baltica</i>	Эталон. послед. в NCBI: NC_018010.1
Cas9 из <i>Psychroflexus torquis</i>	Эталон. послед. в NCBI: NC_018721.1
Cas9 из <i>Streptococcus thermophilus</i>	Эталон. послед. в NCBI: YP_820832.1

Cas9 из <i>Listeria innocua</i>	Эталон. послед. в NCBI: NP_472073.1
Cas9 из <i>Campylobacter jejuni</i>	Эталон. послед. в NCBI: YP_002344900.1
Cas9 из <i>Neisseria meningitidis</i>	Эталон. послед. в NCBI: YP_002342100.1
dCas9 (D10A и H840A)	
Каталитически неактивная Cas9 (dCas9)	
Никаза Cas9 (nCas9)	
Каталитически активная Cas9	
CasY	((ncbi.nlm.nih.gov/protein/APG80656.1) >APG80656.1 CRISPR-ассоциированный белок CasY [некультивируемая бактерия группы <i>Parcubacteria</i>])
CasX	uniprot.org/uniprot/F0NN87; uniprot.org/uniprot/F0NH53
CasX	>tr F0NH53 F0NH53_SULIR CRISPR-ассоциированный белок, Casx OS= <i>Sulfolobus islandicus</i> (штамм REY15A) GN=SiRe_0771 PE=4 SV=1
CasX Deltaproteobacteria	
Cas12b/C2c1	((uniprot.org/uniprot/T0D7A2#2) sp T0D7A2 C2C1_ALIAG CRISPR-ассоциированная эндонуклеаза C2c1 OS= <i>Alicyclobacillus acidoterrestris</i> (штамм ATCC 49025/DSM 3922/ CIP 106132/NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1)
BhCas12b (<i>Bacillus hisashii</i>)	Эталонная последовательность в NCBI: WP_095142515
BvCas12b (<i>Bacillus</i> sp. V3-13)	Эталонная последовательность в NCBI: WP_101661451.1
Cpf1 дикого типа из <i>Francisella novicida</i>	
Cpf1 D917A из <i>Francisella novicida</i>	
Cpf1 E1006A из <i>Francisella novicida</i>	
Cpf1 D1255A из <i>Francisella novicida</i>	

Cpf1 D917A/E1006A из Francisella novicida	
Cpf1 D917A/D1255A из Francisella novicida	
Cpf1 E1006A/D1255A из Francisella novicida	
Cpf1 D917A/E1006A из Francisella novicida	
SaCas9	
SaCas9n	
PAM-связывающий SpCas9	
PAM-связывающий SpCas9n	
PAM-связывающий SpEQR Cas9	
PAM-связывающий SpVQR Cas9	
PAM-связывающий SpVRER Cas9	
PAM-связывающий SpVRQR Cas9	
SpyMacCas9	

В некоторых вариантах осуществления часть или фрагмент средства, которое способствует активности системы модификации гена (например, модулятор ответа хозяина или эпигенетический модификатор), слиты с капсидным белком AAV. В некоторых вариантах осуществления средство представляет собой молекулу, которая способствует иммунотолерантности. В некоторых вариантах осуществления средство представляет собой фермент, который обеспечивает снижение иммунного ответа хозяина посредством расщепления антител хозяина, в том числе нейтрализующих антител к AAV. В некоторых вариантах осуществления фермент представляет собой эндопептидазу, например, Ig-расщепляющую эндопептидазу, например, IdeS или ее вариант.

Варианты Gene Writer, полученные методом направленной эволюции

В некоторых вариантах осуществления в настоящем изобретении предусмотрены варианты Gene Writer, полученные методом направленной эволюции. В некоторых вариантах осуществления варианты, полученные методом направленной эволюции, могут быть получены посредством мутагенеза эталонного Gene Writer или одного из

содержащихся в нем фрагментов или доменов. В некоторых вариантах осуществления метод направленной эволюции применяют в отношении одного или нескольких доменов (например, домена обратной транскриптазы, ДНК-связывающего домена (в том числе, например, направляемых последовательностью ДНК-связывающих элементов), РНК-связывающего домена или домена эндонуклеазы). В некоторых вариантах осуществления один или несколько таких вариантных доменов, полученных методом направленной эволюции, могут быть получены в результате применения метода направленной эволюции к ним отдельно или совместно с другими доменами. Вариантный домен или домены, полученные методом направленной эволюции, могут в некоторых вариантах осуществления быть объединены с когнатным компонентом(компонентами), к которому не был применен метод направленной эволюции, или вариантами когнатного компонента(компонентов), полученными методом направленной эволюции, например, при получении которых методом направленной эволюции мог применяться либо параллельный, либо последовательный подход.

В некоторых вариантах осуществления процесс мутагенеза эталонного Gene Writer, его фрагмента или домена предусматривает мутагенез эталонного Gene Writer, его фрагмента или домена. В вариантах осуществления мутагенез предусматривает способ обеспечения непрерывной эволюции (например, PACE) или способ обеспечения прерывистой эволюции (например, PANCE), например, как описано в данном документе. В некоторых вариантах осуществления Gene Writer, полученный методом направленной эволюции, или его фрагмент или домен содержат одну или несколько аминокислотных вариаций, введенных в их аминокислотную последовательность, по сравнению с аминокислотной последовательностью эталонного Gene Writer или его фрагмента или домена. В вариантах осуществления вариации аминокислотной последовательности могут предусматривать один или несколько мутированных остатков (например, консервативные замены, неконсервативные замены или их комбинацию) в аминокислотной последовательности эталонного Gene Writer, полученные, например, в результате осуществления изменения в последовательности нуклеотидов, кодирующей Gene Writer, которое приводит, например, к изменению кодона в любом конкретном положении в кодирующей последовательности, делеции одной или нескольких аминокислот (например, в случае усеченного белка), вставке одной или нескольких аминокислот или любой комбинации изложенного выше. Вариант Gene Writer, полученный методом направленной эволюции, может предусматривать наличие вариантов в одном или нескольких компонентах или доменах Gene Writer (например, варианты, введенные в домен обратной транскриптазы, домен эндонуклеазы, ДНК-связывающий домен, РНК-связывающий домен или их комбинации).

В некоторых аспектах в настоящем изобретении предусмотрены Gene Writer, системы, наборы и способы, в которых используется или содержится вариант Gene Writer, полученный методом направленной эволюции, например, в которых используется вариант Gene Writer, полученный методом направленной эволюции, или Gene Writer, полученный

или который может быть получен с помощью PACE или PANCE. В вариантах осуществления эталонный Gene Writer, к которому не был применен метод направленной эволюции, представляет собой Gene Writer, раскрытый в данном документе.

Термин “непрерывная эволюция, осуществляемая с использованием фагов (PACE)”, используемый в данном документе, обычно относится к методу непрерывной направленной эволюции, в котором используются фаги в качестве вирусных векторов. Примеры технологии PACE были описаны, например, в международной заявке согласно РСТ № РСТ/US 2009/056194, поданной 8 сентября 2009 г., опубликованной в виде WO 2010/028347 11 марта 2010 г.; международной заявке согласно РСТ, РСТ/US2011/066747, поданной 22 декабря 2011 г., опубликованной в виде WO 2012/088381 28 июня 2012 г.; патенте США № 9023594, выданном 5 мая 2015 г.; патенте США № 9771574, выданном 26 сентября 2017 г.; патенте США № 9394537, выданном 19 июля 2016 г.; международной заявке согласно РСТ, РСТ/US2015/012022, поданной 20 января 2015 г., опубликованной в виде WO 2015/134121 11 сентября 2015 г.; патенте США № 10179911, выданном 15 января 2019 г.; и международной заявке согласно РСТ, РСТ/US2016/027795, поданной 15 апреля 2016 г., опубликованной в виде WO 2016/168631 20 октября 2016 г., полное содержание каждой из которых включено в данный документ посредством ссылки.

Термин “прерывистая эволюция, осуществляемая с использованием фагов (PANCE)”, используемый в данном документе, обычно относится к методу прерывистой направленной эволюции, в котором используются фаги в качестве вирусных векторов. Примеры технологии PANCE описаны, например, в работе Suzuki T. et al, Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase, Nat Chem Biol. 13(12): 1261-1266 (2017), включенной в данный документ посредством ссылки во всей своей полноте. Вкратце, PANCE представляет собой методику быстрой направленной эволюции *in vivo* с использованием серийных переносов в колбах фага для селекции (SP), подвергающегося действию направленной эволюции, который содержит ген, представляющий интерес, который необходимо подвергнуть действию направленной эволюции, в свежих клетках-хозяевах (например, клетках *E. coli*). Гены внутри клетки-хозяина могут поддерживаться в неизменном состоянии, в то время как гены, содержащиеся в SP, - непрерывно эволюционировать. После обеспечения роста фага аликвоту инфицированных клеток можно использовать для трансфекции следующей колбы, содержащей *E. coli* хозяина. Этот процесс можно повторять и/или продолжать до тех пор, пока в результате направленной эволюции не будет получен необходимый фенотип, например, столько переносов, сколько необходимо.

Для быстрого анализа способов применения PACE и PANCE к Gene Writer специалист в данной области техники может обратиться, в числе прочего, к предшествующим литературным источникам. Дополнительные иллюстративные способы управления непрерывной направленной эволюцией белков или систем, модифицирующих геном, например, в популяции клеток-хозяев, например, с использованием фаговых частиц, могут применяться для создания вариантов Gene Writer, полученных методом

направленной эволюции, или их фрагментов или субдоменов. Неограничивающие примеры таких способов описаны в международной заявке согласно PCT, PCT/US2009/056194, поданной 8 сентября 2009 г., опубликованной в виде WO 2010/028347 11 марта 2010 г.; международной заявке согласно PCT, PCT/US2011/066747, поданной 22 декабря 2011 г., опубликованной в виде WO 2012/088381 28 июня 2012 г.; патенте США № 9023594, выданном 5 мая 2015 г.; патенте США № 9771574, выданном 26 сентября 2017 г.; патенте США № 9394537, выданном 19 июля 2016 г.; международной заявке согласно PCT, PCT/US2015/012022, поданной 20 января 2015 г., опубликованной в виде WO 2015/134121 11 сентября 2015 г.; патенте США № 10179911, выданном 15 января 2019 г.; международной заявке № PCT/US2019/37216, поданной 14 июня 2019 г., международной публикации заявки на патент WO 2019/023680, опубликованной 31 января 2019 г., международной заявке согласно PCT, PCT/US2016/027795, поданной 15 апреля 2016 г., опубликованной в виде WO 2016/168631 20 октября 2016 г., и международной публикации заявки на патент № PCT/US2019/47996, поданной 23 августа 2019 г., каждое из которых включено в данный документ посредством ссылки во всей своей полноте.

В некоторых неограничивающих иллюстративных вариантах осуществления способ направленной эволюции в случае варианта Gene Writer, полученного методом направленной эволюции, его фрагмента или домена включает: (a) приведение популяции клеток-хозяев в контакт с популяцией вирусных векторов, содержащих ген, представляющий интерес (исходный Gene Writer или его фрагмент или домен), где (1) клетка-хозяин поддается инфицированию вирусным вектором; (2) в клетке-хозяине экспрессируются вирусные гены, необходимые для образования вирусных частиц; (3) экспрессия по меньшей мере одного вирусного гена, необходимого для продуцирования инфекционной вирусной частицы, зависит от функции гена, представляющего интерес; и/или (4) вирусный вектор обеспечивает экспрессию белка в клетке-хозяине и может реплицироваться и упаковываться в вирусную частицу клеткой-хозяином; в некоторых вариантах осуществления способ включает (b) приведение клеток-хозяев в контакт с мутагеном, использование клеток-хозяев с мутациями, которые повышают скорость мутаций (например, за счет наличия мутационной плазмиды или некоторой модификации генома, например, ДНК-полимеразы с нарушенной функцией исправления ошибок, генов SOS, таких как UmuC, UmuD' и/или RecA, мутации которых, если они связаны с плазмидой, могут находиться под контролем индуцируемого промотора), или их комбинации; в некоторых вариантах осуществления способ включает (c) инкубирование популяции клеток-хозяев в условиях, обеспечивающих репликацию вируса и продуцирование вирусных частиц, где клетки-хозяева удаляют из популяции клеток-хозяев, а свежие, неинфицированные клетки-хозяева вводят в популяции клеток-хозяев, за счет чего пополняется популяция клеток-хозяев и создается поток клеток-хозяев; в некоторых вариантах осуществления клетки инкубируют в условиях, при которых в гене, представляющем интерес, может возникнуть мутация; в некоторых вариантах осуществления способ дополнительно включает (d) выделение мутантной версии

вирусного вектора, кодирующей продукт гена, полученный методом направленной эволюции (например, вариант Gene Writer, полученный методом направленной эволюции, или его фрагмент или домен), из популяции клеток-хозяев.

Опытному специалисту в данной области техники будет известно множество признаков, которые можно использовать в рамках вышеописанной структуры. Например, в некоторых вариантах осуществления вирусный вектор или фаг представляет собой нитчатый фаг, например, фаг M13, например, фаг M13 для селекции. В некоторых вариантах осуществления ген, необходимый для продуцирования инфекционных вирусных частиц, представляет собой ген III M13 (gIII). В вариантах осуществления фаг может не содержать функционального gIII, но при этом содержать gI, gII, gIV, gV, gVI, gVII, gVIII, gIX и gX. В некоторых вариантах осуществления образование инфекционных частиц VSV предусматривает наличие белка оболочки VSV-G. В различных вариантах осуществления могут использоваться разные ретровирусные векторы, например, векторы на основе вируса лейкоза мышей или лентивирусные векторы. В вариантах осуществления ретровирусные векторы могут быть эффективно упакованы с использованием белка оболочки VSV-G, например, в качестве замены нативного белка оболочки вируса.

В некоторых вариантах осуществления клетки-хозяева инкубируют в соответствии с подходящим количеством жизненных циклов вируса, например, по меньшей мере 10, по меньшей мере 20, по меньшей мере 30, по меньшей мере 40, по меньшей мере 50, по меньшей мере 100, по меньшей мере 200, по меньшей мере 300, по меньшей мере 400, по меньшей мере 500, по меньшей мере 600, по меньшей мере 700, по меньшей мере 800, по меньшей мере 900, по меньшей мере 1000, по меньшей мере 1250, по меньшей мере 1500, по меньшей мере 1750, по меньшей мере 2000, по меньшей мере 2500, по меньшей мере 3000, по меньшей мере 4000, по меньшей мере 5000, по меньшей мере 7500, по меньшей мере 10000 или больше последовательных жизненных циклов вируса, что в иллюстративных и неограничивающих примерах с использованием фага M13 составляет 10-20 минут на жизненный цикл вируса. Аналогичным образом, условия можно модулировать для регулирования времени, в течение которого клетка-хозяин остается в популяции клеток-хозяев, например, приблизительно 10, приблизительно 11, приблизительно 12, приблизительно 13, приблизительно 14, приблизительно 15, приблизительно 16, приблизительно 17, приблизительно 18, приблизительно 19, приблизительно 20, приблизительно 21, приблизительно 22, приблизительно 23, приблизительно 24, приблизительно 25, приблизительно 30, приблизительно 35, приблизительно 40, приблизительно 45, приблизительно 50, приблизительно 55, приблизительно 60, приблизительно 70, приблизительно 80, приблизительно 90, приблизительно 100, приблизительно 120, приблизительно 150 или приблизительно 180 минут. Популяции клеток-хозяев можно частично регулировать посредством плотности клеток-хозяев или в некоторых вариантах осуществления плотности клеток-хозяев в поступающей среде, например, 10^3 клеток/мл, приблизительно 10^4 клеток/мл, приблизительно 10^5 клеток/мл, приблизительно $5 \cdot 10^5$ клеток/мл, приблизительно 10^6

клеток/мл, приблизительно $5 \cdot 10^6$ клеток/мл, приблизительно 10^7 клеток/мл, приблизительно $5 \cdot 10^7$ клеток/мл, приблизительно 10^8 клеток/мл, приблизительно $5 \cdot 10^8$ клеток/мл, приблизительно 10^9 клеток/мл, приблизительно $5 \cdot 10^9$ клеток/мл, приблизительно 10^{10} клеток/мл или приблизительно $5 \cdot 10^{10}$ клеток/мл.

Интеины

В некоторых вариантах осуществления, как более подробно описано ниже, интеин-N может быть слит с N-концевой частью первого домена, описанного в данном документе, а интеин-С может быть слит с С-концевой частью второго домена, описанного в данном документе, для обеспечения соединения N-концевой части с С-концевой частью с обеспечением таким образом соединения первого и второго доменов. В некоторых вариантах осуществления каждый из первого и второго доменов независимо выбран из ДНК-связывающего домена, РНК-связывающего домена, домена RT и домена эндонуклеазы.

Используемый в данном документе термин "интеин" относится к интрону белка (например, пептиду), способному к самовырезанию путем сплайсинга, например, который лигирует фланкирующие N-концевые и С-концевые экстеины (например, фрагменты, подлежащие соединению). В некоторых случаях интеин может содержать фрагмент белка, который способен удалить себя и соединить остающиеся фрагменты (экстеины) пептидной связью в процессе, известном как сплайсинг белка. Интеины также называют "белковыми интронами". Осуществляемый интеином процесс удаления самого себя и соединения остающихся частей белка в данном документе называется "белковым сплайсингом" или "интеин-опосредованным белковым сплайсингом". В некоторых вариантах осуществления интеин белка-предшественника (белка, содержащего интеин, до интеин-опосредованного сплайсинга белка) происходит из двух генов. Такой интеин упоминается в данном документе как расщепленный интеин (например, расщепленный интеин-N и расщепленный интеин-С). Например, у цианобактерий DnaE, каталитическая субъединица α ДНК-полимеразы III, кодируется двумя отдельными генами, dnaE-n и dnaE-c. Интеин, кодируемый геном dnaE-n, может быть обозначен в данном документе как "интеин-N". Интеин, кодируемый геном dnaE-c, может быть обозначен в данном документе как "интеин-С".

Использование интеинов для соединения фрагментов гетерологичного белка описано, например, в работе Wood et al., J. Biol. Chem. 289(21); 14512-9 (2014) (включена в данный документ посредством ссылки во всей своей полноте). Например, будучи слитыми с отдельными фрагментами белка, интеины IntN и IntC могут распознавать друг друга, осуществлять вырезание самих себя путем сплайсинга и/или одновременно лигировать фланкирующие N- и С-концевые экстеины фрагментов белка, с которыми они были слиты, за счет чего восстанавливается полноразмерный белок из двух фрагментов белка.

В некоторых вариантах осуществления используется синтетический интеин на основе интеина dnaE, пары интеинов Cfa-N (например, расщепленный интеин-N) и Cfa-C (например, расщепленный интеин-С). Примеры таких интеинов описаны, например, в

работе Stevens et al., J Am Chem Soc. 2016 Feb. 24; 138(7):2162-5 (включена в данный документ посредством ссылки во всей своей полноте). Неограничивающие примеры пар интеинов, которые можно использовать в соответствии с настоящим изобретением, включают интеин Cfa DnaE, интеин Ssp GyrB, интеин Ssp DnaX, интеин Ter DnaE3, интеин Ter ThyX, интеин Rma DnaB и интеин Cne Prp8 (например, как описано в патенте США № 8394604, включенном в данный документ посредством ссылки).

В некоторых вариантах осуществления интеин-N и интеин-C могут быть слиты с N-концевой частью расщепленной Cas9 и C-концевой частью расщепленной Cas9 соответственно для обеспечения соединения N-концевой части расщепленной Cas9 и C-концевой части расщепленной Cas9. Например, в некоторых вариантах осуществления интеин-N слит с C-концом N-концевой части расщепленной Cas9, т. е. с образованием структуры N-[N-концевой участок расщепленной Cas9]-[интеин-N]~C. В некоторых вариантах осуществления интеин-C слит с N-концом C-концевой части расщепленной Cas9, т. е. с образованием структуры N-[интеин-C]~[C-концевой участок расщепленной Cas9]-C. Механизм интеин-опосредованного сплайсинга белков для связывания белков, с которыми слиты интеины (например, расщепленной Cas9), описан в работе Shah et al., Chem Sci. 2014;5(1):446-461, включенной в данный документ посредством ссылки. Способы конструирования и применения интеинов известны в данной области техники и описаны, например, в документах WO2020051561, WO2014004336, WO2017132580, US20150344549 и US20180127780, каждый из которых включен в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления расщепление относится к разделению на два или более фрагментов. В некоторых вариантах осуществления расщепленный белок Cas9 или расщепленная Cas9 предусматривают белок Cas9, который представлен в виде N-концевого фрагмента и C-концевого фрагмента, кодируемых двумя отдельными нуклеотидными последовательностями. Полипептиды, соответствующие N-концевой части и C-концевой части белка Cas9, могут подвергаться сплайсингу с образованием восстановленного белка Cas9. В вариантах осуществления белок Cas9 разделен на два фрагмента в неупорядоченном участке белка, например, как описано в работе Nishimasu et al., Cell, Volume 156, Issue 5, pp. 935-949, 2014, или как описано в работе Jiang et al. (2016) Science 351: 867-871 и файле PDB: 5F9R (каждое из которых включено в данный документ посредством ссылки во всей своей полноте). Неупорядоченный участок можно определить с помощью одной или нескольких методик определения структуры белка, известных в данной области техники, в том числе без ограничения рентгеновской кристаллографии, ЯМР-спектроскопии, электронной микроскопии (например, криоэлектронной микроскопии) и/или моделирования белков *in silico*. В некоторых вариантах осуществления белок разделен на два фрагмента в любом положении C, T, A или S, например, в пределах области SpCas9 между аминокислотами A292-G364, F445-K483 или E565-T637, или в соответствующих положениях в любой другой Cas9, варианте Cas9 (например, nCas9, dCas9) или другом парDNAbp. В некоторых вариантах осуществления белок разделен на

два фрагмента в положениях T310, T313, A456, S469 или C574 SpCas9. В некоторых вариантах осуществления процесс разделения белка на два фрагмента называется расщеплением белка.

В некоторых вариантах осуществления длина фрагмента белка составляет от приблизительно 2 до 1000 аминокислот (например, 2-10, 10-50, 50-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900 или 900-1000 аминокислот). В некоторых вариантах осуществления длина фрагмента белка составляет от приблизительно 5 до 500 аминокислот (например, 5-10, 10-50, 50-100, 100-200, 200-300, 300-400 или 400-500 аминокислот). В некоторых вариантах осуществления длина фрагмента белка составляет от приблизительно 20 до 200 аминокислот (например, 20-30, 30-40, 40-50, 50-100 или 100-200 аминокислот).

В некоторых вариантах осуществления часть или фрагмент Gene Writer (например, Cas9-R2Tg) слиты с интеином. Нуклеаза может быть слита с N-концом или C-концом интеина. В некоторых вариантах осуществления часть или фрагмент слитого белка слиты с интеином и слиты с капсидным белком AAV. Интеин, нуклеаза и капсидный белок могут быть слиты вместе в любой конфигурации (например, нуклеаза-интеин-капсид, интеин-нуклеаза-капсид, капсид-интеин-нуклеаза и т. д.). В некоторых вариантах осуществления N-конец интеина слит с C-концом слитого белка, а C-конец интеина слит с N-концом капсидного белка AAV.

В некоторых вариантах осуществления домен эндонуклеазы (например, домен никазы Cas9) слит с интеином-N, а полипептид, содержащий домен RT, слит с интеином-C.

Иллюстративные нуклеотидные и аминокислотные последовательности интеинов представлены ниже.

Интеин-N DnaE, ДНК:

TGCCTGTCATACGAAACCGAGATACTGACAGTAGAATATGGCCTTCTGCCAA
TCGGGAAGATTGTGGAGAAACGGATAGAATGCACAGTTTACTCTGTTCGATAACAAT
GGTAACATTTATACTCAGCCAGTTGCCAGTGGCACGACCGGGGAGAGCAGGAAGT
ATTCGAATACTGTCTGGAGGATGGAAGTCTCATTAGGGCCACTAAGGACCACAAAT
TTATGACAGTCGATGGCCAGATGCTGCCTATAGACGAAATCTTTGAGCGAGAGTTGG
ACCTCATGCGAGTTGACAACCTTCCTAAT (SEQ ID NO: 29)

Интеин-N DnaE, белок:

CLSYETEILTVEYGLLPYKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEV
FEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN (SEQ ID NO: 30)

Интеин-C DnaE, ДНК:

ATGATCAAGATAGCTACAAGGAAGTATCTTGGCAAACAAAACGTTTATGATA
TTGGAGTCGAAAGAGATCACAACCTTTGCTCTGAAGAACGGATTCATAGCTTCTAAT
(SEQ ID NO: 31)

Интеин-C:

MIKIATRKYLGKQNVYDIGVERDHNFKNGFIASN (SEQ ID NO: 32)

Cfa-N, ДНК:

TGCCTGTCTTATGATACCGAGATACTTACCGTTGAATATGGCTTCTTGCCTAT
TGGAAAGATTGTCTGAAGAGAGAATTGAATGCACAGTATATACTGTAGACAAGAATG
GTTTCGTTTACACACAGCCCATTTGCTCAATGGCACAATCGCGGCCGAACAAGAAGTAT
TTGAGTACTGTCTCGAGGATGGAAGCATCATACGAGCAACTAAAGATCATAAATTC
ATGACCACTGACGGGCAGATGTTGCCAATAGATGAGATATTCGAGCGGGGCTTGA
TCTCAAACAAGTGGATGGATTG CCA (SEQ ID NO: 33)

Cfa-N, белок:

CLSYDTEILTVEYGFLLPIGKIVEERIECTVYTVDKNGFVYTQPIAQWHNRGEQEVF
EYCLEDGSIIRATKDHKFMTTDQGMLPIDEIFERGLDLKQVDGLP (SEQ ID NO: 34)

Cfa-C, ДНК:

ATGAAGAGGACTGCCGATGGATCAGAGTTTGAATCTCCCAAGAAGAAGAGG
AAAGTAAAGATAATATCTCGAAAAAGTCTTGGTACCCAAAATGTCTATGATATTGG
AGTGGAGAAAGATCACAACTTCCTTCTCAAGAACGGTCTCGTAGCCAGCAAC (SEQ
ID NO: 35)

Cfa-C, белок:

MKRTADGSEFESPKKKRKVKIISRKSLGTQNVYDIGVEKDHNFLKNGLVASN
(SEQ ID NO: 36)

Нуклеиновые кислоты-матрицы

В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит одну или несколько последовательностей (например, 2 последовательности), которые связывают полипептид Gene Writer. В некоторых вариантах осуществления нуклеиновая кислота-матрица, например, РНК-матрица, ковалентно связана или слита со средством, которое способствует активности системы модификации гена (например, модулятором ответа хозяина или эпигенетическим модификатором). В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит 5'-UTR, который связывает полипептид Gene Writer, и/или 3'-UTR, который связывает полипептид Gene Writer. В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит первую последовательность инвертированных повторов и вторую последовательность инвертированных повторов, каждая из которых связывает полипептид Gene Writer.

В некоторых вариантах осуществления нуклеиновая кислота-матрица предусматривает РНК. В некоторых вариантах осуществления нуклеиновая кислота-матрица предусматривает ДНК (например, однонитевую или двухнитевую ДНК).

В некоторых вариантах осуществления нуклеиновая кислота-матрица содержит один или несколько (например, 2) доменов гомологии, которые гомологичны последовательности-мишени. В некоторых вариантах осуществления длина доменов гомологии составляет приблизительно 10-20, 20-50 или 50-100 нуклеотидов.

Системы Gene Writer™, описанные в данном документе, могут осуществлять модификацию в сайте-мишени ДНК хозяина с использованием последовательности нуклеиновой кислоты-матрицы. В некоторых вариантах осуществления системы Gene Writer™, описанные в данном документе, транскрибируют матрицу последовательности

РНК в сайты-мишени ДНК хозяина с помощью праймируемой мишенью обратной транскрипции (TPRT). Путем записывания последовательности(последовательностей) ДНК посредством обратной транскрипции последовательности РНК-матрицы непосредственно в геном хозяина система Gene Writer™ может осуществлять вставку последовательности-объекта в геном-мишень без необходимости введения экзогенных последовательностей ДНК в клетку-хозяина (в отличие от, например, систем CRISPR), а также исключая стадию вставки экзогенной ДНК. Система Gene Writer™ также может удалять последовательность из генома-мишени или вводить замену с использованием последовательности объекта. Таким образом, система Gene Writer™ представляет собой платформу для использования индивидуально адаптированных последовательностей РНК-матриц, содержащих последовательности объекта, например, последовательности, предусматривающие кодирование гетерологичных генов и/или информацию о функции.

В некоторых вариантах осуществления РНК-матрица может содержать последовательность gRNA, например, для направления GeneWriter к сайту-мишени, представляющему интерес. В некоторых вариантах осуществления РНК-матрица содержит (например, в направлении от 5'- к 3'-концу) следующее: (i) необязательно последовательность (например, спейсер CRISPR), которая связывает сайт-мишень (например, вторую нить сайта в геноме-мишени), (ii) необязательно последовательность, которая связывает полипептид, описанный в данном документе (например, полипептид GeneWriter или Cas), (iii) последовательность гетерологичного объекта и (iv) 3'-концевой домен гомологии с мишенью.

В некоторых вариантах осуществления РНК-матрица может содержать последовательность gRNA, например, для направления GeneWriter к сайту-мишени, представляющему интерес. В некоторых вариантах осуществления РНК-матрица содержит (например, в направлении от 5'- к 3'-концу) следующее: (i) необязательно последовательность (например, спейсер CRISPR), которая связывает сайт-мишень (например, вторую нить сайта в геноме-мишени), (ii) необязательно последовательность, которая связывает полипептид, описанный в данном документе (например, полипептид GeneWriter или Cas), (iii) последовательность гетерологичного объекта и (iv) 5'-концевой домен гомологии и/или 3'-концевой домен гомологии с мишенью.

В некоторых вариантах осуществления молекула нуклеиновой кислоты-матрицы содержит 5'-концевой домен гомологии и/или 3'-концевой домен гомологии. В некоторых вариантах осуществления 5'-концевой домен гомологии содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью нуклеиновой кислоты, содержащейся в молекуле нуклеиновой кислоты-мишени. В вариантах осуществления последовательность нуклеиновой кислоты в молекуле нуклеиновой кислоты-мишени расположена в пределах приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, или 100 нуклеотидов от сайта-мишени для вставки (например, в направлении 5'-конца от него), например, для последовательности

гетерологичного объекта, например, содержащейся в молекуле нуклеиновой кислоты-матрицы.

В некоторых вариантах осуществления 3'-концевой домен гомологии содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичностью последовательности с последовательностью нуклеиновой кислоты, содержащейся в молекуле нуклеиновой кислоты-мишени. В вариантах осуществления последовательность нуклеиновой кислоты в молекуле нуклеиновой кислоты-мишени расположена в пределах приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, или 100 нуклеотидов от сайта-мишени для вставки (например, в направлении 3'-конца от него), например, для последовательности гетерологичного объекта, например, содержащейся в молекуле нуклеиновой кислоты-матрицы. В некоторых вариантах осуществления 5'-концевой домен гомологии является гетерологичным относительно остальной части молекулы нуклеиновой кислоты-матрицы. В некоторых вариантах осуществления 3'-концевой домен гомологии является гетерологичным относительно остальной части молекулы нуклеиновой кислоты-матрицы.

