

(19)



Евразийское
патентное
ведомство

(21) 202293503 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.03.29

(51) Int. Cl. C12N 9/22 (2006.01)
C07K 19/00 (2006.01)
C12N 9/10 (2006.01)
A61K 47/65 (2017.01)

(22) Дата подачи заявки
2021.06.01

(54) КОМПОЗИЦИИ И СПОСОБЫ РЕДАКТИРОВАНИЯ ГЕНОВ

(31) 63/033,397; 63/114,850; 63/171,698

(32) 2020.06.02; 2020.11.17; 2021.04.07

(33) US

(86) PCT/US2021/035244

(87) WO 2021/247570 2021.12.09

(71) Заявитель:

ДЗЕ РИДЖЕНТС ОФ ДЗЕ
ЮНИВЕРСИТИ ОФ КАЛИФОРНИЯ
(US)

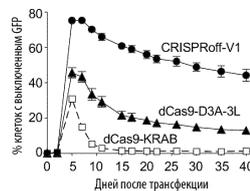
(72) Изобретатель:

Гилберт Люк, Нуэз Джеймс,
Вейссман Джонатан, Чэнь Цзинь (US)

(74) Представитель:

Медведев В.Н. (RU)

(57) В изобретении предлагаются, в частности, слитые белки, композиции и способы манипулирования геномами живых организмов.



202293503

A1

A1

202293503

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-576607EA/032

КОМПОЗИЦИИ И СПОСОБЫ РЕДАКТИРОВАНИЯ ГЕНОВ

ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА РОДСТВЕННЫЕ ЗАЯВКИ

[0001] Данная заявка претендует на приоритет заявки США № 63/171698, поданной 7 апреля 2021 г., заявки США № 63/114850, поданной 17 ноября 2020 г. и заявки США № 63/033397, поданной 2 июня 2020 г., описания которых включены в данный документ посредством ссылок в полном объеме.

ЗАЯВЛЕНИЕ О ПРАВАХ НА ИЗОБРЕТЕНИЯ, СДЕЛАННЫЕ В РАМКАХ ИССЛЕДОВАНИЙ И РАЗРАБОТОК, СПОНСИРУЕМЫХ ФЕДЕРАЛЬНЫМ ПРАВИТЕЛЬСТВОМ

[0002] Данное изобретение было сделано при поддержке правительства в рамках гранта № HR0011-17-2-0043, предоставленного Агентством перспективных оборонных исследовательских проектов. Правительство имеет определенные права на настоящее изобретение.

ССЫЛКА НА «ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ», ТАБЛИЦУ ИЛИ ПРИЛОЖЕНИЕ С ЛИСТИНГОМ КОМПЬЮТЕРНОЙ ПРОГРАММЫ, ПРЕДСТАВЛЕННОЕ В ВИДЕ ASCII-ФАЙЛА

[0003] Перечень последовательностей, записанный в файле 048536-682001WO_SequenceListing_ST25.txt, созданном 31 мая 2021 г., 563294 байт, машинный формат IBM-PC, операционная система MS Windows, настоящим включен посредством ссылки.

УРОВЕНЬ ТЕХНИКИ

[0004] Хотя редактирование генома считается перспективным терапевтическим подходом для лечения заболеваний, оно сопряжено с рисками, связанными с возможностью генотоксичности в результате двухцепочечных разрывов. Кроме того, редактирование генома часто ассоциируется с воздействием на целевой ген по принципу «все или ничего» (т.е. приводит к полному нокауту). В отличие от этого, направленная эпигеномная инженерия не несет риска индуцированной двухцепочечными разрывами (DSB) генотоксичности; кроме того, она дает возможность осуществлять более градуированное воздействие на экспрессию генов и, таким образом, функционировать в диапазоне от полного сайленсинга до менее выраженного эффекта. В данном документе представлены решения для этих и других задач в данной области техники.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[0005] Настоящее изобретение предлагает слитые белки, содержащие домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, и домен K \ddot{u} rrel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных

остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы представляет собой CRISPR-ассоциированный белок.

[0006] Настоящее изобретение предлагает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая варианты осуществления и их аспекты, в клетку, содержащую целевую нуклеиновую кислоту, и доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК. В некоторых аспектах, целевая нуклеиновая кислота содержит CpG-островок. В некоторых аспектах, целевая нуклеиновая кислота не содержит CpG-островок. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы представляет собой CRISPR-ассоциированный белок. Способы сайленсинга последовательности нуклеиновой кислоты-мишени могут использоваться для лечения многих болезней, таких как инфекционные болезни.

[0007] Настоящее изобретение предлагает слитые белки, содержащие домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент эндонуклеазы (например, домен цинкового пальца, TALE (эффектор, подобный активатору транскрипции)), второй линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, и домен K μ rrel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков.

[0008] Настоящее изобретение предлагает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая варианты осуществления и их аспекты, в клетку, содержащую целевую нуклеиновую кислоту. Способы сайленсинга последовательности нуклеиновой кислоты-мишени могут использоваться для лечения многих болезней, таких как инфекционные болезни.

[0009] Эти и другие варианты осуществления и аспекты настоящего изобретения подробно описаны в данном документе.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0010] **Фиг. 1A-1G** показывают устойчивый и мультиплексный сайленсинг генов с помощью CRISPRoff. **Фиг. 1A:** схема слитых белков эпигенетического редактора dCas9, которые были протестированы на активность сайленсинга генов. 3L обозначает Dnmt3L. **Фиг. 1B:** плазмиды, кодирующие продукты слияния dCas9 и sgРНК, котрансфицировали в клетки HEK293T, стабильно экспрессирующие чувствительный к ДНК-метилованию репортер Snrpn-GFP. Трансфицированные клетки сортировали через 2 дня после трансфекции и отслеживали сайленсинг GFP во времени. **Фиг. 1C:** временная зависимость, сравнивающая активность сайленсинга GFP для CRISPRoff-V1, dCas9-3A-3L

и dCas9-KRAB. **Фиг. 1D:** анализ методом бисульфитной ПЦР локуса *Snrpn* до или после нацеливания CRISPRoff. Белые кружки указывают неметилованные динуклеотиды CpG, и черные кружки обозначают метилированные динуклеотиды CpG. Каждая строка представляет одно чтение секвенирования. Красный квадрат обозначает сайт связывания sgPНК. **Фиг. 1E:** сравнение редакторов CRISPRoff-V1 (черный) и CRISPRoff-V2 (синий) при сайленсинге эндогенно меченного GFP гена *H2B*. Пунктирные линии обозначают экспрессию белков CRISPRoff-V1 и -V2. **Фиг. 1F:** репрезентативный график проточной цитометрии экспрессии *H2B-GFP* клетками через 50 дней после трансфекции CRISPRoff V2. **Фиг. 1G:** анализ методом бисульфитного секвенирования области 126 п.о. CpG-островка *H2B*. Красный квадрат обозначает сайт связывания sgPНК. **Фиг. 1H:** количественная оценка клеток с выключенными *ITGB1*, *CD81*, или *CD151* через 3 недели после трансфекции (п.т.) CRISPRoff-V1 или -V2 индивидуальными sgPНК (a-c) или пулом трех sgPНК (a, b, c). **Фиг. 1I:** количественная оценка клеток с выключенными *ITGB1*, *CD81*, *CD151* через 30 дней после трансфекции в экспериментах с нацеливанием на один или два гена. **Фиг. 1J:** количественная оценка мультиплексного сайленсинга трех генов посредством либо гейтирования клеток с выключенным *ITGB1*, а затем гейтирования клеток с выключенными *CD81* и *CD151* (левый столбик), или посредством сначала гейтирования клеток с выключенным *ITGB1*, затем клеток с выключенным *CD151*, и наконец клеток с выключенным *CD81* (правый столбик). Звездочки обозначают популяцию клеток с выключенным отмеченным геном. **Фиг. 1K:** репрезентативный график проточной цитометрии клеток-мишеней сайленсинга *ITGB1*, *CD81* и *CD151*. Клетки сначала гейтировали на сайленсинг *ITGB1* и представленная популяция продемонстрировала сайленсинг *CD81* и *CD151*. **Фиг. 1L:** гистограммный график экспрессии *CLTA* (легкая цепь А клатрина) через 15 месяцев п.т., показывающий 38 клонов, которые сохраняют репрессию *CLTA*, и один клон с реактивированной экспрессией *CLTA*. Средние значения на Фиг. 1C, 1E, 1H, 1I, 1J и 1K были измерены в трех независимых экспериментах. Планки ошибок обозначают СО среднего.

[0011] Фиг. 2A-2H показывают высокоспецифичный и устойчивый транскрипционный сайленсинг с помощью CRISPRoff. **Фиг. 2A-2D:** Графики PНК-секвенирования клеток HEK293T, трансфицированных CRISPRoff и ненацеливающими (NT) sgPНК по сравнению с нацеленными sgPНК (**Фиг. 2B** - на *ITGB1*), (**Фиг. 2C** - на *CD81*), или (**Фиг. 2D** - на *CD151*). Сравнение нетрансфицированных клеток и CRISPRoff с ненацеливающей sgPНК показано на (**Фиг. 2A**). Вулканные диаграммы (внизу) отображают гены-мишени как наиболее значительно подавленные транскрипты в глобальном масштабе. Данные представляют собой средние значения двух независимых реплик. **Фиг. 2E:** манхэттенский график, демонстрирующий дифференциально метилированные CpG в клетках, обработанных CRISPRoff и *CLTA*-нацеливающими или ненацеливающими sgPНК (через 30 дней после трансфекции) при анализе методом WGBS (полногеномное бисульфитное секвенирование). Красные точки обозначают CpG, которые подверглись ДНК-метилованию в клетках с нацеливающими sgPНК, и синие точки

обозначают CpG, которые подверглись ДНК-метилированию в клетках с ненацеливаемыми (NT) sgPНК. Стрелка указывает положение CLTA в геноме. **Фиг. 2F:** сравнение метилирования CpG в окне 55 кб, включающем локус CLTA. Дорожки, обозначенные «Нетр.» обозначают нетрансфицированные клетки; дорожки «NT» (ненацеливающая) обозначают клетки, трансфицированные CRISPRoff и ненацеливающей sgPНК; дорожки «Т» (нацеливающая) обозначают клетки, трансфицированные CRISPRoff и нацеливающей sgPНК. R1 и R2 обозначают две технические реплики. Красные метки обозначают метилированные (бета-значение $>0,5$) и синие метки обозначают неметилированные ($<0,5$) динуклеотиды CpG. CpG-островки обозначены зеленым цветом. **Фиг. 2G:** сравнение сигнала H3K9me3 методом ChIP-seq (технология иммунопреципитации хроматина с последующим секвенированием) для гена H2B в клетках, трансфицированных CRISPRoff и H2B-нацеливающей (фиолетовый) или NT (ненацеливающей) sgPНК (синий), проведенное через 5 дней и 30 дней после трансфекции. Обозначен сайт связывания sgPНК вместе с CpG-островками и соседними генами. Ген BOLA1 содержит два аннотированных сайта начала транскрипции, обозначенных TSS1 и TSS2. **Фиг. 2H:** вулканная диаграмма, сравнивающая данные ChIP-seq для H3K9me3 между трансфицированными CRISPRoff с либо H2B-нацеливающей, либо NT sgPНК. Красные точки указывают гены, проксимальные к H2B-мишени.

[0012] Фиг. 3A-3D показывают, что *флавивирусные* инфекции могут быть блокированы путем эпигеномного редактирования. **Фиг. 3A** представляет собой схему, показывающую эпигеномное редактирование клеток HEK293T для блокирования вирусных инфекций. **Фиг. 3B** инфекция ВЖЛ (вирус желтой лихорадки) и обнаруженное снижение на 50% инфекции ВЖЛ в клетках с выключенной CLTA (правая панель) по сравнению с ДТ HEK293T (левая панель). **Фиг. 3C** показывает, что инфекция DENV-2 значительно снижается в клетках-мишенях SPCS1 и STT3A. **Фиг. 3D** показывает, что сайленсинг генов SPCS1 и STT3A отражает инфекцию DENV-2 в отредактированных клетках.

[0013] Фиг. 4A-4G показывают полногеномный сайленсинг генов с помощью CRISPRoff. **Фиг. 4A:** схема лентивирусного вектора с двойной sgPНК, используемого в полногеномных скринингах CRISPRoff, который содержит две уникальные sgPНК, нацеленные на один и тот же ген. **Фиг. 4B:** схема скрининга геномного пула для определения ландшафта нацеливания CRISPRoff. **Фиг. 4C:** временная зависимость экспрессии CLTA в HEK293T после трансфекции dCas9-KRAB (серый), CRISPRoff-V2 (черный) или мутантом CRISPRoff-D3A^{E765A} (оранжевый). **Фиг. 4D:** сравнение оценок фенотипов (γ) между скринингами CRISPRoff (ось y) и мутантов CRISPRoff (ось x). Выделены три типа ожидаемых отрицательных контролей как псевдо-гены отрицательного контроля (синий), ольфакторные гены (оранжевый) и гены Y-хромосомы (зеленый). **Фиг. 4E:** скрипичный график оценок фенотипов (γ) для генов, определенных как существенные или несущественные из DepMap. Каждый скрининг реплик представлен в виде графика для CRISPRoff (зеленый) и мутанта CRISPRoff (оранжевый). **Фиг. 4F:**

график истинных и ложных коэффициентов позитивности генов, определенных DerMap как существенные. **Фиг. 4G:** график, иллюстрирующий расстояние совпадения существенного гена, определяемое как имеющее $\gamma \leq -0,2$ от ближайшего совпадения существенного гена. Каждая точка соответствует совпадению гена с ближайшим соседним существенным геном, при этом ось x показывает расстояние между двумя генами, а ось y - оценку фенотипа соседнего гена.

[0014] Фиг. 5A-5K показывают опосредованный CRISPRoff сайленсинг генов без аннотаций промоторных CpG-островков. **Фиг. 5A:** график сравнения оценки фенотипа генов между скринингами CRISPRoff и мутантов CRISPRoff, при этом гены, не имеющие аннотации CGI, выделены красным цветом. **Фиг. 5B:** гистограммы флуоресценции mNeonGreen пяти клеточных линий HEK293T, каждая из которых эндогенно мечена указанным геном с расщепленным mNeonGreen. **Фиг. 5C:** количественная оценка клеток с CALD1, DYNC2LI1, LAMP2, MYL6 или VPS25, выключенными после обработки CRISPRoff или мутантом CRISPRoff. Данные измеряли через 14 дней после трансфекции, за исключением VPS25, которые собирали через 11 дней после трансфекции из-за дефекта роста вследствие нокдауна гена. **Фиг. 5D:** количественная оценка процента клеток с DYNC2LI1 или LAMP2, реактивированных после обработки TETv4 нацеливающими или ненацеливающими sgPНК, полученных через 14 дней после трансфекции. **Фиг. 5E:** Профиль метилирования CpG промоторов LAMP2, DYNC2LI1 и MYL6 после обработки CRISPRoff. Белые кружки обозначают статус метилирования CpG нетрансфицированных клеток HEK293T. Каждая точка представляет собой среднее значение для восьми независимых клонов. **Фиг. 5F-5H:** графики временной зависимости экспрессии DYNC2LI1 (**Фиг. 5F**), LAMP2 (**Фиг. 5G**) и MYL6 (**Фиг. 5H**) после трансфекции CRISPRoff или мутантом CRISPRoff. Планки ошибок обозначают СО трех независимых реплик. **Фиг. 5I:** гистограмма экспрессии DYNC2LI1 в 33 клональных линиях, измеренной через 50 дней после трансфекции. Положительный контроль нетрансфицированных клеток помечен. **Фиг. 5J:** манхэттенский график, демонстрирующий дифференциально метилированные CpG между клетками, обработанными CRISPRoff и либо DYNC2LI1-нацеливающей, либо ненацеливающей (NT) sgPНК, при анализе методом WGBS. Стрелка указывает положение DYNC2LI1 в геноме. **Фиг. 5K:** изображение геномного окна размером 10 кб, содержащего локус DYNC2LI1, показывающее усиление метилирования CpG (красный) промотора в клетках, трансфицированных CRISPRoff и DYNC2LI1-нацеливающими sgPНК.

[0015] Фиг. 6A-6L показывают мозаичные скрининги (tiling screens) пула sgPНК, которые выявили широкое целевое окно CRISPRoff-опосредованной репрессии генов. **Фиг. 6A:** схема библиотеки sgPНК с мозаично расположенными PAM-содержащими sgPНК в окне +/- 1 кб от аннотированных сайтов начала транскрипции (TSS). **Фиг. 6B:** сводная информация о количестве генов указанных категорий, составляющих мозаичную библиотеку sgPНК. **Фиг. 6C:** сравнение оценок фенотипа (γ) для генов с аннотированным CGI между CRISPRoff (ось y) и мутантом CRISPRoff (ось x). Каждая точка представляет

собой среднее значение трех наиболее активных sgРНК для каждого гена. Красными точками выделены гены, не имеющие аннотации промотора CGI. **Фиг. 6D:** агрегированный график, сравнивающий нормированную оценку фенотипа для каждой sgРНК, нацеленной на гены с одним аннотированным CGI. Зеленая линия представляет данные скрининга CRISPRoff в HEK293T, оранжевая - мутанта CRISPRoff в HEK293T, и фиолетовая - CRISPRi в K562. **Фиг. 6E-6G:** репрезентативные профили оценки активности sgРНК для DKC1, GPN2 и ZCCHC9 в указанном скрининге (ось y). Зеленый столбик показывает аннотированный CGI, полученный из UCSC Genome Browser. **Фиг. 6H:** репрезентативный профиль оценки активности sgРНК для ORC5 в указанном скрининге (ось y). **Фиг. 6I:** агрегированный график, сравнивающий нормированные оценки фенотипа каждой sgРНК для генов без аннотированных CGI. **Фиг. 6J:** наложение нормированной оценки фенотипа sgРНК по результатам скрининга CRISPRoff (зеленый) на сигнал МНКазы, отражающий занятость нуклеосом (серый). График представляет собой совокупность генов с одним аннотированным CGI. **Фиг. 6K:** наложение нормированной оценки фенотипа sgРНК по результатам скрининга CRISPRoff (зеленый) на сигнал МНКазы, отражающий занятость нуклеосом (серый). График представляет собой совокупность 39 генов, не имеющих аннотированных CGI. **Фиг. 6L:** график активности sgРНК вместе с сигналом МНКазы для H2B, полученный в результате мозаичного скрининга sgРНК, представленного на Фиг. S6F.

[0016] Фиг. 7A-7J: показывают сайленсинг генов CRISPRoff в iPSC (индуцированные плюрипотентные стволовые клетки), iPSC-производных нейронах и энхансерах. **Фиг. 7A:** экспериментальный процесс нокдауна CD81 с помощью CRISPRoff в iPSC с последующей опосредованной NGN2 дифференцировкой отредактированных клеток в нейроны. **Фиг. 7B:** количественная оценка клеток с CD81, выключенным посредством CRISPRi или CRISPRoff с CD81-нацеливающими или ненацеливающими (NT) sgРНК, измеренная через 30 дней после трансфекции. Планки погрешностей представляют СО по результатам трех независимых экспериментов. **Фиг. 7C:** количественная оценка клеток с выключенным CD81 в моменты времени, указанные на (A). Серые столбики показывают процент отредактированных с помощью iPSC клеток с выключенным CD81, которые не дифференцировались во время эксперимента. Красные столбики представляют клетки, прошедшие через протокол дифференцировки нейронов. Планки погрешностей представляют СО по результатам трех независимых экспериментов. **Фиг. 7D:** репрезентативная гистограмма экспрессии CD81 в день 8 нейрональной дифференцировки родительских не отредактированных (серый) или отредактированных с помощью CD81 iPSC (красный). **Фиг. 7E:** бисульфитная ПЦР области 140 п.о. промотора CD81 в клетках, трансфицированных CRISPRoff и ненацеливающей (NT) или CD81-нацеливающей sgРНК. **Фиг. 7F:** репрезентативные изображения микроскопии яркого поля дифференцированных нейронов, полученных из iPSC, трансфицированных CRISPRoff и МАРТ-нацеливающей или ненацеливающей (NT) sgРНК. **Фиг. 7G:** количественная оценка клеток с подавленной экспрессией тау-белка (Tau-off) в клетках, трансфицированных

CRISPRoff и ненацеливающей (NT) или MAPT-нацеливающей sgPHK, измеренная через 10 дней после дифференцировки. **Фиг. 7H:** репрезентативные графики проточной цитометрии окрашивания тау-белка в нейронах, полученных из iPSC, после трансфекции CRISPRoff с ненацеливающей (NT) или MAPT-нацеливающей sgPHK. Гейты основаны на невозмущенных нейронах, полученных из iPSC. **Фиг. 7I:** схема локуса PVT1 с промотором и четырьмя энхансерными элементами (E1-E4), с метками на удалении от TSS. **Фиг. 7J:** графики нормированных уровней транскрипта PVT1 по результатам количественной ОТ-кПЦР клеток, обработанных CRISPRoff (слева) или мутантом CRISPRoff D3A (справа) и sgPHK, нацеливающими либо на промотор (Пр.), либо на четыре энхансерных элемента (E1-E4), нормированных по контрольным sgPHK. Звездочки обозначают статистическую значимость по t-тесту, а каждая техническая реплика представлена красными точками.

[0017] Фиг. 8A-8K показывают оптимизацию конструкции CRISPRoff для долговременного сайленсинга генов, относящуюся к Фиг. 1. **Фиг. 8A:** схема конструктора CRISPRoff-V1 и различные линкерные последовательности, используемые для получения вариантов белков. **Фиг. 8B:** временная зависимость сайленсинга CLTA-GFP после трансфекции вариантов CRISPRoff-V1 или контролей dCas9-KRAB (серый) и dCas9-D3A-D3L (оранжевый). **Фиг. 8C:** кристаллическая структура DNMT3A (оранжевый) и DNMT3L (желтый) в комплексе с CpG-содержащей ДНК (PDB 5YX2). Стрелки указывают позиции присоединения dCas9 для CRISPRoff-V1 и CRISPRoff-V2. **Фиг. 8D:** схема четырех конструкторов CRISPRoff-V2, которые отличаются тем, что BFP используется в качестве линкера между dCas9 и KRAB или отделен от CRISPRoff последовательностью P2A. Конструкторы V2.3 и V2.4 кодируют последовательности NLS (сигнал ядерной локализации) на амино и карбоксильном концах dCas9. **Фиг. 8E:** вестерн-блоттинг экспрессии белка конструктора dCas9, dCas9-KRAB, CRISPRoff. **Фиг. 8F:** временная зависимость сайленсинга CLTA-GFP после трансфекции конструкторов V1 и V2.1, V2.3, V2.4, наряду с контролями, содержащими dCas9-KRAB и только dCas9. **Фиг. 8G:** временная зависимость сайленсинга Snrpn-GFP после трансфекции dCas9-D3A-3L (оранжевый), dCas9-KRAB (серый), или CRISPRoff-V1 (черный) и V2 (синий). **Фиг. 8H:** репрезентативный график проточной цитометрии клеток HEK293T CLTA-GFP через 6 дней после трансфекции мPHK, кодирующей CRISPRoff. **Фиг. 8I:** репрезентативные графики проточной цитометрии мультиплексного генного таргетинга двух генов одновременно, измеренные через 30 дней после трансфекции. **Фиг. 8J:** количественная оценка сайленсинга гена, измеренная через 31 день после трансфекции CRISPRoff одновременно четырьмя sgPHK, нацеленными на ITGB1, CLTA, CD81 и CD151. **Фиг. 8K:** количественная оценка клеток с выключенными CLTA, CD81, и CD151 в клетках, имеющих либо выключенный (синий), либо невыключенный (серый) ITGB1 в эксперименте с нокдауном четырех генов. Планки погрешностей на **Фиг. 8B, F, G, J и K** представляют СО из трех независимых экспериментов.

[0018] Фиг. 9A-9F: показывают, что CRISPRoff применим в различных клеточных

линиях и с ортогональными РНК-направляемыми белками CRISPR, относящимися к Фиг. 1. **Фиг. 9А:** гистограммы проточной цитометрии экспрессии CRISPRoff (BFP) до и после обработки доксициклином (dox). Через 24 часа после введения доксициклина, среду заменяют на среду без доксициклина (1 день и 2 дня после промывки доксициклином) для выключения экспрессии CRISPRoff. **Фиг. 9В:** количественная оценка клеток K562 с выключенным CD81 через 10 дней после начальной индукции доксициклином экспрессии CRISPRoff. Доксициклин был включен в состав среды в течение либо 3 дней (посередине), либо 4 дней (справа) до отмывки клеток для удаления доксициклина. **Фиг. 9С:** количественная оценка клеток с выключенными ITGB1, CD81 или CD151 в HeLa и U2OS, измеренная через 18 дней после трансфекции CRISPRoff-V1 с sgРНК, нацеливающими на указанные гены. **Фиг. 9D:** репрезентативный график проточной цитометрии экспрессии CD81 в клетках iPS после трансфекции CRISPRoff с либо ненацеливающими sgРНК, либо sgРНК, нацеливающими на CD81. **Фиг. 9E-9F:** сравнение клеток с выключенными CLTA (Фиг. 9E) и H2B (Фиг. 9F) через 10 дней после трансфекции CRISPRoff с dCas9 *S. pyogenes* (dSpyCas9) или *S. aureus* (dSauCas9), или dCas12a *Lachnospiraceae bacterium* (dLbCas12a). Планки погрешностей на **Фиг. 9B, 9C, 9E и 9F** представляют СО трех независимых экспериментов.