В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) содержит 3'-концевой домен гомологии с мишенью. В некоторых вариантах осуществления 3'-концевой домен гомологии с мишенью расположен в направлении 3'-конца от последовательности гетерологичного объекта и комплементарен последовательности, прилегающей к сайту, подлежащему модификации с помощью системы, описанной в данном документе, или содержит не более 1, 2, 3, 4 или 5 несовпадений с последовательностью, комплементарной последовательности, прилегающей к сайту, подлежащему модификации с помощью системы/Gene Writer™. В некоторых вариантах осуществления 3'-концевой домен гомологии связывается в пределах 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 нуклеотидов от сайта ника в молекуле нуклеиновой кислоты-мишени. В некоторых вариантах осуществления связывание 3'-концевого домена гомологии с молекулой нуклеиновой кислоты-мишени позволяет инициировать праймируемую мишенью обратную транскрипцию (TPRT), например, при этом 3'-концевой домен гомологии выступает в качестве праймера для TPRT.

В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) содержит последовательность гетерологичного объекта. В некоторых вариантах осуществления последовательность гетерологичного объекта может быть транскрибирована с помощью домена RT полипептида Gene Writer™, например, с обеспечением таким образом введения изменения в сайт-мишень в геномной ДНК. В некоторых вариантах осуществления длина последовательности гетерологичного объекта составляет по меньшей мере 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, 120, 140, 160, 180, 200, 500, или 1,000 нуклеотидов (нт) или по меньшей мере 1, 1,5, 2, 2,5,

3, 3,5, 4, 4,5, 5, 5,5, 6, 6,5, 7, 7,5, 8, 8,5, 9, 9,5, или 10 тысяч оснований. В некоторых вариантах осуществления длина последовательности гетерологичного объекта составляет не более 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, 120, 140, 160, 180, 200, 500, 1,000, или 2000 нуклеотидов (нт) или не более 20, 15, 10, 9, 8, 7, 6, 5, 4 или 3 тысяч оснований. В некоторых вариантах осуществления длина последовательности гетерологичного объекта составляет 30-1000, 40-1000, 50-1000, 60-1000, 70-1000, 74-1000, 75-1000, 76-1000, 77-1000, 78-1000, 79-1000, 80-1000, 85-1000, 90-1000, 100-1000, 120-1000, 140-1000, 160-1000, 180-1000, 200-1000, 500-1000, 30-500, 40-500, 50-500, 60-500, 70-500, 74-500, 75-500, 76-500, 77-500, 78-500, 79-500, 80-500, 85-500, 90-500, 100-500, 120-500, 140-500, 160-500, 180-500, 200-500, 30-200, 40-200, 50-200, 60-200, 70-200, 74-200, 75-200, 76-200, 77-200, 78-200, 79-200, 80-200, 85-200, 90-200, 100-200, 120-200, 140-200, 160-200, 180-200, 30-100, 40-100, 50-100, 60-100, 70-100, 74-100, 75-100, 76-100, 77-100, 78-100, 79-100, 80-100, 85-100, или 90-100 нуклеотидов (нт) или 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3, 1-2, 2-20, 2-15, 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4, 2-3, 3-20, 3-15, 3-10, 3-9, 3-8, 3-7, 3-6, 3-5, 3-4, 4-20, 4-15, 4-10, 4-9, 4-8, 4-7, 4-6, 4-5, 5-20, 5-15, 5-10, 5-9, 5-8, 5-7, 5-6, 6-20, 6-15, 6-10, 6-9, 6-8, 6-7, 7-20, 7-15, 7-10, 7-9, 7-8, 8-20, 8-15, 8-10, 8-9, 9-20, 9-15, 9-10, 10-15, 10-20, или 15-20 тысяч оснований. В некоторых вариантах осуществления длина последовательности гетерологичного объекта составляет 10-100, 10-90, 10-80, 10-70, 10-60, 10-50, 10-40, 10-30 или 10-20 нуклеотидов, например, 10-80, 10-50 или 10-20 нуклеотидов, например, приблизительно 10-20 нуклеотидов.

Нуклеиновая кислота-матрица (например, РНК-матрица) может характеризоваться некоторой степенью гомологии с ДНК-мишенью. В некоторых вариантах осуществления для нуклеиновой кислоты-матрицы (например, РНК-матрицы) 3'-концевой домен гомологии с мишенью может выступать в качестве области отжига для ДНК-мишени, обеспечивая расположение ДНК-мишени, необходимое для праймирования обратной транскрипции нуклеиновой кислоты-матрицы (например, РНК-матрицы). В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) содержит по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200 или больше оснований со строгой гомологией с ДНК-мишенью на 3'-конце РНК. В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) содержит по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 175, 180, или 200 или больше оснований с по меньшей мере 50%, 60%, 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99% или 100% гомологией с ДНК-мишенью, например, на 5'-конце нуклеиновой кислоты-матрицы (например, РНК-матрицы).

Компонент системы редактирования генома Gene Writer™, представляющий собой нуклеиновую кислоту-матрицу (например, РНК-матрицу), описанный в данном документе, как правило, способен связывать белок системы редактирования генома Gene Writer™. В

некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) содержит 3'-концевую область, которая способна связывать белок редактирования генома Gene Writer™. Область связывания, например, 3'-концевая область, может представлять собой структурированную область РНК, например, имеющую по меньшей мере 1, 2 или 3 шпилечные петли, способные связываться с белком системы редактирования генома Gene Writer™. Область связывания может обеспечивать ассоциацию нуклеиновой кислоты-матрицы (например, РНК-матрицы) с любым из полипептидных модулей. В некоторых вариантах осуществления область связывания нуклеиновой кислоты-матрицы (например, РНК-матрицы) может ассоциировать с РНК-связывающим доменом в полипептиде. В некоторых вариантах осуществления область связывания нуклеиновой кислоты-матрицы (например, РНК-матрицы) может ассоциировать с доменом обратной транскрипции полипептида (например, специфично связываться с доменом RT). Например, в случае, если домен обратной транскрипции получен из ретротранспозона, не содержащего LTR, нуклеиновая кислота-матрица (например, РНК-матрица) может содержать область связывания, полученную из ретротранспозона, не содержащего LTR, например, 3'-UTR из ретротранспозона, не содержащего LTR. В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) может ассоциировать с ДНК-связывающим доменом полипептида, например, gRNA ассоциирует с ДНК-связывающим доменом, полученным из Cas9. В некоторых вариантах осуществления область связывания может также обеспечивать распознавание ДНК-мишени, например, gRNA гибридизируется с последовательностью ДНК-мишени и связывает полипептид, например, домен Cas9. В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) может ассоциировать с несколькими компонентами полипептида, например, с ДНК-связывающим доменом и доменом обратной транскрипции. Например, нуклеиновая кислота-матрица (например, РНК-матрица) может содержать область gRNA, которая ассоциирует с ДНК-связывающим доменом, полученным из Cas9, и 3'-UTR из ретротранспозона, не содержащего LTR, которая ассоциирует с доменом обратной транскрипции, полученным из ретротранспозона, не содержащего LTR.

Нуклеиновая кислота-матрица (например, РНК-матрица) может быть сконструирована таким образом, чтобы она приводила в результате к инсерциям, мутациям или делециям в локусе ДНК-мишени. В некоторых вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) может быть сконструирована так, чтобы обуславливать осуществление вставки в ДНК-мишень. Например, нуклеиновая кислота-матрица (например, РНК-матрица) может содержать гетерологичную последовательность, где обратная транскрипция приведет в результате к вставке гетерологичной последовательности в ДНК-мишень. В других вариантах осуществления РНК-матрица может быть сконструирована для записи делеции в ДНК-мишень. Например, нуклеиновая кислота-матрица (например, РНК-матрица) может соответствовать ДНК-мишени выше и ниже участка необходимой делеции, при этом обратная транскрипция будет приводить в

результате к копированию последовательностей нуклеиновой кислоты-матрицы (например, РНК-матрицы) выше и ниже, без промежуточной последовательности, например, обуславливая делецию промежуточной последовательности. В других вариантах осуществления нуклеиновая кислота-матрица (например, РНК-матрица) может быть сконструирована для записи редактирования в ДНК-мишень. Например, РНК-матрица может соответствовать последовательности ДНК-мишени, за исключением одного или нескольких нуклеотидов, где обратная транскрипция будет приводить в результате к копированию этих изменений в ДНК-мишень, например, приводя в результате к мутациям, например, мутациям типа транзиции или трансверсии.

Предполагается, что применимым может быть использование кольцевых и/или линейных состояний РНК во время составления, доставки или реакции Gene Writing в клетке-мишени. Таким образом, в некоторых вариантах осуществления любого из аспектов, описанных в данном документе, система Gene Writing содержит одну или несколько кольцевых РНК (circRNA). В некоторых вариантах осуществления любого из аспектов, описанных в данном документе, система Gene Writing содержит одну или несколько линейных РНК. В некоторых вариантах осуществления нуклеиновая кислота, описанная в данном документе (например, нуклеиновая кислота-матрица, молекула нуклеиновой кислоты, кодирующая полипептид Gene Writer, или и то, и другое), представляет собой circRNA. В некоторых вариантах осуществления кольцевая молекула РНК кодирует полипептид Gene Writer. В некоторых вариантах осуществления молекула circRNA, кодирующая полипептид Gene Writer, доставляется в клетку-хозяина. В некоторых вариантах осуществления кольцевая молекула РНК кодирует рекомбиназу, например, описанную в данном документе. В некоторых вариантах осуществления молекула circRNA, кодирующая рекомбиназу, доставляется в клетку-хозяина. В некоторых вариантах осуществления молекула circRNA, кодирующая полипептид Gene Writer, линейаризуется (например, в клетке-хозяине, например, в ядре клетки-хозяина) перед трансляцией.

Было обнаружено, что кольцевые РНК (circRNA) в естественных условиях встречаются в клетках и выполняют разнообразные функции, включающие как не связанные с кодированием, так и связанные с кодированием белка роли, в клетках человека. Было показано, что circRNA может быть сконструирована путем включения самовырезающегося путем сплайсинга интрона в молекулу РНК (или ДНК, кодирующую молекулу РНК), что приводит в результате к циркуляризации РНК, и сконструированная таким образом circRNA может характеризоваться повышенным продуцированием белка и стабильностью (Wesselhoeft et al. *Nature Communications* 2018). В некоторых вариантах осуществления полипептид Gene Writer™ кодируется в виде circRNA. В некоторых вариантах осуществления нуклеиновая кислота-матрица представляет собой ДНК, такую как dsDNA или ssDNA. В некоторых вариантах осуществления circDNA предусматривает РНК-матрицу.

В некоторых вариантах осуществления circRNA предусматривает одну или несколько последовательностей рибозимов. В некоторых вариантах осуществления

последовательность рибозима активируется для саморасщепления, например, в клетке-хозяине, например, что таким образом приводит к линейаризации *circRNA*. В некоторых вариантах осуществления рибозим активируется, когда концентрация магния достигает уровня, достаточного для расщепления, например, в клетке-хозяине. В некоторых вариантах осуществления *circRNA* сохраняют в среде с низким содержанием магния до доставки в клетку-хозяина. В некоторых вариантах осуществления рибозим представляет собой рибозим, отвечающий на белки. В некоторых вариантах осуществления рибозим представляет собой рибозим, отвечающий на нуклеиновые кислоты. В некоторых вариантах осуществления *circRNA* содержит сайт расщепления. В некоторых вариантах осуществления *circRNA* содержит второй сайт расщепления.

В некоторых вариантах осуществления *circRNA* линейаризуется в ядре клетки-мишени. В некоторых вариантах осуществления линейаризация *circRNA* в ядре клетки задействует компоненты, присутствующие в ядре клетки, например, для активации события расщепления. Например, ретротранспозоны B2 и ALU содержат саморасщепляющиеся рибозимы, активность которых усиливается за счет взаимодействия с белком Polycomb, EZH2 (Hernandez et al. PNAS 117(1):415-425 (2020)). Таким образом, в некоторых вариантах осуществления рибозим, например, рибозим из элемента B2 или ALU, который отвечает на ядерный элемент, например, ядерный белок, например, белок, взаимодействующий с геномом, например, эпигенетический модификатор, например, EZH2, включен в *circRNA*, например, системы Gene Writing. В некоторых вариантах осуществления ядерная локализация *circRNA* приводит в результате к увеличению автокаталитической активности рибозима и линейаризации *circRNA*.

В некоторых вариантах осуществления рибозим гетерологичен одному или нескольким другим компонентам системы Gene Writing. В некоторых вариантах осуществления индуцируемый рибозим (например, в *circRNA*, описанной в данном документе) создается синтетически, например, с использованием конструкции аптамера, отвечающего на белковый лиганд. Описана система для использования сателлитной РНК рибозима типа “головка молота” вируса кольцевой пятнистости табака с аптамером белка оболочки MS2 (Kennedy et al. Nucleic Acids Res 42(19):12306-12321 (2014), включена в данный документ посредством ссылки во всей своей полноте), что приводит в результате к активации активности рибозима в присутствии белка оболочки MS2. В вариантах осуществления такая система отвечает на белок-лиганд, локализованный в цитоплазме или ядре. В некоторых вариантах осуществления белковый лиганд не представляет собой MS2. Способы получения РНК-аптамеров для лигандов-мишеней были описаны, например, на основе систематической эволюции лигандов экспоненциальным обогащением (SELEX) (Tuerk and Gold, Science 249(4968):505-510 (1990); Ellington and Szostak, Nature 346(6287):818-822 (1990); при этом способы из каждого из них включены в данный документ посредством ссылки), и в некоторых случаях им способствовало конструирование *in silico* (Bell et al. PNAS 117(15):8486-8493, способы из которого включены в данный документ посредством ссылки). Таким образом, в некоторых вариантах осуществления

аптамер для лиганда-мишени создается и включается в синтетическую рибозимную систему, например, для запуска опосредованного рибозимом расщепления и линеаризации *circRNA*, например, в присутствии белкового лиганда. В некоторых вариантах осуществления линеаризация *circRNA* запускается в цитоплазме, например, с использованием аптамера, который ассоциирует с лигандом в цитоплазме. В некоторых вариантах осуществления линеаризация *circRNA* запускается в ядре, например, с использованием аптамера, который ассоциирует с лигандом в ядре. В вариантах осуществления лиганд в ядре предусматривает эпигенетический модификатор или фактор транскрипции. В некоторых вариантах осуществления лиганд, который запускает линеаризацию, присутствует на более высоких уровнях в клетках-мишенях, чем в клетках, не являющихся мишенями.

Дополнительно предполагается, что для линеаризации *circRNA* можно использовать рибозимную систему, отвечающую на нуклеиновые кислоты. Например, биосенсоры, которые воспринимают определенные молекулы нуклеиновой кислоты-мишени для запуска активации рибозима, описаны, например, в работе Penchovsky (Biotechnology Advances 32(5):1015-1027 (2014), включенной в данный документ посредством ссылки). С помощью этих способов рибозим естественным образом сворачивается в неактивное состояние и активируется только в присутствии определенной молекулы нуклеиновой кислоты-мишени (например, молекулы РНК). В некоторых вариантах осуществления *circRNA* системы Gene Writing содержит рибозим, отвечающий на нуклеиновые кислоты, который активируется в присутствии определенной нуклеиновой кислоты-мишени, например, РНК, например, мРНК, miRNA, направляющей РНК, gRNA, sgRNA, ncRNA, lncRNA, tРНК, snRNA или mtRNA. В некоторых вариантах осуществления нуклеиновая кислота, которая запускает линеаризацию, присутствует на более высоких уровнях в клетках-мишенях, чем в клетках, не являющихся мишенями.

В некоторых вариантах осуществления любого из аспектов в данном документе система Gene Writing содержит один или несколько рибозимов с индуцируемой специфичностью к ткани-мишени или клетке-мишени, представляющих интерес, например, рибозим, который активируется лигандом или нуклеиновой кислотой, присутствующими при более высоких уровнях в ткани-мишени или клетке-мишени, представляющих интерес. В некоторых вариантах осуществления система Gene Writing содержит рибозим с индуцируемой специфичностью в отношении субклеточного компартмента, например, ядра, ядрышка, цитоплазмы или митохондрий. В некоторых вариантах осуществления рибозим, который активируется лигандом или нуклеиновой кислотой, присутствует при более высоких уровнях в субклеточном компартменте-мишени. В некоторых вариантах осуществления РНК-компонент системы Gene Writing предусмотрен в виде *circRNA*, например, которая активируется посредством линеаризации. В некоторых вариантах осуществления линеаризация *circRNA*, кодирующей полипептид Gene Writing, активирует молекулу для трансляции. В некоторых вариантах осуществления сигнал, который активирует компонент *circRNA* системы Gene Writing, присутствует при

более высоких уровнях в клетках-мишенях или тканях-мишенях, например, таким образом, что система специфически активируется в этих клетках.

В некоторых вариантах осуществления РНК-компонент системы Gene Writing предусмотрен в виде circRNA, которая инактивируется посредством линеаризации. В некоторых вариантах осуществления circRNA, кодирующая полипептид Gene Writer, инактивируется посредством расщепления и деградации. В некоторых вариантах осуществления circRNA, кодирующая полипептид Gene Writing, инактивируется путем расщепления, при котором сигнал трансляции отделяется от последовательности, кодирующей полипептид. В некоторых вариантах осуществления сигнал, который инактивирует компонент circRNA системы Gene Writing, присутствует при более высоких уровнях в клетках или тканях, не являющихся мишенями, таким образом, что система специфически инактивируется в этих клетках.

Получение композиций и систем

Специалисту в данной области техники будет понятно, что способы конструирования и создания конструкций на основе нуклеиновых кислот и белков или полипептидов (таких как системы, конструкции и полипептиды, описанные в данном документе) являются стандартными в данной области техники. Как правило, можно применять способы на основе рекомбинации. См. в целом Smales & James (Eds.), *Therapeutic Proteins: Methods and Protocols* (Methods in Molecular Biology), Humana Press (2005); и Crommelin, Sindelar & Meibohm (Eds.), *Pharmaceutical Biotechnology: Fundamentals and Applications*, Springer (2013). Способы конструирования, получения, оценивания, очистки композиций на основе нуклеиновых кислот и произведения манипуляций с ними описаны в Green и Sambrook (Eds.), *Molecular Cloning: A Laboratory Manual* (четвертое издание), Cold Spring Harbor Laboratory Press (2012).

В настоящем изобретении частично предусмотрена нуклеиновая кислота, например, вектор, кодирующий полипептид Gene Writer, описанный в данном документе, нуклеиновая кислота-матрица, описанная в данном документе, или и то, и другое. В некоторых вариантах осуществления вектор содержит селективный маркер, например, маркер устойчивости к антибиотику. В некоторых вариантах осуществления маркер устойчивости к антибиотикам представляет собой маркер устойчивости к канамицину. В некоторых вариантах осуществления маркер устойчивости к антибиотикам не придает устойчивости к бета-лактамам антибиотикам. В некоторых вариантах осуществления вектор не содержит маркер устойчивости к ампициллину. В некоторых вариантах осуществления вектор содержит маркер устойчивости к канамицину и не содержит маркера устойчивости к ампициллину. В некоторых вариантах осуществления вектор, кодирующий полипептид Gene Writer, интегрирован в геном клетки-мишени (например, при введении в клетку-мишень, ткань-мишень, орган-мишень или субъекту-мишени). В некоторых вариантах осуществления вектор, кодирующий полипептид Gene Writer, не интегрирован в геном клетки-мишени (например, при введении в клетку-мишень, ткань-мишень, орган-мишень или субъекту-мишени). В некоторых вариантах осуществления вектор, кодирующий

нуклеиновую кислоту-матрицу (например, РНК-матрицу), не интегрирован в геном клетки-мишени (например, при введении в клетку-мишень, ткань-мишень, орган-мишень или субъекту-мишени). В некоторых вариантах осуществления, если вектор интегрирован в сайт-мишень в геноме клетки-мишени, то селективный маркер не интегрирован в геном. В некоторых вариантах осуществления, если вектор интегрирован в сайт-мишень в геноме клетки-мишени, гены или последовательности, участвующие в сохранении вектора (например, гены сохранения плазмиды), не интегрированы в геном. В некоторых вариантах осуществления, если вектор интегрирован в сайт-мишень в геноме клетки-мишени, последовательности, регулирующие перенос (например, инвертированные концевые повторы, например, из AAV), не интегрированы в геном. В некоторых вариантах осуществления введение вектора (например, кодирующего полипептид Gene Writer, описанный в данном документе, нуклеиновую кислоту-матрицу, описанную в данном документе, или и то, и другое) в клетку-мишень, ткань-мишень, орган-мишень или субъекту-мишени приводит в результате к интеграции части вектора в один или несколько сайтов-мишеней в геноме(геномах) указанной клетки-мишени, ткани-мишени, органа-мишени или субъекта-мишени. В некоторых вариантах осуществления менее 99, 95, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, 4, 3, 2 или 1% сайтов-мишеней (например, ни один из сайтов-мишеней), содержащих интегрированный материал, содержат селективный маркер (например, ген устойчивости к антибиотикам), последовательность, регулирующую перенос (например, инвертированный концевой повтор, например, из AAV), или и то, и другое из вектора.

Иллюстративные способы получения терапевтического фармацевтического белка или полипептида, описанных в данном документе, включают экспрессию в клетках млекопитающих, хотя рекомбинантные белки также могут быть получены с использованием клеток насекомых, дрожжей, бактерий или других клеток под контролем соответствующих промоторов. Векторы экспрессии для клеток млекопитающих могут содержать нетранскрибируемые элементы, такие как точка начала репликации, подходящий промотор и другие 5'- или 3'- фланкирующие нетранскрибируемые последовательности и 5'- или 3'- нетранслируемые последовательности, такие как необходимые сайты связывания рибосомы, сайт полиаденилирования, донорные и акцепторные сайты сплайсинга и последовательности терминации. Последовательности ДНК, полученные из генома вируса SV40, например, точку начала репликации, ранний промотор, сайты сплайсинга и полиаденилирования SV40, можно использовать для получения других генетических элементов, требуемых для экспрессии гетерологичной последовательности ДНК. Подходящие векторы клонирования и экспрессии для применения с клетками-хозяевами, являющимися клетками бактерий, грибов, дрожжей и млекопитающих, описаны в Green & Sambrook, *Molecular Cloning: A Laboratory Manual* (четвертое издание), Cold Spring Harbor Laboratory Press (2012).

Различные системы культивирования клеток млекопитающих можно использовать для экспрессии и получения рекомбинантного белка. Примеры систем экспрессии на основе

клеток млекопитающих включают линии клеток CHO, COS, HEK293, HeLa и BHK. Способы культивирования клеток-хозяев для получения белковых терапевтических средств описаны в работе Zhou and Kantardjieff (Eds.), *Mammalian Cell Cultures for Biologics Manufacturing (Advances in Biochemical Engineering/Biotechnology)*, Springer (2014). Композиции, описанные в данном документе, могут содержать вектор, такой как вирусный вектор, например, лентивирусный вектор, кодирующий рекомбинантный белок. В некоторых вариантах осуществления вектор, например, вирусный вектор, может содержать нуклеиновую кислоту, кодирующую рекомбинантный белок.

Очистка белковых терапевтических средств описана в работе Franks, *Protein Biotechnology: Isolation, Characterization, and Stabilization*, Humana Press (2013); и в Cutler, *Protein Purification Protocols (Methods in Molecular Biology)*, Humana Press (2010).

Пути применения

В некоторых вариантах осуществления система Gene Writer™, описанная в данном документе, может использоваться для модификации клетки животного, клетки растения или клетки гриба. В некоторых вариантах осуществления система Gene Writer™, описанная в данном документе, может использоваться для модификации клетки млекопитающего (например, клетки человека). В некоторых вариантах осуществления система Gene Writer™, описанная в данном документе, может использоваться для модификации клетки животного, относящегося к домашнему скоту (например, коровы, лошади, овцы, козы, свиньи, ламы, альпаки, верблюда, яка, курицы, утки, гуся или страуса). В некоторых вариантах осуществления система Gene Writer™, описанная в данном документе, может использоваться в качестве лабораторного инструмента или исследовательского инструмента или использована в лабораторном методе или методе исследования, например, для модификации клетки животного, например, клетки млекопитающего (например, клетки человека), клетки растения или клетки гриба.

Способы модификации растений

Системы Gene Writer, описанные в данном документе, можно применять для модификации растения или частей растения (например, листьев, корней, цветков, плодов или семян), например, чтобы повышать приспособленность растения.

А. Доставка в растение

В данном документе предусмотрены способы доставки системы Gene Writer, описанной в данном документе, в растение. Включены способы доставки системы Gene Writer в растение путем приведения растения или его части в контакт с системой Gene Writer. Способы применимы для модификации растения, например, чтобы повышать приспособленность растения.

Более конкретно, в некоторых вариантах осуществления нуклеиновая кислота, описанная в данном документе (например, нуклеиновая кислота, кодирующая GeneWriter), может быть закодирована в векторе, например, вставлена вблизи с промотором растения, например, промотором убиквитина кукурузы (ZmUBI) в векторе для растения (например, pHUC411). В некоторых вариантах осуществления нуклеиновые кислоты, описанные в

данном документе, вводят в растение (например, рис подвида *japonica*) или часть растения (например, каллюс растения) посредством агробактерий. В некоторых вариантах осуществления системы и способы, описанные в данном документе, можно применять в растениях путем замены гена растения (например, гигромицинофосфотрансферазы (HPT)) нулевым аллелем (например, содержащим замену основания в стартовом кодоне). Системы и способы модификации генома растения описаны в работе Xu et. al. *Development of plant prime-editing systems for precise genome editing*, 2020, Plant Communications.

В одном аспекте в данном документе представлен способ повышения приспособленности растения, где способ включает доставку в растение системы Gene Writer, описанной в данном документе (например, в эффективном количестве и в течение эффективного периода времени), для повышения приспособленности растения по сравнению с необработанным растением (например, растением, в которое не доставили систему Gene Writer).

Повышение приспособленности растения как следствие доставки системы Gene Writer может проявляться многими путями, например, приводя тем самым к более высокой производительности растения, например, улучшенной урожайности, улучшенной жизнеспособности растения или качества продукта, собираемого с растения, улучшению признаков, имеющихся до или после сбора урожая, которые считаются необходимыми для сельского хозяйства или садоводства (например, вкуса, внешнего вида, срока хранения), или улучшению признаков, которые иным образом приносят пользу людям (например, понижение продуцирования аллергенов). Улучшение урожайности растения относится к повышению выхода продукта (например, определяемого по выходу биомассы, зерна, семян или плодов растения, содержанию белка, углеводов или масел или площади листьев) растения на измеримую величину по сравнению с выходом того же продукта растения, полученного в тех же условиях, но без применения композиций по настоящему изобретению, или по сравнению с применением обычных средств для модификации растений. Например, урожайность может быть повышена на по меньшей мере приблизительно 0,5%, приблизительно 1%, приблизительно 2%, приблизительно 3%, приблизительно 4%, приблизительно 5%, приблизительно 10%, приблизительно 20%, приблизительно 30%, приблизительно 40%, приблизительно 50%, приблизительно 60%, приблизительно 70%, приблизительно 80%, приблизительно 90%, приблизительно 100% или более 100%. В некоторых случаях способ является эффективным для повышения урожайности в приблизительно 2, 5, 10, 25, 50, 75, 100 или более 100 раз по сравнению с необработанным растением. Урожайность может быть представлена в виде количества по весу или объему растения или продукта растения относительно некоторого параметра. Такой параметр может быть представлен в виде времени, площади выращивания, веса полученных растений или количества используемого сырья. Например, такие способы могут обеспечивать повышение выхода тканей растения, включая без ограничения семена, плоды, ядра, семенные коробочки, клубни, корни и листья.

Повышение жизнеспособности растения как следствие доставки системы Gene

Writer также можно измерять другими способами, как, например, по повышению или улучшению показателя жизнеспособности, густоты (количества растений на единицу площади), высоты растения, окружности стебля, длины стебля, количества листьев, размера листа, листового полога, внешнего вида (например, более зеленый цвет листьев), показателя корней, всхожести, содержания белка, повышенному кущению, более крупным листьям, большему количеству листьев, меньшему количеству мертвых прикорневых листьев, более сильным росткам, потребности в меньшем количестве удобрений, потребности в меньшем количестве семян, более продуктивным росткам, более раннему цветению, более раннему созреванию зерна или семян, меньшему полеганию растений (вылеганию), усиленному росту побегов, более раннему прорастанию или любой комбинации этих факторов на измеримую или заметную величину по сравнению с тем же фактором растения, производимого в тех же условиях, но без введения композиций по настоящему изобретению или с применением общепринятых средств для модификации растений (например, средств для модификации растений, доставляемых без использования РМР).

Соответственно, в данном документе представлен способ модификации растения, при этом способ включает доставку в растение эффективного количества любой из систем Gene Writer, предусмотренной в данном документе, где с помощью данного способа осуществляют модификацию растения и за счет этого осуществляют введение полезного признака в растение или повышение его уровня (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением. В частности, способ может обеспечивать повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В некоторых случаях повышение приспособленности растения представляет собой повышение (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) устойчивости к болезням, засуховыносливости, жаровыносливости, холодовыносливости, солевыносливости, выносливости к металлам, выносливости к гербицидам, выносливости к химическим веществам, эффективности использования воды, использования азота, устойчивости к азотному стрессу, фиксации азота, устойчивости к вредителям, устойчивости к травоядным животным, устойчивости к патогенам, урожайности, урожайности в условиях ограниченного количества воды, жизнеспособности, роста, фотосинтетического потенциала, питания, содержания белка, содержания углеводов, содержания масла, биомассы, длины побегов, длины корней, показателя структуры корневой системы, веса семян или количества собираемого урожая.

В некоторых случаях повышение приспособленности представляет собой повышение (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) скорости развития, роста, урожайности, устойчивости к абиотическим стрессовым факторам или устойчивости к биотическим

стрессовым факторам. Абиотический стресс относится к стрессовому воздействию окружающей среды, которому подвергается растение или часть растения, которое включает, например, стресс, вызываемый засухой, солевой стресс, тепловой стресс, холодовой стресс и стресс, вызываемый низким содержанием питательных веществ. Биотический стресс относится к стрессовому воздействию окружающей среды, которому подвергается растение или часть растения, которое включает, например, стресс, вызываемый нематодами, стресс, вызываемый травоядными насекомыми, стресс, вызываемый грибковыми патогенами, стресс, вызываемый бактериальными патогенами, или стресс, вызываемый вирусными патогенами. Стресс может быть временным, например, длится несколько часов, несколько дней, несколько месяцев, или постоянным, например, длится в течение жизни растения.

В некоторых случаях повышение приспособленности растения представляет собой повышение (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) качества продуктов, собираемых с растения. Например, повышение приспособленности растения может представлять собой улучшение в отношении коммерчески благоприятных признаков (например, вкуса или внешнего вида) продукта, собираемого с растения. В других случаях повышение приспособленности растения представляет собой повышение срока хранения продукта, собираемого с растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%).

В качестве альтернативы повышение приспособленности может представлять собой изменение признака, который полезен для здоровья человека или животного, как, например, снижение продуцирования аллергена. Например, повышение приспособленности может представлять собой снижение (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) продуцирования аллергена (например, пыльцы), который стимулирует иммунный ответ у животного (например, человека).

Модификация растения (например, повышение приспособленности) может возникать в результате модификации одной или нескольких частей растения. Например, растение можно модифицировать посредством приведения в контакт с листом, семенем, пыльцой, корнем, плодом, побегом, цветком, клетками, протопластами или тканью (например, меристематической тканью) растения. По этой причине в другом аспекте в данном документе представлен способ повышения приспособленности растения, при этом способ включает приведение пыльцы растения в контакт с эффективным количеством любой из композиций для модификации растений, представленных в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В еще одном аспекте в данном документе представлен способ повышения приспособленности растения, при этом способ включает приведение семени растения в

контакт с эффективным количеством любой из систем Gene Writer, раскрытых в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В другом аспекте в данном документе представлен способ, включающий приведение протопласта растения в контакт с эффективным количеством любой из систем Gene Writer, описанных в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В дополнительном аспекте в данном документе представлен способ повышения приспособленности растения, при этом способ включает приведение растительной клетки растения в контакт с эффективным количеством любой из систем Gene Writer, описанных в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В другом аспекте в данном документе представлен способ повышения приспособленности растения, при этом способ включает приведение меристематической ткани растения в контакт с эффективным количеством любой из композиций для модификации растений, представленных в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В другом аспекте в данном документе представлен способ повышения приспособленности растения, при этом способ включает приведение зародыша растения в контакт с эффективным количеством любой из композиций для модификации растений, представленных в данном документе, где способ обеспечивает повышение приспособленности растения (например, на приблизительно 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более 100%) по сравнению с необработанным растением.

В. Способы применения

Растение, описанное в данном документе, можно подвергать воздействию любой из композиций на основе системы Gene Writer, описанных в данном документе, с помощью любого подходящего способа, который обеспечивает доставку или введение композиции в растение. Систему Gene Writer можно доставлять либо отдельно, либо в комбинации с другими активными (например, удобрениями) или неактивными веществами и можно наносить, например, посредством распыления, инъекции (например, микроинъекции), через растения, посредством полива, погружения, в форме концентрированных жидкостей, гелей, растворов, суспензий, спреев, порошков, пеллет, брикетов, плиток и т. п., составленных для доставки эффективной концентрации композиции для модификации растений. Количества и места для применения композиций, описанных в данном

документе, как правило, определяются условиями обитания растения, стадией жизненного цикла, на которой на растение можно целенаправленно воздействовать композицией для модификации растений, участком, в котором следует выполняться применение, а также физическими и функциональными характеристиками композиции для модификации растений.

В некоторых случаях композицию распыляют непосредственно на растение, например, сельскохозяйственные культуры, например, путем распыления из ранца, распыления с воздуха, опрыскивания/опыления сельскохозяйственных культур и т. д. В тех случаях, когда систему Gene Writer доставляют в растение, растение, получающее систему Gene Writer, может находиться на любой стадии роста растений. Например, составленные композиции для модификации растений можно применять в виде покрытия для семян или средства для обработки корней на ранних стадиях роста растения или в виде средства для полной обработки растения на более поздних стадиях цикла сельскохозяйственного растения. В некоторых случаях композицию для модификации растений можно применять по отношению к растению в виде средства для местного применения.

Кроме того, систему Gene Writer можно применять (например, в отношении почвы, в которой растет растение, или в отношении воды, которую используют для полива растения) в качестве системного средства, которое поглощается и распределяется по тканям растения. В некоторых случаях растения или употребляемые в пищу организмы могут быть генетически трансформированы для экспрессии системы Gene Writer.

Замедленного или непрерывного высвобождения также можно достигнуть за счет покрытия системы Gene Writer или композиции с композицией(композициями) для модификации растений растворимым или биodeградируемым покровным слоем, таким как желатин, при этом данное покрытие растворяется или разрушается в среде его использования, что делает тогда композицию на основе системы Gene Writer для модификации растений доступной, или за счет распределения средства в растворимой или разрушаемой матрице. Таки средства/устройства с непрерывным высвобождением и/или распределением могут преимущественно использоваться для устойчивого поддержания эффективной концентрации одной или нескольких композиций для модификации растений, описанных в данном документе.

В некоторых случаях систему Gene Writer доставляют в часть растения, например, лист, семя, пыльцу, корень, плод, побег или цветок, или ткань, клетку или протопласт. В некоторых случаях систему Gene Writer доставляют в клетку растения. В некоторых случаях систему Gene Writer доставляют в протопласт растения. В некоторых случаях систему Gene Writer доставляют в ткань растения. Например, композицию можно доставлять в меристематическую ткань растения (например, апикальную меристему, латеральную меристему или интеркалярную меристему). В некоторых случаях композицию доставляют в постоянную ткань растения (например, простые ткани (например, паренхиму, колленхиму или склеренхиму) или сложную постоянную ткань (например, ксилему или флоэму)). В некоторых случаях систему Gene Writer доставляют в зародыш растения.

С. Растения

Систему Gene Writer, описанную в данном документе, можно доставлять во множество растений или обрабатывать их ею. Растения, в которые может быть доставлена система Gene Writer (т. е. "подвергаемые обработке") в соответствии со способами по настоящему изобретению, предусматривают целые растения и их части, в том числе без ограничения органы/структуры вегетативных побегов (например, листья, стебли и клубни), корни, цветки и органы/структуры цветков (например, прицветники, чашелистики, лепестки, тычинки, плодолистики, пыльники и семяпочки), семена (в том числе зародыш, эндосперм, семядоли и оболочку семени) и плоды (зрелую завязь), ткань растения (например, сосудистую ткань, основную паренхиму и т. п.) и клетки (например, замыкающие клетки, яйцеклетки и т. п.) и их потомство. Дополнительно части растения могут относиться к частям растения, таким как побег, корень, стебель, семена, прилистники, листья, лепестки, цветки, семяпочки, прицветники, ветви, черешки, междоузлия, кора, опушение, ростки, корневища, вайи, пластинки, пыльца, тычинки и т. п.

Класс растений, которые можно обрабатывать способом, раскрытым в данном документе, включает класс высших и низших растений, в том числе покрытосеменные (однодольные и двудольные растения), голосеменные, папоротники, хвощи, псилофиты, ликофиты, мохообразные и водоросли (например, многоклеточные или одноклеточные водоросли). Растения, которые можно обрабатывать в соответствии со способами по настоящему изобретению, дополнительно включают любое сосудистое растение, например, однодольные, или двудольные, или голосеменные, в том числе без ограничения люцерну, яблоню, растение рода *Arabidopsis*, банан, ячмень, канолу, клещевину, цикорий, хризантему, клевер, какао, кофе, хлопчатник, семена хлопчатника, кукурузу, крамбе, клюкву, огурец, дендробиум, диоскорею, эвкалипт, овсяницу, лен, гладиолус, растения семейства Лилейные, семя льна, просо, дыню, горчицу, овес, масличную пальму, масличный рапс, папайю, арахис, ананас, декоративные растения, фасоль, картофель, рапс, рис, рожь, райграс, сафлор, кунжут, сорго, сою, сахарную свеклу, сахарный тростник, подсолнечник, клубнику, табак, томат, газонную траву, пшеницу и овощные культуры, такие как салат, сельдерей, брокколи, цветная капуста, тыквенные культуры; фруктовые и ореховые деревья, такие как яблоня, груша, персик, апельсин, грейпфрут, лимон, лайм, миндаль, пекан, грецкий орех, лещина; вьющиеся растения, такие как виноград (например, виноградник), киви, хмель; плодовые кустарники и колючие кустарники, такие как малина, ежевика, крыжовник; лесные деревья, такие как ясень, сосна, пихта, клен, дуб, каштан, тополь; такое как люцерна, канола, клещевина, кукуруза, хлопчатник, крамбе, лен, семя льна, горчицу, масличную пальму, масличный рапс, арахис, картофель, рис, сафлор, кунжут, сою, сахарную свеклу, подсолнечник, табак, томат или пшеницу. Растения, которые можно обрабатывать в соответствии со способами по настоящему изобретению, включают любое культурное растение, например, кормовую культуру, масличную культуру, зерновую культуру, плодовую культуру, овощную культуру, волокнистую культуру, пряную культуру, ореховую культуру, газонную культуру, сахарную культуру,

культуру для производства напитков и лесную культуру. В определенных случаях культурное растение, которое обрабатывают в этом способе, представляет собой растение сои. В других определенных случаях культурное растение представляет собой пшеницу. В определенных случаях культурное растение представляет собой кукурузу. В определенных случаях культурное растение представляет собой хлопчатник. В определенных случаях культурное растение представляет собой люцерну. В определенных случаях культурное растение представляет собой сахарную свеклу. В определенных случаях культурное растение представляет собой рис. В определенных случаях культурное растение представляет собой томат. В определенных случаях культурное растение представляет собой картофель.

В определенных случаях растение представляет собой сельскохозяйственную культуру. Примеры таких культурных растений включают без ограничения однодольные и двудольные растения, в том числе без ограничения кормовые растения или кормовые бобовые растения, декоративные растения, продовольственные культуры, деревья или кустарники, выбранные из *Acer* spp., *Allium* spp., *Amaranthus* spp., *Ananas comosus*, *Apium graveolens*, *Arachis* spp., *Asparagus officinalis*, *Beta vulgaris*, *Brassica* spp. (например, *Brassica napus*, *Brassica rapa* ssp. (канола, масличный рапс, репа масличная), *Camellia sinensis*, *Canna indica*, *Cannabis saliva*, *Capsicum* spp., *Castanea* spp., *Cichorium endivia*, *Citrullus lanatus*, *Citrus* spp., *Cocos* spp., *Coffea* spp., *Coriandrum sativum*, *Corylus* spp., *Crataegus* spp., *Cucurbita* spp., *Cucumis* spp., *Daucus carota*, *Fagus* spp., *Ficus carica*, *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (например, *Glycine max*, *Soja hispida* или *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (например, *Helianthus annuus*), *Hibiscus* spp., *Hordeum* spp. (например, *Hordeum vulgare*), *Ipomoea batatas*, *Juglans* spp., *Lactuca sativa*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus* spp., *Luffa acutangula*, *Lupinus* spp., *Lycopersicon* spp. (например, *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*), *Malus* spp., *Medicago sativa*, *Mentha* spp., *Miscanthus sinensis*, *Morus nigra*, *Musa* spp., *Nicotiana* spp., *Olea* spp., *Oryza* spp. (например, *Oryza sativa*, *Oryza latifolia*), *Panicum miliaceum*, *Panicum virgatum*, *Passiflora edulis*, *Petroselinum crispum*, *Phaseolus* spp., *Pinus* spp., *Pistacia vera*, *Pisum* spp., *Poa* spp., *Populus* spp., *Prunus* spp., *Pyrus communis*, *Quercus* spp., *Raphanus sativus*, *Rheum rhabarbarum*, *Ribes* spp., *Ricinus communis*, *Rubus* spp., *Saccharum* spp., *Salix* sp., *Sambucus* spp., *Secale cereale*, *Sesamum* spp., *Sinapis* spp., *Solanum* spp. (например, *Solanum tuberosum*, *Solanum integrifolium* или *Solanum lycopersicum*), *Sorghum bicolor*, *Sorghum halepense*, *Spinacia* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Triticosecale rimpaui*, *Triticum* spp. (например, *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* или *Triticum vulgare*), *Vaccinium* spp., *Vicia* spp., *Vigna* spp., *Viola odorata*, *Vitis* spp. и *Zea mays*. В определенных вариантах осуществления культурное растение представляет собой рис, масличный рапс, рапс, сою, кукурузу (маис), хлопчатник, сахарный тростник, люцерну, сорго или пшеницу.

Растение или часть растения для применения в настоящем изобретении включают растения на любой стадии развития растения. В определенных случаях доставка может

происходить в ходе стадий прорастания, роста проростков, вегетативного роста и репродуктивного роста. В определенных случаях доставка в растение происходит в ходе стадий вегетативного и репродуктивного роста. В некоторых случаях композицию доставляют в пыльцу растения. В некоторых случаях композицию доставляют в семя растения. В некоторых случаях композицию доставляют в протопласт растения. В некоторых случаях композицию доставляют в ткань растения. Например, композицию можно доставлять в меристематическую ткань растения (например, апикальную меристему, латеральную меристему или интеркалярную меристему). В некоторых случаях композицию доставляют в постоянную ткань растения (например, простые ткани (например, паренхиму, колленхиму или склеренхиму) или сложную постоянную ткань (например, ксилему или флоэму)). В некоторых случаях композицию доставляют в зародыш растения. В некоторых случаях композицию доставляют в растительную клетку. Стадии вегетативного и репродуктивного роста также обозначаются в данном документе как "взрослые" или "зрелые" растения.

В тех случаях, когда систему Gene Writer доставляют в часть растения, часть растения может быть модифицирована с помощью средства для модификации растений. В качестве альтернативы система Gene Writer, может распределяться в другие части растения (например, по сосудистой системе растения), которые впоследствии модифицируются средством для модификации растений.

Введение AAV

В некоторых вариантах осуществления аденоассоциированный вирус (AAV) используется в сочетании с системой, нуклеиновой кислотой-матрицей и/или полипептидом, описанными в данном документе. В некоторых вариантах осуществления AAV используется для доставки, введения или упаковки системы, нуклеиновой кислоты-матрицы и/или полипептида, описанных в данном документе. В некоторых вариантах осуществления AAV представляет собой рекомбинантный AAV (rAAV).

В некоторых вариантах осуществления система содержит следующее: (a) полипептид, описанный в данном документе, или нуклеиновую кислоту, кодирующую его, (b) нуклеиновую кислоту-матрицу (например, РНК-матрицу), описанную в данном документе, и (c) одну или несколько первых тканеспецифических последовательностей контроля экспрессии, специфичных в отношении ткани-мишени, где одна или несколько первых тканеспецифических последовательностей контроля экспрессии, специфичных в отношении ткани-мишени, находятся в функциональной ассоциации с (a), (b) или (a) и (b), где при ассоциации с (a), (a) предусматривает нуклеиновую кислоту, кодирующую полипептид.

В некоторых вариантах осуществления система, описанная в данном документе, дополнительно содержит первый капсидный белок рекомбинантного аденоассоциированного вируса (rAAV); где по меньшей мере одно из (a) или (b) ассоциировано с первым капсидным белком rAAV, где по меньшей мере одно из (a) или (b) фланкировано инвертированными концевыми повторами (ITR) AAV.

В некоторых вариантах осуществления (a) и (b) ассоциированы с первым капсидным белком gAAV.

В некоторых вариантах осуществления (a) и (b) находятся на одной нуклеиновой кислоте.

В некоторых вариантах осуществления система дополнительно содержит второй капсидный белок gAAV, где по меньшей мере одно из (a) или (b) ассоциировано со вторым капсидным белком gAAV, и где по меньшей мере одно из (a) или (b), ассоциированного со вторым капсидным белком gAAV, отличается от по меньшей мере одного из (a) или (b), ассоциированного с первым капсидным белком gAAV.

В некоторых вариантах осуществления по меньшей мере одно из (a) или (b) ассоциировано с первым или вторым капсидным белком gAAV, диспергировано внутри первого или второго капсидного белка gAAV, где первый или второй капсидный белок gAAV находится в форме капсидной частицы AAV.

В некоторых вариантах осуществления система дополнительно содержит наночастицу, где наночастица ассоциирована с по меньшей мере одним из (a) или (b).

В некоторых вариантах осуществления (a) и (b) соответственно ассоциированы со следующим: а) первым капсидным белком gAAV и вторым капсидным белком gAAV; b) наночастицей и первым капсидным белком gAAV; c) первым капсидным белком gAAV; d) первым капсидным белком аденовируса; e) первой наночастицей и второй наночастицей; или f) первой наночастицей.

Вирусные векторы применимы для доставки всей или части системы, предусмотренной в настоящем изобретении, например, для применения в способах, предусмотренных в настоящем изобретении. Системы, полученные из различных вирусов, использовались для доставки полипептидов, нуклеиновых кислот или транспозонов; например: лентивирус с недостаточностью интегразы, аденовирус, аденоассоциированный вирус (AAV), вирус простого герпеса и бакуловирус (описано в работе Hodge et al. *Hum Gene Ther* 2017; Narayanavari et al. *Crit Rev Biochem Mol Biol* 2017; Boehme et al., *Curr Gene Ther* 2015).

Аденовирусы представляют собой распространенные вирусы, которые использовались в качестве средств доставки генов благодаря наличию точных данных по их биологии, генетической стабильности, высокой эффективности трансдукции и простоте крупномасштабного получения (см., например, обзор Lee et al. *Genes & Diseases* 2017). Они имеют линейные геномы в виде dsDNA и представлены множеством серотипов, различающихся вариантами тканевого и клеточного тропизма. Чтобы предупредить репликацию инфекционного вируса в клетках-реципиентах, в геномах аденовирусов, используемых для упаковки, удалены некоторые или все эндогенные вирусные белки, которые обеспечиваются в клетках, продуцирующих вирус, в транс-положении. Это делает геномы зависимыми от хелперов, что означает, что они могут быть реплицированы и упакованы в вирусные частицы только в присутствии недостающих компонентов, обеспечиваемых так называемыми хелперными функциями. Хелпер-зависимая

аденовирусная система с удаленными всеми вирусными ORF может быть совместимой с упаковкой чужеродной ДНК размером до ~37 т. о. (Parks et al. J Virol 1997). В некоторых вариантах осуществления аденовирусный вектор используется для доставки ДНК, соответствующей полипептидному или матричному компоненту системы Gene Writing™, или оба содержатся в отдельных аденовирусных векторах или одном и том же аденовирусном векторе. В некоторых вариантах осуществления аденовирус представляет собой хелпер-зависимый аденовирус (HD-AdV), который не способен к самостоятельной упаковке. В некоторых вариантах осуществления аденовирус представляет собой аденовирус высокой емкости (HC-AdV), в котором удалены все эндогенные вирусные ORF или их значительная часть, при этом сохранены необходимые компоненты последовательности для упаковки в аденовирусные частицы. Для этого типа вектора единственными аденовирусными последовательностями, требующимися для упаковки генома, являются не кодирующие последовательности: инвертированные концевые повторы (ITR) на обоих концах и сигнал упаковки на 5'-конце (Jager et al. Nat Protoc 2009). В некоторых вариантах осуществления геном аденовируса также содержит ДНК-заполнитель для обеспечения соответствия минимальному размеру генома для оптимального продуцирования и стабильности (см., например, Hausl et al. Mol Ther 2010). Аденовирусы использовались в данной области техники для доставки транспозонов в различные ткани. В некоторых вариантах осуществления аденовирус используется для доставки системы Gene Writing™ в печень.

В некоторых вариантах осуществления аденовирус используется для доставки системы Gene Writing™ в гемопоэтические стволовые клетки, например, HDAd5/35++. HDAd5/35++ представляет собой аденовирус с модифицированными волокнами серотипа 35, которые нацеливают вектор на мишень, отличную от печени (Wang et al. Blood Adv 2019). В некоторых вариантах осуществления аденовирус, который доставляет систему Gene Writing™ в гемопоэтические стволовые клетки, использует рецептор, который специфически экспрессируется на примитивных гемопоэтических стволовых клетках, например, CD46.

Аденоассоциированные вирусы (AAB) принадлежат к семейству Parvoviridae и более конкретно составляют род Dependoparvovirus. Геном AAV состоит из линейной одностранных молекулы ДНК, которая содержит примерно 4,7 тысяч оснований (т. о.) и состоит из двух основных открытых рамок считывания (ORF), кодирующих неструктурные белки Rep (репликации) и структурные белки Cap (капсидные). Была идентифицирована вторая ORF внутри гена cap, которая кодирует белок, активирующий сборку (AAP). ДНК, фланкируемые кодирующими областями AAV, представляют собой две действующие в cis-положении последовательности инвертированных концевых повторов (ITR) длиной примерно 145 нуклеотидов с прерывистыми палиндромными последовательностями, которые могут быть уложены в энергетически стабильные структуры шпилек, которые функционируют в качестве праймеров репликации ДНК. Было показано, что в дополнение к их роли в репликации ДНК последовательности ITR участвуют в интеграции вирусной

ДНК в клеточный геном, высвобождении из генома хозяина или плазмиды и инкапсулировании вирусной нуклеиновой кислоты в зрелые вирионы (Muzyczka, (1992) Curr. Top. Micro. Immunol. 158:97-129). В некоторых вариантах осуществления один или несколько компонентов нуклеиновой кислоты Gene Writing™ фланкированы с помощью ITR, полученных из AAV, для обеспечения вирусной упаковки. См., например, WO2019113310.

В некоторых вариантах осуществления один или несколько компонентов системы Gene Writing™ переносят посредством по меньшей мере одного вектора на основе AAV. В некоторых вариантах осуществления по меньшей мере один вектор на основе AAV выбран, исходя из его тропизма к конкретной клетке, ткани, организму. В некоторых вариантах осуществления вектор на основе AAV является псевдотипированным, например, AAV2/8, где AAV2 описывает дизайн конструкции, но капсидный белок заменен капсидным белком из AAV8. Понятно, что любой из описанных векторов может представлять собой псевдотипированное производное, в котором капсидный белок, используемый для упаковки генома AAV, получен из AAV другого серотипа. Перечень иллюстративных серотипов AAV можно найти в таблице 36, что однако не предусматривает ограничения в выборе вектора. В некоторых вариантах осуществления AAV, предназначенный для использования для Gene Writing™, может быть получен методом направленной эволюции для обеспечения тропизма к новым клеткам или тканям, как было продемонстрировано в литературе (например, Davidsson et al. Proc Natl Acad Sci U S A 2019).

В некоторых вариантах осуществления вектор на основе AAV для доставки представляет собой вектор, который содержит два инвертированных концевых повтора (ITR) AAV и нуклеотидную последовательность, представляющую интерес (например, последовательность, кодирующую полипептид Gene Writer™ или ДНК-матрицу, или и то, и другое), при этом каждый из указанных ITR содержит прерывистую (или несплошную) палиндромную последовательность, т. е. последовательность, состоящую из трех сегментов: первого сегмента и последнего сегмента, которые идентичны при чтении 5'→3', но гибридизуются при размещении друг против друга, и сегмента, который отличается, который разделяет одинаковые сегменты. Такие последовательности, особенно ITR, образуют шпильчатые структуры. См., например, WO2012123430.

Общепринято, что вирионы AAV с капсидами получают путем введения плазмиды или плазмид, кодирующих геном gAAV или scAAV, белки Rep и белки Cap (Grimm et al., 1998). При введении этих хелперных плазмид в транс-положении геном AAV "освобождается" (т. е. высвобождается и затем восстанавливается) из генома хозяина и в дальнейшем инкапсулируется с получением инфекционного AAV. В некоторых вариантах осуществления одну или несколько нуклеиновых кислот Gene Writing™ упаковывают в частицы AAV посредством введения нуклеиновых кислот, фланкированных ITR, в упаковывающую клетку в сочетании с хелперными функциями.

В некоторых вариантах осуществления геном AAV представляет собой так называемый самокомплементарный геном (называемый scAAV), в котором

последовательность, расположенная между ITR, содержит как необходимую последовательность нуклеиновой кислоты (например, ДНК, кодирующую полипептид или матрицу Gene Writer™, или и то, и другое), так и, в дополнение, последовательность, обратную комплементарную необходимой последовательности нуклеиновой кислоты, так что эти два компонента могут накладываться друг на друга и самогибридизоваться. В некоторых вариантах осуществления самокомплементарные модули разделены промежуточной последовательностью, которая позволяет ДНК укладываться на свою выше расположенную часть, например, образуя структуру стебель-петля. Преимущество scAAV заключается в том, что он готов к транскрипции при попадании в ядро, а не зависит от первоначального праймирования с помощью ITR и синтеза второй нити для образования dsDNA. В некоторых вариантах осуществления один или несколько компонентов Gene Writing™ сконструированы в виде scAAV, где последовательность между ITR AAV содержит два обратных комплементарных модуля, которые могут гибридизоваться между собой с образованием dsDNA.

В некоторых вариантах осуществления нуклеиновая кислота (например, кодирующая полипептид, или матрицу, или и то, и другое), доставляемая в клетки, представляет собой линейную дуплексную ДНК с закрытыми концами (CELiD-ДНК или ceDNA). В некоторых вариантах осуществления ceDNA получена из репликативной формы генома AAV (Li et al. PLoS One 2013). В некоторых вариантах осуществления нуклеиновая кислота (например, кодирующая полипептид, или ДНК-матрицу, или и то, и другое) фланкирована с помощью ITR, например, ITR AAV, где по меньшей мере один из ITR содержит сайт концевой разрешения и сайт связывания белка репликации (иногда его называют сайтом связывания репликативного белка). В некоторых вариантах осуществления ITR получены из аденоассоциированного вируса, например, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12 или их комбинации. В некоторых вариантах осуществления ITR являются симметричными. В некоторых вариантах осуществления ITR являются асимметричными. В некоторых вариантах осуществления предусмотрен по меньшей мере один белок Rep, обеспечивающий репликацию конструкции. В некоторых вариантах осуществления по меньшей мере один белок Rep получен из аденоассоциированного вируса, например, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12 или их комбинации. В некоторых вариантах осуществления ceDNA генерируют посредством получения клетки-продуцента со следующим: (i) ДНК, фланкированной ITR, например, ITR AAV, и (ii) компонентами, требующимися для ITR-зависимой репликации, например, белками AAV Rep78 и Rep52 (или нуклеиновой кислотой, кодирующей белки). В некоторых вариантах осуществления ceDNA не содержит какого-либо капсидного белка, например, не упакована в инфекционную частицу AAV. В некоторых вариантах осуществления ceDNA составляют с включением в LNP (см., например, WO2019051289A1).

В некоторых вариантах осуществления вектор на основе ceDNA состоит из двух самокомплементарных последовательностей, например, асимметричных, или

симметричных, или по сути симметричных ITR, как определено в данном документе, фланкирующих указанную экспрессионную кассету, при этом вектор *ceDNA* не ассоциирован с капсидным белком. В некоторых вариантах осуществления вектор на основе *ceDNA* содержит две самокомплементарные последовательности, находящиеся в геноме AAV, при этом по меньшей мере одна ITR содержит функциональный Rep-связывающий элемент (RBE) (также иногда называемый в данном документе "RBS") и сайт концевой разрешения (*trs*) AAV или функциональный вариант RBE. См., например, WO2019113310.

В некоторых вариантах осуществления геном AAV содержит два гена, которые кодируют четыре белка репликации и три капсидных белка соответственно. В некоторых вариантах осуществления гены фланкированы с обеих сторон инвертированными концевыми повторами (ITR) длиной 145 п. о. В некоторых вариантах осуществления вирион содержит до трех капсидных белков (Vp1, Vp2 и/или Vp3), например, продуцируемых в соотношении 1:1:10. В некоторых вариантах осуществления капсидные белки получают из одной и той же открытой рамки считывания и/или из сайтов дифференциального сплайсинга (Vp1) и сайтов с альтернативным началом трансляции (Vp2 и Vp3 соответственно). Как правило, Vp3 является наиболее численной субъединицей вириона и участвует в распознавании рецепторов на поверхности клетки, определяя тропизм вируса. В некоторых вариантах осуществления Vp1 содержит домен фосфолипазы, например, который функционирует в контексте вирусной инфекционности, на N-конце Vp1.

В некоторых вариантах осуществления упаковочная способность вирусных векторов ограничивает размер редактора оснований, который может быть упакован в вектор. Например, упаковочная способность AAV может составлять приблизительно 4,5 т. о. (например, приблизительно 3,0, 3,5, 4,0, 4,5, 5,0, 5,5 или 6,0 т. о.), например, включая один или два инвертированных концевых повтора (ITR), например, ITR размером 145 оснований.

В некоторых вариантах осуществления рекомбинантный AAV (rAAV) содержит *цис*-действующие ITR размером 145 п. о., фланкирующие кассеты трансгена в векторе, например, обеспечивая до 4,5 т. о. для упаковки чужеродной ДНК. После инфицирования rAAV в некоторых случаях может экспрессировать слитый белок по настоящему изобретению и персистировать без интеграции в геном хозяина, существуя эписомально в виде кольцевых конкатемеров "голова к хвосту". rAAV можно использовать, например, *in vitro* и *in vivo*. В некоторых вариантах осуществления опосредованная AAV доставка гена требует, чтобы длина кодирующей последовательности гена была равна геному AAV дикого типа или превышала его по размеру.

Доставка с помощью AAV генов, которые превышают этот размер, и/или применение больших физиологических регуляторных элементов могут быть осуществлены, например, посредством разделения белка(белков), который(которые) должен(должны) быть доставлен(доставлены), на два или более фрагмента. В некоторых вариантах осуществления N-концевой фрагмент слит с расщепленным интеином-N. В некоторых вариантах осуществления C-концевой фрагмент слит с расщепленным

интеином-С. В вариантах осуществления фрагменты упакованы в два или более вектора на основе AAV.

В некоторых вариантах осуществления двойные векторы на основе AAV создаются с помощью разделения крупной кассеты экспрессии трансгена на две отдельные половины (5'- и 3'-концы или голову и хвост), например, где каждая половина кассеты упакована в отдельный вектор на основе AAV (размером < 5 т. о.). Повторная сборка полноразмерной кассеты экспрессии трансгена в некоторых вариантах осуществления может быть затем достигнута при совместном инфицировании одной и той же клетки обоими двойными векторами на основе AAV. В некоторых вариантах осуществления после совместного инфицирования следует одно или несколько из следующего: (1) гомологичной рекомбинации (HR) между 5'- и 3'-концевыми геномами (двойные перекрывающиеся векторы на основе AAV); (2) ITR-опосредованной конкатемеризации 5'- и 3'-концевых геномов "голова к хвосту" (двойные транс-сплайсинговые векторы на основе AAV) и/или (3) комбинации этих двух механизмов (двойные гибридные векторы на основе AAV). В некоторых вариантах осуществления применение двойных векторов на основе AAV *in vivo* приводит в результате к экспрессии полноразмерных белков. В некоторых вариантах осуществления использование платформы двойного вектора на основе AAV представляет собой эффективную и жизнеспособную стратегию переноса генов для трансгенов размером более приблизительно 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8, 4,9 или 5,0 т. о. В некоторых вариантах осуществления векторы на основе AAV также можно использовать для трансдукции клеток нуклеиновыми кислотами-мишенями, например, при продуцировании нуклеиновых кислот и пептидов *in vitro*. В некоторых вариантах осуществления векторы на основе AAV можно использовать для процедур генной терапии *in vivo* и *ex vivo* (см., например, West et al., *Virology* 160:38-47 (1987); патент США № 4797368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994), Muzyczka, *J.Clin.Invest.*94:1351 (1994), каждое из которых включено в данный документ посредством ссылки во всей своей полноте). Конструирование рекомбинантных векторов на основе AAV описано в ряде публикаций, в том числе в патенте 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) (включенных в данный документ посредством ссылки во всей своей полноте).

В некоторых вариантах осуществления Gene Writer, описанный в данном документе (например, с одной или несколькими направляющими нуклеиновыми кислотами или без них), может быть доставлен с применением AAV, лентивируса, аденовируса или других типов плазмидных или вирусных векторов, в частности, с использованием составов и доз из, например, патента США № 8454972 (составы, дозы для аденовируса), патента США № 8404658 (составы, дозы для AAV) и патента США № 5846946 (составы, дозы для ДНК-плазмид), а также из клинических испытаний и публикаций, касающихся клинических испытаний с использованием лентивирусов, AAV и аденовирусов. Например, для AAV путь введения, состав и доза могут быть такими, как описано в патенте США № 8454972 и в

клинических испытаниях с использованием AAV. Для аденовируса путь введения, состав и доза могут быть такими, как описано в патенте США № 8404658 и в клинических испытаниях с использованием аденовируса. Для доставки плазмиды путь введения, состав и доза могут быть такими, как описано в патенте США № 5846946 и в клинических испытаниях с использованием плазмид. Дозы могут быть рассчитаны на основе средней массы индивидуума 70 кг (например, взрослого человека мужского пола) или экстраполированы на нее и могут быть скорректированы для пациентов, субъектов, млекопитающих разной массы и разных видов. Частота введения регламентируется практикующим врачом или ветеринаром (например, терапевтом, ветеринаром) в зависимости от обычных факторов, включая возраст, пол, общее состояние здоровья, другие условия пациента или субъекта и конкретное состояние или симптомы, на которые направлено лечение. В некоторых вариантах осуществления вирусные векторы можно инъектировать в ткань, представляющую интерес. Для Gene Writing, специфичной для типа клеток, экспрессия Gene Writer и необязательной направляющей нуклеиновой кислоты в некоторых вариантах осуществления может находиться под управлением промотора, специфичного для типа клеток.