[0019] Фиг. 10A-10L показывают транскрипционную специфичность сайленсинга CLTA, H2B, и RAB11A, относящихся к Фиг. 2. **Фиг. 10А:** график РНК-секвенирования TPM (транскриптов на килобазу на миллион картированных прочтений) для клеток HEK293T, трансфицированных CRISPRoff и NT (ненацеливающими) sgРНК по сравнению с нетрансфицированными клетками HEK293T. **Фиг. 10B-10D:** графики РНК-секвенирования TPM (транскриптов на килобазу на миллион картированных прочтений) для клеток HEK293T, трансфицированных CRISPRoff и либо ненацеливающей (NT) sgРНК по сравнению с sgРНК, нацеливающей на CLTA (Фиг. 10B), HIST2H2BE (Фиг. 10C), либо RAB11A (Фиг. 10D). Данные представляют собой средние значения двух независимых реплик. **Фиг. 10E-10F:** Представление изменений генной экспрессии в пределах +/- 1000 кб от аннотированного гена-мишени для CD81, CD151, ITGB1 (Фиг. 10E) и RAB11, HIST2H2BE, CLTA (Фиг. 10F). Каждый бокс представляет собой ген. **Фиг. 10G:** гистограмма среднего полногеномного покрытия CpG, полученная в результате бисульфитного секвенирования всего генома. ДТ (дикий тип) указывает на нетрансфицированные клетки; NT указывает CRISPRoff, доставленные ненацеливающими sgРНК; Т указывает CRISPRoff, доставленные sgРНК, нацеливающими на CLTA или DYNC2LI1. Эксперименты проводились в двух репликах. Эксперименты с CLTA представлены на Фиг. 2, эксперименты с DYNC2LI1 представлены на Фиг. 5. **Фиг. 10H:** гистограмма среднего полногеномного метилирования CpG (бета-значения) в образцах, представленных на Фиг. 10G. **Фиг. 10I:** графики тепловых карт, сравнивающие среднее метилирование CpG (бета-значения) скользящих окон 20-CpG между нетрансфицированными (ДТ) репликами (слева), ДТ и ненацеливающими (NT) (посередине), и ДТ и нацеливающими sgРНК (справа) для экспериментов с CLTA

(верхний ряд) и DYNC2LI1 (нижний ряд). Красный цвет указывает наибольшую плотность и синий цвет указывает наименьшую плотность. Белые области указывают на отсутствие окон с соответствующими средними уровнями метилирования. **Фиг. 10J:** тепловая карта, показывающая парные корреляции между профилями полногеномного метилирования CpG образцов, описанных на Фиг. S3G, включая реплики. Образцы отсортированы методом неконтролируемой иерархической кластеризации. Темно-коричневый цвет обозначает самую высокую корреляцию, светло-желтый цвет - самую низкую корреляцию. Этот анализ подчеркивает различия в глобальном метилировании CpG между образцами в экспериментах с CLTA и DYNC2LI1. **Фиг. 10K:** крупный план геномной области размером 17 кб, содержащей ген ZSCAN16. Показаны профили метилирования ДНК для нетрансфицированных клеток, клеток с нацеливающими (NT) и нацеливающими sgРНК из эксперимента с CLTA. В боксе показана дифференциально метилированная область промотора гена, что указывает на усиление метилирования CpG в клетках с нацеливающей sgРНК по сравнению с контрольными клетками. **Фиг. 10L:** крупный план геномной области размером 30 кб, содержащей ген RPS6KA6. В боксе показано усиление метилирования CpG промотора в клетках с нацеливающей sgРНК. CpG-островок, расположенный в промоторе гена, выделен зеленым цветом.

[0020] Фиг. 11A-11E показывают, что полногеномный сайленсинг с помощью CRISPRoff является воспроизводимым и специфическим (по отношению к Фиг. 4). **Фиг. 11A:** график, сравнивающий оценку фенотипа (γ) генов между полногеномными скринингами технических реплик CRISPRoff (слева) и CRISPRoff мутанта D3A-метилтрансферазы (справа). Отрицательный контроль sgРНК выделен синим цветом. **Фиг. 11B:** скрипичные графики оценки фенотипа (γ) всех генов каждого скрининга. **Фиг. 11C:** гистограмма числа генов с указанной оценкой фенотипа (γ) из скринингов CRISPRoff и мутанта CRISPRoff. Светло-зеленая и светло-оранжевая линии соответствуют оценке фенотипов sgРНК отрицательного контроля. **Фиг. 11D-11E:** анализ обогащения набора генов (GSEA) для генов, связанных с репликацией ДНК и рибосомой, подтверждающий обогащение ожидаемых существенных генов. Гены ранжированы от низшей (красный) к высшей (синий) оценке фенотипов.

[0021] Фиг. 12A-12I показывают дизайн и валидацию мозаичных скринингов sgРНК, демонстрирующих геномное окно гибкой нацеливающей активности CRISPRoff (относящихся к Фиг. 6). **Фиг. 12A:** гены, выбранные для мозаичных скринингов sgРНК, выделены на вулканной диаграмме, изображающей оценки фенотипов генов, полученные в результате предшествующих полногеномных скринингов CRISPRi в клетках K562 (Horlbeck et al. 2016). Цветом обозначены гены с одним (оранжевый), несколькими (фиолетовый) аннотированными CGI или с их отсутствием (зеленый). **Фиг. 12B:** агрегированный график, сравнивающий нормированную оценку фенотипа каждой sgРНК для генов с несколькими CGI. Зеленая линия представляет данные скрининга, полученные при трансфекции CRISPRoff в клетки HEK293T, оранжевая - при трансфекции мутанта CRISPRoff в HEK293T, и фиолетовая - CRISPRi в K562. **Фиг. 12C:** репрезентативный

профиль оценки активности sgРНК для *TBCD* (*тубулин-специфический шаперон D*) по результатам трех скринингов. Зеленый столбик показывает аннотированные CGI, полученные из UCSC Genome Browser. **Фиг. 12D-12E:** графики с наложением оценок фенотипов sgРНК и сигналов МНКазы для GFER (фактор роста, усилитель регенерации печени) и ИММТ (белок внутренней мембраны митохондрий). **Фиг. 12F:** экспериментальный процесс мозаичного скрининга sgРНК с целью определения оптимальных sgРНК для четырех генов с эндогенной GFP-меткой: CLTA, H2B, RAB11A и VIM. Указанная популяция на гистограммах представляет популяцию клеток, которые поддерживают сайленсинг гена через 4 недели после трансфекции CRISPRoff. **Фиг. 12G-12I:** графики наложения активности sgРНК и сигнала МНКазы для CLTA, RAB11A, и VIM по результатам мозаичного скрининга sgРНК.

[0022] Фиг. 13A-13I: показывают локализацию H3K9me3 и метилирование CpG, несмотря на дистальные сайты эпигенетического становления (относящиеся к Фиг. 6). **Фиг. 13A:** схема промотора H2B с двумя аннотированными сайтами sgРНК: sg-A в TSS (сайт начала транскрипции) и sg-B, расположенный около за 2 кб перед TSS. CpG-островок охватывает 1,4 кб. Сайты 1, 2 и 3 представляют области зондируемые на метилирование CpG методом бисульфитной ПЦР, как представлено на Фиг. 13E-13I. **Фиг. 13B:** временная зависимость сайленсинга H2B после трансфекции CRISPRoff с sg-A или sg-B. **Фиг. 13C:** сравнение профилей H3K9me3 в промоторе H2B (HIST2H2BE, окрашен в красный цвет) на 5-й (зеленые дорожки) и 30-й (фиолетовые дорожки) день после трансфекции CRISPRoff с помощью sg-A, sg-B или ненацеливающей (NT) sgРНК. Помечены сайты связывания sgРНК, а также аннотации островков CpG (зеленый) и базальная неметилованная область CpG промотора H2B до трансфекции, полученная путем полногеномного бисульфитного секвенирования (WGBS) нетрансфицированных клеток ДТ на Фиг. 2. **Фиг. 13D:** сравнение профилей H3K9me3 в промоторе H2B, как описано в С, за исключением того, что в экспериментах использовали CRISPRoff с мутацией D3A-метилтрансферазы. **Фиг. 13E-13F:** количественная оценка метилирования CpG в сайте 1 и сайте 3 промотора H2B (отмечен на Фиг. 13A) в клетках, трансфицированных CRISPRoff и либо sg-A (синий), sg-B (оранжевый), либо ненацеливающей (серый) sgРНК. Клетки собирали для проведения бисульфитной ПЦР в день 5 после трансфекции. **Фиг. 13G-13I:** количественная оценка метилирования CpG в сайтах 1, 2 и 3 промотора H2B, полученная через 30 дней после трансфекции.

[0023] На фигурах и в примерах, CRISPRoff-V1 относится к SEQ ID NO:1, и CRISPRoff-V2 относится к SEQ ID NO:107 (V2.1), если не указано иное.

ПОДРОБНОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

[0024] Определения

[0025] Если не указано иное, все технические и научные термины, используемые в данном документе, имеют значения, обычно понятные специалисту в области техники, к которой относится настоящее изобретение. Следующие ссылки дают специалистам общее определение многих терминов, используемых в настоящем изобретении: Singleton et al.,

Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., Rieger et al. (eds.), Springer Verlag (1991); и Hale & Marham, The Harper Collins Dictionary of Biology (1991). В контексте данной заявки следующие термины имеют приписываемые им значения, если не указано иное.

[0026] Использование в настоящем изобретении и в приведенной ниже формуле изобретения терминов в единственном числе (например, с неопределенным или определенным артиклем «a», «an», «the» и т.д.) соответствует традиционному для патентов значению «по меньшей мере один», если только в конкретном случае из контекста не следует, что данный термин предназначен для обозначения конкретно единственного и только единственного числа. Аналогичным образом, термин «включающий» является открытым, не исключая дополнительные элементы, признаки, компоненты и т.д.

[0027] Термины «состоять», «включать» и «иметь», а также их производные, используются в данном документе взаимозаменяемо как всеобъемлющие, открытые термины. Например, использование слов «состоящий», «включающий» или «имеющий» означает, что какой бы элемент ни входил в состав, содержался или был включен, он не является единственным элементом, охватываемым предметом выражения, содержащего указанный глагол.

[0028] Для конкретных белков, описанных в данном документе (например, KRAB, dCas9, Dnmt3A, Dnmt3L), названный белок включает любые встречающиеся в природе формы белка, или варианты или гомологи, которые сохраняют активность белка (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с нативным белком). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности (например, участку из 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с встречающейся в природе формой. В некоторых аспектах, указанный белок представляет собой белок, идентифицированный по ссылке на последовательность NCBI. В некоторых аспектах, указанный белок представляет собой белок, идентифицированный по ссылке на последовательность NCBI, или его функциональный фрагмент или гомолог.

[0029] Термин «дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы» и т.п. относится, в обычном и общепринятом смысле, к РНК-направляемой ДНК-эндонуклеазе (например, мутированной форме встречающейся в природе РНК-направляемой ДНК-эндонуклеазы), которая нацелена на конкретную фосфодиэфирную связь фосфодиэфирную связь в полинуклеотиде ДНК, при этом распознаванию фосфодиэфирной связи способствует отдельная полинуклеотидная последовательность (например, последовательность РНК (например, одиночная направляющая РНК (sgРНК)), но не способна расщеплять целевую фосфодиэфирную связь в значительной степени

(например, не происходит измеримого расщепления фосфодиэфирной связи в физиологических условиях). Дефицитная по нуклеазе РНК-направляемая ДНК-эндонуклеаза, таким образом, сохраняет ДНК-связывающую способность (например, специфическое связывание с последовательностью-мишенью) в комплексе с полинуклеотидом (например, sgРНК), но не обладает значительной эндонуклеазной активностью (например, любой количественно обнаруживаемой эндонуклеазной активностью). В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, dCpf1, ddCpf1, Cas-phi, дефицитный по нуклеазе вариант Cas9, дефицитную по нуклеазе эндонуклеазу CRISPR класса II, домен цинкового пальца, эффектор, подобный активатору транскрипции (TALE), домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса (бокса связывания тяжелых металлов), домен Wor3, домен ОВ-складки (олигонуклеотид-связывающей складки), домен иммуноглобулина или домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен цинкового пальца, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен ОВ-складки, домен иммуноглобулина или домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен лейциновой молнии. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен крылатой спирали. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой мотив спираль-поворот-спираль. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен спираль-петля-спираль. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен НМВ-бокса. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен Wor3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен ОВ-складки. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен иммуноглобулина. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, ddCpf1, Cas-phi, дефицитный по нуклеазе вариант Cas9 или дефицитную по нуклеазе эндонуклеазу CRISPR класса II. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9 *S. pyogenes*. В некоторых аспектах, дефицитный по нуклеазе

фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9 *S. aureus*. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas12a *Lachnospiracea* (dLbCas12a). В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas12a *Lachnospiracea bacterium*. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas12a. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой ddCas12a. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой Cas-phi.

[0030] Термин «CRISPR-ассоциированный белок» или «белок CRISPR» относится к любому белку CRISPR, который функционирует как дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы, т.е., белок CRISPR, в котором каталитические сайты эндонуклеазной активности являются дефектными или лишены активности. Примеры CRISPR-ассоциированных белков включают dCas9, dCpf1, dCas12, Cas-phi, дефицитный по нуклеазе вариант Cas9, дефицитную по нуклеазе эндонуклеазу CRISPR класса II и т.п.

[0031] «CRISPR-ассоциированный белок 9», «Cas9», «Csn1» или «белок Cas9», как указано в данном документе, включает любую из рекомбинантных или встречающихся в природе форм эндонуклеазы Cas9 или ее вариантов или гомологов, которые сохраняют активность фермента эндонуклеазы Cas9 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Cas9). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности (например, участка из 50, 100, 150 или 200 непрерывно расположенных аминокислот) по сравнению с встречающимся в природе белком Cas9. В некоторых аспектах, белок Cas9 по существу идентичен белку, идентифицированному ссылочным номером UniProt Q99ZW2, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, белок Cas9 имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q99ZW2. В некоторых аспектах, белок Cas9 имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q99ZW2. В некоторых аспектах, белок Cas9 имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q99ZW2. В некоторых аспектах, белок Cas9 имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q99ZW2. В некоторых аспектах,

белок Cas9 имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q99ZW2.

[0032] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9. Термины «dCas9» или «белок dCas9», как указано в данном документе, представляет собой белок Cas9, в котором оба каталитических сайта эндонуклеазной активности являются дефектными или лишены активности. В некоторых аспектах, белок dCas9 имеет мутации в позициях, соответствующих D10A и H840A Cas9 *S. pyogenes*. В некоторых аспектах, белок dCas9 лишен эндонуклеазной активности вследствие точечных мутаций в обоих каталитических сайтах эндонуклеазы (RuvC и HNH) Cas9 дикого типа. Точечные мутации могут представлять собой D10A и H840A. В некоторых аспектах, dCas9 по существу не обладает детектируемой эндонуклеазной (*например*, эндодезоксирибонуклеазной) активностью. В некоторых аспектах, dCas9 включает аминокислотную последовательность SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:23. В некоторых аспектах, dCas9 имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:23. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9 *S. pyogenes*. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9 *S. aureus*.

[0033] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой «ddCpf1» или «ddCas12a». Термины «Cpf1 с инактивацией ДНКазы» или «ddCpf1» относятся к мутированному Cpf1 *Acidaminococcus* sp. (AsCpf1), образуемому в результате инактивации ДНКазной активности Cpf1. В некоторых аспектах, ddCpf1 включает мутацию E993A в домене RuvC AsCpf1. В некоторых аспектах, ddCpf1 по существу не обладает детектируемой эндонуклеазной (*например*, эндодезоксирибонуклеазной) активностью. В некоторых аспектах, ddCpf1 включает аминокислотную последовательность SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность SEQ ID NO:34. В

некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, ddCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:34. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas12a *Lachnospiracea bacterium*.

[0034] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dLbCpf1. Термин «dLbCpf1» относится к мутированному Cpf1 *Lachnospiraceae bacterium* ND2006 (LbCpf1), лишенному активности ДНКазы. В некоторых аспектах, dLbCpf1 включает мутацию D832A. В некоторых аспектах, dLbCpf1 по существу не обладает детектируемой эндонуклеазной (например, эндодезоксирибонуклеазной) активностью. В некоторых аспектах, dLbCpf1 включает аминокислотную последовательность SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:35. В некоторых аспектах, dLbCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:35.

[0035] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dFnCpf1. Термин «dFnCpf1» относится к мутированному Cpf1 *Francisella novicida* U112 (FnCpf1), лишенному активности ДНКазы. В некоторых аспектах, dFnCpf1 включает мутацию D917A. В

некоторых аспектах, the dFnCpf1 по существу не обладает детектируемой эндонуклеазной (*например*, эндодезоксирибонуклеазной) активностью. В некоторых аспектах, dFnCpf1 включает аминокислотную последовательность SEQ ID NO: 36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность SEQ ID NO: 36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:36. В некоторых аспектах, dFnCpf1 имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:36.

[0036] «Cpf1» или «белок Cpf1», как указано в данном документе, включает любые рекомбинантные или встречающиеся в природе формы эндонуклеазы Cpf1 (CRISPR *Prevotella* и *Francisella* 1) или ее вариантов или гомологов, которые сохраняют активность фермента эндонуклеазы Cpf1 (*например*, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Cpf1). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности (*например*, участку из 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с встречающимся в природе белком Cpf1. В некоторых аспектах, белок Cpf1 по существу идентичен белку, идентифицированному ссылочным номером UniProt U2UMQ6, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, белок Cpf1 идентичен белку, идентифицированному ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 идентичен белку, идентифицированному ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 идентичен белку,

идентифицированному ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 идентичен белку, идентифицированному ссылочным номером UniProt U2UMQ6. В некоторых аспектах, белок Cpf1 имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt U2UMQ6.

[0037] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой дефицитный по нуклеазе вариант Cas9. Термин «дефицитный по нуклеазе вариант Cas9» относится к белку Cas9, имеющему одну или несколько мутаций, которые увеличивают его специфичность связывания с PAM по сравнению с Cas9 дикого типа, и дополнительно включающему мутации, которые делают белок неспособным к проявлению эндонуклеазной активности или сильно ухудшают ее. Не желая ограничивать себя теорией, укажем, что, как считается, последовательность-мишень должна быть связана с PAM (мотив, примыкающий к протоспейсеру); т.е., короткой последовательностью, распознаваемой комплексом CRISPR. Точные требования к последовательности и длине PAM отличаются в зависимости от используемого фермента CRISPR, но обычно PAM представляют собой последовательность из 2-5 пар оснований, примыкающую к протоспейсеру (то есть к целевой последовательности). Специфичность связывания дефицитных по нуклеазе вариантов Cas9 с PAM может быть определена любым способом известным в данной области техники. Описания и применение известных вариантов Cas9 приведены, например, в Shmakov et al., Diversity and evolution of class 2 CRISPR-Cas systems. Nat. Rev. Microbiol. 15, 2017, и Cebrian-Serrano et al., CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. Mamm. Genome 7-8, 2017. Примеры вариантов Cas9 перечислены в Таблице 1 ниже.

[0038] Таблица 1

Варианты Cas9	Домены PAM	Ссылки
Cas9 <i>Strep pyogenes</i> (Sp)	NGG	Hsu et al. 2014 Cell
Cas9 <i>Staph aureus</i> (Sa)	NNGRRT или NNGRR NNGGGT, NNGAAT, NNGAGT (Zetsche)	Ran et al. 2015 Nature
Мутант SpCas9 VQR (D1135V, R1335Q, T1337R)	NGAG>NGAT=NGAA>NGAC NGCG	Kleinstiver et al. 2015 Nature
Мутант SpCas9 VRER (D1135V/G1218R/R1335E/T1337R)	NGCG	Kleinstiver et al. 2015 Nature

SpCas9 D1135E	NGG, большая верность, меньше разрезаний по сайтам NAG и NGA	Kleinstiver et al. 2015 Nature
Мутант eSpCas9 1.1 (K848A/K1003A/R1060A)	NGG	Slymaker et al. Science 2015
SpCas9 HF1 (Q695A, Q926A, N497A, R661A)	NGG	Kleinstiver et al. 2016 Nature
AsCpf1	TTTN (5' от sgPHK)	Zetsche et al. 2015 Cell
НураCas9 (N692A, M694A, Q695A, H698A)		Chen et al., Nature, volume 550, pages 407-410 (19 October 2017)

[0039] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой дефицитную по нуклеазе эндонуклеазу CRISPR класса II. Термин «дефицитная по нуклеазе эндонуклеаза CRISPR класса II», используемый в данном документе, относится к любой эндонуклеазе CRISPR класса II, имеющей мутации, приводящие к сниженной, нарушенной или инактивированной эндонуклеазной активности.

[0040] Термин «дефицитный по нуклеазе фермент ДНК-эндонуклеазы» относится к ДНК-эндонуклеазе (например, мутированной форме встречающейся в природе ДНК-эндонуклеазы), которая нацелена на конкретную фосфодиэфирную связь в полинуклеотиде ДНК, но не требует направляющей РНК. В некоторых вариантах осуществления, «дефицитный по нуклеазе фермент ДНК-эндонуклеазы» представляет собой домен цинкового пальца или TALE.

[0041] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой «домен цинкового пальца.» Термины «домен цинкового пальца» или «связывающий домен цинкового пальца» или «ДНК-связывающий домен цинкового пальца» используются взаимозаменяемо и относятся к белку или домену в составе более крупного белка, который связывает ДНК специфическим в отношении последовательности образом посредством одного или нескольких цинковых пальцев, представляющих собой участки аминокислотной последовательности в связующем домене, структура которых стабилизирована посредством координации иона цинка. В некоторых вариантах осуществления, домен цинкового пальца не встречается в природе, поскольку он разработан для связывания с выбранным сайтом-мишенью. В некоторых аспектах, связывающий домен цинкового пальца относится к белку, домену в составе более крупного белка или дефицитному по нуклеазе ферменту РНК-направляемой ДНК-эндонуклеазы, способного связываться с любым цинковым пальцем, известным в данной области техники, таким как цинковые пальцы типа C2H2, типа CCHC, типа PHD или типа RING.

[0042] При использовании в данном документе, «цинковый палец» обозначает полипептидный структурный мотив, сложенный вокруг связанного катиона цинка. В некоторых вариантах осуществления, полипептид цинкового пальца имеет последовательность формы $X_3\text{-Cys-}X_{2-4}\text{-Cys-}X_{12}\text{-His-}X_{3-5}\text{-His-}X_4$, где X обозначает любую аминокислоту (например, X_{2-4} указывает олигопептид длиной в 2-4 аминокислоты). Как правило, существует широкий спектр вариантов последовательности в 28-31 аминокислотах известных полипептидов цинкового пальца. Инвариантными являются только два консенсусных остатка гистидина и два консенсусных остатка цистеина, связанные с центральным атомом цинка. Из остальных остатков три-пять являются высококонсервативными, в то время как среди других остатков могут быть значительные вариации. Несмотря на широкий спектр вариантов последовательности в полипептиде, цинковые пальцы этого типа имеют схожую трехмерную структуру. Однако существует широкий диапазон специфичностей связывания между различными цинковыми пальцами, т.е. различные цинковые пальцы связывают двухцепочечные полинуклеотиды, имеющие широкий спектр последовательностей нуклеотидов. В некоторых аспектах, цинковый палец относится к типу C2H2. В некоторых аспектах, цинковый палец относится к типу CCHC. В некоторых аспектах, цинковый палец относится к типу PHD. В некоторых аспектах, цинковый палец относится к типу RING.

[0043] В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE. «TALE» или «эффектор, подобный активатору транскрипции» относятся к искусственным рестрикционным ферментам, созданным путем слияния ДНК-связывающего домена эффектора TAL с доменом расщепления ДНК. TALE обеспечивают эффективное, программируемое и специфическое расщепление ДНК и представляют собой мощные инструменты для редактирования генома *in situ*. Эффекторы, подобные активатору транскрипции (TALE) могут быть быстро сконструированы для связывания практически любой последовательности ДНК. Термин TALE, при использовании в данном документе, используется в широком смысле и включает мономерный TALE, который может расщеплять двухцепочечную ДНК без помощи другого TALE. Термин TALE также используется для обозначения одного или обоих членов пары TALE, которые сконструированы для совместной работы по расщеплению ДНК в одном и том же сайте. TALE, которые работают вместе, могут быть названы «левый TALE» и «правый TALE», что указывает на хиральность ДНК. TALE представляют собой белки, секретлируемые бактериями *Xanthomonas*. ДНК-связывающий домен содержит высококонсервативную последовательность 33-34 аминокислот, за исключением 12-й и 13-й аминокислот. Эти два положения представляют собой два повторяющихся переменных остатка (RVD) и демонстрируют сильную корреляцию со специфическим распознаванием нуклеотидов. Такая простая связь между аминокислотной последовательностью и распознаванием ДНК позволяет конструировать специфические ДНК-связывающие домены путем выбора комбинации сегментов повторов, содержащих соответствующие RVD.

[0044] Термин «домен Krüppel-ассоциированного бокса» или «KRAB-домен» в данном документе относится к категории доменов транскрипционной репрессии, присутствующих в приблизительно 400 транскрипционных факторах человека на основе белков цинковых пальцев. KRAB-домены типично включают от около 45 до около 75 аминокислотных остатков. Описание KRAB-доменов, включая их функцию и применение, приведено, например, в Ecco, G., Imbeault, M., Trono, D., KRAB zinc finger proteins, *Development* 144, 2017; Lambert et al., *The human transcription factors*, *Cell* 172, 2018; Gilbert et al., *Cell* (2013); и Gilbert et al., *Cell* (2014). В некоторых аспектах, KRAB-домен представляет собой KRAB-домен Koh 1. В некоторых аспектах, KRAB-домен включает последовательность, представленную SEQ ID NO:16. В некоторых аспектах, KRAB-домен представляет собой последовательность SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:16. В некоторых вариантах осуществления, KRAB-домен представляет собой KRAB-домен ZIM3 или аминокислотную последовательность, имеющую 85%, 90% или 95% идентичности последовательности с ним. В некоторых вариантах осуществления, KRAB-домен представляет собой KOX1 или аминокислотную последовательность, имеющую 85%, 90% или 95% идентичности последовательности с ним. Gilbert et al, *Cell*, 159:647-661 (2014); Alerasool et al, *Nature Methods*, (2020).

[0045] Термин «ДНК-метилтрансфераза», используемый в данном документе, относится к ферменту, который катализирует перенос метильной группы к ДНК. Неограничивающие примеры ДНК-метилтрансфераз включают Dnmt1, Dnmt3A и Dnmt3B. В некоторых аспектах, ДНК-метилтрансфераза представляет собой ДНК-метилтрансферазу млекопитающего. В некоторых аспектах, ДНК-метилтрансфераза представляет собой ДНК-метилтрансферазу человека. В некоторых аспектах, ДНК-метилтрансфераза представляет собой мышиную ДНК-метилтрансферазу. В некоторых аспектах, ДНК-метилтрансфераза представляет собой бактериальную цитозинметилтрансферазу и/или бактериальную нецитозинную метилтрансферазу. В зависимости от конкретной ДНК-метилтрансферазы, метилируются разные области ДНК.

Например, Dnmt3A обычно нацелен на метилирование динуклеотидов CpG. Благодаря метилированию ДНК, ДНК-метилтрансферазы могут модифицировать активность сегмента ДНК (например, экспрессию генов) без изменения последовательности ДНК. В некоторых аспектах, метилирование ДНК приводит к репрессии транскрипции гена и/или модуляции чувствительных к метилированию факторов транскрипции или CTCF. Как описано в данном документе, слитые белки могут содержать одну или несколько ДНК-метилтрансфераз (например, две). Если ДНК-метилтрансфераза входит в состав слитого белка, то указанная ДНК-метилтрансфераза может называться «доменом ДНК-метилтрансферазы». В некоторых аспектах, домен ДНК-метилтрансферазы включает одну или несколько ДНК-метилтрансфераз. В некоторых аспектах, домен ДНК-метилтрансферазы включает две ДНК-метилтрансферазы. В некоторых аспектах, домен ДНК-метилтрансферазы дополнительно содержит каталитически неактивный регуляторный фактор ДНК-метилтрансферазы (например, Dnmt3L), который является существенным для функционирования Dnmt1, Dnmt3A и Dnmt3B. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt1. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt3B. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы дополнительно содержит Dnmt3L. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы представляет собой Dnmt3L. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-

метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt3A и Dnmt3L. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы дополнительно содержит регуляторный фактор Dnmt3L, как описано, например, в Siddique et al., Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity, *J. Mol. Biol.* 425, 2013, и Stepper et al., Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase, *Nucleic Acids Res.* 45, 2017.

[0046] Белок «Dnmt3A», «Dnmt3a», «ДНК (цитозин-5)-метилтрансфераза 3A» или «ДНК-метилтрансфераза 3a», как указано в данном документе, включает любые рекомбинантные или встречающиеся в природе формы фермента Dnmt3A или его варианты или гомологи, которые сохраняют активность фермента Dnmt3A (например, в пределах по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Dnmt3A). В

некоторых аспектах, варианты или гомологи имеют по меньшей мере 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности (например, участку из 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с встречающимся в природе белком Dnmt3A. В некоторых аспектах, белок Dnmt3A по существу идентичен белку, идентифицированному ссылочным номером UniProt Q9Y6K1, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, полипептид Dnmt3A кодируется последовательностью нуклеиновой кислоты, идентифицированной номером доступа референсной последовательности NCBI NM_022552, ее гомологами или функциональными фрагментами. В некоторых аспектах, Dnmt3A включает последовательность, представленную SEQ ID NO:26. В некоторых аспектах, Dnmt3A представляет собой последовательность, представленную SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, Dnmt3A представляет собой транскрипт 1 Dnmt3A (Dnmt3A1). В некоторых аспектах, Dnmt3A представляет собой транскрипт 2 Dnmt3A (Dnmt3A2). В некоторых аспектах, Dnmt3A представляет собой транскрипт 3 Dnmt3A (Dnmt3A3). В некоторых аспектах, Dnmt3A представляет собой транскрипт 4 Dnmt3A (Dnmt3A4).

[0047] Белок «Dnmt3L», «ДНК (цитозин-5)-метилтрансфераза 3L» или «ДНК-метилтрансфераза 3L», как указано в данном документе, включает любые рекомбинантные или встречающиеся в природе формы регуляторного фактора Dnmt3L или его вариантов или гомологов, которые сохраняют регуляторную активность Dnmt3L (например, в пределах по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Dnmt3L). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности по сравнению с встречающимся в

аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, полипептид Dnmt3L кодируется последовательностью нуклеиновой кислоты, идентифицированной номером доступа референсной последовательности NCBI NM_001081695, или ее гомологами или функциональными фрагментами. В некоторых аспектах, Dnmt3L включает последовательность, представленную SEQ ID NO:28. В некоторых аспектах, Dnmt3L представляет собой последовательность, представленную SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 50% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 55% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 60% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 65% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 97% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:28.

[0049] Термин «РНК-направляемая ДНК-эндонуклеаза» и т.п., относится, в обычном и общепринятом смысле, к ферменту, который расщепляет фосфодиэфирную связь в полинуклеотидной цепи ДНК, при этом распознавание фосфодиэфирной связи облегчается отдельной последовательностью РНК (например, одиночной направляющей РНК).

[0050] Термин «эндонуклеаза CRISPR класса II» относится к эндонуклеазам, которые обладают сходной с Cas9 эндонуклеазной активностью и участвуют в системе

CRISPR класса II. Примером системы CRISPR класса II является локус CRISPR типа II *Streptococcus pyogenes* SF370, который содержит кластер четырех генов Cas9, Cas1, Cas2 и Csn1, а также два некодирующих РНК-элемента, tracrPНК и характерный набор повторяющихся последовательностей (прямых повторов), перемежающихся короткими участками неповторяющихся последовательностей (спейсерами, около 30 п.о. каждый). Фермент Cpf1 принадлежит к предполагаемой системе CRISPR-Cas типа V. Класс II системы CRISPR-Cas включает системы как типа II, так и типа V.

[0051] «Последовательность ядерной локализации» или «сигнал ядерной локализации» или «NLS» представляет собой пептид, который направляет белки в ядро. В некоторых аспектах, NLS включает пять основных, положительно заряженных аминокислот. NLS может быть расположен в любом месте пептидной цепи. В некоторых аспектах, NLS представляет собой NLS, полученный из SV40. В некоторых аспектах, NLS включает последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS представляет собой последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO: 25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность SEQ ID NO:25.

[0052] В некоторых вариантах осуществления, домен ДНК-метилтрансферазы представляет собой домен Dnmt3A-3L. «Домен Dnmt3A-3L», используемый в данном документе, относится к белку, включающему как Dnmt3A (т.е., ДНК-метилтрансферазу), так и Dnmt3L (т.е., каталитически неактивный регуляторный фактор ДНК-метилтрансферазы). В некоторых аспектах, Dnmt3A и Dnmt3L ковалентно связаны. В некоторых аспектах, Dnmt3A ковалентно связан с Dnmt3L через пептидный линкер. В некоторых аспектах, пептидный линкер включает последовательность, представленную SEQ ID NO:27. В некоторых аспектах, пептидный линкер представляет собой последовательность, представленную SEQ ID NO:27. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:27. В некоторых аспектах,

пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:27. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:27. В некоторых аспектах, домен Dnmt3A-3L включает последовательность, представленную SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L представляет собой последовательность, представленную SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:33. В некоторых аспектах, домен Dnmt3A-3L имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:33.

[0053] Термин «домен Krüppel-ассоциированного бокса» или «KRAB-домен» в данном документе относится к категории доменов транскрипционной репрессии, присутствующих в приблизительно 400 транскрипционных факторах человека на основе белков цинковых пальцев. KRAB-домены типично включают от около 45 до около 75 аминокислотных остатков. Описание KRAB-доменов, включая их функцию и применение, приведено, например, в Ecco, G., Imbeault, M., Trono, D., KRAB zinc finger proteins, *Development* 144, 2017; Lambert et al., *The human transcription factors*, *Cell* 172, 2018; Gilbert et al., *Cell* (2013); и Gilbert et al., *Cell* (2014). В некоторых аспектах, KRAB-домен представляет собой KRAB-домен Koh 1. В некоторых аспектах, KRAB-домен включает последовательность, представленную SEQ ID NO:16. В некоторых аспектах, KRAB-домен представляет собой последовательность SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах,

KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:16. В некоторых аспектах, KRAB-домен включает аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:16.

[0054] Термин «ДНК-метилтрансфераза», используемый в данном документе, относится к ферменту, который катализирует перенос метильной группы к ДНК. Неограничивающие примеры ДНК-метилтрансфераз включают Dnmt1, Dnmt3A и Dnmt3B. В некоторых аспектах, ДНК-метилтрансфераза представляет собой ДНК-метилтрансферазу млекопитающего. В некоторых аспектах, ДНК-метилтрансфераза представляет собой ДНК-метилтрансферазу человека. В некоторых аспектах, ДНК-метилтрансфераза представляет собой мышиную ДНК-метилтрансферазу. В некоторых аспектах, ДНК-метилтрансфераза представляет собой бактериальную цитозинметилтрансферазу и/или бактериальную нецитозинную метилтрансферазу. В зависимости от конкретной ДНК-метилтрансферазы, метилируются разные области ДНК. Например, Dnmt3A обычно нацелен на метилирование динуклеотидов CpG. Благодаря метилированию ДНК, ДНК-метилтрансферазы могут модифицировать активность сегмента ДНК (например, экспрессию генов) без изменения последовательности ДНК. В некоторых аспектах, метилирование ДНК приводит к репрессии транскрипции гена и/или модуляции чувствительных к метилированию факторов транскрипции или CTCF. Как описано в данном документе, слитые белки могут включать одну или несколько ДНК-метилтрансфераз (например, две). Если ДНК-метилтрансфераза входит в состав слитого белка, то указанная ДНК-метилтрансфераза может называться «доменом ДНК-метилтрансферазы». В некоторых аспектах, домен ДНК-метилтрансферазы включает одну или несколько ДНК-метилтрансфераз. В некоторых аспектах, домен ДНК-метилтрансферазы включает две ДНК-метилтрансферазы. В некоторых аспектах, домен ДНК-метилтрансферазы дополнительно содержит каталитически неактивный регуляторный фактор ДНК-метилтрансферазы (например, Dnmt3L), который является существенным для функционирования Dnmt1, Dnmt3A и Dnmt3B. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt1. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt3B. В некоторых аспектах, домен ДНК-метилтрансферазы включает Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:26. В некоторых

последовательности с SEQ ID NO:33. В некоторых аспектах, домен ДНК-метилтрансферазы дополнительно содержит регуляторный фактор Dnmt3L, как описано, например, в Siddique et al., Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity, J. Mol. Biol. 425, 2013, и Stepper et al., Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase, Nucleic Acids Res. 45, 2017.

[0055] Белок «Dnmt3A», «Dnmt3a», «ДНК (цитозин-5)-метилтрансфераза 3A» или «ДНК-метилтрансфераза 3a», как указано в данном документе, включает любые рекомбинантные или встречающиеся в природе формы фермента Dnmt3A или его варианты или гомологи, которые сохраняют активность фермента Dnmt3A (например, в пределах по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Dnmt3A). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 90%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности (например, участку из 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с встречающимся в природе белком Dnmt3A. В некоторых аспектах, белок Dnmt3A по существу идентичен белку, идентифицированному ссылочным номером UniProt Q9Y6K1, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, полипептид Dnmt3A кодируется последовательностью нуклеиновой кислоты, идентифицированной номером доступа референсной последовательности NCBI NM_022552, ее гомологами или функциональными фрагментами. В некоторых аспектах, Dnmt3A включает последовательность, представленную SEQ ID NO:26. В некоторых аспектах, Dnmt3A представляет собой последовательность, представленную SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, домен ДНК-метилтрансферазы имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:26. В некоторых аспектах, Dnmt3A имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:26.

[0056] Белок «Dnmt3L», «ДНК (цитозин-5)-метилтрансфераза 3L» или «ДНК-

метилтрансфераза 3L», как указано в данном документе, включает любые рекомбинантные или встречающиеся в природе формы регуляторного фактора Dnmt3L или его вариантов или гомологов, которые сохраняют регуляторную активность Dnmt3L (например, в пределах по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с Dnmt3L). В некоторых аспектах, варианты или гомологи имеют по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности аминокислотной последовательности по всей последовательности или части последовательности по сравнению с встречающимся в природе белком Dnmt3L. В некоторых аспектах, белок Dnmt3L по существу идентичен белку, идентифицированному ссылочным номером UniProt Q9CWR8, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, белок Dnmt3L идентичен белку, идентифицированному ссылочным номером UniProt Q9CWR8. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9CWR8. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9CWR8. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9CWR8. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9CWR8.

[0057] В некоторых аспектах, белок Dnmt3L по существу идентичен белку, идентифицированному ссылочным номером UniProt Q9UJW, или варианту или гомологу, имеющему существенную степень идентичности с ним. В некоторых аспектах, белок Dnmt3L идентичен белку, идентифицированному ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 50% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 55% идентичности последовательности с белком, идентифицированным ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 60% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 65% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 70% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером

UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, белок Dnmt3L имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного ссылочным номером UniProt Q9UJW. В некоторых аспектах, полипептид Dnmt3L кодируется последовательностью нуклеиновой кислоты, идентифицированной номером доступа референсной последовательности NCBI NM_001081695, или ее гомологами или функциональными фрагментами. В некоторых аспектах, Dnmt3L включает последовательность, представленную SEQ ID NO:28. В некоторых аспектах, Dnmt3L представляет собой последовательность, представленную SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 50% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 55% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 60% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 65% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:28. В некоторых аспектах, Dnmt3L имеет

аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:28.

[0058] Термин «РНК-направляемая ДНК-эндонуклеаза» и т.п., относится, в обычном и общепринятом смысле, к ферменту, который расщепляет фосфодиэфирную связь в полинуклеотидной цепи ДНК, при этом распознавание фосфодиэфирной связи облегчается отдельной последовательностью РНК (например, одиночной направляющей РНК).

[0059] «Последовательность ядерной локализации» или «сигнал ядерной локализации» или «NLS» представляет собой пептид, который направляет белки в ядро. В некоторых аспектах, NLS включает пять основных, положительно заряженных аминокислот. NLS может быть расположен в любом месте пептидной цепи. В некоторых аспектах, NLS представляет собой NLS, полученный из SV40. В некоторых аспектах, NLS включает последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS представляет собой последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO: 25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность SEQ ID NO:25.

[0060] «Направляющая РНК» или «гРНК» в данном документе относится к любой полинуклеотидной последовательности, имеющей достаточную комплементарность с полинуклеотидной последовательностью-мишенью для гибридизации с последовательностью-мишенью и направления последовательность-специфического связывания комплекса CRISPR с последовательностью-мишенью. В некоторых аспектах, степень комплементарности между направляющей последовательностью и ее соответствующей последовательностью-мишенью, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания, составляет около или более чем около 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99%, или более.

[0061] В некоторых вариантах осуществления, полинуклеотид (например, гРНК) представляет собой одноцепочечную рибонуклеиновую кислоту. В некоторых аспектах, полинуклеотид (например, гРНК) имеет длину от около 10 до около 200 остатков

нуклеиновых кислот. В некоторых аспектах, полинуклеотид (например, гРНК) имеет длину от около 50 до около 150 остатков нуклеиновых кислот. В некоторых аспектах, полинуклеотид (например, гРНК) имеет длину от около 80 до около 140 остатков нуклеиновых кислот. В некоторых аспектах, полинуклеотид (например, гРНК) имеет длину от около 90 до около 130 остатков нуклеиновых кислот. В некоторых аспектах, полинуклеотид (например, гРНК) имеет длину от около 100 до около 120 остатков нуклеиновых кислот. В некоторых аспектах, длина полинуклеотида (например, гРНК) составляет около 113 остатков нуклеиновых кислот.

[0062] В общем, направляющая последовательность (т.е., ДНК-нацеливающая последовательность) представляет собой любую полинуклеотидную последовательность, имеющую достаточную комплементарность с полинуклеотидной последовательностью-мишенью для гибридизации с последовательностью-мишенью (например, последовательностью-мишенью геномной или митохондриальной ДНК) и направления последовательность-специфического связывания комплекса (например, комплекса CRISPR) с последовательностью-мишенью. В некоторых аспектах, степень комплементарности между направляющей последовательностью и ее соответствующей последовательностью-мишенью, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания, составляет около или более чем около 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99%, или более. В некоторых аспектах, степень комплементарности между направляющей последовательностью и ее соответствующей последовательностью-мишенью, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания, составляет по меньшей мере около 80%, 85%, 90%, 95% или 100%. В некоторых аспектах, степень комплементарности составляет по меньшей мере 90%. Оптимальное выравнивание может быть определено с помощью любого пригодного алгоритма выравнивания последовательностей, неограничивающие примеры которых включают алгоритм Смита-Уотермана, алгоритм Нидлмана-Вунша, алгоритмы, основанные на преобразовании Барроуза-Уилера (например, программный пакет Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (доступен на сайте soap.genomics.org.cn) и Maq (доступен на сайте maq.sourceforge.net). В некоторых аспектах, направляющая последовательность имеет длину около или более чем около 10, 20, 30, 35, 40, 45, 50, 75 или более нуклеотидов. В некоторых аспектах, направляющая последовательность имеет длину от около 10 до около 50, от около 15 до около 30, или от около 20 до около 25 нуклеотидов. В некоторых аспектах, направляющая последовательность имеет длину менее около 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 или меньше нуклеотидов. В некоторых аспектах, направляющая последовательность имеет длину около 20 нуклеотидов или более. Способность направляющей последовательности к прямому последовательность-специфическому связыванию комплекса (например, комплекса CRISPR) с последовательностью-мишенью можно оценить с помощью любого пригодного анализа. Например, компоненты системы CRISPR, достаточные для образования комплекса

(например, комплекса CRISPR), включая тестируемую направляющую последовательность, могут быть обеспечены в клетке-хозяине, имеющей соответствующую последовательность-мишень, например, путем трансфекции векторами, кодирующими компоненты последовательности CRISPR, с последующей оценкой преимущественного расщепления в пределах последовательности-мишени, например, с помощью известного в данной области техники анализа Surveyor. Аналогично, расщепление полинуклеотидной последовательности-мишени можно оценить экспериментально путем обеспечения последовательности-мишени, компонентов комплекса (например, комплекса CRISPR), включая тестируемую направляющую последовательность и контрольную направляющую последовательность, отличную от тестируемой направляющей последовательности, и сравнения связывания или скорости расщепления последовательности-мишени между реакциями с тестируемой и контрольной направляющими последовательностями. Возможны также другие анализы, как понятно специалистам в данной области техники.

[0063] Термины «sgРНК», «одиночная направляющая РНК» и «одиночная направляющая РНК-последовательность» используются взаимозаменяемо и относятся к полинуклеотидной последовательности, включающей последовательность crРНК и последовательность tracrРНК. Последовательность crРНК включает направляющую последовательность (т.е., «гид» или «спейсер») и парную tracr-последовательность (т.е., «прямой повтор(ы)'). Термин «направляющая последовательность» относится к последовательности, которая определяет сайт-мишень. В некоторых аспектах, две РНК могут кодироваться отдельно в crРНК и tracrРНК как 2 молекулы РНК, которые затем образуют комплекс РНК/РНК путем комплементарного спаривания оснований между crРНК и tracrРНК (т.е., до того, как они станут компетентными для связывания с дефицитным по нуклеазе ферментом РНК-направляемой ДНК-эндонуклеазы). В некоторых аспектах, первая нуклеиновая кислота включает последовательность tracrРНК, и отдельная вторая нуклеиновая кислота включает последовательность гРНК с отсутствующей последовательностью tracrРНК. В некоторых аспектах, первая нуклеиновая кислота, включающая последовательность tracrРНК и вторая нуклеиновая кислота, включающая последовательность гРНК, взаимодействуют друг с другом и, необязательно, входят в состав комплекса (например, комплекса CRISPR). Примеры sgРНК и их последовательностей-мишеней приведены в Таблицах 2, 3 и 4. Примеры sgРНК, описанных в примерах, представлены в Таблице 5.