В некоторых вариантах осуществления AAV обеспечивает низкую токсичность, например, благодаря способу очистки, не требующему ультрацентрифугирования клеточных частиц, которые могут активировать иммунный ответ. В некоторых вариантах осуществления AAV обеспечивает низкую вероятность приведения к инсерционному мутагенезу, например, потому, что он по сути не интегрируется в геном хозяина.

В некоторых вариантах осуществления AAV имеет предел упаковки, составляющий приблизительно 4,4, 4,5, 4,6, 4,7 или 4,75 т. о. В некоторых вариантах осуществления Gene Writer, промотор и терминатор транскрипции могут входить в один вирусный вектор. SpCas9 (4,1 т. о.) в некоторых случаях может быть сложно упаковать в AAV. Следовательно, в некоторых вариантах осуществления применяется Gene Writer, который короче по длине, чем другие Gene Writer или редакторы оснований. В некоторых вариантах осуществления Gene Writer составляет менее приблизительно 4,5 т. о., 4,4 т. о., 4,3 т. о., 4,2 т. о., 4,1 т. о., 4 т. о., 3,9 т. о., 3,8 т. о., 3,7 т. о., 3,6 т. о., 3,5 т. о., 3,4 т. о., 3,3 т. о., 3,2 т. о., 3,1 т. о., 3 т. о., 2,9 т. о., 2,8 т. о., 2,7 т. о., 2,6 т. о., 2,5 т. о., 2 т. о. или 1,5 т. о.

AAV может представлять собой AAV1, AAV2, AAV5 или любую их комбинацию. В некоторых вариантах осуществления тип AAV выбран с учетом клеток, подлежащих нацеливанию; например, серотипы 1, 2, 5 AAV или гибридный капсид AAV1, AAV2, AAV5, или любая их комбинация могут быть выбраны для нацеливания на клетки головного мозга или нервные клетки; или AAV4 может быть выбран для нацеливания на сердечную ткань. В некоторых вариантах осуществления AAV8 выбран для доставки в печень. Иллюстративные серотипы AAV в отношении этих клеток, описаны, например, в работе Grimm, D. et al., J. Virol. 82: 5887-5911 (2008) (включена в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления AAV относится ко всем серотипам, подтипам и встречающимся в природе AAV, а также к рекомбинантному

AAV. AAV может использоваться для обозначения самого вируса или его производного. В некоторых вариантах осуществления AAV включает AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV 12, rh10 и их гибриды, AAV птиц, AAV крупного рогатого скота, AAV собак, AAV лошадей, AAV приматов, AAV животных, отличных от приматов, и AAV овец. Геномные последовательности различных серотипов AAV, а также последовательности нативных концевых повторов (TR), белков Rep и капсидных субъединиц известны в данной области техники. Такие последовательности можно найти в литературе или в общедоступных базах данных, таких как GenBank. Дополнительные иллюстративные серотипы AAV приведены в таблице 36.

В некоторых вариантах осуществления средство, которое способствует активности системы модификации гена, слито с компонентом средства доставки. В некоторых вариантах осуществления компонент слит с AAV, например, с капсидом AAV. В некоторых вариантах осуществления средство представляет собой нуклеиновую кислоту, например, РНК, например, ингибирующую РНК, малую молекулу, большую молекулу, например, биологический препарат, например, полипептид, например, антитело (в том числе конъюгаты антитело-лекарственное средство) или фермент или их функциональный фрагмент, например, домен. В некоторых вариантах осуществления средство модулирует, например, ингибирует или стимулирует, процесс у хозяина. В некоторых вариантах осуществления средство представляет собой фермент, например, эндопептидазу, например, Ig-расщепляющую эндопептидазу, например, IdeS, которая разрушает антитела хозяина, в том числе нейтрализующие антитела к AAV. В некоторых вариантах осуществления средство представляет собой молекулу, которая способствует иммунотолерантности. В некоторых вариантах осуществления средство представляет собой ингибитор комплемента. В некоторых вариантах осуществления средство содержится в средстве доставки с системой модификации гена. В некоторых вариантах осуществления средство встроено в систему доставки с системой модификации гена. В некоторых вариантах осуществления средство представлено снаружи средства доставки, например, слито с капсидным белком AAV или слито с липидом LNP. В некоторых вариантах осуществления средство встраивается в капсид до создания средства доставки, например, экспрессируется в виде слитого белка для AAV. В некоторых вариантах осуществления средство встраивается в капсид после создания средства доставки, например, на капсиде AAV экспрессируется домен, который можно использовать для последующего присоединения, например, ковалентного или нековалентного присоединения, средства (например, фермента) после образования частиц, например, системы SpyTag-SpyCatcher или биотин-стрептавидин. В некоторых вариантах осуществления средство может быть ковалентно присоединено к средству доставки, например, ковалентно присоединено к капсиду AAV. В некоторых вариантах осуществления средство составляют совместно с системой модификации гена. В некоторых вариантах осуществления средство включено в структуру средства доставки, например, включено в структуру LNP. В некоторых вариантах осуществления средство

может содержаться в средстве доставки.

Таблица 36. Иллюстративные серотипы AAV

Ткань-мишень	Среда-носитель	Ссылка
Печень	AAV (AAV8 ¹ , AAVrh.8 ¹ , AAVhu.37 ¹ , AAV2/8, AAV2/rh10 ² , AAV9, AAV2, NP40 ³ , NP59 ^{2,3} , AAV3B ⁵ , AAV-DJ ⁴ , AAV-LK01 ⁴ , AAV-LK02 ⁴ , AAV-LK03 ⁴ , AAV-LK19 ⁴ , AAV5 ⁷ Аденовирус (Ad5, HC-AdV ⁶)	1. Wang <i>et al.</i> , <i>Mol. Ther.</i> 18 , 118-25 (2010) 2. Ginn <i>et al.</i> , <i>JHEP Reports</i> , 100065 (2019) 3. Paulk <i>et al.</i> , <i>Mol. Ther.</i> 26 , 289-303 (2018). 4. L. Lisowski <i>et al.</i> , <i>Nature</i> . 506 , 382-6 (2014). 5. L. Wang <i>et al.</i> , <i>Mol. Ther.</i> 23 , 1877-87 (2015). 6. Hausl <i>Mol Ther</i> (2010) 7. Davidoff <i>et al.</i> , <i>Mol. Ther.</i> 11 , 875-88 (2005)
Легкое	AAV (AAV4, AAV5, AAV6 ¹ , AAV9, H22 ²) Аденовирус (Ad5, Ad3, Ad21, Ad14) ³	1. Duncan <i>et al.</i> , <i>Mol Ther Methods Clin Dev</i> (2018) 2. Cooney <i>et al.</i> , <i>Am J Respir Cell Mol Biol</i> (2019) 3. Li <i>et al.</i> , <i>Mol Ther Methods Clin Dev</i> (2019)
Кожа	AAV (AAV6 ¹ , AAV-LK19 ²)	1. Petek <i>et al.</i> , <i>Mol. Ther.</i> (2010) 2. L. Lisowski <i>et al.</i> , <i>Nature</i> . 506 , 382-6 (2014).
Гемопозитические стволовые клетки	Аденовирус (HDA5/35 ⁺⁺)	Wang <i>et al.</i> <i>Blood Adv</i> (2019)

В некоторых вариантах осуществления фармацевтическая композиция (например, содержащая AAV, описанный в данном документе) содержит менее 10% пустых капсидов, менее 8% пустых капсидов, менее 7% пустых капсидов, менее 5% пустых капсидов, менее 3% пустых капсидов или менее 1% пустых капсидов. В некоторых вариантах осуществления фармацевтическая композиция содержит менее приблизительно 5% пустых капсидов. В некоторых вариантах осуществления количество пустых капсидов ниже предела обнаружения. В некоторых вариантах осуществления преимущественно, чтобы фармацевтическая композиция содержала небольшое количество пустых капсидов,

например, поскольку пустые капсиды могут вызывать неблагоприятный ответ (например, иммунный ответ, воспалительный ответ, ответ со стороны печени и/или ответ со стороны сердца), например, при незначительном терапевтическим эффекте или без него.

В некоторых вариантах осуществления содержание остаточного белка клеток-хозяев (гНСП) в фармацевтической композиции меньше или равняется 100 нг/мл гНСП на 1×10^{13} vg/мл, например, меньше или равняется 40 нг/мл гНСП на 1×10^{13} vg/мл или 1-50 нг/мл гНСП на 1×10^{13} vg/мл. В некоторых вариантах осуществления фармацевтическая композиция содержит менее 10 нг гНСП на $1,0 \times 10^{13}$ vg или менее 5 нг гНСП на $1,0 \times 10^{13}$ vg, менее 4 нг гНСП на $1,0 \times 10^{13}$ vg или менее 3 нг гНСП на $1,0 \times 10^{13}$ vg или любое промежуточное значение концентрации. В некоторых вариантах осуществления содержание остаточной ДНК клеток-хозяев (hcDNA) в фармацевтической композиции меньше или равняется 5×10^6 пг/мл hcDNA на 1×10^{13} vg/мл, меньше или равняется $1,2 \times 10^6$ пг/мл hcDNA на 1×10^{13} vg/мл или 1×10^5 пг/мл hcDNA на 1×10^{13} vg/мл. В некоторых вариантах осуществления содержание остаточной ДНК клеток-хозяев в указанной фармацевтической композиции составляет менее $5,0 \times 10^5$ пг на 1×10^{13} vg, менее $2,0 \times 10^5$ пг на $1,0 \times 10^{13}$ vg, менее $1,1 \times 10^5$ пг на $1,0 \times 10^{13}$ vg, менее $1,0 \times 10^5$ пг hcDNA на $1,0 \times 10^{13}$ vg, менее $0,9 \times 10^5$ пг hcDNA на $1,0 \times 10^{13}$ vg, менее $0,8 \times 10^5$ пг hcDNA на $1,0 \times 10^{13}$ vg или любое промежуточное значение концентрации.

В некоторых вариантах осуществления содержание остаточной плазмидной ДНК в фармацевтической композиции меньше или равняется $1,7 \times 10^5$ пг/мл на $1,0 \times 10^{13}$ vg/мл, или 1×10^5 пг/мл на $1,0 \times 10^{13}$ vg/мл, или $1,7 \times 10^6$ пг/мл на $1,0 \times 10^{13}$ vg/мл. В некоторых вариантах осуществления содержание остаточной плазмидной ДНК в фармацевтической композиции составляет менее $10,0 \times 10^5$ пг на $1,0 \times 10^{13}$ vg, менее $8,0 \times 10^5$ пг на $1,0 \times 10^{13}$ vg или менее $6,8 \times 10^5$ пг на $1,0 \times 10^{13}$ vg. В вариантах осуществления фармацевтическая композиция содержит менее 0,5 нг на $1,0 \times 10^{13}$ vg, менее 0,3 нг на $1,0 \times 10^{13}$ vg, менее 0,22 нг на $1,0 \times 10^{13}$ vg или менее 0,2 нг на $1,0 \times 10^{13}$ vg или любое промежуточное значение концентрации бычьего сывороточного альбумина (BSA). В вариантах осуществления содержание бензоназы в фармацевтической композиции составляет менее 0,2 нг на $1,0 \times 10^{13}$ vg, менее 0,1 нг на $1,0 \times 10^{13}$ vg, менее 0,09 нг на $1,0 \times 10^{13}$ vg, менее 0,08 нг на $1,0 \times 10^{13}$ vg или любое промежуточное значение концентрации. В вариантах осуществления содержание полоксамера 188 в фармацевтической композиции составляет приблизительно 10-150 ppm, приблизительно 15-100 ppm или приблизительно 20-80 ppm. В вариантах осуществления содержание цезия в фармацевтической композиции составляет менее 50 пг/г (ppm), менее 30 пг/г (ppm) или менее 20 пг/г (ppm) или любое промежуточное значение концентрации.

В вариантах осуществления фармацевтическая композиция содержит общее количество примесей, например, как определено с помощью SDS-PAGE, составляющее менее 10%, менее 8%, менее 7%, менее 6%, менее 5%, менее 4%, менее 3%, менее 2% или любую промежуточную процентную долю. В вариантах осуществления общее количество примесей, например, как определено с помощью SDS-PAGE, составляет более 90%, более 92%, более 93%, более 94%, более 95%, более 96%, более 97%, более 98% или любую промежуточную процентную долю. В вариантах осуществления содержание никакой

отдельной неназванной родственной примеси, например, как измерено с помощью SDS-PAGE, не составляет более 5%, более 4%, более 3% или более 2% или любой промежуточной процентной доли. В вариантах осуществления фармацевтическая композиция содержит процентную долю заполненных капсидов по отношению к общему количеству капсидов (например, пик 1+пик 2 согласно измерению с помощью аналитического ультрацентрифугирования), составляющую более 85%, более 86%, более 87%, более 88%, более 89%, более 90%, более 91%, более 91,9%, более 92%, более 93% или любую промежуточную процентную долю. В вариантах осуществления фармацевтической композиции процентная доля заполненных капсидов, измеренная в пике 1 с помощью аналитического ультрацентрифугирования, составляет 20-80%, 25-75%, 30-75%, 35-75% или 37,4-70,3%. В вариантах осуществления фармацевтической композиции процентная доля заполненных капсидов, измеренная в пике 2 с помощью аналитического ультрацентрифугирования, составляет 20-80%, 20-70%, 22-65%, 24-62% или 24,9-60,1%.

В одном варианте осуществления фармацевтическая композиция содержит титр геномов, составляющий от 1,0 до $5,0 \times 10^{13}$ vg/мл, от 1,2 до $3,0 \times 10^{13}$ vg/мл или от 1,7 до $2,3 \times 10^{13}$ vg/мл. В одном варианте осуществления фармацевтическая композиция демонстрирует биологическую нагрузку, составляющую менее 5 КОЕ/мл, менее 4 КОЕ/мл, менее 3 КОЕ/мл, менее 2 КОЕ/мл или менее 1 КОЕ/мл или любое промежуточное значение концентрации. В вариантах осуществления количество эндотоксина в соответствии с USP, например, USP <85> (включенной посредством ссылки во всей своей полноте), составляет менее 1,0 EU/мл, менее 0,8 EU/мл или менее 0,75 EU/мл. В вариантах осуществления осмолярность фармацевтической композиции в соответствии с USP, например, USP <785> (включенной посредством ссылки во всей своей полноте), составляет от 350 до 450 мОсм/кг, от 370 до 440 мОсм/кг или от 390 до 430 мОсм/кг. В вариантах осуществления фармацевтическая композиция содержит менее 1200 частиц размером более 25 мкм на контейнер, менее 1000 частиц размером более 25 мкм на контейнер, менее 500 частиц размером более 25 мкм на контейнер или любое промежуточное значение. В вариантах осуществления фармацевтическая композиция содержит менее 10000 частиц размером более 10 мкм на контейнер, менее 8000 частиц размером более 10 мкм на контейнер или менее 600 частиц размером более 10 пм на контейнер.

В одном варианте осуществления фармацевтическая композиция имеет титр геномов, составляющий от 0,5 до $5,0 \times 10^{13}$ vg/мл, от 1,0 до $4,0 \times 10^{13}$ vg/мл, от 1,5 до $3,0 \times 10^{13}$ vg/мл или от 1,7 до $2,3 \times 10^{13}$ vg/мл. В одном варианте осуществления фармацевтическая композиция, описанная в данном документе, характеризуется одним или несколькими из следующего: менее чем приблизительно 0,09 нг бензоназы на $1,0 \times 10^{13}$ vg, менее чем приблизительно 30 пг/г (ppm) цезия, приблизительно 20-80 ppm полоксамера 188, менее чем приблизительно 0,22 нг BSA на $1,0 \times 10^{13}$ vg, менее чем приблизительно $6,8 \times 10^5$ пг остаточной плазмидной ДНК на $1,0 \times 10^{13}$ vg, менее чем приблизительно $1,1 \times 10^5$ пг остаточной hcDNA на $1,0 \times 10^{13}$ vg, менее чем приблизительно 4 нг гНСП на $1,0 \times 10^{13}$ vg, pH от 7,7 до 8,3, приблизительно 390-430 мОсм/кг, менее чем приблизительно 600 частиц

размером > 25 мкм на контейнер, менее чем приблизительно 6000 частиц размером > 10 мкм на контейнер, титр геномов от приблизительно $1,7 \times 10^{13}$ до $2,3 \times 10^{13}$ vg/мл, инфекционный титр от приблизительно $3,9 \times 10^8$ до $8,4 \times 10^{10}$ ME на $1,0 \times 10^{13}$ vg, содержание общего белка приблизительно 100-300 пг на $1,0 \times 10^{13}$ vg, средняя выживаемость > 24 дней у мышей A7SMA с дозой вирусного вектора приблизительно $7,5 \times 10^{13}$ vg/кг, относительная активность приблизительно 70-130% на основании клеточного анализа *in vitro* и/или менее чем приблизительно 5% пустых капсидов. В различных вариантах осуществления фармацевтические композиции, описанные в данном документе, содержат любую из вирусных частиц, обсуждаемых в данном документе, сохраняют активность в пределах $\pm 20\%$, $\pm 15\%$, $\pm 10\%$ или $\pm 5\%$ от эталонного стандарта. В некоторых вариантах осуществления активность измеряют с использованием подходящего клеточного анализа *in vitro* или модели на животных *in vivo*.

Дополнительные способы получения, определения характеристик и введения доз частиц AAV описаны в документе WO2019094253, который включен в данный документ посредством ссылки во всей своей полноте.

Дополнительные конструкции гAAV, которые можно использовать в соответствии с настоящим изобретением, включают конструкции, описанные в работе Wang et al. 2019, доступной по адресу: [//doi.org/10.1038/s41573-019-0012-9](https://doi.org/10.1038/s41573-019-0012-9), в том числе ее таблице 1, которая включена посредством ссылки во всей своей полноте.

Липидные наночастицы

В способах и системах, предусмотренных в настоящем изобретении, может использоваться любой подходящий носитель или способ доставки, в том числе, в некоторых вариантах осуществления, липидные наночастицы (LNP). Липидные наночастицы в некоторых вариантах осуществления содержат один или несколько ионных липидов, таких как неcatiонные липиды (например, нейтральные, или анионные, или цвиттерионные липиды); один или несколько конъюгированных липидов (таких как липиды, конъюгированные с PEG, или липиды, конъюгированные с полимерами, описанные в таблице 5 документа WO2019217941; включенного в данный документ посредством ссылки во всей своей полноте); один или несколько стеринов (например, холестерин) и необязательно одну или несколько нацеливающих молекул (например, конъюгированные рецепторы, лиганды рецепторов, антитела) или комбинации вышеперечисленного.

Липиды, которые можно применять для образования наночастиц (например, липидных наночастиц), включают, например, липиды, описанные в таблице 4 документа WO2019217941, который включен посредством ссылки, например, липидосодержащая наночастица может содержать один или несколько липидов из таблицы 4 документа WO2019217941. Липидные наночастицы могут содержать дополнительные элементы, такие как полимеры, такие как полимеры, описанные в таблице 5 документа WO2019217941, включенного посредством ссылки.

В некоторых вариантах осуществления конъюгированные липиды в случае их присутствия могут включать одно или несколько из PEG-диацилглицерина (DAG) (такого

как 1-(монометоксиполиэтиленгликоль)-2,3-димиристоилглицерин (PEG-DMG)), PEG-диалкилоксипропила (DAA), PEG-фосфолипида, PEG-церамида (Cer), пегилированного фосфатидилэтаноламина (PEG-PE), PEG-сукцинатдиацилглицерина (PEG-DAG) (такого как 4-0-(2',3'-ди(тетрадеканоиокси)пропил-1-0-(ω-метокси(полиэтокси)этил)бутандиоат (PEG-S-DMG)), PEG-диалкоксипропилкарбама, натриевой соли N-(карбонилметоксиполиэтиленгликоль-2000)-1,2-дистеароил-sn-глицеро-3-фосфоэтаноламина и соединений, описанных в таблице 2 документа WO2019051289 (включенного посредством ссылки), а также комбинаций вышеперечисленного.

В некоторых вариантах осуществления стерина, которые могут быть включены в липидные наночастицы, включают одно или несколько из холестерина или производных холестерина, например, таких как в документах WO2009/127060 или US2010/0130588, которые включены посредством ссылки. Дополнительные иллюстративные стерина включают фитостерина, в том числе описанные в Eygeris et al (2020), [dx.doi.org/10.1021/acs.nanolett.0c01386](https://doi.org/10.1021/acs.nanolett.0c01386), включенном в данный документ посредством ссылки.

В некоторых вариантах осуществления липидная частица содержит ионизируемый липид, некатионный липид, конъюгированный липид, который ингибирует агрегацию частиц, и стерин. Количества этих компонентов можно варьировать независимо и для достижения необходимых свойств. Например, в некоторых вариантах осуществления липидная наночастица содержит ионизируемый липид в количестве от приблизительно 20 мол. % до приблизительно 90 мол. % от общего количества липидов (в других вариантах осуществления это может составлять 20-70% (мол.), 30-60% (мол.) или 40-50% (мол.); от приблизительно 50 мол. % до приблизительно 90 мол. % от общего количества липидов, присутствующих в липидной наночастице), некатионный липид в количестве от приблизительно 5 мол. % до приблизительно 30 мол. % от общего количества липидов, конъюгированный липид в количестве от приблизительно 0,5 мол. % до приблизительно 20 мол. % от общего количества липидов и стерин в количестве от приблизительно 20 мол. % до приблизительно 50 мол. % от общего количества липидов. Соотношение общего количества липидов и нуклеиновой кислоты (например, кодирующей Gene Writer или нуклеиновой кислоты-матрицы) можно варьировать по необходимости. Например, соотношение общего количества липидов и нуклеиновой кислоты (по массе или весу) может составлять от приблизительно 10: 1 до приблизительно 30: 1.

В некоторых вариантах осуществления ионизируемый липид может представлять собой катионный липид, ионизируемый катионный липид, например, катионный липид, который может существовать в положительно заряженной или нейтральной форме в зависимости от pH, или аминсодержащий липид, который может быть легко протонирован. В некоторых вариантах осуществления катионный липид представляет собой липид, способный быть положительно заряженным, например, в физиологических условиях. Иллюстративные катионные липиды содержат одну или несколько аминокислотных групп, которые несут положительный заряд. В некоторых вариантах осуществления липидная частица

содержит катионный липид в составе с одним или несколькими нейтральными липидами, ионизируемыми аминокислотами липидами, биоразлагаемыми алкиновыми липидами, стероидами, фосфолипидами, содержащими полиненасыщенные липиды, структурными липидами (например, стеринами), PEG, холестерином и липидами, конъюгированными с полимером. В некоторых вариантах осуществления катионный липид может представлять собой ионизируемый катионный липид. Иллюстративный катионный липид, раскрытый в данном документе, может характеризоваться эффективным значением pK_a , составляющим более 6,0. В вариантах осуществления липидная наночастица может содержать второй катионный липид, характеризующийся другим эффективным значением pK_a (например, более высоким, чем первое эффективное значение pK_a) по сравнению с первым катионным липидом. Липидная наночастица может содержать от 40 до 60 мольных процентов катионного липида, нейтрального липида, стероида, полимер-конъюгированного липида и терапевтического средства, например, нуклеиновой кислоты (например, РНК), описанной в данном документе (например, нуклеиновой кислоты-матрицы или нуклеиновой кислоты, кодирующей GeneWriter), инкапсулированной внутри липидной наночастицы или ассоциированной с ней. В некоторых вариантах осуществления нуклеиновая кислота составлена совместно с катионным липидом. Нуклеиновая кислота может быть адсорбирована на поверхности LNP, например, LNP, содержащей катионный липид. В некоторых вариантах осуществления нуклеиновая кислота может быть инкапсулирована в LNP, например, в LNP, содержащую катионный липид. В некоторых вариантах осуществления липидная наночастица может содержать нацеливающий фрагмент, например, покрытый нацеливающим средством. В вариантах осуществления состав на основе LNP является биоразлагаемым. В некоторых вариантах осуществления липидная наночастица, содержащая один или несколько липидов, описанных в данном документе, например, формулы (i), (ii), (iii), (vii) и/или (ix), инкапсулирует по меньшей мере 1%, по меньшей мере 5%, по меньшей мере 10%, 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 92%, по меньшей мере 95%, по меньшей мере 97%, по меньшей мере 98% или 100% молекулы РНК, например, РНК-матрицы и/или мРНК, кодирующей полипептид Gene Writer.

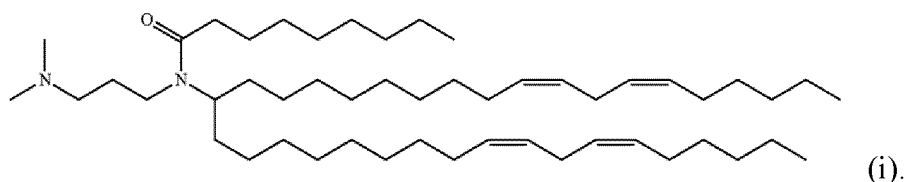
В некоторых вариантах осуществления соотношение липидов и нуклеиновой кислоты (соотношение масса/масса; соотношение вес/вес) может находиться в диапазоне от приблизительно 1: 1 до приблизительно 25: 1, от приблизительно 10: 1 до приблизительно 14: 1, от приблизительно 3: 1 до приблизительно 15: 1, от приблизительно 4: 1 до приблизительно 10: 1, от приблизительно 5: 1 до приблизительно 9: 1 или от приблизительно 6: 1 до приблизительно 9: 1. Количества липидов и нуклеиновой кислоты можно регулировать для обеспечения требуемого соотношения N/P, например, соотношения N/P, составляющего 3, 4, 5, 6, 7, 8, 9, 10 или выше. В целом, общее содержание липидов в составе на основе липидных наночастиц может находиться в диапазоне от приблизительно 5 мг/мл до приблизительно 30 мг/мл.

Иллюстративные ионизируемые липиды, которые можно применять в составах на основе липидных наночастиц, включают без ограничения липиды, перечисленные в таблице 1 документа WO2019051289, включенного в данный документ посредством ссылки. Дополнительные иллюстративные липиды включают без ограничения одну или несколько из следующих формул: X из US2016/0311759; I из US20150376115 или из US2016/0376224; I, II или III из US20160151284; I, IA, II или IIA из US20170210967; I-с из US20150140070; A из US2013/0178541; I из US2013/0303587 или US2013/0123338; I из US2015/0141678; II, III, IV или V из US2015/0239926; I из US2017/0119904; I или II из WO2017/117528; A из US2012/0149894; A из US2015/0057373; A из WO2013/116126; A из US2013/0090372; A из US2013/0274523; A из US2013/0274504; A из US2013/0053572; A из WO2013/016058; A из WO2012/162210; I из US2008/042973; I, II, III или IV из US2012/01287670; I или II из US2014/0200257; I, II или III из US2015/0203446; I или III из US2015/0005363; I, IA, IB, IC, ID, II, IIA, IIB, IIC, IID или III-XXIV из US2014/0308304; из US2013/0338210; I, II, III или IV из WO2009/132131; A из US2012/01011478; I или XXXV из US2012/0027796; XIV или XVII из US2012/0058144; из US2013/0323269; I из US2011/0117125; I, II или III из US2011/0256175; I, II, III, IV, V, VI, VII, VIII, IX, X, XI, XII из US2012/0202871; I, II, III, IV, V, VI, VII, VIII, X, XII, XIII, XIV, XV или XVI из US2011/0076335; I или II из US2006/008378; I из US2013/0123338; I или X-A-Y-Z из US2015/0064242; XVI, XVII или XVIII из US2013/0022649; I, II или III из US2013/0116307; I, II или III из US2013/0116307; I или II из US2010/0062967; I-X из US2013/0189351; I из US2014/0039032; V из US2018/0028664; I из US2016/0317458; I из US2013/0195920; 5, 6 или 10 из US10,221,127; III-3 из WO2018/081480; I-5 или I-8 из WO2020/081938; 18 или 25 из US9867888; A из US2019/0136231; II из WO2020/219876; 1 из US2012/0027803; OF-02 из US2019/0240349; 23 из US10086013; cKK-E12/A6 из работы Miao et al (2020); C12-200 из WO2010/053572; 7C1 из работы Dahlman et al (2017); 304-O13 или 503-O13 из работы Whitehead et al; TS-P4C2 из US9708628; I из WO2020/106946; I из WO2020/106946.

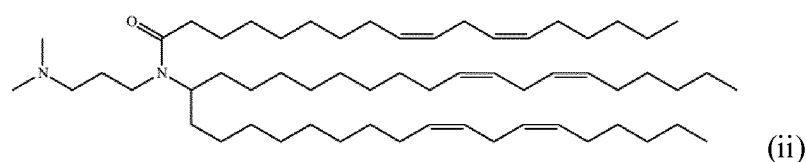
В некоторых вариантах осуществления ионизируемый липид представляет собой MC3 (6Z,9Z,28Z,31Z)-гептатриаконта-6,9,28,31-тетраен-19-ил-4-(диметиламино)бутаноат (DLin-MC3-DMA или MC3), например, описанный в примере 9 документа WO2019051289A9 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой липид ATX-002, например, описанный в примере 10 документа WO2019051289A9 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой (13Z,16Z)-A, A-диметил-3-нонилдокоза-13-16-диен-1-амин (соединение 32), например, описанное в примере 11 документа WO2019051289A9 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой соединение 6 или соединение 22, например, описанные в примере 12 документа WO2019051289A9 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый

липид представляет собой гептадекан-9-ил-8-((2-гидроксиэтил)(6-оксо-6-(ундецилокси)гексил)амино)октаноат (SM-102); например, описанный в примере 1 документа US9867888 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой (9Z,12Z)-3-((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропилоктадека-9,12-диеноат (LP01), например, синтезированный в примере 13 документа WO2015/095340 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой ди((Z)-нон-2-ен-1-ил)-9-(((4-диметиламино)бутаноил)окси)гептадекандиоат (L319), например, синтезированный в примерах 7, 8 или 9 документа US2012/0027803 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой 1,1'-((2-(4-(2-((2-(бис(2-гидроксидодецил)амино)этил)(2-гидроксидодецил)амино)этил)пиперазин-1-ил)этил)азандиил)бис(додекан-2-ол) (C12-200), например, синтезированный в примерах 14 и 16 документа WO2010/053572 (включенного в данный документ посредством ссылки во всей своей полноте). В некоторых вариантах осуществления ионизируемый липид представляет собой липид на основе сложного эфира холестерина, содержащего имидазол (ICE) - (3S,10R,13R,17R)-10,13-диметил-17-((R)-6-метилгептан-2-ил)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-тетрадекагидро-1H-циклопента[а]фенантрен-3-ил-3-(1H-имидазол-4-ил)пропаноат, например, структуру (I) из документа WO2020/106946 (включенного в данный документ посредством ссылки во всей своей полноте).

Некоторые неограничивающие примеры липидных соединений, которые можно применять (например, в комбинации с другими липидными компонентами) для образования липидных наночастиц для доставки композиций, описанных в данном документе, например, нуклеиновой кислоты (например, РНК), описанной в данном документе (например, нуклеиновой кислоты-матрицы или нуклеиновой кислоты, кодирующей Gene Writer), включают

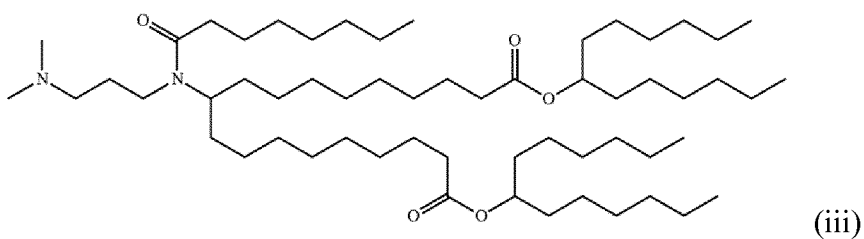


В некоторых вариантах осуществления LNP, содержащую формулу (i), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.