[0064] Таблица 2

Название	Последовательность-мишень (от 5' к 3')	Последовательность sgРНК (от 5' к 3')
A (JKNg156)	ACTGCGGAAATTTGAGCGT (SEQ ID NO: 37)	ACGCUCAAAUUCCGCAGU (SEQ ID NO: 38)
B (JKNg158)	AGGCAATGGCTGCACATGC	GCAUGUGCAGCCAUUGCCU

Название	Последовательность-мишень (от 5' к 3')	Последовательность sgPHK (от 5' к 3')
	(SEQ ID NO: 39)	(SEQ ID NO: 40)
C (JKNg160)	GACGCTTGGTTCTGAGGAG (SEQ ID NO: 41)	CUCCUCAGAACCAAGCGUC (SEQ ID NO: 42)

[0065] Таблица 3

Название	Последовательность-мишень (от 5' к 3')	Последовательность sgPHK (от 5' к 3')
CD29, sgPHK-A	TCCGGAAACGCATTCCTCT (SEQ ID NO: 43)	AGAGGAAUGCGUUUCCGGA (SEQ ID NO: 44)
CD29, sgPHK-B	CCGCGTCAGCCCGGCCCGG (SEQ ID NO: 45)	CCGGGCCGGGCGUGACGCGG (SEQ ID NO: 46)
CD29, sgPHK-C	CGACTCCCGCTGGGCCTCT (SEQ ID NO: 47)	AGAGGCCAGCGGGAGUCG (SEQ ID NO: 48)
CD81, sgPHK-A	CCGTTGCGCGCTCGCTCTC (SEQ ID NO: 49)	GAGAGCGAGCGCGCAACGG (SEQ ID NO: 50)
CD81, sgPHK-B	CCGCGCATCCTGCCAGGCC (SEQ ID NO: 51)	GGCCUGGCAGGAUGCGCGG (SEQ ID NO: 52)
CD81, sgPHK-C	CCAACCTGGCGCGTTTCGG (SEQ ID NO: 53)	CCGAAACGCGCCAAGUUGG (SEQ ID NO: 54)
CD151, sgPHK-A	ACCACGCGTCCGAGTCCGG (SEQ ID NO: 55)	CCGGACUCGGACGCGUGGU (SEQ ID NO: 56)
CD151, sgPHK-B	TGCTCATTGTCCCTGGACA (SEQ ID NO: 57)	UGUCCAGGGACAAUGAGCA (SEQ ID NO: 58)
CD151, sgPHK-C	GGACACCCTGCTCATTGTC (SEQ ID NO: 59)	GACAAUGAGCAGGGUGUCC (SEQ ID NO: 60)

[0066] Таблица 4

Название	Последовательность-мишень (от 5' к 3')	Последовательность sgPHK (от 5' к 3')
Pcsk9 sgPHK-1	TCCGGAAACGCATTCCTCT (SEQ ID NO: 43)	AGAGGAAUGCGUUUCCGGA (SEQ ID NO: 44)
Pcsk9 sgPHK-2	ACCGGCAGCCTGCGCGTCC (SEQ ID NO: 61)	GGACGCGCAGGCUGCCGGU (SEQ ID NO: 62)
Pcsk9 sgPHK-3	CGATGGGCACCCACTGCTC	GAGCAGUGGGUGCCCAUCG

	(SEQ ID NO: 63)	(SEQ ID NO: 64)
Pcsk9 sgPHK-4	CCTTCACGTGGACGCGCAG (SEQ ID NO: 65)	CUGC GCGUCCACGUGAAGG (SEQ ID NO: 66)
Pcsk9 sgPHK-5	CGTGAAGGTGGAAGCCTTC (SEQ ID NO: 67)	GAAGGCUUCCACCUUCACG (SEQ ID NO: 68)
Npc1 sgPHK-1	CTCCTTGGTCAGGCGCCGG (SEQ ID NO: 69)	CCGGCGCCUGACCAAGGAG (SEQ ID NO: 70)
Npc1 sgPHK-2	TGGTCAGGCGCCGGTTCCG (SEQ ID NO: 71)	CGGAACCGGCGCCUGACCA (SEQ ID NO: 72)
Npc1 sgPHK-3	TAGAGGTCGCCTTCTCCTC (SEQ ID NO: 73)	GAGGAGAAGGCGACCUCUA (SEQ ID NO: 74)
Npc1 sgPHK-4	CGACGCTCGGGTCGCGGTG (SEQ ID NO:75)	CACCGCGACCCGAGCGUCG (SEQ ID NO:76)
Npc1 sgPHK-5	ATGCTGTGCCGCGCGGGG (SEQ ID NO:77)	CCCCGCGCGGCGACAGCAU (SEQ ID NO:78)
Spes1 sgPHK-1	CTCACCTCACCGGAGCCA (SEQ ID NO:79)	UGGCUCCGGUGAGGGUGAG (SEQ ID NO:80)
Spes1 sgPHK-2	CCGCAAACCTTACTCCTTA (SEQ ID NO:81)	UAAGGAGUAAAGUUUGCGG (SEQ ID NO:82)
Spes1 sgPHK-3	CTCGGAGACATCCGCTTCC (SEQ ID NO: 60)	GGAAGCGGAUGUCUCCGAG (SEQ ID NO: 60)
Spes1 sgPHK-4	CTCCTAAGATTGGCTTAC (SEQ ID NO:83)	GUGAAGCCAAUCUUAGGAG (SEQ ID NO:84)
Spes1 sgPHK-5	CCGGAGCCACTCCTAAGAT (SEQ ID NO:85)	AUCUUAGGAGUGGCUCCGG (SEQ ID NO:86)
Cd81 sgPHK-1	TTCTCTACCCTACGTCTCA (SEQ ID NO:87)	UGAGACGUAGGGUAGAGAA (SEQ ID NO:88)
Cd81 sgPHK-2	TACGTCTCATTCTCCGCAA (SEQ ID NO:89)	UUGC GGAGAAUGAGACGUA (SEQ ID NO:90)
Cd81 sgPHK-3	GCTAGGCCTCCAGCCCTTC (SEQ ID NO:91)	GAAGGGCUGGAGGCCUAGC (SEQ ID NO:92)
Cd81 sgPHK-4	ACAGGTGGCGCCGCAACTT (SEQ ID NO:93)	AAGUUGC GGCGCCACCUGU (SEQ ID NO:94)
Cd81 sgPHK-5	AGCCGGAGGCGCGAGAGTC	GACUCUCGCGCCUCCGGCU

	(SEQ ID NO:95)	(SEQ ID NO:96)
--	----------------	----------------

[0067] Таблица 5

Название	Протоспейсер (от 5' к 3')	SEQ ID NO:
Промотор Snrpn	CTCCTCAGAACCAAGCGTC	SEQ ID NO:127
H2B sgPHK-a	GTAAGACACAGTACAAACG	SEQ ID NO:128
H2B sgPHK-b	GAACCGGCAAAATCCGCTC	SEQ ID NO:129
H2B sgPHK-c	GGGCCGGAGCGGATTTTGC	SEQ ID NO:130
CLTA sgPHK-a	GAAGCCCTACCCGTGTATC	SEQ ID NO:131
CLTA sgPHK-b	CTTGTCCTCTCTCCCAGT	SEQ ID NO:132
CLTA sgPHK-c	CTCCCAGTCGGCACCACAG	SEQ ID NO:133
ITGB1 sgPHK-a	AGAGGAATGCGTTTCCGGA	SEQ ID NO:134
ITGB1 sgPHK-b	CCGGGCCGGGCTGACGCGG	SEQ ID NO:135
ITGB1 sgPHK-c	AGAGGCCAGCGGGAGTCG	SEQ ID NO:136
CD81 sgPHK-a	GAGAGCGAGCGCGCAACGG	SEQ ID NO:137
CD81 sgPHK-b	GGCCTGGCAGGATGCGCGG	SEQ ID NO:138
CD81 sgPHK-c	CCGAAACGCGCCAAGTTGG	SEQ ID NO:139
CD151 sgPHK-a	CCGACTCGGACGCGTGGT	SEQ ID NO:140
CD151 sgPHK-b	TGTCCAGGGACAATGAGCA	SEQ ID NO:141
CD151 sgPHK-c	GACAATGAGCAGGGTGTCC	SEQ ID NO:142
CALD1 sgPHK-a	TTGGCTGGGGTGTCTCAT	SEQ ID NO:143
CALD1 sgPHK-b	TCTGGGAGGGGCCAAGGAA	SEQ ID NO:144
CALD1 sgPHK-c	GAGATGCCTGTCATTCCCT	SEQ ID NO:145
CALD1 sgPHK-d	CCTCCGACTGTAAACATA	SEQ ID NO:146
CALD1 sgPHK-e	AGGCAGGCTGCATCCACCT	SEQ ID NO:147
DYNC2LI1 sgPHK-a	GCCGAGAATGAGATGTAAA	SEQ ID NO:148
DYNC2LI1 sgPHK-b	CGCGGTTGCGTGGGGAGAC	SEQ ID NO:149
DYNC2LI1 sgPHK-c	GTGGGGAGACGGGCATCAT	SEQ ID NO:150
DYNC2LI1 sgPHK-d	GTCGCGGCCGCGGTTGCGT	SEQ ID NO:151
DYNC2LI1 sgPHK-e	GGGAGCGACGTCGCGGCCG	SEQ ID NO:152
LAMP2 sgPHK-a	TCTCTCAGGAGCATAGGAA	SEQ ID NO:153
LAMP2 sgPHK-b	GAGAGGGGTGGTGATGTAG	SEQ ID NO:154
LAMP2 sgPHK-c	CCGTCCGCGCATATCTCTC	SEQ ID NO:155
LAMP2 sgPHK-d	CGACCACGCCCTGGCTTTT	SEQ ID NO:156
LAMP2 sgPHK-e	CGCTTGAAAGCGGCGAGAG	SEQ ID NO:157
MYL6 sgPHK-a	CGGAGGTGGGGGGGGTCTGT	SEQ ID NO:158
MYL6 sgPHK-b	GGGCACAGAACCCGCTCGG	SEQ ID NO:159
MYL6 sgPHK-c	ACCCCCCCCACCTCCGAGC	SEQ ID NO:160
MYL6 sgPHK-d	AACCCGCTCGGAGGTGGGG	SEQ ID NO:161
MYL6 sgPHK-e	CACAGAACCCGCTCGGAGG	SEQ ID NO:162
VPS25 sgPHK-a	TAGGATAACAGCGAGGGAC	SEQ ID NO:163
VPS25 sgPHK-b	TGGAAGAGAGAGACCGGC	SEQ ID NO:164
VPS25 sgPHK-c	AAGGCCGGAAGTCCTTAGC	SEQ ID NO:165
VPS25 sgPHK-d	CCAAAAGATAGGGCTTTTT	SEQ ID NO:166
VPS25 sgPHK-e	AGGCAATGCAGAAAGGTC	SEQ ID NO:167
PVT1 E1-a	GAGGTCATGATTCATCCCAC	SEQ ID NO:168
PVT1 E1-b	GGCTACTGGTCTCCACCGGC	SEQ ID NO:169
PVT1 E1-c	GAAGCTCGCTTACAGGAGGG	SEQ ID NO:170
PVT1 E2-a	GCTGCTGAGCGATGACTCAG	SEQ ID NO:171

PVT1 E2-b		GTCCCTGCGCCAGGGTAAAC	SEQ ID NO:172
PVT1 E2-c		GCAATCTCGGTCCGCCTTAC	SEQ ID NO:173
PVT1 E3-a		GCCGGCTGTTGAGTCAGCTC	SEQ ID NO:174
PVT1 E3-b		GCTGGATGACACCACTCTGT	SEQ ID NO:175
PVT1 E3-c		GCGGTGGAGACCTCATGATC	SEQ ID NO:176
PVT1 E4-a		GCAGCCTCTTGTTCGCTGCT	SEQ ID NO:177
PVT1 E4-b		GTTAAGATAGAGAATGCCAT	SEQ ID NO:178
PVT1 E4-c		GTTCCGTACACAGTCCTTTC	SEQ ID NO:179
Промотор PVT1		GCCTCCGGGCAGAGCGCGTG	SEQ ID NO:180
Лямбда контроль-1	негативный	GATCGTCGGTCGGGTCATACG	SEQ ID NO:181
Лямбда контроль-2	негативный	GTTCGTCGCGATAGATGATCG	SEQ ID NO:182
Лямбда контроль-3	негативный	GCGCGTATCGTTTCATCGGCG	SEQ ID NO:183

[0068] Последовательности в Таблицах представляют собой нацеливающие последовательности sgPНК. В качестве примера, полная одиночная направляющая PНК (sgPНК) для SEQ ID NO:38 представляет собой: GACGCUCAAAUUUCCGCAGUGUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGU UUAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUU UUUUU (SEQ ID NO:100). Общая tracr-последовательность для каждого одиночного гида Sp Cas9 представляет собой GUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGUUUUAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:101). Квалифицированному специалисту будет понятно, что последовательности sgPНК в Таблицах состоят из 19 пар оснований и не отражают того, что каждая sgPНК начинается с G, который необходим при экспрессии с промотора pol-III для инициации транскрипции. Таким образом, для SEQ ID NO:38, последовательность будет представлять собой GACGCUCAAAUUUCCGCAGU (SEQ ID NO:102), а не ACGCUCAAUUUCCGCAGU (SEQ ID NO:38). В некоторых вариантах осуществления, SEQ ID NO:38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94 и 96 содержат каждая G в качестве первого нуклеотида. В некоторых вариантах осуществления, SEQ ID NO:127-183 содержат каждая G в качестве первого нуклеотида.

[0069] В общем, парная к tracr последовательность включает любую последовательность, которая обладает достаточной комплементарностью с последовательностью tracrPНК, чтобы способствовать чему-либо одному или нескольким из: (1) эксцизии направляющей последовательности, фланкированной парными к tracr последовательностями в клетке, содержащей соответствующую последовательность tracr; и (2) образования комплекса (например, комплекса CRISPR) на последовательности-мишени, причем комплекс (например, комплекс CRISPR) содержит парную к tracr последовательность, гибридованную с последовательностью tracr. В общем, степень комплементарности относится к оптимальному выравниванию парной к tracr последовательности и последовательности tracrPНК, по длине более короткой из двух

последовательностей. Оптимальное выравнивание может быть определено с помощью любого пригодного алгоритма выравнивания, и может дополнительно учитывать вторичные структуры, такие как самокомплементарность в пределах либо последовательности *tracrRNA*, либо парной к *tracr* последовательности. В некоторых аспектах, степень комплементарности между последовательностью *tracrRNA* и парной к *tracr* последовательностью по длине более короткой из двух указанных последовательностей при оптимальном выравнивании, составляет около или более чем около 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99%, или больше. В некоторых аспектах, степень комплементарности составляет около или по меньшей мере около 80%, 90%, 95% или 100%. В некоторых аспектах, последовательность *tracrRNA* имеет длину около или более чем около 5, 10, 15, 20, 30, 40, 50 или более нуклеотидов. В некоторых аспектах, последовательность *tracrRNA* и парная к *tracr* последовательность содержатся в одном транскрипте, так что при гибридизации между ними образуется транскрипт, имеющий вторичную структуру, например, шпильку.

[0070] «Нуклеиновая кислота» относится к нуклеотидам (например, дезоксирибонуклеотидам или рибонуклеотидам) и их полимерам в одно-, двух- или многоцепочечной форме, или их компонентам. Термины «полинуклеотид», «олигонуклеотид», «олигомер» и т.п. относятся, в обычном и общепринятом смысле, к линейной последовательности нуклеотидов. Термин «нуклеотид» относится, в обычном и общепринятом смысле, к единичному звену полинуклеотида, *m.e.*, мономеру. Нуклеотиды могут быть рибонуклеотидами, дезоксирибонуклеотидами, или их модифицированными вариантами. Примеры полинуклеотидов, предусматриваемых в данном документе, включают одно- и двухцепочечную ДНК, одно- и двухцепочечную РНК, и гибридные молекулы, содержащие смеси одно- и двухцепочечных ДНК и РНК. Примеры нуклеиновых кислот, например, полинуклеотидов, предусматриваемых в данном документе, включают, без ограничений, любой тип РНК, например, мРНК, миРНК, микроРНК, sgРНК и направляющую РНК, и любой тип ДНК, геномную ДНК, плазмидную ДНК и миникольцевую ДНК, и любые их фрагменты. В некоторых аспектах, нуклеиновая кислота представляет собой матричную РНК. В некоторых аспектах, матричная РНК представляет собой матричный рибонуклеопротеин (РНП). Термин «дуплекс» в контексте полинуклеотидов относится, в обычном и общепринятом смысле, к двухцепочечности. Нуклеиновые кислоты могут быть линейными или разветвленными. Например, нуклеиновые кислоты могут быть линейной цепью нуклеотидов или нуклеиновые кислоты могут быть разветвленными, например, так чтобы нуклеиновые кислоты содержали одну или несколько цепочек или ответвлений нуклеотидов. Необязательно, разветвленные нуклеиновые кислоты имеют повторяющиеся разветвления с образованием структур с более высокой степенью упорядоченности, таких как дендримеры и т.п.

[0071] В данном документе, термины «нуклеиновая кислота», «молекула нуклеиновой кислоты», «олигомер нуклеиновой кислоты», «олигонуклеотид», «последовательность нуклеиновой кислоты», «фрагмент нуклеиновой кислоты» и

«полинуклеотид» могут использоваться взаимозаменяемо и должны включать, без ограничений, полимерную форму нуклеотидов, ковалентно связанных друг с другом, которые могут иметь различную длину, представлять собой дезоксирибонуклеотиды или рибонуклеотиды или их аналоги, производные или модификации. Разные полинуклеотиды могут иметь разные трехмерные структуры, и могут выполнять различные функции, известные или неизвестные. Неограничивающие примеры полинуклеотидов включают ген, фрагмент гена, экзон, интрон, межгенную ДНК (включая, без ограничений, гетерохроматическую ДНК), матричную РНК (мРНК), транспортную РНК, рибосомальную РНК, рибозим, кДНК, рекомбинантный полинуклеотид, разветвленный полинуклеотид, плазмиду, вектор, выделенную ДНК последовательности, выделенную РНК последовательности, sgРНК, направляющую РНК, зонд нуклеиновой кислоты и праймер. Полинуклеотиды, пригодные для использования в способах настоящего изобретения, могут включать природные последовательности нуклеиновых кислот и их варианты, искусственные последовательности нуклеиновых кислот, или комбинацию таких последовательностей.

[0072] Полинуклеотид обычно состоит из определенной последовательности четырех нуклеотидных оснований: аденина (A), цитозина (C), гуанина (G) и тимина (T) (урацила (U) вместо тимина (T), если полинуклеотид представляет собой РНК). Таким образом, термин «полинуклеотидная последовательность» является алфавитным представлением полинуклеотидной молекулы; альтернативно, указанный термин может быть применен по отношению к самой полинуклеотидной молекуле. Это алфавитное представление может быть введено в базы данных в компьютере с центральным процессором и использовано в приложениях биоинформатики, таких как функциональная геномика и поиск гомологии. Полинуклеотиды могут необязательно включать один или несколько нестандартных нуклеотидов, нуклеотидных аналогов и/или модифицированных нуклеотидов.

[0073] Нуклеиновые кислоты, включая, например, нуклеиновые кислоты с фосфоротиоатным остовом, могут включать один или несколько реакционноспособных фрагментов. Используемый в данном документе термин реакционноспособный фрагмент включает любую группу, способную реагировать с другой молекулой, например, нуклеиновой кислотой или полипептидом посредством ковалентных, нековалентных или других взаимодействий. В качестве примера, нуклеиновая кислота может включать аминокислотный реакционноспособный фрагмент, который реагирует с аминокислотой белка или полипептида посредством ковалентного, нековалентного или другого взаимодействия.

[0074] Указанные термины также охватывают нуклеиновые кислоты, содержащие известные нуклеотидные аналоги или модифицированные остатки остова или связи, которые являются синтетическими, встречающимися в природе, и не встречающимися в природе, которые обладают свойствами связывания, аналогичными референсной нуклеиновой кислоте, и которые метаболизируются аналогично референсным

нуклеотидам. Примеры таких аналогов включают, без ограничения, фосфодиэфирные производные, включая, *например*, фосфорамидатные, фосфородиамидатные, фосфоротиоатные (также известные как фосфоротиоатные с серой с двойной связью, замещающей кислород в фосфате), фосфородитиоатные, фосфонокарбоновые кислоты, фосфонокарбоксилаты, фосфоноксусную кислоту, фосфомуравьиную кислоту, метилфосфонат, борфосфонат или O-метилфосфороамидитные связи (см. Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press) а также модификации нуклеотидных оснований, такие как в 5-метилцитидин или псевдоуридин; и остовы и связи пептидных нуклеиновых кислот. Другие аналоги нуклеиновых кислот включают имеющие положительно заряженные остовы; неионные остовы, модифицированные сахара и нерибозные остовы (например, фосфородиамидатморфолиновые олигомеры или замкнутые нуклеиновые кислоты (ЗНК), как известно в данной области техники), включая описанные в патентах США №№ 5235033 и 5034506, и в главах 6 и 7, ASC Symposium Series 580, *Carbohydrate Modifications in Antisense Research*, Sanghui & Cook, eds. Нуклеиновые кислоты, содержащие один или несколько карбоциклических сахаров, также включены в одно из определений нуклеиновых кислот. Модификации рибозофосфатного остова могут быть введены по различным причинам, *например*, для увеличения стабильности и периода полувыведения таких молекул в физиологических средах или при использовании в качестве зондов на биочипе. Могут быть приготовлены смеси встречающихся в природе нуклеиновых кислот и аналогов; альтернативно, могут быть приготовлены смеси разных аналогов нуклеиновых кислот и смеси встречающихся в природе нуклеиновых кислот и аналогов. В некоторых аспектах, межнуклеотидные связи в ДНК являются фосфодиэфирными, производными фосфодиэфирных, или комбинацией обоих указанных видов.

[0075] Нуклеиновые кислоты могут включать неспецифические последовательности. Используемый в данном документе термин «неспецифическая последовательность» относится к последовательности нуклеиновой кислоты, содержащей ряд остатков, которые не должны быть комплементарными, или являются только частично комплементарными, к любой другой последовательности нуклеиновой кислоты. В качестве примера, неспецифическая последовательность нуклеиновой кислоты представляет собой последовательность остатков нуклеиновых кислот, которая не функционирует как ингибирующая нуклеиновая кислота при контакте с клеткой или организмом.

[0076] Термин «комплементарный» или «комплементарность» относится к способности нуклеиновой кислоты образовывать водородную связь (связи) с другой последовательностью нуклеиновой кислоты либо по традиционному уотсон-криковскому, либо по другим нетрадиционным типам. Например, последовательность А-Г-Т является комплементарной к последовательности Т-С-А. Процент комплементарности указывает процентную долю остатков в молекуле нуклеиновой кислоты, которые могут образовывать водородные связи (например, уотсон-криковское спаривание оснований) со

второй последовательностью нуклеиновой кислоты (например, 5, 6, 7, 8, 9, 10 из 10 составляют 50%, 60%, 70%, 80%, 90% и 100% комплементарности, соответственно). «Идеально комплементарный» означает, что все смежные остатки последовательности нуклеиновой кислоты будут образовывать водородные связи с таким же числом смежных остатков во второй последовательности нуклеиновой кислоты. «Существенно комплементарный», при использовании в данном документе, относится к степени комплементарности, составляющей по меньшей мере 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% или 100% в области, состоящей из 10, 15, 20, 25, 30, 35, 40, 45, 50 или более нуклеотидов, или относится к двум нуклеиновым кислотам, которые гибридизуются в жестких условиях (т.е., в жестких условиях гибридизации).

[0077] Выражение «жесткие условия гибридизации» относится к условиям, в которых зонд будет гибридизоваться со своей субпоследовательностью-мишенью, обычно в сложной смеси нуклеиновых кислот, но не с другими последовательностями. Жесткие условия зависят от последовательности и будут различаться в разных обстоятельствах. Более длинные последовательности гибридизуются специфично при более высоких температурах. Обширное руководство по гибридизации нуклеиновых кислот приведено в Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, «Overview of principles of hybridization and the strategy of nucleic acid assays» (1993). Обычно жесткие условия выбирают так, чтобы они соответствовали температуре на 5-10°C ниже точки плавления (T_m) для конкретной последовательности при определенных значениях ионной силы, pH. T_m может представлять собой температуру (при определенных ионной силе, pH и концентрации нуклеиновой кислоты), при которой 50% зондов, комплементарных с мишенью, гибридизуются с последовательностью-мишенью в равновесном состоянии (поскольку последовательности-мишени присутствуют в избытке, при равновесии при T_m будет занято 50% зондов). Жесткие условия также можно обеспечивать добавлением дестабилизирующих агентов, таких как формамид. Для селективной или специфической гибридизации положительный сигнал превышает фоновую гибридизацию по крайней мере в два раза, предпочтительно в 10 раз. Примерные жесткие условия гибридизации могут быть следующими: 50% формамида, 5x SSC и 1% ДСН, инкубация при 42 °C, или 5x SSC (стандартный солевой раствор), 1% ДСН, инкубация при 65°C с промывкой в 0,2x SSC и 0,1% ДСН при 65 °C.

[0078] Нуклеиновые кислоты, которые не гибридизуются друг с другом в жестких условиях, все равно являются существенно идентичными, если кодируемые ими полипептиды являются существенно идентичными. Это происходит, например, когда копия нуклеиновой кислоты создается с использованием максимального вырождения кодонов, допускаемого генетическим кодом. В таких случаях, нуклеиновые кислоты обычно гибридизуются в умеренно жестких условиях гибридизации. Пример «умеренно жестких условий гибридизации» включает гибридизацию в буфере, содержащем 40% формамида, 1 M NaCl, 1% ДСН при 37 °C, и промывку в 1X SSC при 45 °C. Положительная гибридизация по меньшей мере в два раза превышает фоновую.

Специалисту в данной области техники будет понятно, что можно использовать альтернативные условия гибридизации и промывки для обеспечения условий аналогичной жесткости. Дополнительные указания по определению параметров гибридизации приведены в многочисленных источниках, например, *Current Protocols in Molecular Biology*, ed. Ausubel, et al., supra.

[0079] Термин «ген» означает сегмент ДНК, участвующий в производстве белка; он включает области, предшествующие и следующие за кодирующей областью (лидерную и трейлер), а также промежуточные последовательности (интроны) между отдельными кодирующими сегментами (экзонами). Лидерная последовательность, трейлер, а также интроны включают регуляторные элементы, необходимые во время транскрипции и трансляции гена. Кроме того, «белковый генный продукт» представляет собой белок, экспрессируемый из определенного гена.

[0080] Слова «экспрессия» или «экспрессируемый», используемые в данном документе по отношению к гену, означают продукт транскрипции и/или трансляции этого гена. Уровень экспрессии молекулы ДНК в клетке может быть определен на основе либо количества соответствующей мРНК, которая присутствует в клетке, либо количества вырабатываемого клеткой белка, кодируемого этой ДНК. Уровень экспрессии некодирующих молекул нуклеиновой кислота (например, sgРНК) может быть определен стандартными методами ПЦР или нозерн-блоттинга, хорошо известными в данной области техники. См. Sambrook et al., 1989 *Molecular Cloning: A Laboratory Manual*, 18.1-18.88.

[0081] Термин «транскрипционная регуляторная последовательность», используемый в данном документе, относится к сегменту ДНК, способному увеличивать или уменьшать транскрипцию (например, экспрессию) конкретного гена в организме. Неограничивающие примеры транскрипционных регуляторных последовательностей включают промоторы, энхансеры и сайленсеры.

[0082] Термины «сайт начала транскрипции» и «сайт инициации транскрипции» могут использоваться взаимозаменяемо в данном документе для обозначения 5'-конца последовательности гена (например, последовательности ДНК), где РНК-полимераза (например, ДНК-направленная РНК-полимераза) начинает синтез транскрипта РНК. Сайт начала транскрипции может быть первым нуклеотидом транскрибируемой ДНК-последовательности, где РНК-полимераза начинает синтез транскрипта РНК. Квалифицированный специалист может определить сайт начала транскрипции с помощью рутинных экспериментов и анализа, например, путем проведения анализа прерванной транскрипции или путем определения по базе данных FANTOM5.

[0083] Термин «промотор», используемый в данном документе, относится к области ДНК, которая иницирует транскрипцию конкретного гена. Промоторы обычно расположены поблизости от сайта начала транскрипции гена, перед геном и в той же цепи (т.е., 5' в смысловой цепи) ДНК. Промоторы могут иметь длину от около 100 до около 1000 пар оснований.

[0084] Термин «энхансер», используемый в данном документе, относится к области ДНК, которая может связываться белками (например, факторами транскрипции) для увеличения вероятности транскрипции гена. Энхансеры могут иметь длину от около 50 до около 1500 пар оснований. Энхансеры могут быть расположены перед или после сайта инициации транскрипции, который они регулируют, и могут быть расположены на расстоянии нескольких сотен пар оснований, нескольких тысяч пар оснований, или нескольких миллионов пар оснований от сайта инициации транскрипции.

[0085] Термин «сайленсер», используемый в данном документе, относится к последовательности ДНК, способной связывать факторы регуляции транскрипции, известные как репрессоры, тем самым негативно влияя на транскрипцию гена. ДНК-последовательности сайленсеров могут присутствовать во многих различных положениях в ДНК, включая, без ограничений, перед геном-мишенью, на который они воздействуют для подавления транскрипции гена (например, сайленсинга экспрессии гена).

[0086] Термин «аминокислота» относится к встречающимся в природе и синтетическим аминокислотам, а также к аминокислотным аналогам и аминокислотным миметикам, которые функционируют аналогично природным аминокислотам. Встречающиеся в природе аминокислоты представляют собой аминокислоты, кодируемые генетическим кодом, а также аминокислоты, которые были впоследствии модифицированы, *например*, гидроксипролин, γ -карбоксиглутамат и O-фосфосерин. Аминокислотные аналоги относятся к соединениям, которые имеют такую же основную химическую структуру, что и встречающаяся в природе аминокислота, то есть α -углерод, связанный с водородом, карбоксильную группу, аминогруппу и группу R, например, гомосерин, норлейцин, метионинсульфоксид, метионинметилсульфоний. Такие аналоги имеют модифицированные группы R (*например*, норлейцин) или модифицированные пептидные основные остовы, но сохраняют ту же основную химическую структуру, что и у встречающейся в природе аминокислоты. Аминокислотные миметики относятся к химическим соединениям, которые имеют структуру, отличную от общей химической структуры аминокислоты, но функционируют подобно встречающейся в природе аминокислоте. Термины «не встречающаяся в природе аминокислота» и «неприродная аминокислота» относятся к аминокислотным аналогам, синтетическим аминокислотам и аминокислотным миметикам, которые не встречаются в природе.

[0087] Аминокислоты могут обозначаться в данном документе либо общеизвестными трехбуквенными символами, либо однобуквенными символами, рекомендованными Комиссией по биохимической номенклатуре IUPAC-IUB. Нуклеотиды также могут обозначаться их общепринятыми однобуквенными кодами.

[0088] Термины «полипептид», «пептид» и «белок» используются в данном документе взаимозаменяемо по отношению к полимеру аминокислотных остатков, причем указанный полимер может, в некоторых аспектах, быть конъюгирован с фрагментом, не состоящим из аминокислот. Эти термины применимы к аминокислотным полимерам, в которых один или несколько аминокислотных остатков являются

искусственным химическим миметиком соответствующей встречающейся в природе аминокислоты, а также к полимерам встречающихся в природе аминокислот и полимерам не встречающихся в природе аминокислот. «Слитый белок» относится к химерному белку, кодирующему две или более отдельных белковых последовательностей, которые рекомбинантно экспрессируются в виде единого фрагмента.

[0089] «Консервативно модифицированные варианты» относятся как к аминокислотным последовательностям, так и к последовательностям нуклеиновых кислот. Что касается конкретных последовательностей нуклеиновых кислот, «консервативно модифицированные варианты» относятся к тем нуклеиновым кислотам, которые кодируют идентичные или по существу идентичные аминокислотные последовательности. Из-за вырожденности генетического кода ряд последовательностей нуклеиновых кислот будет кодировать любой данный белок. Например, кодоны GCA, GCC, GCG и GCU все кодируют аминокислоту аланин. Таким образом, в каждом положении, в котором аланин определяется кодоном, указанный кодон может быть изменен на любой из соответствующих описанных кодонов без изменения кодируемого полипептида. Такие варианты нуклеиновых кислот являются «молчащими вариантами», которые представляют собой один из видов консервативно модифицированных вариантов. Каждая последовательность нуклеиновой кислоты в данном документе, которая кодирует полипептид, также охватывает все возможные молчащие варианты нуклеиновой кислоты. Специалисту в данной области понятно, что каждый кодон в нуклеиновой кислоте (за исключением AUG, который обычно является единственным кодоном для метионина, и TGG, который обычно является единственным кодоном для триптофана), может быть модифицирован для получения функционально идентичной молекулы. Соответственно, каждая описанная последовательность подразумевает каждый молчащий вариант нуклеиновой кислоты, кодирующей полипептид.

[0090] Что касается аминокислотных последовательностей, специалисту в данной области техники должно быть понятно, что отдельные замены, делеции или добавления в последовательности нуклеиновой кислоты, пептида, полипептида или белка, которые изменяют, добавляют или удаляют отдельную аминокислоту или небольшой процент аминокислот в кодированной последовательности, представляют собой «консервативно модифицированный вариант», в котором изменение приводит к замене аминокислоты на химически схожую аминокислоту. Таблицы консервативных замен, описывающие функционально схожие аминокислоты, хорошо известны в данной области техники. Такие консервативно модифицированные варианты являются дополнительными и не исключают полиморфные варианты, межвидовые гомологи и аллели по настоящему изобретению. Каждая из следующих восьми групп содержит аминокислоты, которые являются консервативными заменами друг для друга: (1) аланин (A), глицин (G); (2) аспарагиновая кислота (D), глутаминовая кислота (E); (3) аспарагин (N), глутамин (Q); (4) аргинин (R), лизин (K); (5) изолейцин (I), лейцин (L), метионин (M), валин (V); (6) фенилаланин (F), тирозин (Y), триптофан (W); (7) серин (S), треонин (T); и (8) цистеин (C), метионин (M)

(см., например, Creighton, Proteins (1984)).

[0091] «Процент идентичности последовательности» определяется путем сравнения двух оптимально выравненных последовательностей в окне сравнения, причем часть последовательности полинуклеотида или полипептида в окне сравнения может содержать добавления или делеции (*m.e.*, гэпы) по сравнению с референсной последовательностью (которая не содержит добавлений или делеций) для оптимального выравнивания двух последовательностей. Процентную величину рассчитывают путем определения числа позиций, в которых в обеих последовательностях находятся идентичные основания нуклеиновой кислоты или аминокислотные остатки, для получения числа совпадающих позиций, деления числа совпадающих позиций на общее число позиций в окне сравнения, и умножения результата на 100 для получения процента идентичности последовательностей.

[0092] Термины «идентичный» или процент «идентичности», в контексте двух или более последовательностей нуклеиновых кислот или полипептидов, относятся к двум или более последовательностям или субпоследовательностям, которые являются одинаковыми или имеют определенный процент аминокислотных остатков или нуклеотидов, являющихся одинаковыми (т.е., около 60% идентичности, предпочтительно 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, или больше, идентичности в указанной области, при сравнении и выравнивании с максимальным соответствием в окне сравнения или указанной области), измеренный с использованием алгоритмов сравнения последовательностей BLAST или BLAST 2.0 с параметрами по умолчанию, описанными ниже, или путем выравнивания вручную и визуальной оценки (см., например, веб-сайт Национального центра биотехнологической информации (NCBI) - ncbi.nlm.nih.gov/BLAST/ и т.п.). В таком случае говорят, что такие последовательности «по существу идентичны». Это определение также относится к, или может быть применено к, комплементу тестируемой последовательности. Указанное определение также включает последовательности, имеющие делеции и/или дополнения, а также имеющие замены. Как описано ниже, предпочтительные алгоритмы могут учитывать гэпы и т.п. Предпочтительно, идентичность существует на участке, длина которого составляет по меньшей мере около 25 аминокислот или нуклеотидов, или, более предпочтительно, на участке, длина которого составляет 50-100 аминокислот или нуклеотидов.

[0093] «Позиция» аминокислоты или нуклеотидного основания обозначается номером, который последовательно идентифицирует каждую аминокислоту (или нуклеотидное основание) в референсной последовательности на основе ее положения относительно N-конца (или 5'-конца). Вследствие делеций, вставок, усечений, слияний и т.п., которые должны учитываться при определении оптимального выравнивания, в общем, номер аминокислотного остатка в тестируемой последовательности, определенный простым отсчетом от N-конца, не обязательно будет совпадать с номером его соответствующей позиции в референсной последовательности. Например, в случае, когда вариант имеет делецию по сравнению с выравненной референсной

последовательностью, в варианте будет отсутствовать аминокислота, соответствующая позиции сайта делеции в референсной последовательности. При наличии вставки в выравненной референсной последовательности, эта вставка не будет соответствовать пронумерованной позиции аминокислоты в референсной последовательности. В случае усечений или слияний либо в референсной, либо в выравниваемой последовательности могут присутствовать участки аминокислот, не соответствующих любым аминокислотам в соответствующей последовательности.

[0094] Термины «пронумерованный по отношению к» или «соответствующий», при использовании в контексте нумерации данной аминокислотной или полинуклеотидной последовательности, относятся к нумерации остатков указанной референсной последовательности, когда данная аминокислотная или полинуклеотидная последовательность сравнивается с референсной последовательностью.

[0095] «Клетка» при использовании в данном документе, относится к клетке, осуществляющей метаболическую или другую функцию, достаточную для сохранения или репликации ее геномной ДНК. Клетка может быть идентифицирована хорошо известными в данной области методами, включая, например, наличие неповрежденной мембраны, окрашивание определенным красителем, способность производить потомство или, в случае гаметы, способность объединяться со второй гаметой для получения жизнеспособного потомства. Клетки могут включать прокариотические и эукариотические клетки. Прокариотические клетки включают, без ограничений, бактерии. Эукариотические клетки включают, без ограничений, дрожжевые клетки и клетки, полученные от растений и животных, например, клетки млекопитающих, насекомых (*например*, совки (*Spodoptera*)) и человека. Клетки могут быть полезны, если они по своей природе не прилипают к поверхности или были обработаны, чтобы не прилипали к поверхности, например, путем трипсинизации.

[0096] Используемый в данном документе термин «вектор» относится к молекуле нуклеиновой кислоты, способной транспортировать другую нуклеиновую кислоту, с которой она была связана. Одним из типов вектора является «плазмида», которая относится к линейной или кольцевой двухцепочечной петле ДНК, в которую могут быть лигированы дополнительные сегменты ДНК. Другим типом вектора является вирусный вектор, в котором дополнительные сегменты ДНК могут быть лигированы в вирусный геном. Некоторые векторы способны к автономной репликации в клетке-хозяине, в которую они введены (*например*, бактериальные векторы, имеющие бактериальные векторы, имеющие бактериальную точку начала репликации, и эписомальные векторы млекопитающих). Другие векторы (*например*, неэписомальные векторы млекопитающих) интегрируются в геном клетки-хозяина после введения в клетку-хозяина и, таким образом, реплицируются вместе с геномом хозяина. Более того, определенные векторы способны направлять экспрессию генов, с которыми они функционально связаны. Такие векторы называются в данном документе «экспрессионными векторами». В общем, векторы экспрессии, применимые в методиках рекомбинантной ДНК, часто имеют форму плазмид.

В настоящем изобретении, «плазмида» и «вектор» могут использоваться взаимозаменяемо, поскольку плазмида является наиболее часто используемой формой вектора. Однако, настоящее изобретение предусматривает включение таких других форм экспрессионных векторов, как вирусные векторы (*например*, ретровирусы с дефектом репликации, аденовирусы и аденоассоциированные вирусы), выполняющие эквивалентные функции. Дополнительно, некоторые вирусные векторы способны нацеливаться на клетки определенного типа специфически или неспецифически. Некомпетентные по репликации вирусные векторы или вирусные векторы с дефектом репликации относятся к вирусным векторам, способным инфицировать свои клетки-мишени и доставлять свою вирусную нагрузку, но не способны затем продолжать типичный литический путь, приводящий к лизису и гибели клеток.

[0097] Термины «трансфекция», «трандукция», «трансфицирующий» или «трансдуцирующий» могут использоваться взаимозаменяемо и определяются как процесс введения молекулы нуклеиновой кислоты (например, мРНК, ДНК, РНП) и/или белка в клетку. Нуклеиновые кислоты могут быть введены в клетку с помощью невирусных или вирусных методов. Молекула нуклеиновой кислоты может быть последовательностью, кодирующей полные белки или их функциональные части. Обычно вектор нуклеиновой кислоты включает элементы, необходимые для экспрессии белка (например, промотор, сайт начала транскрипции и т.д.). Невирусные методы трансфекции включают любые пригодные методы, не использующие вирусной ДНК или вирусных частиц в качестве системы доставки для введения молекулы нуклеиновой кислоты в клетку. Примеры невирусных методов трансфекции включают инкапсулирование в наночастицы нуклеиновых кислот, которые кодируют слитый белок (например, липидные наночастицы, наночастицы золота и т.п.), трансфекцию фосфатом кальция, липосомальную трансфекцию, нуклеофекцию, сонопорацию, трансфекцию с тепловым шоком, магнитофекцию и электропорацию. В вирусных методах, любой подходящий вирусный вектор может использоваться в способах, описанных в данном документе. Примеры вирусных векторов включают, без ограничений, ретровирусные, аденовирусные, лентивирусные и аденоассоциированные вирусные векторы. В некоторых аспектах, молекулы нуклеиновой кислоты вводят в клетку с помощью ретровирусного вектора в соответствии со стандартными процедурами, хорошо известными в данной области техники. Термины «трансфекция» или «трандукция» также относятся к введению белков в клетку из внешней среды. Обычно, трандукция или трансфекция белка основана на присоединении пептида или белка, способного пересекать клеточную мембрану, к белку, представляющему интерес. См., например, Ford et al. (2001) *Gene Therapy* 8:1-4 и Prochiantz (2007) *Nat. Methods* 4:119-20.

[0098] «Пептидный линкер», при использовании в данном документе, представляет собой линкер, включающий пептидный фрагмент. В некоторых вариантах осуществления, пептидный линкер представляет собой двухвалентный пептид, такой как аминокислотная последовательность, присоединенная на N-конце и C-конце к остальной части соединения

SEQ ID NO:20. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:21. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:22. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:24. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:27. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:29. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:17, 18, 19, 20, 21, 22, 24, 27 или 29. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:17. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:18. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:19. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:20. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:21. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:22. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:24. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:27. В некоторых аспектах, пептидный линкер имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:29.

[0101] Термины «XTEN», «линкер XTEN» или «полипептид XTEN», используемые в данном документе, относятся к рекомбинантному полипептиду (например, неструктурированному рекомбинантному пептиду), не содержащему гидрофобных аминокислотных остатков. Разработка и применение XTEN описаны, например, в Schellenberger et al., *Nature Biotechnology* 27, 1186-1190 (2009). В некоторых аспектах, линкер XTEN включает последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN представляет собой последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN включает последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN представляет собой

последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:32.

[0102] «Эпитопная метка» относится к биологическому фрагменту, такому как пептид, который введен в рекомбинантный белок методами генной инженерии и функционирует как универсальный эпитоп, легко обнаруживаемый с помощью коммерчески доступных анализов или антител и обычно не нарушающий нативную структуру или функцию белка.

[0103] «Детектируемый агент» или «детектируемый фрагмент» представляет собой композицию, определяемую соответствующими средствами, такими как спектроскопические, фотохимические, биохимические, иммунохимические, химические способы, магнитно-резонансная визуализация или другие физические средства. Например, пригодные детектируемые агенты включают ^{18}F , ^{32}P , ^{33}P , ^{45}Ti , ^{47}Sc , ^{52}Fe , ^{59}Fe , ^{62}Cu , ^{64}Cu , ^{67}Cu , ^{67}Ga , ^{68}Ga , ^{77}As , ^{86}Y , ^{90}Y , ^{89}Sr , ^{89}Zr , ^{94}Tc , ^{94}Tc , $^{99\text{m}}\text{Tc}$, ^{99}Mo , ^{105}Pd , ^{105}Rh , ^{111}Ag , ^{111}In , ^{123}I , ^{124}I , ^{125}I , ^{131}I , ^{142}Pr , ^{143}Pr , ^{149}Pm , ^{153}Sm , $^{154-1581}\text{Gd}$, ^{161}Tb , ^{166}Dy , ^{166}Ho , ^{169}Er , ^{175}Lu , ^{177}Lu , ^{186}Re , ^{188}Re , ^{189}Re , ^{194}Ir , ^{198}Au , ^{199}Au , ^{211}At , ^{211}Pb , ^{212}Bi , ^{212}Pb , ^{213}Bi , ^{223}Ra , ^{225}Ac , Cr, V, Mn, Fe, Co, Ni, Cu, La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, ^{32}P , флуорофор (например, флуоресцентные красители), электронно-плотные реагенты, ферменты (например, как обычно используется в ИФА), биотин, дигоксигенин, парамагнитные молекулы, парамагнитные наночастицы, сверхмалые суперпарамагнитные наночастицы оксида железа (USPIO), агрегаты наночастиц USPIO, суперпарамагнитные наночастицы оксида железа (SPIO), агрегаты наночастиц SPIO, монокристаллические наночастицы оксида железа, монокристаллический оксид железа, наночастицы контрастных агентов, липосомы или другие средства доставки, содержащие молекулы хелата гадолиния (Gd-хелат), гадолиний, радиоизотопы, радионуклиды (напр. g., углерод-11, азот-13, кислород-15, фтор-18, рубидий-82), фтордезоксиглюкоза (например, меченную фтором-18), любые радионуклиды, испускающие гамма-лучи, позитрон-излучающий радионуклид, радиомеченную глюкозу, радиомеченную воду, радиомеченный аммиак, биокolloиды,

микропузырьки (например, включая оболочки микропузырьков, включающие альбумин, галактозу, липид и/или полимеры; газовое ядро микропузырьков, включающее воздух, тяжелые газы, перфторуглерод, азот, октафторпропан, липидную микросферу перфлексан, перфлюотрен), йодированные контрастные вещества (например, иогексол, иодиксанол, иоверсол, иопамидол, иоксилан, иопромид, диатризоат, метризоат, иоксаглат), сульфат бария, диоксид тория, золото, наночастицы золота, агрегаты наночастиц золота, флуорофоры, двухфотонные флуорофоры, гаптены и белки или другие вещества, которые можно сделать детектируемыми, *например*, путем включения радиометок в пептид или антитело, специфически реагирующее с целевым пептидом.

[0104] Детектируемый фрагмент представляет собой моновалентный детектируемый агент или детектируемый агент, способный образовывать связь с другой композицией. В некоторых аспектах, детектируемый агент представляет собой эпитопную метку. В некоторых аспектах, эпитопная метка представляет собой метку HA. В некоторых аспектах, метка HA включает последовательность, представленную SEQ ID NO:24. В некоторых аспектах, метка HA представляет собой последовательность, представленную SEQ ID NO:24. В некоторых аспектах, метка HA имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:24. В некоторых аспектах, метка HA имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:24. В некоторых аспектах, метка HA имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:24. В некоторых аспектах, метка HA имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:24.

[0105] В некоторых аспектах, детектируемый агент представляет собой флуоресцентный белок. В некоторых аспектах, флуоресцентный белок представляет собой синий флуоресцентный белок (BFP). В некоторых аспектах, BFP включает последовательность, представленную SEQ ID NO:30. В некоторых аспектах, BFP представляет собой последовательность, представленную SEQ ID NO:30. В некоторых аспектах, BFP имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:30. В некоторых аспектах, BFP имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:30. В некоторых аспектах, BFP имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:30. В некоторых аспектах, BFP имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:30.

[0106] Радиоактивные вещества (например, радиоизотопы), которые могут быть использованы в качестве агентов визуализации и/или меченя в соответствии с аспектами настоящего изобретения, включают, без ограничений, ^{18}F , ^{32}P , ^{33}P , ^{45}Ti , ^{47}Sc ,

⁵²Fe, ⁵⁹Fe, ⁶²Cu, ⁶⁴Cu, ⁶⁷Cu, ⁶⁷Ga, ⁶⁸Ga, ⁷⁷As, ⁸⁶Y, ⁹⁰Y, ⁸⁹Sr, ⁸⁹Zr, ⁹⁴Tc, ⁹⁴Tc, ^{99m}Tc, ⁹⁹Mo, ¹⁰⁵Pd, ¹⁰⁵Rh, ¹¹¹Ag, ¹¹¹In, ¹²³I, ¹²⁴I, ¹²⁵I, ¹³¹I, ¹⁴²Pr, ¹⁴³Pr, ¹⁴⁹Pm, ¹⁵³Sm, ¹⁵⁴⁻¹⁵⁸Gd, ¹⁶¹Tb, ¹⁶⁶Dy, ¹⁶⁶Ho, ¹⁶⁹Er, ¹⁷⁵Lu, ¹⁷⁷Lu, ¹⁸⁶Re, ¹⁸⁸Re, ¹⁸⁹Re, ¹⁹⁴Ir, ¹⁹⁸Au, ¹⁹⁹Au, ²¹¹At, ²¹¹Pb, ²¹²Bi, ²¹²Pb, ²¹³Bi, ²²³Ra и ²²⁵Ac. Парамагнитные ионы, которые могут быть использованы в качестве дополнительных агентов визуализации в соответствии с аспектами настоящего изобретения включают, без ограничений, ионы переходных и лантанидных металлов (например, металлов, имеющих атомные номера 21-29, 42, 43, 44 или 57-71). Такие металлы включают ионы Cr, V, Mn, Fe, Co, Ni, Cu, La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb и Lu.

[0107] «Контактирование» используется в соответствии с его простым обычным значением и относится к процессу, позволяющему по крайней мере двум различным формам сблизиться настолько, чтобы вступить в реакцию, взаимодействовать или физически соприкоснуться. Следует понимать, однако, что образующийся продукт реакции может быть получен непосредственно в результате реакции между добавляемыми реагентами или из промежуточного продукта одного или нескольких добавляемых реагентов, который может быть получен в реакционной смеси.

[0108] Термин «введение в контакт» может включать обеспечение возможности реакции, взаимодействия или физического соприкосновения двух форм, причем эти две формы могут представлять собой, например, слитый белок, как предусмотрено в данном документе, и последовательность нуклеиновой кислоты (например, последовательность ДНК-мишени).

[0109] Как определено в данном документе, термины «ингибирование», «ингибировать», «ингибирующий», «репрессия», «репрессирующий», «сайленсинг», «выключение» и т.п. при использовании в отношении предложенной в данном документе композиции (например, слитого белка, комплекса, нуклеиновой кислоты, вектора) означают негативное влияние (например, снижение) на активность (например, транскрипцию) последовательности нуклеиновой кислоты (например, снижение транскрипции гена) по сравнению с активностью последовательности нуклеиновой кислоты (например, транскрипцией гена) в отсутствие композиции (например, слитого белка, комплекса, нуклеиновой кислоты, вектора). В некоторых аспектах, ингибирование относится к уменьшению заболевания или симптомов заболевания (например, рака). Таким образом, ингибирование включает, по меньшей мере отчасти, частичное или полное блокирование активации (например, транскрипции), или снижение, предотвращение или задержку активации (например, транскрипции) последовательности нуклеиновой кислоты. Ингибируемая активность (например, транскрипция) может составлять 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% или менее от контрольной. В некоторых аспектах, ингибирование является 1,5-кратным, 2-кратным, 3-кратным, 4-кратным, 5-кратным, 10-кратным или более по сравнению с контролем.