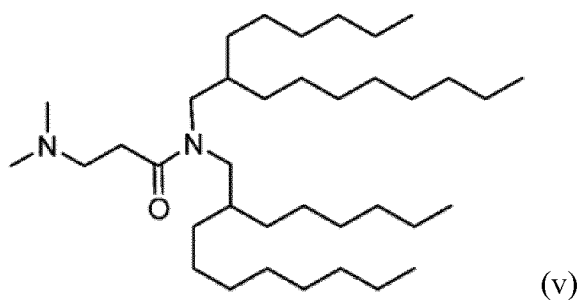
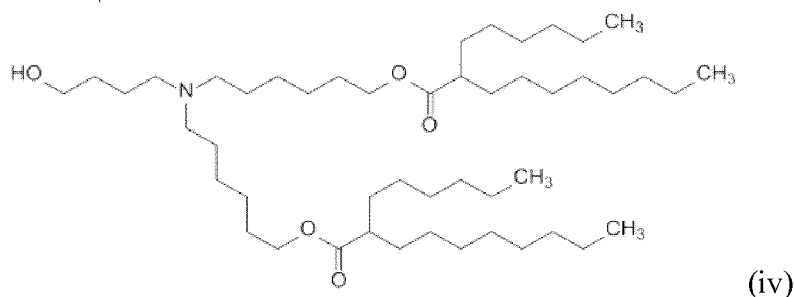


В некоторых вариантах осуществления LNP, содержащую формулу (ii), применяют

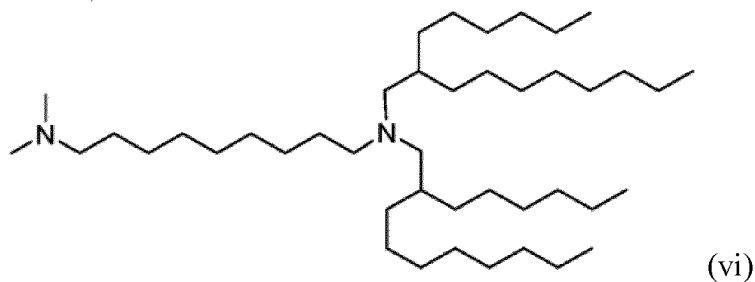
для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.



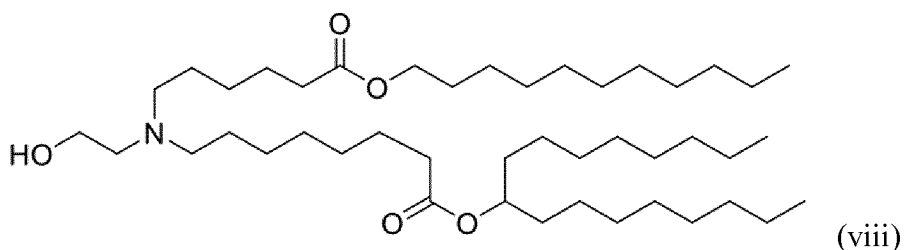
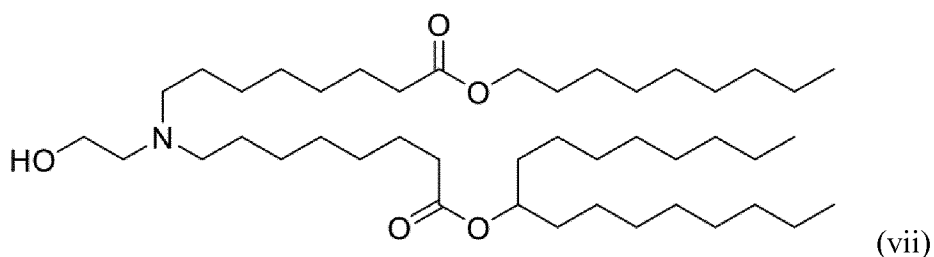
В некоторых вариантах осуществления LNP, содержащую формулу (iii), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.



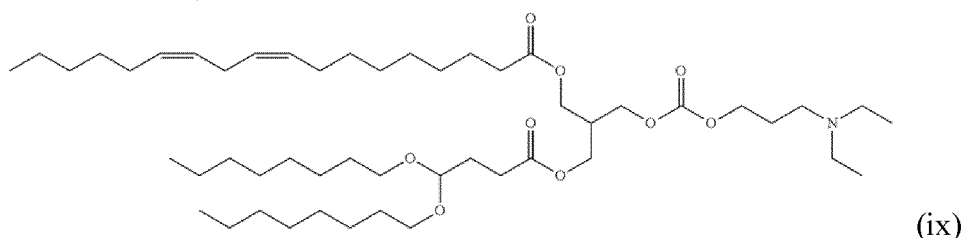
В некоторых вариантах осуществления LNP, содержащую формулу (v), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.



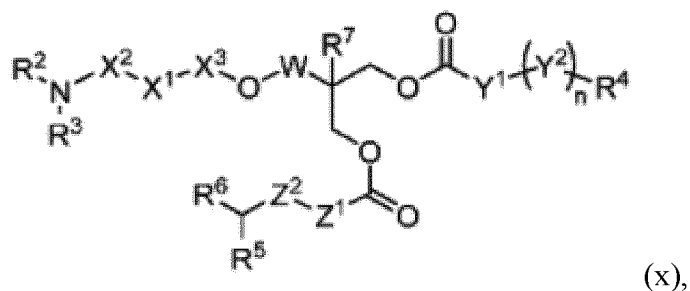
В некоторых вариантах осуществления LNP, содержащую формулу (vi), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.



В некоторых вариантах осуществления LNP, содержащую формулу (viii), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.

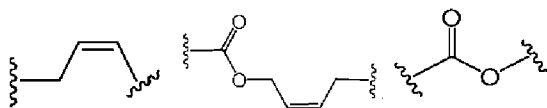


В некоторых вариантах осуществления LNP, содержащую формулу (ix), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.



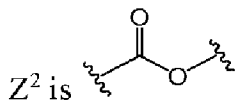
где

X^1 представляет собой O, NR^1 или прямую связь, X^2 представляет собой C2-5-алкилен, X^3 представляет собой C(=O) или прямую связь, R^1 представляет собой H или Me, R^3 представляет собой C-3-алкил, R^2 представляет собой C-3-алкил, или R^2 , взятый вместе с атомом азота, к которому он присоединен, и 1-3 атома углерода из X^2 образуют 4-, 5- или 6-членное кольцо, или X^1 представляет собой NR^1 , R^1 и R^2 , взятые вместе с атомами азота, к которому они присоединены, образуют 5- или 6-членное кольцо, или R^2 , взятый вместе с R^3 и атомом азота, к которому они присоединены, образует 5-, 6- или 7-членное кольцо, Y^1 представляет собой C2-12-алкилен, Y^2 выбран из



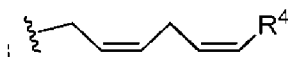
(в любой ориентации), (в любой ориентации), (в любой ориентации),

n равняется 0-3, R^4 представляет собой $Ci-15$ -алкил, Z^1 представляет собой $Ci-6$ -алкилен или прямую связь,



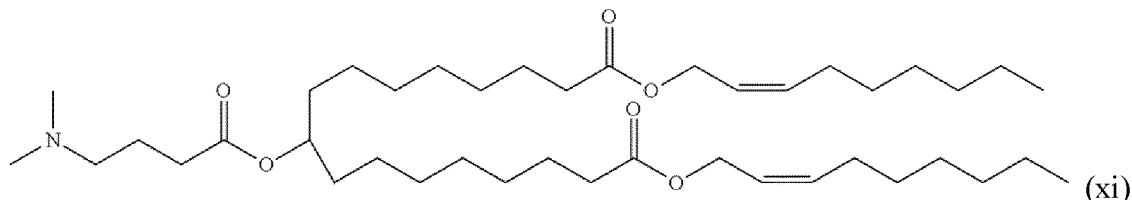
(в любой ориентации) или отсутствует, при условии, что если Z^1 представляет собой прямую связь, то Z^2 отсутствует;

R^5 представляет собой $C5-9$ -алкил или $C6-10$ -алкокси, R^6 представляет собой $C5-9$ -алкил или $C6-10$ -алкокси, W представляет собой метилен или прямую связь, и R^7 представляет собой H или Me или их соль, при условии, что если R^3 и R^2 представляют собой $C2$ -алкилы, X^1 представляет собой O , X^2 представляет собой линейный $C3$ -алкилен, X^3 представляет собой $C(=O)$, Y^1 представляет собой линейный Ce -алкилен, $(Y^2)_n-R^4$ представляет собой

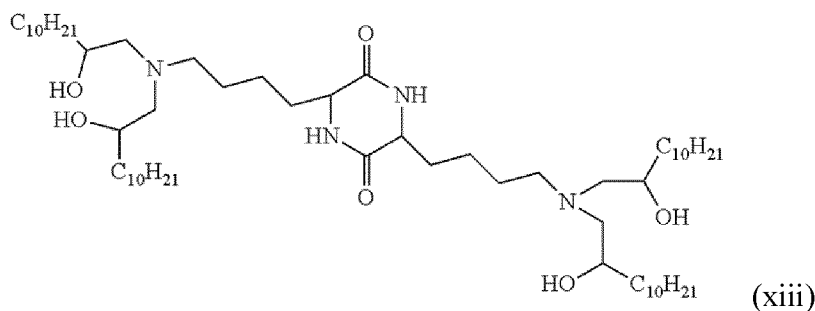
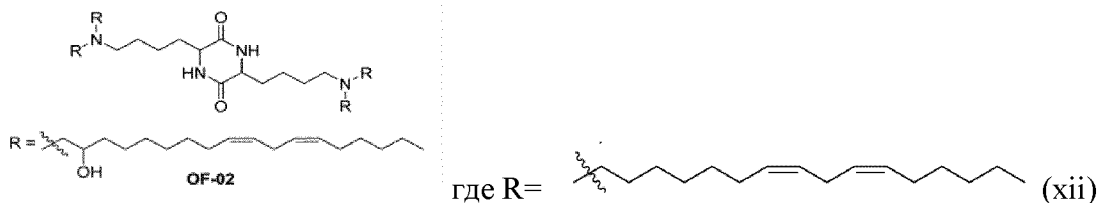


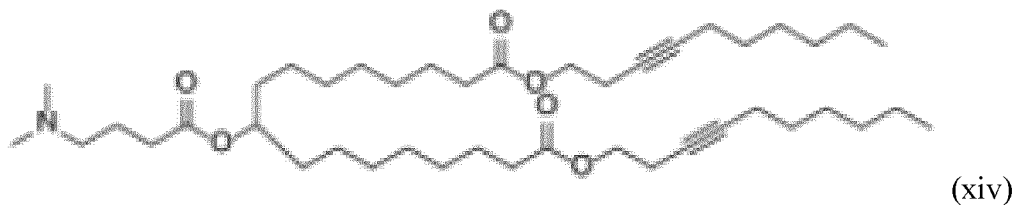
, R^4 представляет собой линейный $C5$ -алкил, Z^1 представляет собой $C2$ -алкилен, Z^2 отсутствует, W представляет собой метилен, и R^7 представляет собой H , то R^5 и R^6 не представляют собой Cx -алкокси.

В некоторых вариантах осуществления LNP, содержащую формулу (xii), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.

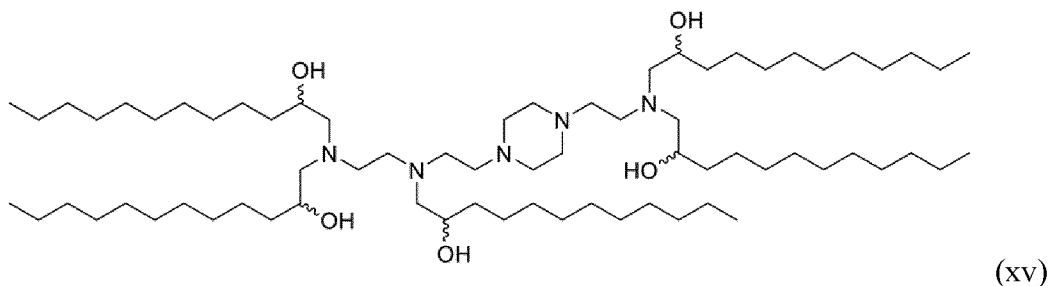


В некоторых вариантах осуществления LNP, содержащую формулу (xi), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.

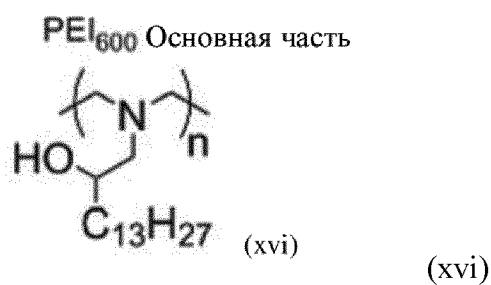




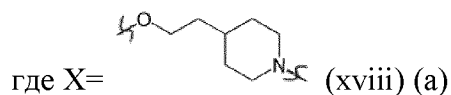
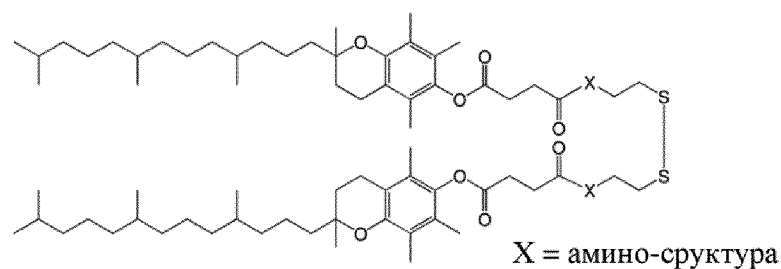
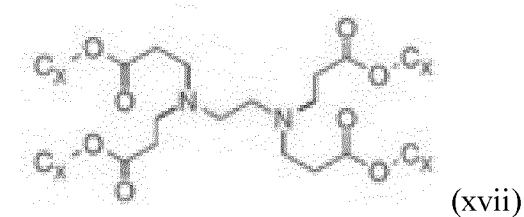
В некоторых вариантах осуществления LNP содержит соединение формулы (xiii) и соединение формулы (xiv).

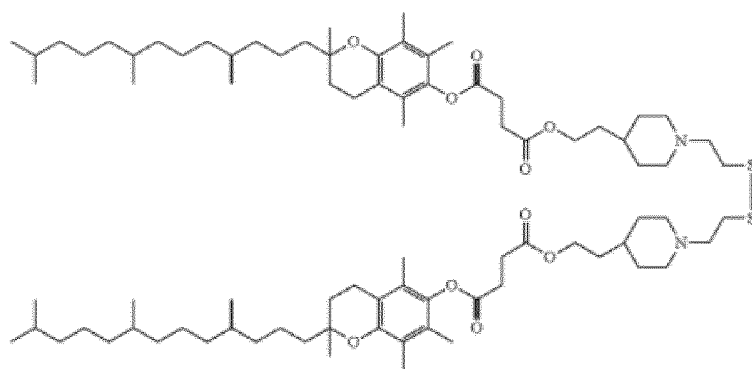


В некоторых вариантах осуществления LNP, содержащую формулу (xv), применяют для доставки композиции Gene Writer, описанной в данном документе, в клетки печени и/или гепатоциты.

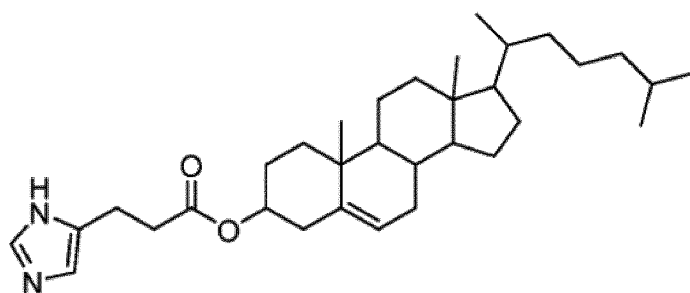


В некоторых вариантах осуществления LNP, содержащую состав формулы (xvi), применяют для доставки композиции Gene Writer, описанной в данном документе, в эндотелиальные клетки легкого.



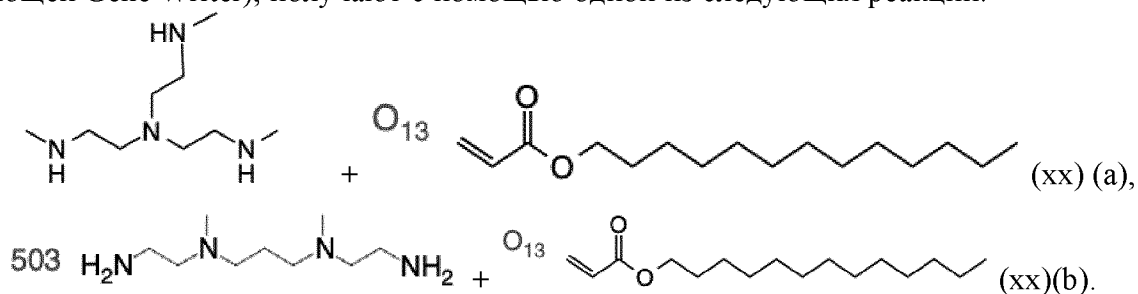


(xviii)(b)



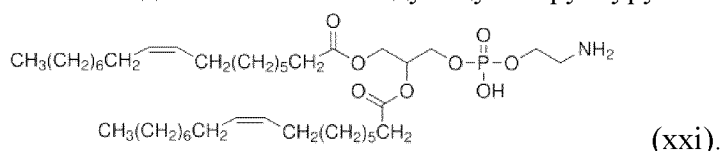
(xix)

В некоторых вариантах осуществления липидное соединение, применяемое для образования липидных наночастиц для доставки композиций, описанных в данном документе, например, нуклеиновой кислоты (например, РНК), описанной в данном документе (например, нуклеиновой кислоты-матрицы или нуклеиновой кислоты, кодирующей Gene Writer), получают с помощью одной из следующих реакций:



Иллюстративные некатيونные липиды включают без ограничения дистеароил-sn-глицерофосфоэтаноламин, дистеароилфосфатидилхолин (DSPC), диолеоилфосфатидилхолин (DOPC), дипальмитоилфосфатидилхолин (DPPC), диолеоилфосфатидилглицерин (DOPG), дипальмитоилфосфатидилглицерин (DPPG), диолеоилфосфатидилэтаноламин (DPPG), 1,2-диолеоил-sn-глицеро-3-фосфоэтаноламин (DOPE), пальмитоилолеоилфосфатидилхолин (POPC), пальмитоилолеоилфосфатидилэтаноламин (POPE), диолеоилфосфатидилэтаноламин-4-(N-малеимидометил)-циклогексан-1-карбоксилат (DOPE-Mal), дипальмитоилфосфатидилэтаноламин (DPPE), димиристоилфосфоэтаноламин (DMPE), дистеароилфосфатидилэтаноламин (DSPE), монометилфосфатидилэтаноламин (такой как 16-О-монометил-PE), диметилфосфатидилэтаноламин (такой как 16-О-диметил-PE), 18-1-транс-PE, 1-стеароил-2-олеоилфосфатидилэтаноламин (SOPE), гидрогенизированный соевый фосфатидилхолин (HSPC), яичный фосфатидилхолин (EPC), диолеоилфосфатидилсерин (DOPS), сфингомиелин (SM), димиристоилфосфатидилхолин

(DMPC), димиристоилфосфатидилглицерин (DMPG), дистеароилфосфатидилглицерин (DSPG), диэрукоилфосфатидилхолин (DEPC), пальмитоилолеоилфосфатидилглицерин (POPG), диэлаидоилфосфатидилэтаноламин (DEPE), лецитин, фосфатидилэтаноламин, лизолецитин, лизофосфатидилэтаноламин, фосфатидилсерин, фосфатидилинозитол, сфингомиелин, яичный сфингомиелин (ESM), цефалин, кардиолипин, фосфатидную кислоту, цереброзиды, дицетилфосфат, лизофосфатидилхолин, дилинолеоилфосфатидилхолин или их смеси. Понятно, что можно также применять другие диацилфосфатидилхолиновые и диацилфосфатидилэтаноламиновые фосфолипиды. Ацильные группы в данных липидах предпочтительно представляют собой ацильные группы, полученные из жирных кислот, имеющих углеродные цепи C10-C24, например, лауроил, миристоил, пальмитоил, стеароил или олеоил. Дополнительные иллюстративные липиды в некоторых вариантах осуществления включают без ограничения липиды, описанные в Kim et al. (2020) [dx.doi.org/10.1021/acs.nanolett.0c01386](https://doi.org/10.1021/acs.nanolett.0c01386), включенном в данный документ посредством ссылки. Такие липиды включают в некоторых вариантах осуществления растительные липиды, которые, как было обнаружено, улучшают трансфекцию мРНК в печени (например, DGTS). В некоторых вариантах осуществления некатионный липид может иметь следующую структуру:



Другие примеры некатионных липидов, пригодных для применения в липидных наночастицах, включают без ограничения липиды, не содержащие фосфор, такие как, например, стеариламин, додециламин, гексадециламин, ацетилпальмитат, глицеринрицинолеат, гексадецилстеарат, изопропилмиристат, амфотерные акриловые полимеры, лаурилсульфат триэтаноламина, алкиларилсульфат, полиэтилоксилированные амиды жирных кислот, бромид диоктадецилдиметиламмония, церамид, сфингомиелин и т. п. Другие некатионные липиды описаны в документе WO2017/099823 или в публикации заявки на патент США US2018/0028664, содержание которых включено в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления некатионный липид представляет собой олеиновую кислоту или соединение формулы I, II или IV из документа US2018/0028664, включенного в данный документ посредством ссылки во всей своей полноте. Некатионный липид может составлять, например, 0-30% (мол.) от общего количества липидов, присутствующих в липидной наночастице. В некоторых вариантах осуществления содержание некатионных липидов составляет 5-20% (мол.) или 10-15% (мол.) от общего количества липидов, присутствующих в липидной наночастице. В вариантах осуществления молярное соотношение ионизируемых липидов и нейтральных липидов находится в диапазоне от приблизительно 2:1 до приблизительно 8:1 (например, составляет приблизительно 2:1, 3:1, 4:1, 5:1, 6:1, 7:1 или 8:1).

В некоторых вариантах осуществления липидные наночастицы не содержат каких-либо фосфолипидов.

В некоторых аспектах липидная наночастица может дополнительно содержать компонент, такой как стерин, для обеспечения целостности мембраны. Одним иллюстративным стерин, который можно применять в липидной наночастице, является холестерин и его производные. Неограничивающие примеры производных холестерина включают полярные аналоги, такие как 5 α -холестанол, 5 β -копростанол, холестерил-(2'-гидроксид)-этиловый эфир, холестерил-(4'-гидроксид)-бутиловый эфир и 6-кетохолестанол; неполярные аналоги, такие как 5 α -холестан, холестенон, 5 α -холестанон, 5 β -холестанон и холестерилдеканоат и их смеси. В некоторых вариантах осуществления производное холестерина представляет собой полярный аналог, например, холестерил-(4'-гидроксид)бутиловый эфир. Иллюстративные производные холестерина описаны в публикации согласно РСТ WO2009/127060 и публикации заявки на патент США US2010/0130588, каждая из которых включена в данный документ посредством ссылки во всей своей полноте.

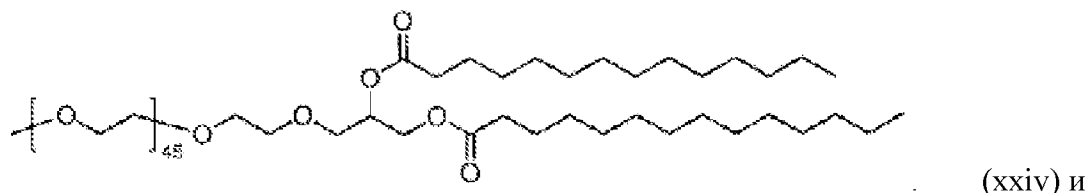
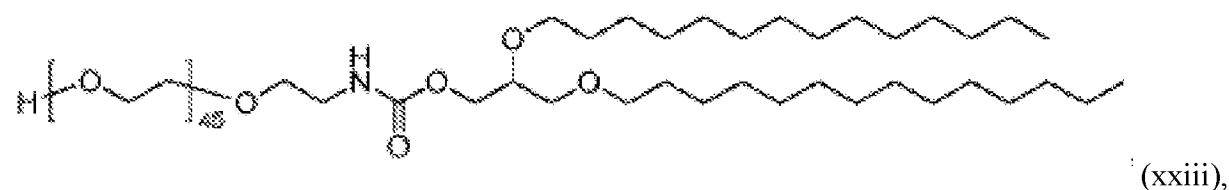
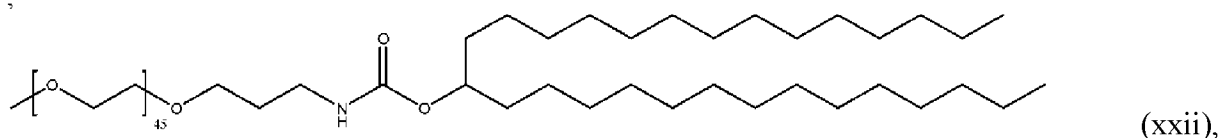
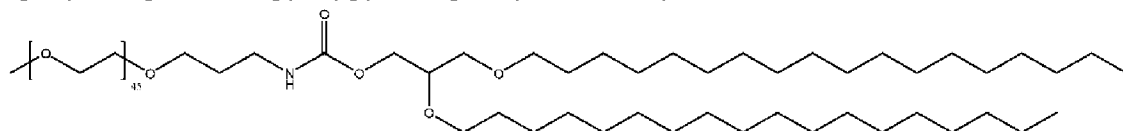
В некоторых вариантах осуществления компонент, обеспечивающий целостность мембраны, такой как стерин, может составлять 0-50% (мол.) (например, 0-10%, 10-20%, 20-30%, 30-40% или 40-50%) от общего количества липидов, присутствующих в липидной наночастице. В некоторых вариантах осуществления такой компонент составляет 20-50% (мол.) или 30-40% (мол.) от общего количества липидов в липидной наночастице.

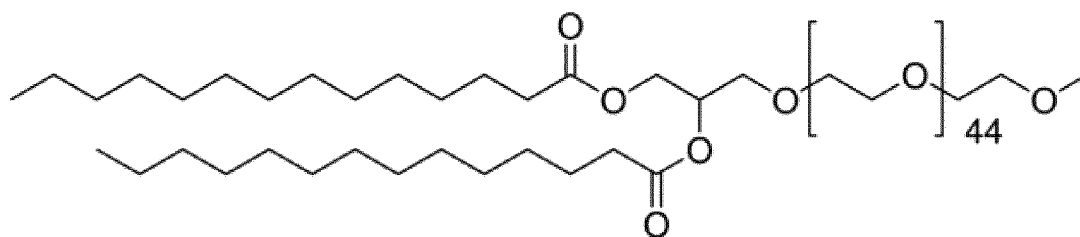
В некоторых вариантах осуществления липидная наночастица может содержать полиэтиленгликоль (PEG) или конъюгированную молекулу липида. Обычно они применяются для ингибирования агрегации липидных наночастиц и/или обеспечения стерической стабилизации. Иллюстративные конъюгированные липиды включают без ограничения конъюгаты PEG-липид, конъюгаты полиоксазолин (POZ)-липид, конъюгаты полиамид-липид (такие как конъюгаты АТТА-липид), конъюгаты катионный полимер-липид (CPL) и их смеси. В некоторых вариантах осуществления молекула конъюгированного липида представляет собой конъюгат PEG-липид, например, липид, конъюгированный с (метоксиполиэтиленгликолем).

Иллюстративные конъюгаты PEG-липид включают без ограничения PEG-диацилглицерин (DAG) (такой как 1-(монометоксиполиэтиленгликоль)-2,3-димиристоилглицерин (PEG-DMG)), PEG-диалкилоксипропил (DAA), PEG-фосфолипид, PEG-церамид (Cer), пегилированный фосфатидилэтаноламин (PEG-PE), 1,2-димиристоил-sn-глицерин, метоксиполиэтиленгликоль (DMG-PEG-2K), PEG-сукцинатдиацилглицерин (PEGS-DAG) (такой как 4-0-(2',3'-ди(тетрадеканоилокси)пропил-1-0-(w-метокси(полиэтокси)этил)бутандиоат (PEG-S-DMG)), PEG-диалкоксипропилкарбам, натриевую соль N-(карбонил-метоксиполиэтиленгликоль 2000)-1,2-дистеароил-sn-глицеро-3-фосфоэтаноламина или их смесь. Дополнительные иллюстративные конъюгаты PEG-липид описаны, например, в US5885613, US6287591,

US2003/0077829, US2003/0077829, US2005/0175682, US2008/0020058,

US2011/0117125, US2010/0130588, US2016/0376224, US2017/0119904 и US/099823, содержание каждого из которых включено в данный документ посредством ссылки во всей своей полноте. В некоторых вариантах осуществления PEG-липид представляет собой соединение формулы III, III-a-I, III-a-2, III-b-1, III-b-2 или V из документа US 2018/0028664, содержание которого включено в данный документ посредством ссылки во всей своей полноте. В некоторых вариантах осуществления PEG-липид характеризуется формулой II из документов US 20150376115 или US 2016/0376224, содержание которых включено в данный документ посредством ссылки во всей своей полноте. В некоторых вариантах осуществления конъюгат PEG-DAA может представлять собой, например, PEG-дилаурилоксипропил, PEG-димиристилоксипропил, PEG-дипальмитилоксипропил или PEG-дистеарилоксипропил. PEG-липид может представлять собой один или несколько из PEG-DMG, PEG-дилаурилглицерина, PEG-дипальмитоилглицерина, PEG-дистерилглицерина, PEG-дилаурилгликамида, PEG-димиристилгликамида, PEG-дипальмитоилгликамида, PEG-дистерилгликамида, PEG-холестерина (1-[8'-(холест-5-ен-3[бета]-окси)карбоксамидо-3',6'-диоксоактанил]карбамоил-[омега]-метилполи(этиленгликоля), PEG-DMB (3,4-дитетрадекоксилбензил-[омега]-метилполи(этиленгликолевого) эфира) и 1,2-димиристоил-sn-глицеро-3-фосфоэтаноламин-N-[метокси(полиэтиленгликоля)-2000]. В некоторых вариантах осуществления PEG-липид включает PEG-DMG, 1,2-димиристоил-sn-глицеро-3-фосфоэтаноламин-N-[метокси(полиэтиленгликоль)-2000]. В некоторых вариантах осуществления PEG-липид предусматривает структуру, выбранную из следующего:





(xxv).

В некоторых вариантах осуществления вместо PEG-липида также можно применять липиды, конъюгированные с молекулой, отличной от PEG. Например, вместо PEG-липида или в дополнение к нему можно использовать конъюгаты полиоксазолин (POZ)-липид, конъюгаты полиамид-липид (такие как конъюгаты АТТА-липид) и конъюгаты катионный полимер-липид (GPL).

Иллюстративные конъюгированные липиды, т. е. конъюгаты PEG-липид, конъюгаты (POZ)-липид, конъюгаты АТТА-липид и катионный полимер-липид, описаны в патентных заявках согласно РСТ и LIS, указанных в таблице 2 документа WO2019051289A9 и в документе WO2020106946A1, содержание каждого из которых включено в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления LNP предусматривает соединение формулы (xix), соединение формулы (xxi) и соединение формулы (xxv). В некоторых вариантах осуществления LNP, содержащую состав формулы (xix), формулы (xxi) и формулы (xxv), применяют для доставки композиции Gene Writer, описанной в данном документе, в легкое или клетки легкого.