[0110] «Контрольный» образец или значение относится к образцу, служащему эталоном, обычно известным эталоном, для сравнения с тестируемым образцом.

Например, тестируемый образец может быть взят в условиях тестирования, например, в присутствии тестируемого соединения, и сравнен с образцами из известных условий, например, в отсутствие тестируемого соединения (отрицательный контроль) или в присутствии известного соединения (положительный контроль). Контроль также может представлять собой среднее значение, определяемое из нескольких тестов или результатов. Специалисту в данной области понятно, что контроли могут быть предназначены для оценки любого количества параметров. Например, можно разработать контроль для сравнения терапевтического эффекта на основе фармакологических данных (например, периода полувыведения) или терапевтических показателей (например, сравнение побочных эффектов). Специалист в данной области понимает, какие контроли являются значимыми в конкретной ситуации, и может анализировать данные на основе сравнения с контрольными значениями. Контроли также важны для определения значимости данных. Например, если значения данного параметра широко варьируются у контролей, то вариации в тестируемых образцах не будут считаться значимыми.

[0111] Слитые белки

[0112] В данном документе представлены, в частности, слитые белки, которые могут долговременно (например, необратимо) и обратимо выключать гены в клетках млекопитающих путем редактирования эпигенома на основе CRISPR. В вариантах осуществления изобретения слитый белок включает единый полипептид слитых белков (например, каталитически неактивного Cas9 (например, dCas9), домена KRAB, Dnmt3A и Dnmt3L), который может быть транзитно доставлен в виде мРНК, ДНК или РНП и транзитно экспрессирован в клетках. Слитый белок может быть нацелен на определенный участок в геноме млекопитающего с помощью sgРНК или cr:tracrРНК. После размещения в правильном положении, и без намерения быть связанными теорией, слитый белок добавляет к целевой нуклеиновой кислоте метилирование ДНК и/или репрессивные метки хроматина, что приводит к сайленсингу гена, наследуемому при последующих клеточных делениях. Таким образом, слитый белок может осуществлять редактирование эпигенома, минуя необходимость генерировать двухцепочные разрывы ДНК в геноме хозяина, что делает его безопасным и обратимым способом манипулирования геномом живого организма.

[0113] В некоторых вариантах осуществления, слитый белок содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы; домен KRAB и домен ДНК-метилтрансферазы. В некоторых аспектах, слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен KRAB. В некоторых аспектах, слитый белок содержит, от N-конца к С-концу, домен KRAB, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен ДНК-метилтрансферазы. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, слитый белок дополнительно содержит один или несколько пептидных

линкеров. В некоторых аспектах, слитый белок дополнительно содержит одну или несколько детектируемых меток. В некоторых аспектах, слитый белок дополнительно содержит одну или несколько последовательностей ядерной локализации. В некоторых аспектах, слитый белок дополнительно содержит один или несколько пептидных линкеров, одну или несколько детектируемых меток, одну или несколько последовательностей ядерной локализации, или комбинацию двух или более из вышеперечисленного. В тех случаях, когда слитый белок содержит один или несколько пептидных линкеров, все пептидные линкеры могут быть одинаковыми или разными. В тех случаях, когда слитый белок содержит одну или несколько детектируемых меток, все детектируемые метки могут быть одинаковыми или разными. В некоторых аспектах, слитый белок содержит от 1 до 10 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 9 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 8 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 7 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 6 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 5 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 4 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 3 детектируемых меток. В некоторых аспектах, слитый белок содержит от 1 до 2 детектируемых меток. В некоторых аспектах, слитый белок содержит 1 детектируемую метку. В некоторых аспектах, слитый белок содержит 2 детектируемые метки. В некоторых аспектах, слитый белок содержит 3 детектируемые метки. В некоторых аспектах, слитый белок содержит 4 детектируемые метки. В некоторых аспектах, слитый белок содержит 5 детектируемых меток.

[0114] В некоторых вариантах осуществления, слитый белок содержит дефицитный по нуклеазе фермент ДНК-эндонуклеазы, домен KRAB и домен ДНК-метилтрансферазы. В некоторых аспектах, слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, дефицитный по нуклеазе фермент ДНК-эндонуклеазы и домен KRAB. В некоторых аспектах, слитый белок содержит, от N-конца к С-концу, домен KRAB, дефицитный по нуклеазе фермент ДНК-эндонуклеазы и домен ДНК-метилтрансферазы. В некоторых вариантах осуществления, слитый белок дополнительно содержит один или несколько пептидных линкеров. В некоторых аспектах, слитый белок дополнительно содержит одну или несколько детектируемых меток. В некоторых аспектах, слитый белок дополнительно содержит одну или несколько последовательностей ядерной локализации. В некоторых аспектах, слитый белок дополнительно содержит один или несколько пептидных линкеров, одну или несколько детектируемых меток, одну или несколько последовательностей ядерной локализации, или комбинацию двух или более из вышеперечисленного. В тех случаях, когда слитый белок содержит один или несколько пептидных линкеров, все пептидные линкеры могут быть одинаковыми или разными. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или TALE. В некоторых

вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE.

[0115] В некоторых вариантах осуществления, настоящее изобретение предусматривает слитый белок, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит больше аминокислотных остатков, чем второй линкер XTEN. В некоторых вариантах осуществления, слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый и второй линкеры XTEN содержат от около 20 до около 100 аминокислотных остатков. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы представляет собой CRISPR-ассоциированный белок. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, dCpf1, ddCpf1, дефицитный по нуклеазе вариант Cas9, дефицитную по нуклеазе эндонуклеазу CRISPR класса II, домен цинкового пальца, домен лейциновой молнии, домен крылатой спирали, эффектор, подобный активатору транскрипции (TALE), мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен HMB-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой Cas-phi. В некоторых аспектах, домен ДНК-метилтрансферазы содержит Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы (Dnmt3A) дополнительно содержит регуляторный фактор Dnmt3L (называемый в данном документе доменом Dnmt3A-3L или доменом Dnmt3B-3L). В некоторых аспектах, слитый белок дополнительно содержит эпитопную метку, 2A-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0116] В некоторых вариантах осуществления, настоящее изобретение предусматривает слитый белок, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит один или несколько аминокислотных остатков, чем второй линкер XTEN. В некоторых вариантах осуществления, слитый белок, содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 5 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN, содержащий от около 5 до около

864 аминокислотных остатков, и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый и второй линкеры XTEN содержат от около 20 до около 100 аминокислотных остатков. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или TALE. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE. В некоторых аспектах, домен ДНК-метилтрансферазы содержит Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы (Dnmt3A) дополнительно содержит регуляторный фактор Dnmt3L (называемый в данном документе доменом Dnmt3A-3L или доменом Dnmt3B-3L). В некоторых аспектах, слитый белок дополнительно содержит эпитопную метку, 2A-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0117] В некоторых вариантах осуществления, настоящее изобретение предусматривает слитый белок, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит больше аминокислотных остатков, чем второй линкер XTEN. В некоторых вариантах осуществления, слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от более чем 50 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN, содержащий от около 5 до 50 аминокислотных остатков, и домен Krüppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до 40 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, dCpf1, ddCpf1, дефицитный по нуклеазе вариант Cas9, дефицитную по нуклеазе эндонуклеазу CRISPR класса II, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Woz3, домен OB-складки, домен иммуноглобулина или домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой Cas-phi. В некоторых аспектах, домен ДНК-метилтрансферазы содержит Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы (Dnmt3A) дополнительно содержит регуляторный фактор Dnmt3L (называемый в данном документе доменом Dnmt3A-3L или

доменом Dnmt3B-3L). В некоторых аспектах, слитый белок дополнительно содержит эпитопную метку, 2А-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0118] В некоторых вариантах осуществления, настоящее изобретение предусматривает слитый белок, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN и домен K \ddot{u} ppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит один или несколько аминокислотных остатков, чем второй линкер XTEN. В некоторых вариантах осуществления, слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от более чем около 50 до около 864 аминокислотных остатков, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN, содержащий от около 5 до 50 аминокислотных остатков, и домен K \ddot{u} ppel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до 40 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или TALE. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE. В некоторых аспектах, домен ДНК-метилтрансферазы содержит Dnmt3A. В некоторых аспектах, домен ДНК-метилтрансферазы (Dnmt3A) дополнительно содержит регуляторный фактор Dnmt3L (называемый в данном документе доменом Dnmt3A-3L или доменом Dnmt3B-3L). В некоторых аспектах, слитый белок дополнительно содержит эпитопную метку, 2А-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0119] В некоторых вариантах осуществления, слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, домен K \ddot{u} ppel-ассоциированного бокса, расщепляемый 2А-пептид и флуоресцентную белковую метку. В некоторых аспектах, первый линкер XTEN содержит больше (one or more) аминокислотных остатков, чем второй линкер XTEN. В некоторых аспектах, первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до 40 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN

содержит от около 10 до около 30 аминокислотных остатков. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, dCpf1, ddCpf1, дефицитный по нуклеазе вариант Cas9, дефицитную по нуклеазе эндонуклеазу CRISPR класса II, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9. В некоторых аспектах, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой Cas-phi. В некоторых аспектах, домен ДНК-метилтрансферазы включает домен Dnmt3A. В некоторых аспектах, домен Dnmt3A связан с регуляторным фактором Dnmt3L (называемым в данном документе доменом Dnmt3A-3L). В некоторых аспектах, домен ДНК-метилтрансферазы включает домен Dnmt3B. В некоторых аспектах, домен Dnmt3B связан с регуляторным фактором Dnmt3L (называемым в данном документе доменом Dnmt3B-3L).

[0120] В некоторых вариантах осуществления, слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, домен Krüppel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку. В некоторых аспектах, первый линкер XTEN содержит больше (one or more) аминокислотных остатков, чем второй линкер XTEN. В некоторых аспектах, первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до 40 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или TALE. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца. В некоторых аспектах, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE. В некоторых аспектах, домен ДНК-метилтрансферазы включает домен Dnmt3A. В некоторых аспектах, домен Dnmt3A связан с регуляторным фактором Dnmt3L (называемым в данном документе доменом Dnmt3A-3L). В некоторых аспектах, домен ДНК-метилтрансферазы включает домен Dnmt3B. В некоторых аспектах, домен Dnmt3B связан с регуляторным фактором Dnmt3L (называемым в данном документе доменом Dnmt3B-3L).

[0121] В некоторых аспектах, слитый белок включает аминокислотную

идентичности последовательности с SEQ ID NO:99. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 98% идентичности последовательности с SEQ ID NO:107. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 98% идентичности последовательности с SEQ ID NO:108. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 98% идентичности последовательности с SEQ ID NO:109. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 98% идентичности последовательности с SEQ ID NO:110.

[0131] В некоторых вариантах осуществления, слитый белок содержит структуру: А-В-С, или В-А-С или С-А-В, или С-В-А, или В-С-А, или А-С-В; где А содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы или дефицитный по нуклеазе фермент ДНК-эндонуклеазы; В содержит домен KRAB, С содержит домен ДНК-метилтрансферазы; и при этом компонент, указанный слева, представляет собой N-конец, и компонент справа представляет собой С-конец. В некоторых аспектах, слитый белок дополнительно содержит один или несколько пептидных линкеров и одну или несколько детектируемых меток. В некоторых аспектах, А-В, В-А, В-С, С-В, А-С и С-А в каждом случае независимо соединены друг с другом с помощью ковалентной связи, пептидного линкера, детектируемой метки, последовательности ядерной локализации, или комбинации двух или более из них. Пептидный линкер может быть любым известным в данной области техники (например, расщепляемым Р2А-пептидом, линкером XTEN и т.п.). В некоторых аспектах, слитый белок содержит другие компоненты, такие как детектируемые метки (например, НА-метку, синий флуоресцентный белок и т.п.). В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой домен цинкового пальца. В некоторых вариантах осуществления, «А» представляет собой TALE.

[0132] В некоторых вариантах осуществления, слитый белок содержит структуру: А- L_1 -В- L_2 -С или С- L_2 -В- L_1 -А или С- L_2 -А- L_1 -В, где А содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы или дефицитный по нуклеазе фермент ДНК-эндонуклеазы; В содержит домен KRAB, С содержит домен ДНК-метилтрансферазы, L_1 отсутствует, представляет собой ковалентную связь или пептидный линкер, и L_2 отсутствует, представляет собой ковалентную связь или пептидный линкер; и при этом компонент, расположенный слева, находится на N-конце, и компонент справа находится на С-конце. В некоторых аспектах, Dnmt3A ковалентно связан с Dnmt3L с помощью пептидного линкера. В некоторых аспектах, А ковалентно связан с В с помощью ковалентной связи. В некоторых аспектах, В ковалентно связан с С с помощью

пептидного линкера. В некоторых аспектах, В ковалентно связан с С с помощью ковалентной связи. Пептидный линкер может быть любым известным в данной области техники (например, расщепляемым Р2А-пептидом, линкером ХТЕН и т.п.). В некоторых аспектах, слитый белок содержит другие компоненты, такие как детектируемые метки, последовательности ядерной локализации и т.п. В некоторых аспектах, L₁ представляет собой ковалентную связь, пептидный линкер, детектируемую метку, последовательности ядерной локализации или их комбинацию. В некоторых аспектах, L₂ представляет собой ковалентную связь, пептидный линкер, детектируемую метку, последовательности ядерной локализации или их комбинацию. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой домен цинкового пальца. В некоторых вариантах осуществления, «А» представляет собой TALE.

[0133] В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 80% идентичности последовательности со слитым белком, имеющим аминокислотную последовательность формулы (А); при этом аминокислотная последовательность формулы (А) представляет собой, от N-конца к С-концу:



где C₁ содержит SEQ ID NO:26 или SEQ ID NO:106; R³ отсутствует или R³ содержит SEQ ID NO:27; C₂ содержит SEQ ID NO:28; R² отсутствует или R² содержит SEQ ID NO:32; А содержит SEQ ID NO:23; R¹ отсутствует или R¹ содержит SEQ ID NO:25; R⁴ отсутствует или R⁴ содержит SEQ ID NO:31; и В содержит SEQ ID NO:16, SEQ ID NO:103, SEQ ID NO:104 или SEQ ID NO:105. В некоторых вариантах осуществления, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ содержит SEQ ID NO:25 и R⁴ отсутствует. В некоторых вариантах осуществления, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ отсутствует и R⁴ отсутствует. В некоторых вариантах осуществления, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ отсутствует и R⁴ содержит SEQ ID NO:31. В некоторых вариантах осуществления, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ содержит SEQ ID NO:25 и R⁴ содержит SEQ ID NO:31. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ содержит SEQ ID NO:25 и R⁴ отсутствует. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ отсутствует и R⁴ отсутствует. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ отсутствует и R⁴ содержит SEQ ID NO:31. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ содержит SEQ ID NO:25 и R⁴ содержит SEQ ID NO:31. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ отсутствует и R⁴ содержит SEQ ID NO:31. В некоторых вариантах осуществления, C₁ содержит SEQ ID NO:26, В содержит SEQ ID NO:16, R² содержит SEQ ID NO:32, R³ содержит SEQ ID NO:27, R¹ содержит SEQ ID NO:25 и R⁴ содержит SEQ ID NO:31.

аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 91% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 92% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 93% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 94% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 96% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 97% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 98% идентичности последовательности с аминокислотной последовательностью формулы (А). В некоторых вариантах осуществления, слитый белок имеет по меньшей мере 99% идентичности последовательности с аминокислотной последовательностью формулы (А).

[0134] В некоторых вариантах осуществления, слитый белок содержит структуру: В-L₁-А-L₂-С или С-L₁-А-L₂-В, где А содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы или дефицитный по нуклеазе фермент ДНК-эндонуклеазы; В содержит домен KRAB, С содержит домен ДНК-метилтрансферазы, L₁ представляет собой ковалентную связь или пептидный линкер, и L₂ представляет собой ковалентную связь или пептидный линкер. В некоторых вариантах осуществления слитый белок содержит структуру: В-L₁-А-L₂-С. В некоторых вариантах осуществления слитый белок содержит структуру: С-L₁-А-L₂-В. В некоторых аспектах, L₁ представляет собой пептидный линкер. В некоторых аспектах, L₁ представляет собой ковалентную связь. В некоторых аспектах, L₂ представляет собой пептидный линкер. В некоторых аспектах, L₂ представляет собой ковалентную связь. Пептидный линкер может быть любым известным в данной области техники или описанным в данном документе (например, расщепляемым Р2А-пептидом, линкером XTEN и т.п.). В некоторых аспектах, слитый белок содержит другие компоненты, такие как детектируемые метки. В некоторых аспектах, L₁ представляет собой ковалентную связь, пептидный линкер, детектируемую метку или их комбинацию. В некоторых аспектах, L₂ представляет собой ковалентную связь, пептидный линкер, детектируемую метку или их комбинацию. В некоторых аспектах, слитый белок дополнительно содержит последовательность ядерной локализации. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе

фермент РНК-направляемой ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой домен цинкового пальца. В некоторых вариантах осуществления, «А» представляет собой TALE.

[0135] В некоторых вариантах осуществления, слитый белок содержит структуру: В-L₃-А-L₄-С-L₅-D или С-L₃-А-L₄-В-L₅-D, где А содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы или дефицитный по нуклеазе фермент ДНК-эндонуклеазы; В содержит домен KRAB, С содержит домен ДНК-метилтрансферазы, D отсутствует или D содержит одну или несколько детектируемых меток, L₃ содержит ковалентную связь, пептидный линкер, детектируемую метку, или комбинацию двух или более из них, L₄ содержит ковалентную связь, пептидный линкер, детектируемую метку, или комбинацию двух или более из них, L₅ отсутствует или L₅ содержит ковалентную связь или пептидный линкер. В некоторых вариантах осуществления, слитый белок содержит структуру: В-L₃-А-L₄-С-L₅-D. В некоторых вариантах осуществления, слитый белок содержит структуру: С-L₃-А-L₄-В-L₅-D. В некоторых аспектах, L₃ представляет собой пептидный линкер. В некоторых аспектах, L₃ представляет собой ковалентную связь. В некоторых аспектах, L₃ содержит пептидный линкер и детектируемую метку. В некоторых аспектах, L₃ содержит детектируемую метку. В некоторых аспектах, L₄ представляет собой пептидный линкер. В некоторых аспектах, L₄ содержит пептидный линкер и детектируемую метку. В некоторых аспектах, L₄ представляет собой ковалентную связь. В некоторых аспектах, L₄ содержит детектируемую метку. В некоторых аспектах, L₅ представляет собой пептидный линкер. В некоторых аспектах, L₅ представляет собой ковалентную связь. В некоторых аспектах, D содержит одну или несколько детектируемых меток. В некоторых аспектах, D содержит одну детектируемую метку. В некоторых аспектах, D содержит две детектируемые метки. В некоторых аспектах, D содержит три детектируемые метки. В некоторых аспектах, D содержит несколько детектируемых меток. D может быть любой детектируемой меткой, известной в данной области техники и/или описанной в данном документе (например, НА-меткой, синим флуоресцентным белком и т.п.). В некоторых аспектах L₅ и D отсутствуют. В тех случаях, когда L₃, L₄, L₅ и D содержат две или более детектируемых меток, все детектируемые метки являются одинаковыми или разными. Пептидный линкер может быть любым известным в данной области техники и/или описанным в данном документе (например, расщепляемым Р2А-пептидом, линкером XTEN и т.п.). В некоторых аспектах, слитый белок дополнительно содержит последовательность ядерной локализации. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой домен

цинкового пальца. В некоторых вариантах осуществления, «А» представляет собой TALE.

[0136] В некоторых вариантах осуществления, слитый белок содержит структуру: С-L₃-А-L₄-В-L₅-D, где А содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы или дефицитный по нуклеазе фермент ДНК-эндонуклеазы; В содержит домен KRAB, С содержит домен ДНК-метилтрансферазы, D отсутствует или D содержит одну или несколько детектируемых меток, L₃ содержит ковалентную связь, пептидный линкер, детектируемую метку или комбинацию двух или более из них, L₄ содержит ковалентную связь, пептидный линкер, детектируемую метку или комбинацию двух или более из них, L₅ отсутствует или L₅ содержит ковалентную связь или пептидный линкер; и при этом С находится на N-конце и D находится на С-конце. В некоторых аспектах, L₃ представляет собой пептидный линкер. В некоторых аспектах, L₃ представляет собой ковалентную связь. В некоторых аспектах, L₃ содержит детектируемую метку. В некоторых аспектах, L₃ содержит пептидный линкер и детектируемую метку. В некоторых аспектах, L₄ представляет собой пептидный линкер. В некоторых аспектах, L₄ представляет собой ковалентную связь. В некоторых аспектах, L₄ содержит детектируемую метку. В некоторых аспектах, L₄ содержит пептидный линкер и детектируемую метку. В некоторых аспектах, L₅ представляет собой пептидный линкер. В некоторых аспектах, L₅ представляет собой ковалентную связь. В некоторых аспектах, D содержит одну или несколько детектируемых меток. В некоторых аспектах, D содержит одну детектируемую метку. В некоторых аспектах, D содержит две детектируемые метки. В некоторых аспектах, D содержит три детектируемые метки. В некоторых аспектах, D содержит несколько детектируемых меток. D может быть любой детектируемой меткой, известной в данной области техники и/или описанной в данном документе (например, НА-меткой, синим флуоресцентным белком и т.п.). В некоторых аспектах L₅ и D отсутствуют. В тех случаях, когда L₃, L₄, L₅ и D содержат две или более детектируемых меток, все детектируемые метки являются одинаковыми или разными. Пептидный линкер может быть любым известным в данной области техники и/или описанным в данном документе (например, расщепляемым Р2А-пептидом, линкером XTEN и т.п.). В некоторых аспектах, слитый белок дополнительно содержит последовательность ядерной локализации. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой CRISPR-ассоциированный белок. В некоторых вариантах осуществления, «А» представляет собой дефицитный по нуклеазе фермент ДНК-эндонуклеазы. В некоторых вариантах осуществления, «А» представляет собой домен цинкового пальца. В некоторых вариантах осуществления, «А» представляет собой TALE.

[0137] В некоторых вариантах осуществления, слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 60 до около 150 аминокислотных остатков, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN, содержащий от около 5 до около 50 аминокислотных остатков, и домен K \ddot{u} rpel-ассоциированного бокса. В некоторых

аспектах, первый линкер XTEN содержит от около 70 до около 90 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 20 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит около 80 аминокислотных остатков, и второй линкер XTEN содержит от около 16 аминокислотных остатков. В некоторых аспектах, слитый белок дополнительно содержит детектируемую метку (например, эпитопную метку, флуоресцентную белковую метку), 2A-пептид (например, P2A-пептид), сигнальный пептид ядерной локализации или комбинацию двух или более из них. В некоторых аспектах, слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 60 до около 150 аминокислотных остатков, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, содержащий от около 5 до около 50 аминокислотных остатков, домен K \ddot{u} rrel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок.

[0138] В некоторых вариантах осуществления, слитый белок, содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 60 до около 150 аминокислотных остатков, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN, содержащий от около 5 до около 50 аминокислотных остатков, и домен K \ddot{u} rrel-ассоциированного бокса. В некоторых аспектах, первый линкер XTEN содержит от около 70 до около 90 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 20 аминокислотных остатков. В некоторых аспектах, первый линкер XTEN содержит около 80 аминокислотных остатков, и второй линкер XTEN содержит от около 16 аминокислотных остатков. В некоторых аспектах, слитый белок дополнительно содержит детектируемую метку (например, эпитопную метку, флуоресцентную белковую метку), 2A-пептид (например, P2A-пептид), сигнальный пептид ядерной локализации или комбинацию двух или более из них. В некоторых аспектах, слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, содержащий от около 60 до около 150 аминокислотных остатков, дефицитный по нуклеазе фермент эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, содержащий от около 5 до около 50 аминокислотны остатки, домен K \ddot{u} rrel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент эндонуклеазы представляет собой домен цинкового пальца или TALE.

[0139] В некоторых вариантах осуществления, пептидный линкер представляет собой линкер XTEN. В некоторых аспектах, линкер XTEN включает от около 16 до около 80 аминокислотных остатков. В некоторых аспектах, линкер XTEN включает от около 17 до около 80 аминокислотных остатков. В некоторых аспектах, линкер XTEN включает от около 18 до около 80 аминокислотных остатков. В некоторых аспектах, линкер XTEN

линкер XTEN содержит от около 60 до около 150 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 60 до около 120 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 60 до около 110 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 60 до около 100 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 70 до около 90 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 75 до около 85 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит около 80 аминокислотных остатков. В некоторых аспектах, если слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер XTEN, который содержит от около 50 до около 200 аминокислотных остатков, называется первым пептидным линкером XTEN.

[0142] В некоторых вариантах осуществления, линкер XTEN содержит от около 5 до около 55 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 5 до около 50 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 5 до около 40 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 10 до около 30 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 10 до около 25 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 10 до около 20 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит от около 14 до около 18 аминокислотных остатков. В некоторых аспектах, линкер XTEN содержит около 16 аминокислотных остатков. В некоторых аспектах, если слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер XTEN, который содержит от около 5 до около 55 аминокислотных остатков, называется вторым пептидным линкером XTEN.

[0143] В некоторых вариантах осуществления, линкер XTEN включает последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN представляет собой последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN включает последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN представляет собой последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 50%, 55%, 60%,

65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:32.

[0144] Слитый белок может включать аминокислотные последовательности, пригодные для нацеливания слитого белка на конкретные области клетки (например, цитоплазму, ядро). Таким образом, в некоторых аспектах, слитый белок дополнительно включает пептид сигнала ядерной локализации (NLS). В некоторых аспектах, NLS включает последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS представляет собой последовательность, представленную SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO: 25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:25. В некоторых аспектах, NLS имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:25.

[0145] В некоторых вариантах осуществления, слитый белок включает, от N-конца к C-концу, домен KRAB, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен ДНК-метилтрансферазы. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, и домен ДНК-метилтрансферазы представляет собой домен Dnmt3A-3L. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9, и домен ДНК-метилтрансферазы представляет собой домен Dnmt3A-3L. В некоторых вариантах осуществления, dCas9 ковалентно связан с доменом KRAB с помощью

пептидного линкера, и при этом dCas9 ковалентно связан с доменом Dnmt3A-3L с помощью пептидного линкера.

[0146] В некоторых вариантах осуществления, слитый белок включает, от N-конца к С-концу, домен KRAB, дефицитный по нуклеазе фермент ДНК-эндонуклеазы и домен ДНК-метилтрансферазы. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца и домен ДНК-метилтрансферазы представляет собой домен Dnmt3A-3L. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой TALE, и домен ДНК-метилтрансферазы представляет собой домен Dnmt3A-3L. В некоторых вариантах осуществления, дефицитный по нуклеазе фермент ДНК-эндонуклеазы ковалентно связан с доменом KRAB с помощью пептидного линкера, и при этом дефицитный по нуклеазе фермент ДНК-эндонуклеазы ковалентно связан с доменом Dnmt3A-3L с помощью пептидного линкера.

[0147] В некоторых вариантах осуществления, пептидный линкер представляет собой линкер XTEN. В некоторых аспектах, линкер XTEN включает последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN представляет собой последовательность, представленную SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:31. В некоторых аспектах, линкер XTEN включает последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN представляет собой последовательность, представленную SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 75% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:32. В некоторых аспектах, линкер XTEN имеет аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:32.

[0148] В некоторых вариантах осуществления, слитый белок включает

аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:12. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:13. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:14. В некоторых аспектах, слитый белок включает аминокислотную последовательность, имеющую по меньшей мере 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:15.

[0150] Комплексы

[0151] Для того чтобы слитый белок мог осуществлять редактирование эпигенома, слитый белок взаимодействует (например, нековалентно связан) с полинуклеотидом (например, sgРНК), который комплементарен полинуклеотидной последовательности-мишени (например, целевой последовательности ДНК, подлежащей редактированию) и дополнительно включает последовательность (т.е. связывающую последовательность), с которой может связываться дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы слитого белка, как описано в данном документе. В некоторых аспектах, полинуклеотид, комплементарный к полинуклеотидной последовательности-мишени (например, целевой последовательности ДНК, подлежащей редактированию) и дополнительно включающий связывающую последовательность, с которой может связываться дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы слитого белка, как описано в данном документе, представляет собой sgРНК. В некоторых аспектах, полинуклеотид, комплементарный к полинуклеотидной последовательности-мишени (например, целевой последовательности ДНК, подлежащей редактированию) и дополнительно включающий связывающую последовательность, к которой может связываться дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы слитого белка, как описано в данном документе, представляет собой cr:tracrРНК. В результате образования этого комплекса слитый белок занимает надлежащее положение для осуществления редактирования эпигенома. Термин «комплекс» относится к композиции, которая включает два или более компонентов, причем компоненты связываются друг с другом с образованием функциональной единицы. В некоторых аспектах, комплекс, описанный в данном документе, включает слитый белок, описанный в данном документе, и полинуклеотид, описанный в данном документе. Таким образом, в одном из аспектов предусматривается слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты, и sgРНК или cr:tracrРНК (т.е.,

полинуклеотид, включающий: (1) нацеливающую на ДНК последовательность, комплементарную к полинуклеотидной последовательности-мишени; и (2) связывающую последовательность дефицитного по нуклеазе фермента РНК-направляемой ДНК-эндонуклеазы, причем дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы связан с полинуклеотидом с помощью связывающей последовательности (например, аминокислотной последовательности, способной связываться с нацеливающей на ДНК последовательностью)).

[0152] Нацеливающая на ДНК последовательность относится к полинуклеотиду, который включает нуклеотидную последовательность, комплементарную к полинуклеотидной последовательности-мишени (ДНК или РНК). В некоторых аспектах, нацеливающая на ДНК последовательность может быть отдельной молекулой РНК (отдельным полинуклеотидом РНК), который может включать «одиочную направляющую РНК» или «sgРНК». В некоторых аспектах, нацеливающая на ДНК последовательность может содержать две молекулы РНК (два полинуклеотида РНК), и называется направляющей РНК (гРНК). В некоторых аспектах, нацеливающая на ДНК последовательность включает две молекулы РНК (например, соединенные посредством гибридизации на связывающей последовательности (например, dCas9-связывающей последовательности)). В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% комплементарной к полинуклеотидной последовательности-мишени. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) связывается с последовательностью клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 75% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 80% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) связывается с последовательностью клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 85% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) связывается с последовательностью клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере на 90% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) связывается с последовательностью клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) является по меньшей мере

на 95% комплементарной к последовательности клеточного гена. В некоторых аспектах, нацеливающая на ДНК последовательность (например, sgРНК) связывается с последовательностью клеточного гена.

[0153] «Нуклеиновая кислота-мишень» или «целевая последовательность нуклеиновой кислоты», как предусмотрено в данном документе, представляет собой последовательность нуклеиновой кислоты, присутствующую в клетке или экспрессируемую клеткой, с которой направляющая последовательность (или нацеливающая на ДНК последовательность) должна обладать комплементарностью, причем гибридизация между последовательностью-мишенью и направляющей последовательностью (или нацеливающей на ДНК последовательностью) способствует образованию комплекса (например, комплекса CRISPR). Полная комплементарность не является необходимой, при условии наличия достаточной комплементарности для обеспечения гибридизации и промотирования образования комплекса (например, комплекса CRISPR). В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой экзогенную последовательность нуклеиновой кислоты. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой эндогенную последовательность нуклеиновой кислоты.

[0154] Полинуклеотидная последовательность-мишень может быть любой областью полинуклеотида (например, ДНК-последовательности), пригодной для эпигеномного редактирования. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью гена. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью регуляторной последовательности транскрипции. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью промотора, энхансера или сайленсера. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью промотора. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью энхансера. В некоторых аспектах, полинуклеотидная последовательность-мишень является частью сайленсера.

[0155] В некоторых вариантах осуществления, полинуклеотидная последовательность-мишень представляет собой гипометилированную последовательность нуклеиновой кислоты. «Гипометилированная последовательность нуклеиновой кислоты» используется в данном документе в соответствии со стандартным значением в данной области техники и относится к потере или отсутствию метильных групп на 5-метилцитозинном нуклеотиде (например, в CpG). Потеря или отсутствие метильных групп может определяться относительно стандартного контроля. Гипометилирование может наблюдаться, например, в стареющих клетках или при раковом заболевании (например, на ранних стадиях новообразований) по сравнению с более молодыми клетками или нераковыми клетками, соответственно. Таким образом, комплекс может быть полезен для восстановления нормальных (например, соответствующих не старым или не больным клеткам) уровней метилирования.

[0156] В некоторых вариантах осуществления, полинуклеотидная последовательность-мишень находится в сайте начала транскрипции или вблизи него. В некоторых аспектах, полинуклеотидная последовательность-мишень находится в пределах около 3000, 2500, 2000, 1500, 500, 100, 80, 70, 60, 50, 40, 30, 20, 10 или менее пар оснований (п.о.), фланкирующих сайт начала транскрипции.

[0157] В некоторых вариантах осуществления, полинуклеотидная последовательность-мишень расположена рядом с, поблизости от, или в промоторной последовательности. В некоторых аспектах, полинуклеотидная последовательность-мишень находится внутри CpG-островка. В некоторых аспектах, известно, что полинуклеотидная последовательность-мишень ассоциирована с заболеванием или состоянием, характеризующимся гипометилированием или гиперметилированием ДНК.

[0158] В некоторых вариантах осуществления, полинуклеотидная последовательность-мишень включает последовательность SEQ ID NO: 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93 или 95. В некоторых аспектах, полинуклеотидная последовательность-мишень включает аминокислотную последовательность, имеющую по меньшей мере 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93 или 95. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:37. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:39. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:41. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:43. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:45. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:47. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:49. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:51. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:53. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:55. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:57. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:59. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:61. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:63. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:65. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:67. В некоторых аспектах, полинуклеотидная последовательность-мишень представляет собой SEQ ID NO:69. В некоторых аспектах, полинуклеотидная

90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:85. В некоторых аспектах, полинуклеотидная последовательность-мишень имеет по меньшей мере 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:87. В некоторых аспектах, полинуклеотидная последовательность-мишень имеет по меньшей мере 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:89. В некоторых аспектах, полинуклеотидная последовательность-мишень имеет по меньшей мере 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:91. В некоторых аспектах, полинуклеотидная последовательность-мишень имеет по меньшей мере 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:93. В некоторых аспектах, полинуклеотидная последовательность-мишень имеет по меньшей мере 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:95.

[0160] В некоторых вариантах осуществления, комплекс включает dCas9, связанный с полинуклеотидом посредством связывания связывающей последовательности полинуклеотида и, тем самым, образования рибонуклеопротеинового комплекса. В некоторых аспектах, связывающая последовательность образует структуру шпильки. В некоторых аспектах, связывающая последовательность имеет длину 10-200 нуклеотидов, 15-150 нуклеотидов, 20-140 нуклеотидов, 30-100 нуклеотидов, 35-50 нуклеотидов, 37-47 нуклеотидов или 42 нуклеотида.

[0161] В некоторых вариантах осуществления, связывающая последовательность (например, Cas9-связывающая последовательность) взаимодействует с или связывается с белком Cas9 (например, белком dCas9), и вместе они связываются с полинуклеотидной последовательностью-мишенью, распознаваемой нацеливающей на ДНК последовательностью. Связывающая последовательность (например, Cas9-связывающая последовательность) включает два комплементарных участка нуклеотидов, которые гибридизуются друг с другом с образованием двухцепочечного дуплекса РНК (дуплекс дцРНК). Эти два комплементарных участка нуклеотидов могут быть ковалентно связаны промежуточными нуклеотидами, известными как линкеры или линкерные нуклеотиды (например, в случае одномолекулярного полинуклеотида), и гибридизуются с образованием двухцепочечного РНК-дуплекса (дуплекс дцРНК или «Cas9-связывающая шпилька») связывающей последовательности (например, Cas9-связывающей последовательности), тем самым создавая структуру стебель-петля. Альтернативно, в некоторых аспектах, два комплементарных участка нуклеотидов могут не быть ковалентно связанными, но, вместо этого, удерживаются вместе вследствие гибридизации между комплементарными последовательностями (например, двухмолекулярный полинуклеотид).

[0162] Связывающая последовательность (например, Cas9-связывающая

последовательность) может иметь длину от 10 нуклеотидов до 200 нуклеотидов, например, от 20 нуклеотидов (nt) до 150 нуклеотидов. В некоторых аспектах, связывающая последовательность имеет длину от 80 нуклеотидов (nt) до 100 нуклеотидов. Дуплекс дцРНК связывающей последовательности (например, Cas9-связывающей последовательности) может иметь длину от 6 пар оснований (п.о.) до 200 п.о. Например, дуплекс дцРНК связывающей последовательности (например, Cas9-связывающей последовательности) может иметь длину от 6 п.о. до 200 п.о., от 10 п.о. до 180 п.о., от 10 п.о. до 150 п.о., от 80 п.о. до 100 п.о. и т.п.

[0163] В некоторых вариантах осуществления, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO: 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92 или 94 или их соответствующую РНК-последовательность. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность, имеющую по меньшей мере 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92 или 94 или их соответствующей РНК-последовательностью. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:38. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:40. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:42. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:44. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:46. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:48. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:50. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:52. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:54. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:56. В некоторых аспектах, полинуклеотид, который образует комплекс со слитым белком, описанным в данном документе, включает последовательность SEQ ID NO:58. В некоторых аспектах, полинуклеотид, который

аспектов предусматривается последовательность нуклеиновой кислоты, кодирующая слитый белок, описанный в данном документе, включая его варианты осуществления и аспекты. В одном из аспектов предусматривается последовательность нуклеиновой кислоты, кодирующая слитый белок, описанный в данном документе (включая нацеливающую на ДНК последовательность), включая его варианты осуществления и аспекты. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитый белок, описанный в данном документе, включая слитые белки, имеющие аминокислотные последовательности с определенным % идентичности последовательностей, описанных в данном документе. В некоторых аспектах, нуклеиновая кислота представляет собой РНК. В некоторых аспектах, нуклеиновая кислота представляет собой матричную РНК. В некоторых аспектах, слитый белок доставляется в виде ДНК, мРНК, белка или РНП. Для РНП белком будет dCas9, а РНК будет кодировать sgРНК. Аналогично, sgРНК может быть доставлена в виде ДНК, кодирующей промотор и sgРНК, РНК, кодирующей промотор и sgРНК. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитые белки, описанные в данном документе, включая их варианты осуществления и аспекты. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитый белок любой из SEQ ID NOS:1-15. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитый белок SEQ ID NO:97. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитый белок SEQ ID NO:98. В некоторых аспектах, последовательность нуклеиновой кислоты кодирует слитый белок SEQ ID NO:99.

[0166] Дополнительно предусматривается, что последовательность нуклеиновой кислоты, кодирующая слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты, может быть включена в вектор. Таким образом, в одном из аспектов предусматривается вектор, включающий последовательность нуклеиновой кислоты, как описано в данном документе, включая его варианты осуществления и аспекты. В некоторых аспектах, вектор содержит последовательность нуклеиновой кислоты, которая кодирует слитый белок, описанный в данном документе, включая слитые белки, имеющие аминокислотные последовательности с определенным % идентичности последовательностей, описанных в данном документе. В некоторых аспектах, нуклеиновая кислота представляет собой матричную РНК. В некоторых аспектах, матричная РНК представляет собой матричный рибонуклеопротеин (РНП). В некоторых аспектах, вектор содержит последовательность нуклеиновой кислоты, которая кодирует слитый белок любой из SEQ ID NOS:1-15. В некоторых аспектах, вектор содержит последовательность нуклеиновой кислоты, которая кодирует слитый белок SEQ ID NO:97. В некоторых аспектах, вектор содержит последовательность нуклеиновой кислоты, которая кодирует слитый белок SEQ ID NO:98. В некоторых аспектах, вектор содержит последовательность нуклеиновой кислоты, которая кодирует слитый белок SEQ ID NO:99.

[0167] В некоторых вариантах осуществления, вектор дополнительно включает

полинуклеотид, причем полинуклеотид включает: (1) нацеливающую на ДНК последовательность, которая комплементарна к полинуклеотидной последовательности-мишени; и (2) связывающую последовательность дефицитного по нуклеазе фермента РНК-направляемой ДНК-эндонуклеазы. В некоторых аспектах, вектор дополнительно включает полинуклеотид, при этом полинуклеотид включает sgРНК. В некоторых аспектах, вектор дополнительно включает полинуклеотид, при этом полинуклеотид включает cr:tracrРНК. Таким образом, один или несколько векторов могут включать все необходимые компоненты для осуществления эпигеномного редактирования.

[0168] Клетки

[0169] Композиции, описанные в данном документе, могут быть включены в клетку. Внутри клетки, композиции, описанные в данном документе, включая их варианты осуществления и аспекты, могут осуществлять эпигеномное редактирование. Соответственно, в одном из аспектов предусматривается клетка, включающая слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты, нуклеиновую кислоту, как описано в данном документе, включая ее варианты осуществления и аспекты, комплекс, как описано в данном документе, включая его варианты осуществления и аспекты, или вектор, как описано в данном документе, включая его варианты осуществления и аспекты. В некоторых аспектах предусматривается клетка, включающая слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты. В некоторых аспектах предусматривается клетка, включающая нуклеиновую кислоту, как описано в данном документе, включая ее варианты осуществления и аспекты. В некоторых аспектах предусматривается клетка, включающая комплекс, как описано в данном документе, включая его варианты осуществления и аспекты. В некоторых аспектах предусматривается клетка, включающая вектор, как описано в данном документе, включая его варианты осуществления и аспекты. В некоторых аспектах, клетка представляет собой эукариотическую клетку. В некоторых аспектах, клетка представляет собой клетку млекопитающего.

[0170] Способы

[0171] Описанные в данном документе слитые белки программируют долговременную память о сайленсинге генов. Как показано в примерах, слитые белки (например, SEQ ID NOS:15, 97, 98, 99, 107, 108) программируют долговременную память сайленсинга генов, причем более 80% трансфицированных клеток обеспечивают сайленсинг репортера Snrpn-GFP и более 90% эндогенно GFP-меченного гена HIST2H2BE (H2B) через 50 дней после трансфекции (**Фиг. 1E, 1F, 8F**). Примечательно, что начиная с 10 дня после трансфекции, слитый белок (т.е. белок CRISPR-off) не детектировался, что указывает на то, что наблюдаемый сайленсинг генов не зависит от конститутивной экспрессии слитого белка (**Фиг. 1E**). Данные на **Фиг. 1E** и **1F** представляют трансфекцию ДНК, кодирующей CRISPRoff. Квалифицированный специалист может, альтернативно, трансфицировать РНП CRISPRoff, состоящий из белка CRISPRoff и sgРНК, образующей

комплекс *in vitro*. **Фиг. 1L** также демонстрирует, что транзистентная экспрессия CRISPRoff в течение менее чем 10 дней приводит к стабильному сайленсингу гена эндогенного CLTA через 15 месяцев после трансфекции. В течение 15 месяцев происходит более 450 клеточных делений, что больше числа делений большинства клеток в организме взрослого человека за всю жизнь. Сайленсинг гена достигается путем трансфекции мРНК, кодирующей слитые белки, описанные в данном документе. Таким образом, транзистентная экспрессия слитого белка приводит к эффективному сайленсингу гена (**Фиг. 8G**). Эпигенетическая память CRISPRoff с использованием слитых белков, описанных в настоящем документе, распространяется клеткой, а не устойчивой экспрессией трансгена.

[0172] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие: (i) доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы), в клетку, содержащую нуклеиновую кислоту-мишень; и (ii) доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы), в клетку, содержащую нуклеиновую кислоту-мишень; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени. В некоторых аспектах, нуклеиновая кислота-мишень содержит CpG-островок. В некоторых аспектах, нуклеиновая кислота-мишень включает не-CpG-островок. В некоторых аспектах, нуклеиновая кислота-мишень содержит CpG-островок и не-CpG-островок. «Содержит CpG-островок» или «содержит не-CpG-островок» относится к одному или нескольким CpG-островкам или не-CpG-островкам, соответственно. В некоторых аспектах, последовательность-мишень нуклеиновой кислоты содержит несколько CpG-островков (например, 2, 3, 4, 5 или больше CpG-островков). В некоторых аспектах, последовательность-мишень нуклеиновой кислоты содержит несколько не-CpG-островков (например, 2, 3, 4, 5 или больше не-CpG-островков). В некоторых аспектах, последовательность-мишень нуклеиновой кислоты не содержит CpG-островок и не содержит не-CpG-островок. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения синдрома Ангельмана у нуждающегося в этом пациента. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения вирусной инфекции у нуждающегося в этом пациента.

В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения инфекционного заболевания у нуждающегося в этом пациента.

[0173] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие: доставку полинуклеотида: (а) кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы), и (b) содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени. В некоторых вариантах осуществления, полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающие доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы), в клетку, содержащую нуклеиновую кислоту-мишень; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени. В некоторых аспектах, нуклеиновая кислота-мишень содержит CpG-островок. В некоторых аспектах, нуклеиновая кислота-мишень содержит не-CpG-островок. В некоторых аспектах, нуклеиновая кислота-мишень содержит CpG-островок и не-CpG-островок. «Содержит CpG-островок» или «содержит не-CpG-островок» относится к одному или нескольким CpG-островкам или не-CpG-островкам, соответственно. В некоторых аспектах, последовательность-мишень нуклеиновой кислоты содержит несколько CpG-островков (например, 2, 3, 4, 5 или больше CpG-островков). В некоторых аспектах, последовательность-мишень нуклеиновой кислоты содержит несколько не-CpG-островков (например, 2, 3, 4, 5 или больше не-CpG-островков). В некоторых аспектах, последовательность-мишень нуклеиновой кислоты не содержит CpG-островок и не содержит не-CpG-островок. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения синдрома Ангельмана у нуждающегося в этом пациента. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения вирусной инфекции у нуждающегося в этом пациента. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения инфекционного заболевания у нуждающегося в этом пациента. В некоторых аспектах, способ сайленсинга последовательности нуклеиновой кислоты-мишени представляет собой способ лечения нейродегенеративного заболевания у нуждающегося в этом пациента.

[0174] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекционного заболевания у нуждающегося в этом

субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение инфекционного заболевания у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекционного заболевания у нуждающегося в этом субъекта, включающие доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение инфекционного заболевания.

[0175] Термин «инфекция» или «инфекционное заболевание» относится к болезни, вызываемой такими организмами, как бактерия, вирус, грибки, или любые другие патогенные микробные агенты. В некоторых вариантах осуществления, инфекционное заболевание вызывается бактериями. В некоторых вариантах осуществления, инфекционное заболевание представляет собой связанное с бактериями заболевание (например, туберкулез, вызываемый *Mycobacterium tuberculosis*). Неограничивающие инфекционные заболевания, вызываемые бактериями, включают пневмонию (например, *Streptococcus*, *Pseudomonas*); или передающиеся с пищей заболевания (например, *Shigella*, *Campylobacter*, *Salmonella*). Инфекционные заболевания, вызываемые бактериями, также включают столбняк, брюшной тиф, дифтерию, сифилис и проказу. В некоторых вариантах осуществления, инфекционное заболевание представляет собой бактериальный вагиноз (т.е. бактерии, изменяющие вагинальную микробиоту в результате избыточного роста бактерий, которые вытесняют виды *Lactobacilli*, поддерживающие здоровые вагинальные микробные популяции) (например, дрожжевая инфекция или *Trichomonas vaginalis*); бактериальный менингит (т.е. бактериальное воспаление мозговых оболочек); бактериальную пневмонию (т.е. бактериальную инфекцию легких); инфекцию мочевыводящих путей; бактериальный гастроэнтерит; или бактериальные инфекции кожи (например, импетиго или целлюлит). В некоторых вариантах осуществления, инфекционное заболевание представляет собой инфекцию *Campylobacter jejuni*, *Enterococcus faecalis*, *Haemophilus influenzae*, *Helicobacter pylori*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Staphylococcus aureus*, *Streptococcus pneumoniae* или *Vibrio cholera*. В некоторых вариантах осуществления, инфекционное заболевание вызывается грибами. В некоторых вариантах осуществления, инфекционное заболевание вызывается вирусом.

[0176] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения бактериальной инфекции или грибковой инфекции у

нуждающегося в этом субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение бактериальной инфекции или грибковой инфекции у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения бактериальной инфекции или грибковой инфекции у нуждающегося в этом субъекта, включающие доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение бактериальной инфекции или грибковой инфекции. В некоторых вариантах осуществления, способы включают лечение бактериальной инфекции. В некоторых вариантах осуществления, способы включают лечение грибковой инфекции.

[0177] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения вирусной инфекции у нуждающегося в этом субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение вирусной инфекции у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения вирусной инфекции у нуждающегося в этом субъекта, включающие доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение вирусной инфекции. В некоторых вариантах осуществления, вирусная инфекция представляет собой флавивирусную инфекцию. В некоторых вариантах осуществления, флавивирусная инфекция вызывается вирусом Западного Нила, вирусом денге, вирусом клещевого энцефалита, вирусом желтой лихорадки или вирусом Зика. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения флавивирусной инфекции у нуждающегося в этом субъекта способами, описанными в данном документе. В некоторых вариантах осуществления, настоящее

изобретение предусматривает способы лечения инфекции вируса Западного Нила у нуждающегося в этом субъекта способами, описанными в данном документе. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекции вируса денге у нуждающегося в этом субъекта способами, описанными в данном документе. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекции вируса клещевого энцефалита у нуждающегося в этом субъекта способами, описанными в данном документе. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекции вируса желтой лихорадки у нуждающегося в этом субъекта способами, описанными в данном документе. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения инфекции вируса Зика у нуждающегося в этом субъекта способами, описанными в данном документе.