В некоторых вариантах осуществления липидная наночастица может содержать один или несколько катионных липидов, выбранных из формулы (i), формулы (ii), формулы (iii), формулы (vii) и формулы (ix). LNP согласно любому из предыдущих вариантов осуществления, дополнительно содержащая один или несколько нейтральных липидов, например, DSPC, DPPC, DMPC, DOPC, POPC, DOPE, SM, стероид, например, холестерин, и/или один или несколько липидов, конъюгированных с полимером, например, пегилированный липид, например, PEG-DAG, PEG-PE, PEG-S-DAG, PEG-Cer или PEG-диалкилоксипропилкарбамат.

В некоторых вариантах осуществления PEG или конъюгированный липид могут составлять 0-20% (мол.) от общего количества липидов, присутствующих в липидной наночастице. В некоторых вариантах осуществления содержание PEG или конъюгированного липида составляет 0,5-10% или 2-5% (мол.) от общего количества липидов, присутствующих в липидной наночастице. Молярные соотношения ионизируемого липида, некатионного липида, стерина и PEG-конъюгированного липида можно изменять по мере необходимости. Например, липидная частица может содержать 30-70% ионизируемого липида по молям или от общего веса композиции, 0-60% холестерина по молям или от общего веса композиции, 0-30% некатионного липида по молям или от общего веса композиции и 1-10% конъюгированного липида по молям или от общего веса композиции. Предпочтительно композиция содержит 30-40% ионизируемого

липидов по молям или от общего веса композиции, 40-50% холестерина по молям или от общего веса композиции и 10-20% некатионного липида по молям или от общего веса композиции. В некоторых других вариантах осуществления композиция содержит 50-75% ионизируемого липида по молям или от общего веса композиции, 20-40% холестерина по молям или от общего веса композиции, 5-10% некатионного липида по молям или от общего веса композиции и 1-10% конъюгированного липида по молям или от общего веса композиции. Композиция может содержать 60-70% ионизируемого липида по молям или от общего веса композиции, 25-35% холестерина по молям или от общего веса композиции и 5-10% некатионного липида по молям или от общего веса композиции. Композиция может также содержать до 90% ионизируемого липида по молям или от общего веса композиции и 2-15% некатионного липида по молям или от общего веса композиции. Состав также может представлять собой состав на основе липидных наночастиц, например, содержащий 8-30% ионизируемого липида по молям или от общего веса композиции, 5-30% некатионного липида по молям или от общего веса композиции и 0-20% холестерина по молям или от общего веса композиции; 4-25% ионизируемого липида по молям или от общего веса композиции, 4-25% некатионного липида по молям или от общего веса композиции, 2-25% холестерина по молям или от общего веса композиции, 10-35% конъюгированного липида по молям или от общего веса композиции и 5% холестерина по молям или от общего веса композиции; или 2-30% ионизируемого липида по молям или от общего веса композиции, 2-30% некатионного липида по молям или от общего веса композиции, 1-15% холестерина по молям или от общего веса композиции, 2-35% конъюгированного липида по молям или от общего веса композиции и 1-20% холестерина по молям или от общего веса композиции; или даже до 90% ионизируемого липида по молям или от общего веса композиции и 2-10% некатионных липидов по молям или от общего веса композиции, или даже 100% катионного липида по молям или от общего веса композиции. В некоторых вариантах осуществления состав на основе липидных частиц содержит ионизируемый липид, фосфолипид, холестерин и пегилированный липид в молярном соотношении 50: 10:38,5: 1,5. В некоторых других вариантах осуществления состав на основе липидных частиц содержит ионизируемый липид, холестерин и пегилированный липид в молярном соотношении 60:38,5: 1,5.

В некоторых вариантах осуществления липидная частица содержит ионизируемый липид, некатионный липид (например, фосфолипид), стерин (например, холестерин) и пегилированный липид, где молярное соотношение липидов находится в диапазоне от 20 до 70 молярных процентов для ионизируемого липида при целевом значении 40-60, молярный процент для некатионного липида находится в диапазоне от 0 до 30 при целевом значении от 0 до 15, молярный процент для стерина находится в диапазоне от 20 до 70 при целевом значении от 30 до 50, и молярный процент для пегилированного липида находится в диапазоне от 1 до 6 при целевом значении от 2 до 5.

В некоторых вариантах осуществления липидная частица содержит ионизируемый липид/некатионный липид/стерин/конъюгированный липид в молярном соотношении 50:

10:38.5: 1,5.

В одном аспекте в настоящем изобретении предусмотрен состав на основе липидных наночастиц, содержащий фосфолипиды, лецитин, фосфатидилхолин и фосфатидилэтаноламин.

В некоторых вариантах осуществления также могут быть включены одно или несколько дополнительных соединений. Данные соединения можно вводить отдельно, или дополнительные соединения могут быть включены в липидные наночастицы по настоящему изобретению. Другими словами, липидные наночастицы могут содержать другие соединения в дополнение к нуклеиновой кислоте или по меньшей мере вторую нуклеиновую кислоту, отличающуюся от первой. Другие дополнительные соединения могут без ограничений быть выбраны из группы, состоящей из малых или больших органических или неорганических молекул, моносахаридов, дисахаридов, трисахаридов, олигосахаридов, полисахаридов, пептидов, белков, аналогов пептидов и их производных, пептидомиметиков, нуклеиновых кислот, аналогов нуклеиновых кислот и производных, экстрактов, полученных из биологических материалов, или любых их комбинаций.

В некоторых вариантах осуществления липидная наночастица (или состав, содержащий липидные наночастицы) не содержит реакционноспособных примесей (например, альдегидов или кетонов) или содержит реакционноспособные примеси (например, альдегиды или кетоны) в количестве, меньшем, чем предварительно выбранный уровень. Без ограничения какой-либо теорией, в некоторых вариантах осуществления липидный реагент применяется для получения состава на основе липидных наночастиц, и липидный реагент может содержать загрязняющую реакционноспособную примесь (например, альдегид или кетон). Липидный реагент может быть выбран для изготовления на основании наличия реакционноспособных примесей (например, альдегидов или кетонов) в количестве, меньшем, чем предварительно выбранный уровень. Без ограничения какой-либо теорией, в некоторых вариантах осуществления альдегиды могут вызывать модификацию и повреждение РНК, например, поперечную сшивку оснований и/или ковалентное конъюгирование липида с РНК (например, образование аддуктов липид-РНК). В некоторых случаях это может привести к неудачной реакции с участием обратной транскриптазы и/или включению неподходящих оснований, например, в месте(местах) повреждения(повреждений), например, мутации во вновь синтезированной ДНК-мишени.

В некоторых вариантах осуществления состав на основе липидных наночастиц получают с использованием липидного реагента, характеризующегося общим содержанием реакционноспособных примесей (например, альдегидов), составляющим менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1%. В некоторых вариантах осуществления состав на основе липидных наночастиц получают с использованием липидного реагента, содержащего менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов). В некоторых вариантах осуществления состав на основе липидных наночастиц получают с использованием липидного реагента,

предусматривающего следующее: (i) содержание менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% общих реакционноспособных примесей (например, альдегидов); и (ii) менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов). В некоторых вариантах осуществления состав на основе липидных наночастиц получают с использованием совокупности липидных реагентов, и каждый липидный реагент из множества независимо соответствует одному или нескольким критериям, описанным в данном абзаце. В некоторых вариантах осуществления каждый липидный реагент из совокупности соответствует одному и тому же критерию, например, критерию из этого абзаца.

В некоторых вариантах осуществления состав на основе липидных наночастиц характеризуется общим содержанием реакционноспособных примесей (например, альдегидов), составляющим менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1%. В некоторых вариантах осуществления состав на основе липидных наночастиц содержит менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов). В некоторых вариантах осуществления состав на основе липидных наночастиц предусматривает следующее: (i) содержание менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% общих реакционноспособных примесей (например, альдегидов); и (ii) менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов).

В некоторых вариантах осуществления один или несколько или необязательно все из липидных реагентов, используемых для получения липидной наночастицы, описанной в данном документе, или состава на ее основе, характеризуются общим содержанием реакционноспособных примесей (например, альдегидов), составляющим менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1%. В некоторых вариантах осуществления один или несколько или необязательно все из липидных реагентов, используемых для получения липидной наночастицы, описанной в данном документе, или состава на ее основе, содержат менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов). В некоторых вариантах осуществления один или несколько или необязательно все липидные реагенты, используемые для получения липидной наночастицы, описанной в данном документе, или состава на ее основе предусматривают следующее: (i) содержание менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% общих реакционноспособных примесей (например, альдегидов); и (ii) менее 5%, 4%, 3%, 2%, 1%, 0,9%, 0,8%, 0,7%, 0,6%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов).

В некоторых вариантах осуществления общее содержание альдегидов и/или

количество какой-либо отдельной разновидности реакционноспособных примесей (например, альдегидов) определяется с помощью жидкостной хроматографии (LC), например, в сочетании с тандемной масс-спектрометрией (MS/MS), например, в соответствии со способом, описанным в примере 5. В некоторых вариантах осуществления содержание реакционноспособных примесей (например, альдегидов) и/или количество разновидностей реакционноспособных примесей (например, альдегидов) определяют посредством обнаружения одной или нескольких химических модификаций молекулы нуклеиновой кислоты (например, молекулы РНК, например, описанной в данном документе), ассоциированных с наличием реакционноспособных примесей (например, альдегидов), например, в липидных реагентах. В некоторых вариантах осуществления содержание реакционноспособных примесей (например, альдегидов) и/или количество разновидностей реакционноспособных примесей (например, альдегидов) определяют посредством обнаружения одной или нескольких химических модификаций нуклеотида или нуклеозида (например, рибонуклеотида или рибонуклеозида, например, содержащихся в молекуле нуклеиновой кислоты-матрицы или выделенных из нее, например, как описано в данном документе), ассоциированных с наличием реакционноспособных примесей (например, альдегидов), например, в липидных реагентах, например, как описано в примере 6. В вариантах осуществления химические модификации молекулы нуклеиновой кислоты, нуклеотида или нуклеозида обнаруживают посредством определения присутствия одного или нескольких модифицированных нуклеотидов или нуклеозидов, например, с применением анализа методом LC-MS/MS, например, как описано в примере 6.

В некоторых вариантах осуществления нуклеиновая кислота (например, РНК), описанная в данном документе (например, нуклеиновая кислота-матрица или нуклеиновая кислота, кодирующая Gene Writer), не содержит альдегидную модификацию или содержит альдегидные модификации в количестве, меньшем, чем предварительно выбранное. В некоторых вариантах осуществления в среднем нуклеиновая кислота имеет менее 50, 20, 10, 5, 2 или 1 альдегидной модификации на 1000 нуклеотидов, например, где одна поперечная сшивка двух нуклеотидов представляет собой одну альдегидную модификацию. В некоторых вариантах осуществления альдегидная модификация представляет собой аддукт РНК (например, аддукт липид-РНК). В некоторых вариантах осуществления нуклеотид с альдегидной модификацией образует поперечные сшивки между основаниями. В некоторых вариантах осуществления нуклеиновая кислота (например, РНК), описанная в данном документе, содержит менее 50, 20, 10, 5, 2 или 1 поперечных сшивок между нуклеотидами.

В некоторых вариантах осуществления LNP направляются к конкретным тканям посредством добавления нацеливающих доменов. Например, биологические лиганды могут быть представлены на поверхности LNP для усиления взаимодействия с клетками, на которых представлены когнатные рецепторы, что тем самым управляет ассоциацией с тканями, где клетки экспрессируют рецептор, и доставкой груза к ним. В некоторых вариантах осуществления биологический лиганд может представлять собой лиганд,

который управляет доставкой в печень, например, LNP, на которых представлены GalNAc, приводят в результате к доставке груза в виде нуклеиновой кислоты в гепатоциты, которые содержат асиалогликопротеиновый рецептор (ASGPR). В работе Akinc et al. *Mol Ther* 18(7):1357-1364 (2010) изложена идея конъюгации трехвалентного лиганда GalNAc и PEG-липид (GalNAc-PEG-DSG) с получением LNP, зависящих от ASGPR, для наблюдаемого эффекта доставки груза с помощью LNP (см., например, фигуру 6). Другие составы на основе LNP, на которых представлен лиганд, например, в которые включены фолат, трансферрин или антитела, обсуждаются в документе WO2017223135, который включен в данный документ посредством ссылки во всей своей полноте, в дополнение к литературным источникам, используемым в данном документе, а именно Kolhatkar et al., *Curr Drug Discov Technol.* 2011 8:197-206; Musacchio and Torchilin, *Front Biosci.* 2011 16:1388-1412; Yu et al., *Mol Membr Biol.* 2010 27:286-298; Patil et al., *Crit Rev Ther Drug Carrier Syst.* 2008 25:1-61; Benoit et al., *Biomacromolecules.* 2011 12:2708-2714; Zhao et al., *Expert Opin Drug Deliv.* 2008 5:309-319; Akinc et al., *Mol Ther.* 2010 18:1357-1364; Srinivasan et al., *Methods Mol Biol.* 2012 820:105-116; Ben-Arie et al., *Methods Mol Biol.* 2012 757:497-507; Peer 2010 *J Control Release.* 20:63-68; Peer et al., *Proc Natl Acad Sci U S A.* 2007 104:4095-4100; Kim et al., *Methods Mol Biol.* 2011 721:339-353; Subramanya et al., *Mol Ther.* 2010 18:2028-2037; Song et al., *Nat Biotechnol.* 2005 23:709-717; Peer et al., *Science.* 2008 319:627-630; и Peer and Lieberman, *Gene Ther.* 2011 18:1127-1133.

В некоторых вариантах осуществления LNP выбирают по тканеспецифической активности посредством добавления молекулы селективного нацеливания на орган (SORT) в состав, содержащий традиционные компоненты, такие как ионизируемые катионные липиды, амфипатические фосфолипиды, холестерин и конъюгированные с поли(этиленгликолем) (PEG) липиды. Идеи Cheng et al. *Nat Nanotechnol* 15(4):313-320 (2020) демонстрируют, что добавление дополнительного компонента "SORT" точно изменяет профиль доставки РНК *in vivo* и опосредует тканеспецифическую (например, в легкие, печень, селезенку) доставку и редактирование генов в зависимости от процентной доли и биофизических свойств молекулы SORT.

В некоторых вариантах осуществления LNP содержат биоразлагаемые ионизируемые липиды. В некоторых вариантах осуществления LNP содержат (9Z,12Z)-3-(((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил)октадека-9,12-диеноат, также называемый 3-(((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил(9Z,12Z)-октадека-9,12-диеноат, или другой ионизируемый липид. См., например, липиды в документах WO2019/067992, WO/2017/173054, WO2015/095340 и WO2014/136086, а также приведенные в них ссылки. В некоторых вариантах осуществления термины "катионный" и "ионизируемый" применительно к липидам LNP являются взаимозаменяемыми, например, где ионизируемые липиды являются катионными в зависимости от значения pH.

В некоторых вариантах осуществления несколько компонентов системы Gene Writer

могут быть получены в виде одного состава на основе LNP, например, состав на основе LNP содержит мРНК, кодирующую полипептид Gene Writer, и РНК-матрицу. Соотношения компонентов, представляющих собой нуклеиновые кислоты, можно изменять с целью усиления до максимума свойств терапевтического средства. В некоторых вариантах осуществления соотношение РНК-матрицы и мРНК, кодирующей полипептид Gene Writer, составляет от приблизительно 1:1 до 100:1, например, от приблизительно 1:1 до 20:1, от приблизительно 20:1 до 40:1, от приблизительно 40:1 до 60:1, от приблизительно 60:1 до 80:1 или от приблизительно 80:1 до 100:1 в молярном соотношении. В других вариантах осуществления система из нескольких нуклеиновых кислот может быть получена с помощью отдельных составов, например, одного состава LNP, содержащего РНК-матрицу, и второго состава LNP, содержащего мРНК, кодирующую полипептид Gene Writer. В некоторых вариантах осуществления система может содержать более двух компонентов, представляющих собой нуклеиновые кислоты, составленных с включением в LNP. В некоторых вариантах осуществления система может содержать белок, например, полипептид Gene Writer, и РНК-матрицу, составленные в виде по меньшей мере одного состава на основе LNP.

В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP может составлять от нескольких 10 до нескольких 100 нм, например, при измерении с помощью динамического светорассеяния (DLS). В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP может составлять от приблизительно 40 нм до приблизительно 150 нм, например, приблизительно 40 нм, 45 нм, 50 нм, 55 нм, 60 нм, 65 нм, 70 нм, 75 нм, 80 нм, 85 нм, 90 нм, 95 нм, 100 нм, 105 нм, 110 нм, 115 нм, 120 нм, 125 нм, 130 нм, 135 нм, 140 нм, 145 нм или 150 нм. В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP может составлять от приблизительно 50 нм до приблизительно 100 нм, от приблизительно 50 нм до приблизительно 90 нм, от приблизительно 50 нм до приблизительно 80 нм, от приблизительно 50 нм до приблизительно 70 нм, от приблизительно 50 нм до приблизительно 60 нм, от приблизительно 60 нм до приблизительно 100 нм, от приблизительно 60 нм до приблизительно 90 нм, от приблизительно 60 нм до приблизительно 80 нм, от приблизительно 60 нм до приблизительно 70 нм, от приблизительно 70 нм до приблизительно 100 нм, от приблизительно 70 нм до приблизительно 90 нм, от приблизительно 70 нм до приблизительно 80 нм, от приблизительно 80 до приблизительно 100 нм, от приблизительно 80 до приблизительно 90 нм или от приблизительно 90 нм до приблизительно 100 нм. В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP может составлять от приблизительно 70 нм до приблизительно 100 нм. В конкретном варианте осуществления средний диаметр LNP в составе на основе LNP может составлять приблизительно 80 нм. В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP может составлять приблизительно 100 нм. В некоторых вариантах осуществления средний диаметр LNP в составе на основе LNP находится в диапазоне от приблизительно 1 мкм до приблизительно 500 мкм, от

приблизительно 5 мм до приблизительно 200 мм, от приблизительно 10 мм до приблизительно 100 мм, от приблизительно 20 мм до приблизительно 80 мм, от приблизительно 25 мм до приблизительно 60 мм, от приблизительно 30 мм до приблизительно 55 мм, от приблизительно 35 мм до приблизительно 50 мм или от приблизительно 38 мм до приблизительно 42 мм.

В некоторых случаях LNP может являться относительно однородной. Индекс полидисперсности может применяться для указания однородности LNP, например, распределения липидных наночастиц по размерам. Небольшой (например, менее 0,3) индекс полидисперсности обычно указывает на узкое распределение частиц по размерам. LNP может характеризоваться индексом полидисперсности, составляющим от приблизительно 0 до приблизительно 0,25, например, 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09, 0,10, 0,11, 0,12, 0,13, 0,14, 0,15, 0,16, 0,17, 0,18, 0,19, 0,20, 0,21, 0,22, 0,23, 0,24 или 0,25. В некоторых вариантах осуществления индекс полидисперсности LNP может составлять от приблизительно 0,10 до приблизительно 0,20.

Дзета-потенциал LNP можно использовать для обозначения электрокинетического потенциала композиции. В некоторых вариантах осуществления дзета-потенциал может описывать поверхностный заряд LNP. Липидные наночастицы с относительно низким зарядом, положительным или отрицательным, обычно являются желательными, поскольку более сильно заряженные соединения могут взаимодействовать с клетками, тканями и другими элементами в организме нежелательным образом. В некоторых вариантах осуществления дзета-потенциал LNP может составлять от приблизительно -10 мВ до приблизительно +20 мВ, от приблизительно -10 мВ до приблизительно +15 мВ, от приблизительно -10 мВ до приблизительно +10 мВ, от приблизительно -10 мВ до приблизительно +5 мВ, от приблизительно -10 мВ до приблизительно 0 мВ, от приблизительно -10 мВ до приблизительно -5 мВ, от приблизительно -5 мВ до приблизительно +20 мВ, от приблизительно -5 мВ до приблизительно +15 мВ, от приблизительно -5 мВ до приблизительно +10 мВ, от приблизительно -5 мВ до приблизительно +5 мВ, от приблизительно -5 мВ до приблизительно 0 мВ, от приблизительно 0 мВ до приблизительно +20 мВ, от приблизительно 0 мВ до приблизительно +15 мВ, от приблизительно 0 мВ до приблизительно +10 мВ, от приблизительно 0 мВ до приблизительно +5 мВ, от приблизительно +5 мВ до приблизительно +20 мВ, от приблизительно +5 мВ до приблизительно +15 мВ или от приблизительно +5 мВ до приблизительно +10 мВ.

Эффективность инкапсуляции белка и/или нуклеиновой кислоты, например, полипептида Gene Writer или мРНК, кодирующей полипептид, описывает количество белка и/или нуклеиновой кислоты, которое инкапсулировано или иным образом ассоциировано с LNP после получения, относительно предоставленного исходного количества. Желательно, чтобы эффективность инкапсуляции была высокой (например, близкой к 100%). Эффективность инкапсуляции можно измерить, например, путем сравнения количества белка или нуклеиновой кислоты в растворе, содержащем липидную наночастицу, до и

после разрушения липидной наночастицы с помощью одного или нескольких органических растворителей или детергентов. Анионообменную смолу можно применять для измерения количества свободного белка или нуклеиновой кислоты (например, РНК) в растворе. Флуоресценцию можно применять для измерения количества свободного белка и/или нуклеиновой кислоты (например, РНК) в растворе. Для липидных наночастиц, описанных в данном документе, эффективность инкапсуляции белка и/или нуклеиновой кислоты может составлять по меньшей мере 50%, например, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100%. В некоторых вариантах осуществления эффективность инкапсуляции может составлять по меньшей мере 80%. В некоторых вариантах осуществления эффективность инкапсуляции может составлять по меньшей мере 90%. В некоторых вариантах осуществления эффективность инкапсуляции может составлять по меньшей мере 95%.

LNP необязательно может содержать одно или несколько покрытий. В некоторых вариантах осуществления LNP может быть составлена в виде капсулы, пленки или таблетки с покрытием. Капсула, пленка или таблетка, содержащие композицию, описанную в данном документе, могут характеризоваться любым полезным размером, прочностью на растяжение, твердостью или плотностью.

Дополнительные иллюстративные липиды, составы, способы и определение характеристик LNP описаны в документе WO2020061457, который включен в данный документ посредством ссылки во всей своей полноте.

В некоторых вариантах осуществления липофекцию клеток *in vitro* или *ex vivo* осуществляют с использованием Lipofectamine MessengerMax (Thermo Fisher) или реагента для трансфекции TransIT-mRNA (Mirus Bio). В определенных вариантах осуществления LNP составляют с использованием смеси ионизируемых липидов GenVoy_ILM (Precision NanoSystems). В некоторых вариантах осуществления LNP составлены с использованием 2,2-дилинолеил-4-диметиламиноэтил-[1,3]-диоксолана (DLin-KC2-DMA) или дилинолеилметил-4-диметиламинобутирата (DLin-MC3-DMA или MC3), состав и применение которого *in vivo* описаны в работе Jayaraman et al. *Angew Chem Int Ed Engl* 51(34):8529-8533 (2012), включенной в данный документ посредством ссылки во всей своей полноте.

Составы на основе LNP, оптимизированные для доставки систем CRISPR-Cas, например, RNP Cas9-gRNA, gRNA, мРНК Cas9, описаны в документах WO2019067992 и WO2019067910, оба из которых включены посредством ссылки.

Дополнительные конкретные составы на основе LNP, пригодные для доставки нуклеиновых кислот, описаны в US8158601 и US8168775, оба из которых включены посредством ссылки, и они включают составы, применяемые в патисиране, продаваемом под наименованием ONPATTRO.

Иллюстративная дозировка LNP с Gene Writer может составлять приблизительно 0,1, 0,25, 0,3, 0,5, 1, 2, 3, 4, 5, 6, 8, 10 или 100 мг/кг (РНК). Иллюстративная дозировка AAV, содержащего нуклеиновую кислоту, кодирующую один или несколько компонентов

системы, может предусматривать MOI, составляющую приблизительно 10^{11} , 10^{12} , 10^{13} и 10^{14} vg/кг.

Наборы, изделия и фармацевтические композиции

В одном аспекте в настоящем изобретении предусмотрен набор, содержащий Gene Writer или систему Gene Writer, например, описанную в данном документе. В некоторых вариантах осуществления набор содержит полипептид Gene Writer (или нуклеиновую кислоту, кодирующую полипептид) и РНК-матрицу (или ДНК, кодирующую РНК-матрицу). В некоторых вариантах осуществления набор дополнительно содержит реагент для введения системы в клетку, например, реагент для трансфекции, LNP и т.п. В некоторых вариантах осуществления набор подходит для любого из способов, описанных в данном документе. В некоторых вариантах осуществления набор содержит один или несколько элементов, композиций (например, фармацевтических композиций), Gene Writer и/или систем Gene Writer или их функциональных фрагментов или компонентов, например, размещенных в изделии. В некоторых вариантах осуществления набор содержит инструкции по его применению.

В одном аспекте в настоящем изобретении предусмотрено изделие, например, в котором находится набор, описанный в данном документе, или его компонент.

В одном аспекте в настоящем изобретении предусмотрена фармацевтическая композиция, содержащая Gene Writer или систему Gene Writer, например, как описано в данном документе. В некоторых вариантах осуществления фармацевтическая композиция дополнительно содержит фармацевтически приемлемый носитель или наполнитель. В некоторых вариантах осуществления фармацевтическая композиция содержит РНК-матрицу и/или РНК, кодирующую полипептид. В вариантах осуществления фармацевтическая композиция обладает одной или несколькими (например, 1, 2, 3 или 4) следующими характеристиками:

(а) менее 1% (например, менее 0,5%, 0,4%, 0,3%, 0,2% или 0,1%) ДНК-матрицы относительно РНК-матрицы и/или РНК, кодирующей полипептид, например, в пересчете на моли;

(b) менее 1% (например, менее 0,5%, 0,4%, 0,3%, 0,2% или 0,1%) некэпированной РНК относительно РНК-матрицы и/или РНК, кодирующей полипептид, например, в пересчете на моли;

(с) менее 1% (например, менее 0,5%, 0,4%, 0,3%, 0,2% или 0,1%) РНК частичной длины относительно РНК-матрицы и/или РНК, кодирующей полипептид, например, в пересчете на моли;

(d) по сути отсутствуют непрореагировавшие динуклеотиды кэпа.

ПРИМЕРЫ

Пример 1. Использование доминантно-отрицательной мРНК для временного ингибирования P53.

В этом примере описано применение мРНК, которая обеспечивает экспрессию доминантно-отрицательной мутантной формы белка в пути ответа хозяина с обеспечением

таким образом эффекта, представляющего собой временное ингибирование пути. В частности, для осуществления этого ингибирования используется доминантно-отрицательная мПНК P53, например, GSE56, как описано в литературе (Schirolli et al. Cell Stem Cell 24, 551-565 (2019)).

В этом примере гемопоэтические стволовые клетки CD34⁺ (HSPC) получают замороженными от Lonza. Вкратце, клетки высевают в концентрации $\sim 5 \times 10^5$ клеток/мл в бессывороточной среде StemSpan (StemCell Technologies), дополненной пенициллином, стрептомицином, глутамином, 1 mM SR-1 (Biovision), 50 нМ UM171 (STEMCell Technologies), 10 mM PGE2, добавляемым только в начале культивирования (Cauman), и цитокинами человека раннего действия (SCF 100 нг/мл, Flt3-L 100 нг/мл, TPO 20 нг/мл и IL-6 20 нг/мл; все приобретали у Peprotech). HSPC культивируют в увлажненной атмосфере с 5% CO₂ при 37°C. После 3 дней стимуляции клетки промывают с помощью PBS и подвергают электропорации с использованием набора P3 Primary Cell 4D-Nucleofector X и программы EO-100 (Lonza). Клетки подвергают электропорации со следующими образцами:

1. мПНК, кодирующая полипептид Gene Writer, нацеливающаяся на AAVS1+ПНК-матрица Gene Writer, несущая репортерный ген GFP;
2. Условие 1 с генетически инактивированным полипептидом Gene Writer;
3. Условие 1+мПНК GSE56 (150 мг/мл);
4. Условие 1+контрольная мПНК (RFP) (150 мг/мл).

Эффективность Gene Writing измеряется в культивируемых клетках через 3 дня после электропорации с помощью проточной цитометрии для определения процентной доли клеток, экспрессирующих GFP, или с помощью цифрового капельного ПЦР-анализа с использованием праймеров и зонда для стыка между матричной последовательностью и локусом-мишенью и для эталонных последовательностей, описанных ранее (см. PCT/US2019/048607). В некоторых вариантах осуществления условие 3 приведет в результате к повышению процентной доли клеток, экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR, по сравнению с условием 4. В некоторых вариантах осуществления условие 3 приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием 4.

Пример 2. Использование siRNA для временного ингибирования пути репарации ДНК для стимуляции интеграции.

В этом примере описано использование siRNA для модулирования пути хозяина. В частности, siRNA, нацеливающаяся на BRCA1 (и, таким образом, на BRCA1-зависимый путь HR) используется для временного ингибирования этого пути для повышения эффективности Gene Writer.

В этом примере клетки HeLa культивируют в DMEM с 10% FBS и 1 mM L-глутамина. После посева клетки трансфицируют следующими образцами:

1. мПНК, кодирующая полипептид Gene Writer, нацеливающаяся на AAVS1+ПНК-

матрица Gene Writer, несущая репортерный ген GFP;

2. Условие 1 с генетически инактивированным полипептидом Gene Writer;
3. Условие 1+siRNA, нацеливающаяся на BRCA1 (siBRCA1);
4. Условие 1+контрольная siRNA (SiScramble).

Эффективность Gene Writing измеряется в культивируемых клетках через 3 дня после трансфекции с помощью проточной цитометрии для определения процентной доли клеток, экспрессирующих GFP, или с помощью цифрового капельного ПЦР-анализа с использованием праймеров и зонда для стыка между матричной последовательностью и локусом-мишенью и для эталонных последовательностей, описанных ранее (см. PCT/US2019/048607). В некоторых вариантах осуществления условие 3 приведет в результате к повышению процентной доли клеток, экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR, по сравнению с условием 4. В некоторых вариантах осуществления условие 3 приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием 4.

Пример 3. Опосредованная малыми молекулами репрессия иммунного ответа с участием РНК

В этом примере описано использование малой молекулы для модулирования пути хозяина. В частности, соединение BAY 11-7082 (CAS 19542-67-7) используется в качестве ингибитора активации комплекса ИКК, за счет чего отделяются пути обнаружения РНК от активации NFκB и внутриклеточного иммунного ответа, который может привести к дестабилизации РНК. Ранее было показано, что BAY11 обеспечивает улучшение в отношении экспрессии OCT4 из синтетической мРНК в клетках кожи человека (Awe et al., Stem Cell Research & Therapy 4 (2013)).

В этом примере первичные фибробласты кожи человека (ATCC PCS-201-012) культивируют в соответствии с инструкциями ATCC. Клетки нуклеофицируют (Lonza Nucleofector®) следующими образцами:

1. мРНК, кодирующая полипептид Gene Writer, нацеливающаяся на AAVS1+РНК-матрица Gene Writer, несущая репортерный ген GFP;
2. Условие (1) с генетически инактивированным полипептидом Gene Writer;
3. Условие (1) + BAY 11-7082;
4. Условие (1) + BAY 11-7082+IFN-бета.

Эффективность Gene Writing измеряется в культивируемых клетках через 3 дня после трансфекции с помощью проточной цитометрии для определения процентной доли клеток, экспрессирующих GFP, или с помощью цифрового капельного ПЦР-анализа с использованием праймеров и зонда для стыка между матричной последовательностью и локусом-мишенью и для эталонных последовательностей, описанных ранее (см. PCT/US2019/048607). В некоторых вариантах осуществления условие (3) приведет в результате к повышению процентной доли клеток, экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR,

по сравнению с условием (1). В некоторых вариантах осуществления условие (3) приведет в результате к повышению процентной доли клеток, экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR, по сравнению с условием (4). В некоторых вариантах осуществления условие (3) приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием (1). В некоторых вариантах осуществления условие (3) приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием (4). В некоторых вариантах осуществления добавление BAY11 будет обеспечивать повышение одного или обоих из экспрессии полипептида Gene Writer и стабильности РНК-матрицы. В некоторых вариантах осуществления добавление PAX11 также будет обеспечивать снижение цитотоксичности, например, цитотоксичности, обусловленной внутриклеточными иммунными путями.

Пример 4. Применение фактора, происходящего из вируса, для улучшения функции Gene Writer.

В этом примере описано использование белка вирусного происхождения, вспомогательного белка лентивируса, вирусного белка X (Vpx), для модулирования пути хозяина. В частности, было обнаружено, что белок Vpx HIV-2 обеспечивает нацеливание белка 1, содержащего домен стерильного мотива альфа и домен HD (SAMHD1), для протеасомного расщепления (Hofmann et al. J Virol 86, 12552-12560 (2012)). Без ограничения какой-либо теорией, считается, что SAMHD1 гидролизует клеточный пул дезоксинуклеотидтрифосфатов до уровня ниже того, который требуется для обратной транскрипции, обеспечивая таким образом ингибирование вирусов и мобильных элементов, требующих стадии обратной транскрипции.

В этом примере миелоидные клетки человека U937 (ATCC CRL-1593.2) культивируют в соответствии с инструкциями ATCC. Клетки U937 трансфицируют одним или комбинацией из следующего:

1. мРНК, кодирующая полипептид Gene Writer, нацеливающаяся на AAVS1+РНК-матрица Gene Writer, несущая репортерный ген GFP;
2. Условие (1) с генетически инактивированным полипептидом Gene Writer;
3. Условие (1) + мРНК Vpx;
4. Условие (1) + мРНК RFP.

Необязательно для условия (4) клетки сначала трансфицируют с использованием мРНК Vpx за день до эксперимента. Эффективность Gene Writing измеряется в культивируемых клетках через 3 дня после трансфекции с помощью проточной цитометрии для определения процентной доли клеток, экспрессирующих GFP, или с помощью цифрового капельного ПЦР-анализа с использованием праймеров и зонда для стыка между матричной последовательностью и локусом-мишенью и для эталонных последовательностей, описанных ранее (см. PCT/US2019/048607). В некоторых вариантах осуществления условие (3) приведет в результате к повышению процентной доли клеток,

экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR, по сравнению с условием (1). В некоторых вариантах осуществления условие (3) приведет в результате к повышению процентной доли клеток, экспрессирующих GFP, по данным проточной цитометрии, и/или эффективности интеграции, измеряемой с помощью ddPCR, по сравнению с условием (4). В некоторых вариантах осуществления условие (3) приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием (1) или условием (4). В некоторых вариантах осуществления условие (3) приводит в результате к снижению цитотоксичности (например, при использовании PrestoBlue) в момент времени через три дня после трансфекции по сравнению с условием (4).

Пример 5. Выбор липидных реагентов со сниженным содержанием альдегидов

В этом примере липиды выбирают для дальнейшего использования в составах на основе липидных наночастиц, содержащих компонент Gene Writing, представляющий собой нуклеиновую(нуклеиновые) кислоту(кислоты), и липиды выбирают, по меньшей мере частично, на основании отсутствия или низкого уровня контаминирующих альдегидов. Реакционноспособные альдегидные группы в липидных реагентах могут обуславливать химические модификации компонента, представляющего собой нуклеиновую(нуклеиновые) кислоту(кислоты), например, РНК, например, РНК-матрицу, во время составления LNP. Следовательно, в некоторых вариантах осуществления содержание альдегидов в липидных реагентах сводят к минимуму.

Жидкостную хроматографию (LC) в сочетании с тандемной масс-спектрометрией (МС/МС) можно использовать для разделения, определения характеристик и количественной оценки содержания альдегидов в реагентах, например, как описано в работе Zurek et al. *The Analyst* 124(9):1291-1295 (1999), включенной в данный документ посредством ссылки. В данном документе каждый липидный реагент подвергают анализу методом LC-MS/MS. Посредством метода LC/MS-MS сначала отделяют липид и одну или несколько примесей на колонке HPLC C8, а затем следует обнаружение и определение структуры этих молекул с помощью масс-спектрометра. Если альдегид присутствует в липидном реагенте, его количественно оценивают с использованием стандарта, меченного основным изотопом (SIL), который структурно идентичен альдегиду, но тяжелее вследствие мечения с помощью C13 и N15. Соответствующее количество SIL-стандарта добавляют в липидный реагент. Затем смесь подвергают анализу методом LC-MS/MS. Количество контаминирующего альдегида определяют путем умножения количества SIL-стандарта на соотношение пиков (неизвестный/SIL). Любой(любые) идентифицированный(идентифицированные) альдегид(альдегиды) в липидных реагентах оценивают количественно, как описано. В некоторых вариантах осуществления было обнаружено, что исходные липидные материалы, выбранные для состава на основе LNP, не содержат каких-либо контаминирующих альдегидов на уровне, выше выбранного. В некоторых вариантах осуществления один или несколько и необязательно все липидные

реагенты, используемые для составления, составляют менее 3% общего содержания альдегидов. В некоторых вариантах осуществления один или несколько и необязательно все липидные реагенты, используемые для составления, составляют менее 0,3% какой-либо отдельной разновидности альдегидов. В некоторых вариантах осуществления один или несколько и необязательно все липидные реагенты, используемые в составе, составляют менее 0,3% какой-либо отдельной разновидности альдегидов и менее 3% общего содержания альдегидов.

Пример 6. Количественная оценка модификации РНК, обусловленной альдегидами, во время составления

В этом примере молекулы РНК анализируют после составления для определения степени любых модификаций, которые могли произойти в процессе составления, например, для обнаружения химических модификаций, обусловленных контаминацией альдегидами липидных реагентов (см., например, пример 5).

Модификации РНК можно обнаружить с помощью анализа рибонуклеозидов, например, согласно способам из работы Su et al. *Nature Protocols* 9:828-841 (2014), включенной в данный документ посредством ссылки во всей своей полноте. В этом способе РНК расщепляется до смеси нуклеозидов, а затем подвергается анализу методом LC-MS/MS. РНК после составления содержится в LNP и должна быть сначала отделена от липидов посредством соосаждения с GlycoBlue в 80% растворе изопропанола. После центрифугирования осадки, содержащие РНК, осторожно переносят в новую пробирку Eppendorf, в которую добавляют коктейль ферментов (бензоназы, фосфодиэстеразы типа 1, фосфатазы) для расщепления РНК до нуклеозидов. Пробирку Eppendorf помещают в предварительно нагретый до 37°C прибор Thermomixer на 1 час. Полученную смесь нуклеозидов анализируют непосредственно методом LC-MS/MS, с помощью которого сначала отделяют нуклеозиды и модифицированные нуклеозиды с помощью колонки C18, а затем обнаруживают их с помощью масс-спектрометрии.

Если альдегид(альдегиды) в липидных реагентах обусловил(обусловили) возникновение химической модификации, анализа данных позволит установить наличие ассоциации между модифицированным(модифицированными) нуклеозид(нуклеозидами) и альдегидом(альдегидами). Модифицированный нуклеозид можно подвергнуть количественной оценке с использованием SIL-стандарта, который структурно идентичен нативному нуклеозиду, за исключением того, что более тяжелый вследствие мечения с помощью C13 и N15. Соответствующее количество стандарта SIL добавляют в гидролизат нуклеозидов, который затем подвергают анализу методом LC-MS/MS. Количество модифицированного нуклеозида получают путем умножения количества SIL-стандарта на соотношение пиков (неизвестный/SIL). LC-MS/MS позволяет осуществлять одновременную количественную оценку всех молекул-мишеней.

В некоторых вариантах осуществления использование липидных реагентов с более высоким содержанием контаминирующих альдегидов приводит в результате к более высоким уровням модификации РНК по сравнению с использованием липидных реагентов

более высокой чистоты в качестве материалов в процессе составления липидных наночастиц. Таким образом, в предпочтительных вариантах осуществления используются липидные реагенты более высокой чистоты, в результате чего уровень модификации РНК находится ниже приемлемого уровня.

Пример 7. Gene Writer™ позволяет осуществлять вставку крупного участка в геномную ДНК

В этом примере описано использование системы редактирования генов Gene Writer™ для изменения геномной последовательности путем вставки крупной последовательности нуклеотидов.

В этом примере полипептид Gene Writer™, gRNA и матрица для записи представлены в виде ДНК, трансфицированной в клетки HEK293T. Полипептид Gene Writer™ использует нуклеазу Cas9 как в качестве ДНК-связывающей, так и эндонуклеазной функциональной единицы. Функциональная единица обратной транскриптазы получена из высокопроцессивного домена RT ретротранспозазы R2. Матрица для записи сконструирована так, что она характеризуется гомологией с последовательностью-мишенью и в то же время содержит генетическую полезную нагрузку в необходимом положении таким образом, что обратная транскрипция РНК-матрицы приводит в результате к образованию новой нити ДНК, содержащей необходимую вставку.

Для создания вставки крупного участка в ДНК клетки HEK293T человека полипептид Gene Writer™ используют в сочетании со специфической gRNA, которая направляет Cas9-содержащий Gene Writer™ к локусу-мишени, и РНК-матрицей для обратной транскрипции, которая содержит RT-связывающий мотив (3'-UTR из элемента R2) для ассоциации с обратной транскриптазой, область гомологии с сайтом-мишенью для праймирования обратной транскрипции и генетическую полезную нагрузку (экспрессионную единицу GFP). Этот комплекс осуществляет ник в сайте-мишени, а затем выполняет TPRT на матрице, иницируя реакцию с использованием праймирующих областей на матрице, которые комплементарны последовательности, непосредственно примыкающей к сайту ника, и копируя полезную нагрузку GFP в геномную ДНК.

После трансфекции клетки инкубируют в течение трех дней для обеспечения экспрессии системы Gene Writing™ и конверсии геномной ДНК-мишени. После инкубационного периода из клеток экстрагируют геномную ДНК. Затем геномную ДНК подвергают амплификации на основе ПЦР с использованием сайт-специфических праймеров, а ампликоны секвенируют на Illumina MiSeq в соответствии с протоколами производителя. Затем выполняют анализ последовательности с определением частоты прочтений, содержащих необходимое редактирование.

Пример 8. Gene Writer могут интегрировать генетическую полезную нагрузку независимо от пути репарации однонитевой матрицы

В этом примере описано использование системы Gene Writer в клетке человека, в которой путь репарации однонитевой матрицы (SSTR) ингибирован.

В этом примере путь SSSTR будет ингибироваться с помощью siRNA, специфических

в отношении основных компонентов пути: FANCA, FANCD2, FANCE, USP1. Также будут включены контрольные siPHK, представляющие собой контроль, не специфический в отношении мишени. В 200 тыс. клетках U2OS будут осуществлять нуклеофекцию с использованием 30 пмоль (1,5 мкМ) siRNA, а также драйвера R2Tg и трансгенных плазмид (транс-конфигурация). В частности, 250 нг плазмид, экспрессирующих R2Tg, контрольный R2Tg с мутацией в домене RT или контрольный R2Tg с мутацией, инактивирующей эндонуклеазу, используют в сочетании с трансгеном в молярном соотношении 1:4 (драйвер к трансгену). Трансфекцию клеток U2OS проводят в буфере SE с использованием программы DN100. После нуклеофекции клетки выращивают в полной среде в течение 3 дней. gDNA собирают в день 3 и проводят ddPCR для оценки интеграции в сайте gDNA. Интеграцию трансгена в gDNA обнаруживают по отсутствию основных компонентов пути SSTR.

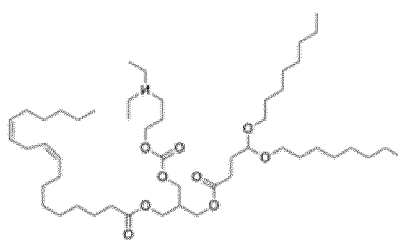
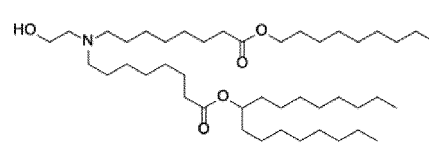
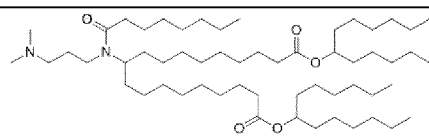
Пример 9. Состав на основе липидных наночастиц, инкапсулирующих мРНК люциферазы светлячка

В этом примере мРНК репортерного гена, кодирующую люциферазу светлячка, составляли с включением в липидные наночастицы, содержащие различные ионизируемые липиды. Компоненты липидных наночастиц (LNP) (ионизируемый липид, вспомогательный липид, стерин, PEG) растворяли в 100% растворе этанола с липидным компонентом. Затем их получали при молярных соотношениях 50:10:38,5:1,5 с использованием ионизируемого липида LIPIDV004 или LIPIDV005 (таблица A1), DSPC, холестерина и DMG-PEG 2000 соответственно. мРНК люциферазы светлячка-LNP, содержащие ионизируемый липид LIPIDV003 (таблица A1), получали при молярном соотношении 45:9:44:2 с использованием LIPIDV003, DSPC, холестерина и DMG-PEG 2000 соответственно. мРНК люциферазы светлячка, используемую в этих составах, получали путем транскрипции *in vitro*, и она кодировала белок люциферазы светлячка, дополнительно содержащий 5'-кэп, 5'- и 3'-UTR и поли(A)-хвост. мРНК синтезировали в стандартных условиях для транскрипции РНК-полимеразой T7 *in vitro* с экзипированием во время транскрипции, однако с 100% заменой нуклеотидтрифосфата UTP на N1-метилпсевдоуридинтрифосфат в реакции. Очищенную мРНК растворяли в 25 мМ цитрате натрия, pH 4, до концентрации 0,1 мг/мл.

мРНК люциферазы светлячка составляли с включением в LNP с молярным соотношением амина липида и фосфата РНК (N:P), составляющим 6. LNP составляли посредством микрофлюидного смешивания растворов липидов и РНК с использованием настольного прибора Precision Nanosystems NanoAssemblerTM с использованием параметров, рекомендованных производителем. Соотношение водного и органического растворителей 3:1 поддерживали во время смешивания с использованием различных скоростей потока. После смешивания LNP собирали и подвергали диализу в 15 мМ Tris, 5% буфере сахарозы при 4°C в течение ночи. Состав на основе мРНК люциферазы светлячка-LNP концентрировали с помощью центрифугирования с использованием центрифужных фильтров Amicon с размерами пор 10 кДа (Millipore). Полученную смесь затем фильтровали

с использованием стерильного фильтра с размером пор 0,2 мкм. Конечный LNP хранили при -80°C до дальнейшего использования.

Таблица А1. Ионизируемые липиды, используемые в примере 9

ID ЛИПИДА	Химическое название	Молекулярная масса	Структура
LIPIDV003	(9Z,12Z)-3-((4,4-бис(октилокси)бутаноил)окси)-2-((((3-(диэтиламино)пропокси)карбонил)окси)метил)пропилотдадека-9,12-диеноат	852,29	
LIPIDV004	Гептадекан-9-ил 8-((2-гидроксиэтил)(8-(нонилокси)-8-оксооктил)амино)октаноат	710,18	
LIPIDV005		919,56	

Полученные LNP анализировали в отношении размера, однородности и % инкапсуляции РНК. Измерения размера и однородности проводили с помощью динамического светорассеяния с использованием прибора для DLS Malvern Zetasizer (Malvern Panalytical). LNP разводили в PBS перед измерением с помощью DLS с определением среднего размера частиц (нанометры, нм) и индекса полидисперсности (pdi). Размеры частиц мРНК люциферазы светлячка-LNP показаны в **таблице А2**.

Таблица А2. Размер и однородность частиц LNP

ID LNP	Ионизируемый липид	Размер частиц (нм)	pdi
LNPV019-002	LIPIDV005	77	0,04
LNPV006-006	LIPIDV004	71	0,08
LNPV011-003	LIPIDV003	87	0,08

Процент инкапсуляции мРНК люциферазы измеряли с помощью анализа для количественной оценки РНК на основе флуоресценции Ribogreen (ThermoFisher Scientific). Образцы LNP разводили в буфере 1× TE и смешивали с реагентом Ribogreen в соответствии с рекомендациями производителя и измеряли на спектрофотометре i3 SpectraMax (Molecular Devices) с использованием длин волн возбуждения 644 нм и испускания 673 нм. Для определения процента инкапсуляции LNP измеряли с помощью анализа Ribogreen с

интактными LNP и разрушенными LNP, где частицы инкубировали с буфером 1× TE, содержащим 0,2% (вес/вес) Triton-X100, с обеспечением разрушения частиц для обеспечения взаимодействия инкапсулированной РНК с реагентом Ribogreen. Образцы измеряли повторно на спектрофотометре i3 SpectraMax для определения общего количества присутствующей РНК. Суммарную РНК вычитали из количества РНК, обнаруженного, когда LNP были интактными, с определением доли, которая была инкапсулирована. Значения умножали на 100 с определением процента инкапсуляции. Значения мРНК люциферазы светлячка-LNP, которые измеряли с помощью Ribogreen, и процент инкапсуляции РНК приведены в таблице А3.

Таблица А3. Инкапсуляция РНК после составления LNP

ID LNP	Ионизируемый липид	% Инкапсуляции мРНК
LNPV019-002	LIPIDV005	98
LNPV006-006	LIPIDV004	92
LNPV011-003	LIPIDV003	97

Пример 10. Тестирование активности мРНК-LNP in vitro в первичных гепатоцитах

В этом примере LNP, содержащие мРНК репортерного гена люциферазы, использовали для доставки полезной нагрузки РНК в клетки в культуре. Первичные гепатоциты мыши или первичные гепатоциты человека размораживали и высевали в покрытые коллагеном 96-луночные планшеты для тканевых культур с плотностью 30000 или 50000 клеток на лунку соответственно. Клетки высевали в 1х среду Е Уильяма без фенолового красного и инкубировали при 37°C с 5% CO₂. Через 4 часа среду заменяли поддерживающей средой (1х среда Е Уильяма без фенола, содержащая Hepatocyte Maintenance Supplement Pack (ThermoFisher Scientific)) и клетки выращивали в течение ночи при 37°C с 5% CO₂. мРНК люциферазы светлячка-LNP размораживали при 4°C и осторожно перемешивали. LNP разводили до соответствующей концентрации в поддерживающей среде, содержащей 7,5% эмбриональной бычьей сыворотки. LNP инкубировали при 37°C в течение 5 минут перед добавлением к высеянным первичным гепатоцитам. Для оценки доставки полезной нагрузки РНК в клетки LNP инкубировали с первичными гепатоцитами в течение 24 часов, и затем клетки собирали и лизировали для анализа активности люциферазы. Вкратце, среду аспирировали из каждой лунки с последующей промывкой с помощью 1х PBS. PBS аспирировали из каждой лунки и обратно добавляли 200 мкл буфера для пассивного лизиса (PLB) (Promega) в каждую лунку, а затем помещали на шейкер для планшетов на 10 минут. Лизированные клетки в PLB замораживали и хранили при -80°C до проведения анализа активности люциферазы.

Для проведения анализа активности люциферазы клеточные лизаты в буфере для пассивного лизиса размораживали, переносили в круглодонный 96-луночный микротитрационный планшет и центрифугировали при 15000g при 4°C в течение 3 минут с удалением клеточного дебриса. Концентрацию белка измеряли для каждого образца с

использованием набора для анализа белка Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) в соответствии с инструкциями производителя. Значения концентрации белка использовали для нормализации по количеству клеток и определения подходящих разбавлений лизатов для анализа с использованием люциферазы. Анализ активности люциферазы проводили в 96-луночных микротитрационных планшетах с белыми стенками с использованием реагента для анализа с использованием люциферазы (Promega) в соответствии с инструкциями производителя, и люминесценцию измеряли с помощью планшет-ридера i3X SpectraMax (Molecular Devices). Результаты в виде зависимости "доза-ответ" для активности люциферазы светлячка, опосредованной мРНК светлячка-LNP, показаны на ФИГ. 1А, и указывают на успешную LNP-опосредованную доставку РНК в первичные клетки в культуре. Как показано на фиг. 1А, LNP, составленные согласно примеру 9, анализировали в отношении доставки полезной нагрузки к первичным гепатоцитам человека (фиг. 1А) и мыши (фиг. 1В) в соответствии с примером 10. В результате анализа с использованием люциферазы было выявлено дозозависимую активность люциферазы в лизатах клеток, что указывает на успешную доставку РНК в клетки и экспрессию люциферазы светлячка из полезной нагрузки мРНК.

Пример 11. LNP-опосредованная доставка РНК в печень мыши.

Для измерения эффективности LNP-опосредованной доставки частиц, содержащих люциферазу светлячка, в печень, LNP составляли и определяли их характеристики, как описано в примере 9, и тестировали *in vitro* до (пример 10) введения мышам. Самцам мышей C57BL/6 (Charles River Labs) в возрасте примерно 8 недель вводили LNP внутривенным (*i.v.*) путем в дозе 1 мг/кг. Животным с введением контроля, представляющего собой средуноситель, вводили *i.v.* 300 мкл фосфатно-солевого буферного раствора. Мышам инъецировали внутривенно дексаметазон в дозе 5 мг/кг за 30 минут до инъекции LNP. Ткани собирали при вскрытии через 6, 24 или 48 часов после введения LNP при размере группы 5 мышей на момент времени. Образцы печени и других тканей собирали, быстро замораживали в жидком азоте и хранили при температуре -80°C до проведения анализа.

Образцы замороженной печени измельчали на сухом льду и переносили в пробирки для гомогенизации, содержащие гранулы лизирующей матрицы D (MP Biomedical). В каждую пробирку добавляли ледяной 1x реагент для лизиса культуры клеток с люциферазой (CCLR) (Promega) и образцы гомогенизировали в гомогенизаторе Fast Prep-24 5G (MP Biomedical) при скорости 6 м/с в течение 40 секунд. Образцы переносили в чистую микроцентрифужную пробирку и очищали с помощью центрифугирования. Перед анализом активности люциферазы для каждого образца определяли концентрацию белка в гомогенатах печени с использованием набора для анализа белка Pierce™ BCA (ThermoFisher Scientific) в соответствии с инструкциями производителя. Активность люциферазы измеряли с использованием 200 мкг (общий белок) гомогената печени с использованием реагента для анализа с использованием люциферазы (Promega) в соответствии с инструкциями производителя с использованием планшет-ридера i3X SpectraMax (Molecular Devices). Для образцов печени была показана успешная доставка

мРНК при всех липидных составах, при этом активность репортерного гена соответствовала ранжированию LIPIDV005>LIPIDV004>LIPIDV003 (ФИГ. 2). Как показано на ФИГ. 2, LNP, содержащие мРНК люциферазы светлячка, составляли и доставляли мышам iv, и образцы печени собирали и анализировали в отношении активности люциферазы через 6, 24 и 48 часов после введения. Активность репортерного гена в различных составах соответствовала ранжированию LIPIDV005>LIPIDV004>LIPIDV003. Экспрессия РНК была временной, и уровни ферментов возвращались к уровню, близкому к фоновому уровню в среде-носителе, через 48 часов после введения. В этом анализе подтверждали использование данных ионизируемых липидов и их соответствующих составов для систем на основе РНК для доставки в печень.

Без ограничения примером, липиды и составы, описанные в этом примере, свидетельствуют об эффективности доставки *in vivo* и других молекул РНК помимо мРНК репортерного гена. Системы Gene Writing, состоящие полностью из РНК, могут быть доставлены с помощью составов, описанных в данном документе. Например, системы, состоящие полностью из РНК, в которых используется мРНК полипептида Gene Writer, РНК-матрица и необязательная gRNA, предусматривающая второй ник, описаны для редактирования генома *in vitro* путем нуклеофекции, с использованием модифицированных нуклеотидов, путем липофекции и для редактирования клеток, например, первичных Т-клеток. Как описано в настоящей заявке, эти системы, состоящие полностью из РНК, характеризуются многими уникальными преимуществами в отношении клеточной иммуногенности и токсичности, что важно при работе с более чувствительными первичными клетками, особенно иммунными клетками, например, Т-клетками, в отличие от линий клеток на основе иммортализованной культуры клеток. Кроме того, предполагается, что все эти системы на основе РНК могут быть нацелены на альтернативные ткани и типы клеток с использованием новых систем доставки липидов, как указано в данном документе, например, для доставки в печень, легкие, мышцы, иммунные клетки и другие, учитывая, что функцию систем Gene Writing проверяли в данном документе *in vitro* на нескольких типах клеток, и функция других систем на основе РНК, доставляемых с помощью нацеленных LNP, известна в данной области техники. Доставка *in vivo* систем Gene Writing может оказать большое влияние на многие области медицины, например, исправление патогенных мутаций, внедрение защитных вариантов и улучшение клеток, эндогенных для организма, например, Т-клеток. Предполагается, что при надлежащем составлении Gene Writing, состоящая полностью из РНК, сделает возможным изготовление средств клеточной терапии *in situ* в организме пациента.

Следует понимать, что для всех числовых границ, описывающих какой-либо параметр в настоящей заявке, таких как “приблизительно”, “по меньшей мере”, “меньше чем” и “больше чем”, описание также обязательно охватывает любой диапазон, ограниченный указанными значениями. Соответственно, например, описание “по меньшей мере 1, 2, 3, 4 или 5” также описывает в числе прочего, диапазоны 1-2, 1-3, 1-4, 1-5, 2-3, 2-

4, 2-5, 3-4, 3-5 и 4-5 и так далее.

Следует понимать, что все патенты, заявки или другие ссылки, цитируемые в данном документе, такие как непатентная литература и информация об эталонных последовательностях, включены посредством ссылки во всей своей полноте для всех целей, а также для цитируемого предложения. Если существует какое-либо противоречие между документом, включенным посредством ссылки, и настоящей заявкой, настоящая заявка будет иметь преимущественную силу. Вся информация, связанная с эталонными последовательностями генов, раскрытыми в настоящей заявке, такая как GeneID или номера доступа (обычно ссылки на номера доступа в NCBI), в том числе, например, геномные локусы, геномные последовательности, функциональные аннотации, аллельные варианты и эталонные мРНК (в том числе, например, границы экзонов или элементы отклика) и последовательности белков (такие как структуры консервативных доменов), а также ссылки, касающиеся химических соединений (например, соединения в PubChem, вещества в PubChem или записи в PubChem Bioassay, в том числе содержащиеся в них аннотации, такие как структуры и анализы, и т.д.), настоящим включены посредством ссылки во всей своей полноте.

Заголовки, используемые в настоящей заявке, предназначены только для удобства и не влияют на интерпретацию настоящей заявки.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ модификации молекулы ДНК-мишени в клетке-хозяине млекопитающего, при этом способ включает:

- а) приведение клетки-хозяина в контакт с системой модификации гена и
- б) приведение клетки-хозяина в контакт с модулятором ответа хозяина,

где система модификации гена содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта.

2. Набор, содержащий:

- а) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и
- б) модулятор ответа хозяина.

3. Композиция, содержащая:

- а) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и
- б) модулятор ответа хозяина.

4. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает антитело, полипептид (например, доминантно-отрицательную мутантную форму полипептида в пути ответа хозяина) или нуклеиновую кислоту (например, молекулу для RNAi).

5. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой ингибитор ответа хозяина.

6. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой стимулятор ответа хозяина.

7. Композиция, набор или способ по любому из предыдущих пунктов, где приведение клетки-хозяина в контакт с полипептидом Gene Writer и модулятором ответа хозяина приводит в результате к повышению уровней последовательности гетерологичного объекта в геноме клетки-хозяина по сравнению с таковыми в аналогичной в ином отношении клетке, не приведенной в контакт с модулятором ответа хозяина, например, где количество копий последовательности гетерологичного объекта в геноме популяции клеток-хозяев является на по меньшей мере 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% или 100% большим или в по меньшей мере 2 раза, 5 раз или 10 раз большим, чем количество копий последовательности гетерологичного объекта в геноме аналогичных в ином отношении клеток, которые были приведены в контакт с системой модификации гена, но не с модулятором ответа хозяина.

8. Композиция, набор или способ по любому из предыдущих пунктов, где один или несколько модуляторов ответа хозяина ингибируют активность одного или нескольких из белков пути ответа на повреждение ДНК, белков пути противовирусного ответа, белков-ингибиторов средства терапии на основе мРНК, белков, участвующих в обнаружении ДНК, белков, участвующих в рестрикции мобильных элементов, провоспалительных белков или их комбинации, например, на по меньшей мере 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99%, или больше, например, в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500, или 1000 раз по сравнению с контролем.

9. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в репарации, направляемой гомологией (HDR) (например, PARP1, PARP2, MRE11, RAD50, NBS1, BARD1, BRCA2, BRCA1, RTS, RECQ5, RPA3, PP4, PALB2, DSS1, RAD51, BACH1, FANCI, Topbp1, TOPO III, FEN1, MUS81, EME1, SLX1, SLX4, RECQ1, WRN, CtIP, EXO1, DNA2, MRN, комплекс MRN), комплементирующей группе анемии Фанкони (FANC) (например, FANCA, FANCB, FANCC, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCI, FANCI, FANCI, FANCL, FANCM, FANCN, FANCO, FANCP, FANCQ, FANCR, FANCS, FANCT), препятствовании HDR (например, FBH1, RECQ5, BLM, FANCI, PARI, RECQ1, WRN, RTEL, RAP80, miR-155, miR-545, miR-107, miR-1255, miR-148, miR-193), однокитевом отжиге (SSA) (например, RPA, RPA1, RPA2, RPA3, RAD52, XPF, ERCC1), каноническом нехомологичном соединении концов (C-NHEJ) (например, DNA-PK, DNA-PKcs, 53BP1, XRCC4, LIG4, XLF, АРТЕМИДА, APLF, PNK, Rif1, PTIP, ДНК-полимераза, Ku70, Ku80), альтернативном нехомологичном соединении концов (Alt-NHEJ) (PARP1, PARP2, CtIP, LIG3, MRE11, Rad50, Nbs1, XPF, ERCC1, LIG1, ДНК-полимераза θ , комплекс MRN, XRCC1), репарации ошибочно спаренных оснований (MMR) (например, EXO1, MSH2, MSH3, MSH6, MLH1, PMS2, MLH3, ДНК-полимераза дельта, RPA, RFC, LIG1), эксцизионной репарации нуклеотидов (NER) (например, XPF, XPG, ERCC1, TTDA, UVSSA, USP7, CETN2, RAD23B, UV-DDB, подкомплекс CAK, RPA, PCNA), эксцизионной репарации оснований (BER) (например, APE1, полимераза β , полимераза δ , полимераза ϵ , XRCC1, LIG3, FEN-1, PCNA, RECQL4, WRN, MYH, PNKP, APTX), репарации однокитевого разрыва (SSBR) (например, PARP1, PARP2, PARG, XRCC1, ДНК-полимераза β , ДНК-полимераза δ , ДНК-полимераза ϵ , PCNA, LIG1, PNK, PNKP, APE1, APTX, APLF, TDP1, LIG3, FEN1, CtIP, MRN, ERCC1), модификации хроматина (например, Ezh2, HDAC класса I, HDAC класса II, KDM4A/JMJD2A, FACT), клеточном цикле (например, CDK1, CDC7, ATM, ATR), синтезе ДНК на поврежденной матрице (TLS) (например, UBC13 или RAD18), клеточном метаболизме (например, mTOR), гибели клетки (например, p53), или разделении РНК:ДНК/R-петли (например, SETX, RNH1 или RNH2), или ответе с задействованием интерферона типа I (например, каспаза-1, IFN α , IFN β , NF- κ B, TNF- α).

10. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в противовирусном ответе, например, ZAP, TREX1, MOV10, hnRNPL, SAMHD1, РНКазы L,

рецептор мелатонина 1, APOBEC3 (A3) (например, ингибитор A3 Vif), SAMHD1 (например, ингибитор SAMHD1 Vpx), BST-2/тетерин (Vpu) или любую их комбинацию.

11. Композиция, набор или способ по любому из предыдущих пунктов, где один или несколько модуляторов ответа хозяина ингибируют один или несколько белков, участвующих в ингибировании средства терапии на основе мРНК.

12. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в обнаружении РНК и ответе на нее, например, TLR3, TLR4, TLR7, TLR8, MyD88, TRIF, IKK, NF- κ B, IRF3, IRF7, IFN- α , IFN- β , TNF α , IL-6, IL-12, JAK-1, TYK-2, STAT1, STAT2, IRF-9, PKR, OAS, ADAR, RIG-I, MDA5, LGP2, MAVS, NLRP3, NOD2 или каспазу 1 или любую их комбинацию.

13. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует RIG-I, например, при этом модулятор ответа хозяина предусматривает протеазу HIV-1 или ее функциональный фрагмент или вариант.

14. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует комплекс IKK, например, модулятор ответа хозяина ингибирует IKK, например, при этом модулятор ответа хозяина предусматривает BAY11.

15. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует TRIF, например, при этом модулятор ответа хозяина предусматривает Pepinh-TRIF.

16. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует комплекс MyD88, например, ингибирует MyD88, при этом модулятор ответа хозяина предусматривает Pepinh-MYD.

17. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует путь IFN, например, ингибирует IFN, при этом модулятор ответа хозяина предусматривает интерферон-связывающий белок, например, B18R вируса осповакцины.

18. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует созревание эндосом, например, при этом модулятор ответа хозяина предусматривает хлорохин или бафиломицин A1 или их комбинацию.

19. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в обнаружении ДНК, например, cGAS, STING, TBK1, IRF3, DNA-PK, HSPA8/HSC70 или любую их комбинацию.

20. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует cGAS, например, при этом модулятор ответа хозяина предусматривает PF-06928215, RU.365, RU.521, RU.521 или G150 или любую их комбинацию.

21. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует STING, например, при этом модулятор ответа

хозяина предусматривает C-176, C-178, H151, циклопептид астин С, астин С, выявленное в результате скрининга соединение 1, соединение 13, E1A (hAd5), E7 (HPV18) или любую их комбинацию.

22. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой siRNA, характеризующуюся последовательностью, представленной под любым из SEQ ID NO: 6-9 из документа WO2018201144A1, который включен в данный документ посредством ссылки.

23. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует TBK1, например, при этом модулятор ответа хозяина предусматривает BX795, тозасертиб, тозасертиб-15a, 20b, азабензимидазол, представляющий собой выявленное соединение 1a, CYT387, Domainex, соединение II Amgen, MRT67307 или AZ13102909 или любую их комбинацию.

24. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует IRF3, например, при этом модулятор ответа хозяина предусматривает одну или несколько siRNA, представленных под любым из SEQ ID NO: 2-5 из документа WO2018201144A1, который включен в данный документ посредством ссылки, или при этом модулятор ответа хозяина предусматривает BX795, тозасертиб, тозасертиб-15a, 20b, азабензимидазол, представляющий собой выявленное соединение 1a, CYT387, Domainex, соединение II Amgen, MRT67307, AZ13102909 или любую их комбинацию.

25. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует DNA-ПК, например, при этом модулятор ответа хозяина предусматривает Nu-7441, E1A hAd5 или ICP0 HSV-1 или любую их комбинацию.

26. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой иммуносупрессивное средство, например, иммуносупрессивное средство, которое обеспечивает снижение иммунного ответа хозяина на вирусный полипептид, например, вирусный полипептид, участвующий в доставке системы модификации гена, например, полипептид AAV, например, капсидный белок AAV.

27. Композиция, набор или способ по п. 26, где иммуносупрессивное средство представляет собой стероид.

28. Композиция, набор или способ по п. 26, где иммуносупрессивное средство представляет собой противовоспалительное средство.

29. Композиция, набор или способ по п. 26, где иммуносупрессивное средство представляет собой циклоспорин (например, циклоспорин А), микофенолат, ритуксимаб или их производное.

30. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в рестрикции мобильных элементов, например, p53, BRCA1 или их комбинацию.

31. Композиция, набор или способ по любому из предыдущих пунктов, где

модулятор ответа хозяина ингибирует p53, например, при этом модулятор ответа хозяина предусматривает нуклеиновую кислоту, кодирующую GSE56, например, мРНК GSE56 (доминантно-отрицательную).

32. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в ответе с задействованием интерферона типа I, например, IFN α , IFN β , NF- κ B, TNF- α .

33. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой иммуносупрессор.

34. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает белок, который изгибает ДНК (например, HMGB1), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

35. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина содержит белок, который стимулирует прогрессирование клеточного цикла (например, PDGF), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

36. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает белок, который обеспечивает повышение уровня биосинтеза дезоксинуклеотидов (например, повышение уровня биосинтеза dNDP из rNDP) (например, рибонуклеотидредуктаза (RNR)), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

37. Композиция, набор или способ по любому из предыдущих пунктов, где полипептид Gene Writer содержит домен обратной транскриптазы и домен эндонуклеазы, при этом домен обратной транскриптазы и домен эндонуклеазы необязательно являются взаимно гетерологичными.

38. Композиция, набор или способ по п. 37, где Gene Writer дополнительно содержит домен, связывающий ДНК-мишень, например, элемент, представляющий собой цинковый палец или его функциональный фрагмент; или элемент, представляющий собой эффектор TAL или его функциональный фрагмент; домен Mub или его функциональный фрагмент; или направляемый последовательностью ДНК-связывающий элемент, необязательно при этом ДНК-связывающий домен является гетерологичным в отношении домена обратной транскриптазы.

39. Композиция, набор или способ по любому из предыдущих пунктов, где направляемый последовательностью ДНК-связывающий элемент предусматривает белок, связанный с CRISPR, например, белок Cas, например, Cas9 или Cpf1, например, белок dCas9.

40. Композиция, набор или способ по любому из предыдущих пунктов, где

нуклеиновая кислота-матрица дополнительно содержит область gRNA, например, область gRNA, которая связывается с сайтом-мишенью.

41. Композиция, набор или способ по любому из предыдущих пунктов, где полипептид Gene Writer предусматривает белок, представляющий собой рекомбиназу, например, тирозиновую рекомбиназу или сериновую рекомбиназу.

42. Композиция, набор или способ по любому из предыдущих пунктов, где полипептид Gene Writer предусматривает белок, представляющий собой ДНК-транспозазу, такую как Sleeping Beauty (в том числе сконструированные производные, такие как SBx100) или piggy Bac.

43. Способ по любому из п. 1 или пп. 4-42, где клетка-хозяин млекопитающего представляет собой клетку примата, такую как клетка человека.

44. Способ по любому из п. 1 или пп. 4-43, где приведение в контакт осуществляют *ex vivo*, например, ДНК клетки-хозяина млекопитающего модифицируют *ex vivo*.

45. Способ по любому из п. 1 или пп. 4-43, где приведение в контакт осуществляют *in vivo*, например, ДНК клетки-хозяина млекопитающего модифицируют *in vivo*.

46. Способ по любому из п. 1 или пп. 4-45, где для системы модификации гена и модулятора ответа хозяина доступ к клетке-хозяину обеспечивают по сути одновременно, например, путем одновременного введения.

47. Способ по любому из п. 1 или пп. 4-45, где для системы модификации гена и модуляторов ответа хозяина доступ к клетке-хозяину обеспечивают последовательно, например, путем последовательного введения, например, при этом модулятор ответа хозяина обеспечивают перед системой модификации гена, или при этом систему модификации гена обеспечивают перед модулятором ответа хозяина.

48. Способ по любому из п. 1 или пп. 4-47, где клетку приводят в контакт с модулятором ответа хозяина несколько раз, например, при этом субъект получает многократные введения модулятора ответа хозяина.

49. Способ по любому из п. 1 или пп. 4-48, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает обеспечение доступа системы модификации гена к клетке-хозяину.

50. Способ по любому из п. 1 или пп. 4-49, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает введение системы модификации гена субъекту, у которого имеется клетка-хозяин.

51. Способ по любому из п. 1 или пп. 4-50, где приведение клетки-хозяина в контакт с модулятором ответа хозяина предусматривает обеспечение доступа системы модификации гена к клетке-хозяину.

52. Способ по любому из п. 1 или пп. 4-51, где приведение клетки-хозяина в контакт с модулятором ответа хозяина предусматривает введение модулятора ответа хозяина субъекту, у которого имеется клетка-хозяин.

53. Способ по любому из п. 1 или пп. 4-52, который включает приведение клетки-хозяина в контакт со вторым модулятором ответа хозяина.

54. Способ по любому из п. 1 или пп. 4-53, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает приведение клетки-хозяина в контакт с нуклеиновой кислотой (например, ДНК или РНК), кодирующей полипептид Gene Writer, в условиях, обеспечивающих возможность продуцирования полипептида Gene Writer.

55. Способ по любому из п. 1 или пп. 4-54, где по сравнению с подобным способом без стадии (b) способ приводит в результате к снижению цитотоксичности в отношении клетки-хозяина млекопитающего или субъекта-млекопитающего, в организме которого расположена клетка-хозяин млекопитающего.

По доверенности

ИЗМЕНЕННАЯ ФОРМУЛА ИЗОБРЕТЕНИЯ, (СТАТЬЯ 34 РСТ)

1. Способ модификации молекулы ДНК-мишени в клетке-хозяине млекопитающего, при этом способ включает:

- а) приведение клетки-хозяина в контакт с системой модификации гена и
- б) приведение клетки-хозяина в контакт с модулятором ответа хозяина,

где система модификации гена содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта;

при этом полипептид Gene Writer предусматривает никазу Cas9 и домен RT ретровируса.

2. Набор, содержащий:

а) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

б) модулятор ответа хозяина;

при этом полипептид Gene Writer предусматривает никазу Cas9 и домен RT ретровируса.

3. Композиция, содержащая:

а) систему модификации гена, которая содержит полипептид Gene Writer или нуклеиновую кислоту, кодирующую полипептид Gene Writer, и нуклеиновую кислоту-матрицу, где нуклеиновая кислота-матрица содержит: i) последовательность, которая связывает полипептид Gene Writer, и ii) последовательность гетерологичного объекта; и

б) модулятор ответа хозяина;

при этом полипептид Gene Writer предусматривает никазу Cas9 и домен RT ретровируса.

4. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает антитело, полипептид (например, доминантно-отрицательную мутантную форму полипептида в пути ответа хозяина) или нуклеиновую кислоту (например, молекулу для RNAi).

5. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой ингибитор ответа хозяина.

6. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой стимулятор ответа хозяина.

7. Композиция, набор или способ по любому из предыдущих пунктов, где приведение клетки-хозяина в контакт с полипептидом Gene Writer и модулятором ответа хозяина приводит в результате к повышению уровней последовательности гетерологичного объекта в геноме клетки-хозяина по сравнению с таковыми в аналогичной в ином

отношении клетке, не приведенной в контакт с модулятором ответа хозяина, например, где количество копий последовательности гетерологичного объекта в геноме популяции клеток-хозяев является на по меньшей мере 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% или 100% большим или в по меньшей мере 2 раза, 5 раз или 10 раз большим, чем количество копий последовательности гетерологичного объекта в геноме аналогичных в ином отношении клеток, которые были приведены в контакт с системой модификации гена, но не с модулятором ответа хозяина.

8. Композиция, набор или способ по любому из предыдущих пунктов, где один или несколько модуляторов ответа хозяина ингибируют активность одного или нескольких из белков пути ответа на повреждение ДНК, белков пути противовирусного ответа, белков-ингибиторов средства терапии на основе мРНК, белков, участвующих в обнаружении ДНК, белков, участвующих в рестрикции мобильных элементов, провоспалительных белков или их комбинации, например, на по меньшей мере 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99% или больше, например, в по меньшей мере 2, 4, 8, 10, 20, 50, 100, 200, 500 или 1000 раз по сравнению с контролем.

9. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в репарации, направляемой гомологией (HDR) (например, PARP1, PARP2, MRE11, RAD50, NBS1, BARD1, BRCA2, BRCA1, RTS, RECQ5, RPA3, PP4, PALB2, DSS1, RAD51, BACH1, FANCI, FANCD1, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FANCN, FANCO, FANCP, FANCQ, FANCR, FANCS, FANCT), препятствовании HDR (например, FBH1, RECQ5, BLM, FANCI, PARI, RECQ1, WRN, RTEL, RAP80, miR-155, miR-545, miR-107, miR-1255, miR-148, miR-193), одонитевом отжиге (SSA) (например, RPA, RPA1, RPA2, RPA3, RAD52, XPF, ERCC1), каноническом негомологичном соединении концов (C-NHEJ) (например, DNA-PK, DNA-PKcs, 53BP1, XRCC4, LIG4, XLF, ARTEMIS, APLF, PNK, Rif1, PTIP, ДНК-полимераза, Ku70, Ku80), альтернативном негомологичном соединении концов (Alt-NHEJ) (PARP1, PARP2, CtIP, LIG3, MRE11, Rad50, Nbs1, XPF, ERCC1, LIG1, ДНК-полимераза θ , комплекс MRN, XRCC1), репарации ошибочно спаренных оснований (MMR) (например, EXO1, MSH2, MSH3, MSH6, MLH1, PMS2, MLH3, ДНК-полимераза дельта, RPA, RFC, LIG1), эксцизионной репарации нуклеотидов (NER) (например, XPF, XPG, ERCC1, TTDA, UVSSA, USP7, CETN2, RAD23B, UV-DDB, подкомплекс САК, RPA, PCNA), эксцизионной репарации оснований (BER) (например, APE1, полимераза β , полимераза δ , полимераза ϵ , XRCC1, LIG3, FEN-1, PCNA, RECQL4, WRN, MYH, PNKP, APTX), репарации одонитевого разрыва (SSBR) (например, PARP1, PARP2, PARG, XRCC1, ДНК-полимераза β , ДНК-полимераза δ , ДНК-полимераза ϵ , PCNA, LIG1, PNK, PNKP, APE1, APTX, APLF, TDP1, LIG3, FEN1, CtIP, MRN, ERCC1), модификации хроматина (например, Ezh2, HDAC класса I, HDAC класса II, KDM4A/JMJD2A, FACT), клеточном цикле (например, CDK1,

CDC7, ATM, ATR), синтезе ДНК на поврежденной матрице (TLS) (например, UBC13 или RAD18), клеточном метаболизме (например, mTOR), гибели клетки (например, p53), или разделении РНК:ДНК/R-петли (например, SETX, RNH1 или RNH2), или ответе с задействованием интерферона типа I (например, каспаза-1, IFN α , IFN β , NF- κ B, TNF- α).

10. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в противовирусном ответе, например ZAP, TREX1, MOV10, hnRNPL, SAMHD1, РНКазу L, рецептор мелатонина 1, APOBEC3 (A3) (например, ингибитор A3 Vif), SAMHD1 (например, ингибитор SAMHD1 Vpx), BST-2/тетерин (Vpu) или любую их комбинацию.

11. Композиция, набор или способ по любому из предыдущих пунктов, где один или несколько модуляторов ответа хозяина ингибируют один или несколько белков, участвующих в ингибировании средства терапии на основе мРНК.

12. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в обнаружении РНК и ответе на нее, например TLR3, TLR4, TLR7, TLR8, MyD88, TRIF, IKK, NF- κ B, IRF3, IRF7, IFN- α , IFN- β , TNF α , IL-6, IL-12, JAK-1, TYK-2, STAT1, STAT2, IRF-9, PKR, OAS, ADAR, RIG-I, MDA5, LGP2, MAVS, NLRP3, NOD2 или каспазу 1 или любую их комбинацию.

13. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует RIG-I, например, при этом модулятор ответа хозяина предусматривает протеазу HIV-1 или ее функциональный фрагмент или вариант.

14. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует комплекс IKK, например, модулятор ответа хозяина ингибирует IKK, например, при этом модулятор ответа хозяина предусматривает BAY11.

15. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует TRIF, например, при этом модулятор ответа хозяина предусматривает Pepinh-TRIF.

16. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует комплекс MyD88, например, ингибирует MyD88, при этом модулятор ответа хозяина предусматривает Pepinh-MYD.

17. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует путь IFN, например, ингибирует IFN, при этом модулятор ответа хозяина предусматривает интерферон-связывающий белок, например B18R вируса осповакцины.

18. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует созревание эндосом, например, при этом модулятор ответа хозяина предусматривает хлорохин или бафиломицин A1 или их комбинацию.

19. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в обнаружении ДНК, например cGAS, STING, TBK1, IRF3, DNA-PK, HSPA8/HSC70 или

любую их комбинацию.

20. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует cGAS, например, при этом модулятор ответа хозяина предусматривает PF-06928215, RU.365, RU.521, RU.521 или G150 или любую их комбинацию.

21. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует STING, например, при этом модулятор ответа хозяина предусматривает C-176, C-178, H151, циклопептид астин С, астин С, выявленное в результате скрининга соединение 1, соединение 13, E1A (hAd5), E7 (HPV18) или любую их комбинацию.

22. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой siRNA, характеризующуюся последовательностью, представленной под любым из SEQ ID NO: 6-9 из документа WO2018201144A1, который включен в данный документ посредством ссылки.

23. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует TBK1, например, при этом модулятор ответа хозяина предусматривает BX795, тозасертиб, тозасертиб-15a, 20b, азабензимидазол, представляющий собой выявленное соединение 1a, CYT387, Domainex, соединение II Amgen, MRT67307 или AZ13102909 или любую их комбинацию.

24. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует IRF3, например, при этом модулятор ответа хозяина предусматривает одну или несколько siRNA, представленных под любым из SEQ ID NO: 2-5 из документа WO2018201144A1, который включен в данный документ посредством ссылки, или при этом модулятор ответа хозяина предусматривает BX795, тозасертиб, тозасертиб-15a, 20b, азабензимидазол, представляющий собой выявленное соединение 1a, CYT387, Domainex, соединение II Amgen, MRT67307, AZ13102909) или любую их комбинацию.

25. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует DNA-ПК, например, при этом модулятор ответа хозяина предусматривает Nu-7441, E1A hAd5 или ICP0 HSV-1 или любую их комбинацию.

26. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой иммуносупрессивное средство, например иммуносупрессивное средство, которое обеспечивает снижение иммунного ответа хозяина на вирусный полипептид, например вирусный полипептид, участвующий в доставке системы модификации гена, например полипептид AAV, например капсидный белок AAV.

27. Композиция, набор или способ по п. 26, где иммуносупрессивное средство представляет собой стероид.

28. Композиция, набор или способ по п. 26, где иммуносупрессивное средство представляет собой противовоспалительное средство.

29. Композиция, набор или способ по п. 26, где иммуносупрессивное средство

представляет собой циклоспорин (например, циклоспорин А), микофенолат, ритуксимаб или их производное.

30. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в рестрикции мобильных элементов, например p53, BRCA1 или их комбинацию.

31. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует p53, например, при этом модулятор ответа хозяина предусматривает нуклеиновую кислоту, кодирующую GSE56, например мРНК GSE56 (доминантно-отрицательную).

32. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина ингибирует один или несколько белков, участвующих в ответе с задействованием интерферона типа I, например IFN α , IFN β , NF- κ B, TNF- α .

33. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина представляет собой иммуносупрессор.

34. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает белок, который изгибает ДНК (например, HMGB1), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

35. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина содержит белок, который стимулирует прогрессирование клеточного цикла (например, PDGF), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

36. Композиция, набор или способ по любому из предыдущих пунктов, где модулятор ответа хозяина предусматривает белок, который обеспечивает повышение уровня биосинтеза дезоксинуклеотидов (например, повышение уровня биосинтеза dNDP из rNDP) (например, рибонуклеотидредуктаза (RNR)), или нуклеиновую кислоту, кодирующую такой белок, или средство, которое обеспечивает повышение экспрессии гена, кодирующего такой белок (например, посредством CRISPRa).

37. Композиция, набор или способ по любому из предыдущих пунктов, где полипептид Gene Writer содержит домен обратной транскриптазы и домен эндонуклеазы, при этом домен обратной транскриптазы и домен эндонуклеазы необязательно являются взаимно гетерологичными.

38. Композиция, набор или способ по любому из предыдущих пунктов, где нуклеиновая кислота-матрица дополнительно содержит область gRNA, например область gRNA, которая связывается с сайтом-мишенью.

39. Способ по любому из п. 1 или пп. 4-38, где клетка-хозяин млекопитающего представляет собой клетку примата, такую как клетка человека.

40. Способ по любому из п. 1 или пп. 4-39, где приведение в контакт осуществляют

ex vivo, например, ДНК клетки-хозяина млекопитающего модифицируют *ex vivo*.

41. Способ по любому из п. 1 или пп. 4-39, где приведение в контакт осуществляют *in vivo*, например, ДНК клетки-хозяина млекопитающего модифицируют *in vivo*.

42. Способ по любому из п. 1 или пп. 4-41, где для системы модификации гена и модулятора ответа хозяина доступ к клетке-хозяину обеспечивают по сути одновременно, например, путем одновременного введения.

43. Способ по любому из п. 1 или пп. 4-41, где для системы модификации гена и модуляторов ответа хозяина доступ к клетке-хозяину обеспечивают последовательно, например, путем последовательного введения, например, при этом модулятор ответа хозяина обеспечивают перед системой модификации гена, или при этом систему модификации гена обеспечивают перед модулятором ответа хозяина.

44. Способ по любому из п. 1 или пп. 4-43, где клетку приводят в контакт с модулятором ответа хозяина несколько раз, например, при этом субъект получает многократные введения модулятора ответа хозяина.

45. Способ по любому из п. 1 или пп. 4-44, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает обеспечение доступа системы модификации гена к клетке-хозяину.

46. Способ по любому из п. 1 или пп. 4-45, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает введение системы модификации гена субъекту, у которого имеется клетка-хозяин.

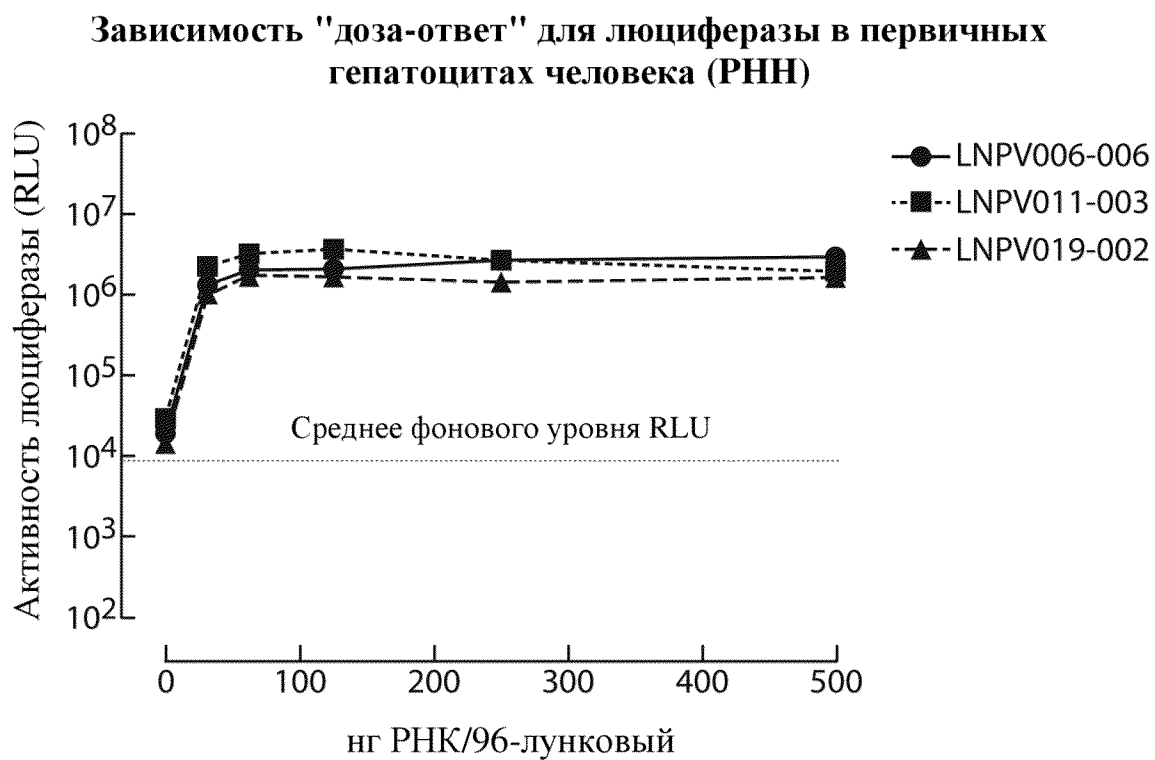
47. Способ по любому из п. 1 или пп. 4-46, где приведение клетки-хозяина в контакт с модулятором ответа хозяина предусматривает обеспечение доступа системы модификации гена к клетке-хозяину.

48. Способ по любому из п. 1 или пп. 4-47, где приведение клетки-хозяина в контакт с модулятором ответа хозяина предусматривает введение модулятора ответа хозяина субъекту, у которого имеется клетка-хозяин.

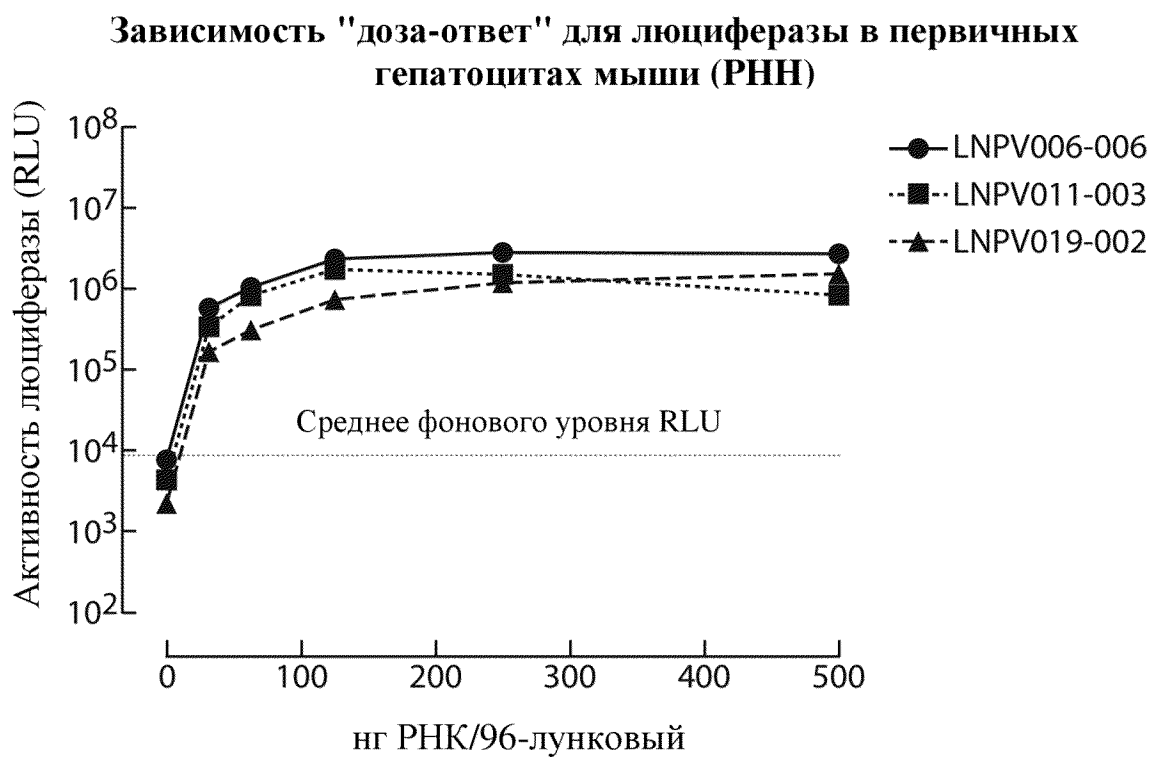
49. Способ по любому из п. 1 или пп. 4-48, который включает приведение клетки-хозяина в контакт со вторым модулятором ответа хозяина.

50. Способ по любому из п. 1 или пп. 4-49, где приведение клетки-хозяина в контакт с системой модификации гена предусматривает приведение клетки-хозяина в контакт с нуклеиновой кислотой (например, ДНК или РНК), кодирующей полипептид Gene Writer, в условиях, обеспечивающих возможность продуцирования полипептида Gene Writer.

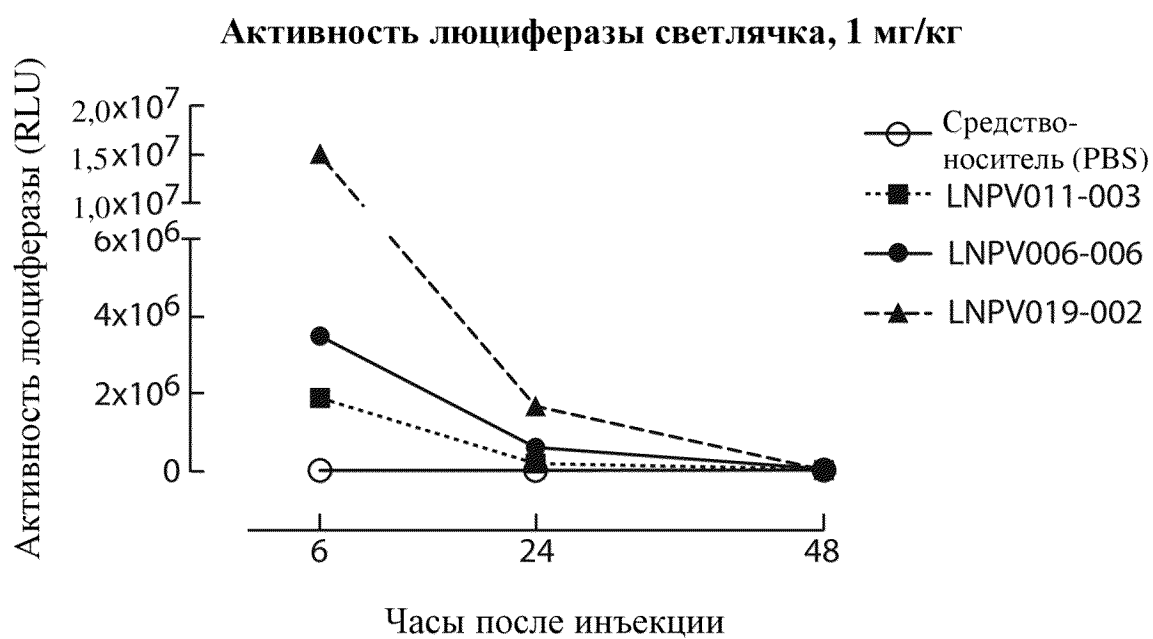
51. Способ по любому из п. 1 или пп. 4-50, где по сравнению с подобным способом без стадии (b) способ приводит в результате к снижению цитотоксичности в отношении клетки-хозяина млекопитающего или субъекта-млекопитающего, в организме которого расположена клетка-хозяин млекопитающего.



ФИГ. 1А



ФИГ. 1В



ФИГ. 2