[0178] Термин «вирусная инфекция» или «вирусное заболевание» относится к заболеванию или состоянию, вызываемому вирусом. Неограничивающие примеры вирусных инфекций включают вирусные заболевания печени (например, гепатит А, В, С, D, E), инфекцию вируса герпеса (например, HSV-1, HSV-2, herpes zoster), флавивирусную инфекцию (например, инфекцию вируса Зика, инфекцию вируса денге, инфекцию вируса желтой лихорадки, инфекцию вируса Западного Нила, инфекцию вируса клещевого энцефалита), цитомегаловирусную инфекцию, респираторную вирусную инфекцию (например, аденовирусную инфекцию, грипп, тяжелый острый респираторный синдром, коронавирусную инфекцию (например, SARS-CoV-1, SARS-CoV-2, MERS-CoV, COVID-19, MERS)), желудочно-кишечную вирусную инфекцию (например, норовирусную инфекцию, ротавирусную инфекцию, астровирусную инфекцию), экзантематозную вирусную инфекцию (например, корь, опоясывающий лишай, оспа, краснуха), вирусную геморрагическую болезнь (например, лихорадку Эбола, лихорадку Ласса, лихорадку денге, желтую лихорадку), неврологическую вирусную инфекцию (например, вирусную инфекцию Западного Нила, полиомиелит, вирусный менингит, вирусный энцефалит, японский энцефалит, бешенство) и вирусную инфекцию папилломы человека.

[0179] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения тау-патологии у нуждающегося в этом субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение тау-патологии у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения тау-патологии у нуждающегося в этом субъекта, включающие доставку субъекту эффективного

количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение тау-патологии. Тау-патология относится к нейродегенеративным заболеваниям, характеризующимся патологическим накоплением тау-белка в нейрофибриллярных клубках (NFT). Заболевания с таким типичным патологическим признаком называются тауопатиями и включают, например, болезнь Альцгеймера, болезнь Паркинсона, прогрессирующий надъядерный паралич, болезнь Хантингтона, боковой амиотрофический склероз, болезнь Пика, пугилистическую деменцию и фронтотемпоральную деменцию. В некоторых вариантах осуществления, тау-патология представляет собой болезнь Альцгеймера. В некоторых вариантах осуществления, тау-патология представляет собой болезнь Паркинсона. В некоторых вариантах осуществления, тау-патология представляет собой болезнь Паркинсона, сцепленную с хромосомой 17. В некоторых вариантах осуществления, тау-патология представляет собой прогрессирующий надъядерный паралич. В некоторых вариантах осуществления, тау-патология представляет собой болезнь Хантингтона. В некоторых вариантах осуществления, тау-патология представляет собой боковой амиотрофический склероз. В некоторых вариантах осуществления, тау-патология представляет собой болезнь Пика. В некоторых вариантах осуществления, тау-патология представляет собой пугилистическую деменцию. В некоторых вариантах осуществления, тау-патология представляет собой фронтотемпоральную деменцию. В некоторых вариантах осуществления, тау-патология представляет собой фронтотемпоральную деменцию, сцепленную с хромосомой 17.

[0180] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения нейродегенеративного заболевания у нуждающегося в этом субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение нейродегенеративного заболевания у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК (например, гРНК). В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения нейродегенеративного заболевания у нуждающегося в этом субъекта, включающие доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение нейродегенеративного заболевания.

[0181] Используемый в данном документе термин «нейродегенеративное

расстройство» или «нейродегенеративное заболевание» относится к заболеванию или состоянию, при котором нарушается функция нервной системы субъекта. Примеры нейродегенеративных заболеваний, которые можно лечить слитыми белками и способами, описанными в данном документе, включают болезнь Александера, болезнь Альперса, болезнь Альцгеймера, боковой амиотрофический склероз, атаксию-телеангиэктазию, болезнь Баттена (также известную как болезнь Шпильмейера-Фогта-Шегрена-Баттена), губчатую энцефалопатию крупного рогатого скота (BSE), болезнь Канавана, синдром хронической усталости, синдром Коккейна, кортикобазальную дегенерацию, болезнь Крейтцфельда-Якоба, лобно-височную деменцию, синдром Герстмана-Штраусслера-Шейнкера, болезнь Хантингтона, ВИЧ-ассоциированную деменцию, болезнь Кеннеди, болезнь Краббе, куру, деменцию с тельцами Леви, болезнь Мачадо-Джозефа (спиноцеребеллярная атаксия типа 3), рассеянный склероз, множественную системную атрофию, миалгический энцефаломиелит, нарколепсию, нейроборрелиоз, болезнь Паркинсона, болезнь Пелицеуса-Мерцбахера, болезнь Пика, первичный боковой склероз, прионные заболевания, болезнь Рефсума, болезнь Сандхоффа, болезнь Шильдера, подострую комбинированную дегенерацию спинного мозга при пернициозной анемии, шизофрению, спиноцеребеллярную атаксию (несколько типов с различными характеристиками), спинальную мышечную атрофию, болезнь Стила-Ричардсона-Ольшевского, прогрессирующий надъядерный паралич и сухотку спинного мозга.

[0182] Синдром Ангельмана (AS) представляет собой неврологическое генетическое расстройство, вызываемое потерей экспрессии материнской копии UBE3A в головном мозге. Из-за специфического для мозга генетического импринтинга в этом локусе, отцовский UBE3A также выключается антисмысловым транскриптом. Способы, описанные в данном документе, могут быть использованы для ингибирования антисмыслового транскрипта и, тем самым, включения отцовского UBE3A, что приводит к лечению синдрома Ангельмана. См., например, Vailus et al., *Mol Ther*, 24(3):548-555 (2016). Другими словами, способы, описанные в данном документе, выключают негативный регулятор выключенного отцовского UBE3A, что равносильно включению отцовского UBE3A.

[0183] В некоторых вариантах осуществления, настоящее изобретение предусматривает способы лечения синдрома Ангельмана у нуждающегося в этом субъекта, включающие: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы); и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение синдрома Ангельмана у субъекта. В некоторых вариантах осуществления, второй полинуклеотид содержит sgРНК. В некоторых вариантах осуществления, второй полинуклеотид содержит две разных sgРНК. В некоторых вариантах осуществления, настоящее изобретение предусматривает способы

лечения синдрома Ангельмана у нуждающегося в этом субъекта, включающие доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все его варианты осуществления и аспекты (например, содержащего дефицитный по нуклеазе фермент ДНК-эндонуклеазы); тем самым, обеспечивая лечение синдрома Ангельмана у субъекта

[0184] В некоторых аспектах, указанная последовательность в последовательности-мишени нуклеиновой кислоты является метилированной. В некоторых аспектах, последовательность, расположенная в пределах около 3000, 2900, 2800, 2700, 2600, 2500, 2400, 2300, 2200, 2100, 2000, 1900, 1800, 1700, 1600, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 50, 20 или 10 пар оснований от последовательности-мишени нуклеиновой кислоты, является метилированной. Без намерения быть ограниченным какой-либо теорией, укажем, что метилирование хроматина означает, что ДНК метилируется по нуклеотиду С последовательностей CG, присутствующих в CpG-островках или не-CpG-островках (т.е. добавляются метильные метки по нуклеотиду С сайтов CG ДНК, присутствующих в CpG-островках).

[0185] Термин «репрессивные хроматиновые маркеры», используемый в данном документе, относится к модификациям хроматина, которые приводят к сайленсингу (например, снижению или ингибированию транскрипции) последовательности нуклеиновой кислоты-мишени (например, гена). Примеры репрессивных хроматиновых маркеров включают, без ограничений, моно-, ди- и/или триметилирование, ацетилирование/деацетилирование, фосфорилирование и убиквитинирование гистонов (например, H3K9, H3K27, H3K79, H2BK5).

[0186] Термин «CpG-островок» используется в его обычном значении по отношению к областям нуклеиновой кислоты, имеющим высокую частоту нуклеотидов G и C друг рядом с другом (т.е., динуклеотидам CpG). В некоторых аспектах, CpG-островок относится к области последовательности нуклеиновой кислоты, имеющей область с содержанием GC более 50%, при соотношении наблюдаемых к ожидаемым CpG более 60%. В некоторых аспектах, CpG-островок относится к области последовательности нуклеиновой кислоты, имеющей по меньшей мере 50 пар оснований, и содержанию GC более 50%, при соотношении наблюдаемых к ожидаемым CpG более 60%. В некоторых аспектах, CpG-островок относится к области последовательности нуклеиновой кислоты, имеющей по меньшей мере 100 пар оснований, и содержанию GC более 50%, при соотношении наблюдаемых к ожидаемым CpG более 60%. В некоторых аспектах, CpG-островок относится к области последовательности нуклеиновой кислоты, имеющей по меньшей мере 150 пар оснований, и содержанию GC более 50%, при соотношении наблюдаемых к ожидаемым CpG более 60%. В некоторых аспектах, CpG-островок относится к области последовательности нуклеиновой кислоты, имеющей по меньшей мере 200 пар оснований, и содержанию GC более 50%, при соотношении наблюдаемых к ожидаемым CpG более 60%. Процент CpG представляет собой отношение количества нуклеотидных оснований CpG (удвоенное число CpG) к длине. Отношение наблюдаемых

к ожидаемым CpG рассчитывается по формуле:

$$\text{Набл./Ожид. CpG} = \text{Число CpG} * N / (\text{Число C} * \text{Число G}),$$

где N=длина последовательности. См. Gardiner-Garden et al., *Journal of Molecular Biology*, 196(2):261-282 (1987)).

[0187] Фраза «нуклеиновая кислота-мишень не содержит CpG-островка» или «нуклеиновая кислота-мишень, которая не содержит CpG-островка» или «не-CpG-островок» относится к нуклеиновой кислоте-мишени, которая не содержит «CpG-островка», как этот термин определен в данном документе. Эта область может быть любой областью, кодируемой в геноме млекопитающего (например, человека). В некоторых аспектах, фраза «нуклеиновая кислота-мишень не содержит CpG-островка» относится к областям в нуклеиновой кислоте-мишени, которые не имеют нуклеотидов G и C друг рядом с другом (т.е., динуклеотидов CpG), или которые имеют низкую частоту расположения нуклеотидов G и C друг рядом с другом. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим область с содержанием динуклеотида GC менее 50%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 50%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 50%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 50%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 50%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 45%, при соотношении наблюдаемых к ожидаемым CpG менее 55%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC менее 40%, при соотношении наблюдаемых к ожидаемым CpG менее 50%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 1% до 45%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 1% до 45%, при соотношении наблюдаемых к ожидаемым CpG менее 55%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени с содержанием динуклеотида GC от 1% до 45%, при соотношении наблюдаемых к ожидаемым CpG менее 50%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 5% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых

аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 5% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 55%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 5% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 50%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 10% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 60%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 10% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 55%. В некоторых аспектах, не-CpG-островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотида GC от 10% до 40%, при соотношении наблюдаемых к ожидаемым CpG менее 50%. В некоторых аспектах, нуклеиновая кислота-мишень, не содержащая CpG-островок, имеет по меньшей мере 10 пар оснований. В некоторых аспектах, нуклеиновая кислота-мишень, не содержащая CpG-островок, имеет по меньшей мере 50 пар оснований. В некоторых аспектах, нуклеиновая кислота-мишень, не содержащая CpG-островок, имеет по меньшей мере 100 пар оснований. В некоторых аспектах, нуклеиновая кислота-мишень, не содержащая CpG-островок, имеет по меньшей мере 150 пар оснований. В некоторых аспектах, нуклеиновая кислота-мишень, не содержащая CpG-островок, имеет по меньшей мере 200 пар оснований.

[0188] В некоторых вариантах осуществления, сайленсинг относится к полному подавлению транскрипции. В некоторых аспектах, сайленсинг относится к значительному снижению транскрипции по сравнению с контрольными уровнями транскрипции.

[0189] В некоторых вариантах осуществления, первый полинуклеотид содержится в первом векторе. В некоторых аспектах, первый полинуклеотид содержится во втором векторе. В некоторых аспектах, первый вектор и второй вектор являются одинаковыми. В некоторых аспектах, первый вектор отличается от второго вектора.

[0190] В некоторых вариантах осуществления, полинуклеотид, описанный в данном документе, доставляют в клетку любым способом, известным в данной области техники, например, трансфекцией, электропорацией или трансдукцией.

[0191] Альтернативно, в некоторых аспектах предусматривается способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий доставку комплекса, как описано в данном документе, включая варианты осуществления и их аспекты, в клетку, содержащую нуклеиновую кислоту-мишень. Без намерения быть ограниченным какой-либо теорией, комплекс выключает последовательность нуклеиновой кислоты-мишени в клетке путем метилирования хроматина, содержащего последовательность нуклеиновой кислоты-мишени, и/или путем введения репрессивных хроматиновых меток в хроматин, содержащий последовательность нуклеиновой кислоты-мишени.

[0192] В некоторых вариантах осуществления, способ имеет специфичность, в 2

раза превышающую специфичность к нецелевой последовательности нуклеиновой кислоты. В некоторых аспектах, способ имеет специфичность, которая по меньшей мере в 2 раза (например, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 15-, 20-, 25-кратно) превышает специфичность к нецелевой последовательности нуклеиновой кислоты. Способы определения специфичности хорошо известны в данной области техники и включают, без ограничений, РНК-секвенирование (RNA-seq), бисульфитное секвенирование, иммунопреципитацию хроматина, проточную цитометрию и кПЦР. Таким образом, в некоторых аспектах, специфичность определяют методом РНК-секвенирования (RNA-seq). В некоторых аспектах, специфичность определяют методом бисульфитного секвенирования. В некоторых аспектах, специфичность определяют методом иммунопреципитации хроматина. В некоторых аспектах, специфичность определяют методом проточной цитометрии. В некоторых аспектах, специфичность определяют методом кПЦР.

[0193] В некоторых аспектах, комплекс доставляют в клетку любым способом, известным в данной области техники. В некоторых аспектах, комплекс доставляют в клетку посредством доставки РНК, ДНК или рибонуклеопротеина (РНП). В некоторых аспектах, комплекс доставляют в клетку посредством РНК. В некоторых аспектах, комплекс доставляют в клетку посредством ДНК. В некоторых аспектах, комплекс доставляют в клетку посредством трансфекции, вируса, липидных наночастиц (LNP) или вирусоподобных частиц. В некоторых аспектах, комплекс доставляют в клетку посредством трансфекции. В некоторых аспектах, комплекс доставляют в клетку посредством вируса. В некоторых аспектах, комплекс доставляют в клетку посредством липидных наночастиц. Способы доставки комплексов в клетку хорошо известны в данной области техники.

[0194] Варианты осуществления 1-76

[0195] Вариант осуществления 1. Слитый белок, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN и домен K \ddot{u} rperl-ассоциированного бокса.

[0196] Вариант осуществления 2. Слитый белок в соответствии с Вариантом осуществления 1, в котором первый линкер XTEN содержит от около 5 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до около 864 аминокислотных остатков.

[0197] Вариант осуществления 3. Слитый белок в соответствии с Вариантом осуществления 2, в котором первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков, и домен K \ddot{u} rperl-ассоциированного бокса.

[0198] Вариант осуществления 4. Слитый белок в соответствии с Вариантом осуществления 3, в котором первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 40

аминокислотных остатков.

[0199] Вариант осуществления 5. Слитый белок в соответствии с Вариантом осуществления 4, в котором первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков.

[0200] Вариант осуществления 6. Слитый белок в соответствии с любым из Вариантов осуществления 1-5, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен ОВ-складки, домен иммуноглобулина или домен В3.

[0201] Вариант осуществления 7. Слитый белок в соответствии с Вариантом осуществления 6, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.

[0202] Вариант осуществления 8. Слитый белок в соответствии с Вариантом осуществления 6, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.

[0203] Вариант осуществления 9. Слитый белок в соответствии с любым из Вариантов осуществления 1-8, в котором домен ДНК-метилтрансферазы содержит домен Dnmt3A.

[0204] Вариант осуществления 10. Слитый белок в соответствии с Вариантом осуществления 9, в котором домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

[0205] Вариант осуществления 11. Слитый белок в соответствии с любым из Вариантов осуществления 1-10, дополнительно содержащий эпитопную метку, 2А-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0206] Вариант осуществления 12. Слитый белок в соответствии с Вариантом осуществления 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, сигнальный пептид ядерной локализации и домен Kẗurpel-ассоциированного бокса.

[0207] Вариант осуществления 13. Слитый белок в соответствии с Вариантом осуществления 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, сигнальный пептид ядерной локализации, второй линкер XTEN и домен Kẗurpel-ассоциированного бокса.

[0208] Вариант осуществления 14. Слитый белок в соответствии с Вариантом осуществления 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер

XTEN, домен Krüppel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку.

[0209] Вариант осуществления 15. Слитый белок, имеющий по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью формулы (A); где аминокислотная последовательность формулы (A) представляет собой, от N-конца к C-концу: C₁-R³-C₂-R²-A-R¹-R⁴-B (A), где: C₁ содержит SEQ ID NO:26 или SEQ ID NO:106; R³ отсутствует или R³ содержит SEQ ID NO:27; C₂ содержит SEQ ID NO:28; R² отсутствует или R² содержит SEQ ID NO:32; A содержит SEQ ID NO:23; R¹ отсутствует или R¹ содержит SEQ ID NO:25; R⁴ отсутствует или R⁴ содержит SEQ ID NO:31; и B содержит SEQ ID NO:16, SEQ ID NO:103, SEQ ID NO:104 или SEQ ID NO:105.

[0210] Вариант осуществления 16. Слитый белок в соответствии с Вариантом осуществления 15, имеющий по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью формулы (A).

[0211] Вариант осуществления 17. Слитый белок в соответствии с Вариантом осуществления 16, имеющий по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью формулы (A).

[0212] Вариант осуществления 18. Слитый белок, имеющий по меньшей мере 85% идентичности последовательности с SEQ ID NO:97, 98, 99, 107, 108, 109 или 110.

[0213] Вариант осуществления 19. Слитый белок в соответствии с Вариантом осуществления 18, имеющий по меньшей мере 90% идентичности последовательности с SEQ ID NO:97, 98, 99, 107, 108, 109 или 110.

[0214] Вариант осуществления 20. Клетка, содержащая слитый белок в соответствии с любым из Вариантов осуществления 1-19.

[0215] Вариант осуществления 21. Клетка в соответствии с Вариантом осуществления 20, представляющая собой эукариотическую клетку, клетку млекопитающего или стволовую клетку.

[0216] Вариант осуществления 22. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий: (i) доставку первого полинуклеотида, кодирующего слитый белок в соответствии с любым из Вариантов осуществления 1-19, в клетку, содержащую нуклеиновую кислоту-мишень; и (ii) доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени.

[0217] Вариант осуществления 23. Способ в соответствии с Вариантом осуществления 22, в котором нуклеиновая кислота-мишень содержит CpG-островок.

[0218] Вариант осуществления 24. Способ в соответствии с Вариантом осуществления 22, в котором нуклеиновая кислота-мишень не содержит CpG-островка.

[0219] Вариант осуществления 25. Способ в соответствии с любым из Вариантов осуществления 22-24, в котором второй полинуклеотид содержит sgРНК.

[0220] Вариант осуществления 26. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий: (i) доставку первого

полинуклеотида, кодирующего слитый белок, в клетку, содержащую нуклеиновую кислоту-мишень, причем нуклеиновая кислота-мишень не содержит CpG-островка; при этом слитый белок содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы, домен K \ddot{u} rrel-ассоциированного бокса и домен ДНК-метилтрансферазы; и (ii) доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени в клетке.

[0221] Вариант осуществления 27. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, содержащий дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы, домен K \ddot{u} rrel-ассоциированного бокса и домен ДНК-метилтрансферазы; и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания.

[0222] Вариант осуществления 28. Способ в соответствии с Вариантом осуществления 26 или 27, в котором второй полинуклеотид содержит sgРНК.

[0223] Вариант осуществления 29. Способ в соответствии с любым из Вариантов осуществления 26-28, в котором слитый белок содержит, от N-конца к С-концу, домен ДНК-метилтрансферазы, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен K \ddot{u} rrel-ассоциированного бокса.

[0224] Вариант осуществления 30. Способ в соответствии с любым из Вариантов осуществления 26-29, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Woz3, домен ОВ-складки, домен иммуноглобулина или домен В3.

[0225] Вариант осуществления 31. Способ в соответствии с Вариантом осуществления 30, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.

[0226] Вариант осуществления 32. Способ в соответствии с Вариантом осуществления 30, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.

[0227] Вариант осуществления 33. Способ в соответствии с любым из Вариантов осуществления 26-32, в котором домен ДНК-метилтрансферазы содержит домен Dnmt3A.

[0228] Вариант осуществления 34. Способ в соответствии с Вариантом осуществления 33, в котором домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

[0229] Вариант осуществления 35. Способ в соответствии с любым из Вариантов

осуществления 26-34, в котором dCas9 ковалентно связан с доменом Dnmt3A с помощью пептидного линкера, и в котором домен Dnmt3A ковалентно связан с доменом Krüppel-ассоциированного бокса с помощью пептидного линкера.

[0230] Вариант осуществления 36. Способ в соответствии с Вариантом осуществления 35, в котором пептидный линкер представляет собой линкер XTEN.

[0231] Вариант осуществления 37. Способ в соответствии с любым из Вариантов осуществления 26-28, в котором слитый белок содержит, от N-конца к C-концу, Krüppel-ассоциированный бокс, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен ДНК-метилтрансферазы.

[0232] Вариант осуществления 38. Способ в соответствии с Вариантом осуществления 37, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3.

[0233] Вариант осуществления 39. Способ в соответствии с Вариантом осуществления 38, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.

[0234] Вариант осуществления 40. Способ в соответствии с Вариантом осуществления 38, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.

[0235] Вариант осуществления 41. Способ в соответствии с любым из Вариантов осуществления 37-39, в котором домен ДНК-метилтрансферазы содержит домен Dnmt3A.

[0236] Вариант осуществления 42. Способ в соответствии с Вариантом осуществления 41, в котором домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

[0237] Вариант осуществления 43. Способ в соответствии с любым из Вариантов осуществления 39-42, в котором dCas9 ковалентно связан с доменом Dnmt3A с помощью пептидного линкера, и в котором домен Krüppel-ассоциированного бокса ковалентно связан с dCas9 с помощью пептидного линкера.

[0238] Вариант осуществления 44. Способ в соответствии с Вариантом осуществления 43, в котором пептидный линкер представляет собой линкер XTEN.

[0239] Вариант осуществления 45. Способ в соответствии с любым из Вариантов осуществления 37-44, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы ковалентно связан с доменом Krüppel-ассоциированного бокса с помощью пептидного линкера.

[0240] Вариант осуществления 46. Способ в соответствии с любым из Вариантов осуществления 37-45, в котором дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы ковалентно связан с доменом ДНК-метилтрансферазы с помощью пептидного линкера.

[0241] Вариант осуществления 47. Способ в соответствии с любым из Вариантов осуществления 37-46, в котором домен Kgrüppel-ассоциированного бокса ковалентно связан с доменом ДНК-метилтрансферазы с помощью пептидного линкера.

[0242] Вариант осуществления 48. Способ в соответствии с любым из Вариантов осуществления 26-47, в котором слитый белок дополнительно содержит эпитопную метку, 2А-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0243] Вариант осуществления 49. Способ в соответствии с любым из Вариантов осуществления 26-47, в котором слитый белок дополнительно содержит сигнальный пептид ядерной локализации.

[0244] Вариант осуществления 50. Способ в соответствии с любым из Вариантов осуществления 26-28, в котором слитый белок имеет (having) по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью формулы (A); где аминокислотная последовательность формулы (A) представляет собой, от N-конца к C-концу: C₁-R³-C₂-R²-A-R¹-R⁴-B (A), где: C₁ содержит SEQ ID NO:26 или SEQ ID NO:106; R³ отсутствует или R³ содержит SEQ ID NO:27; C₂ содержит SEQ ID NO:28; R² отсутствует или R² содержит SEQ ID NO:32; A содержит SEQ ID NO:23; R¹ отсутствует или R¹ содержит SEQ ID NO:25; R⁴ отсутствует или R⁴ содержит SEQ ID NO:31; и B содержит SEQ ID NO:16, SEQ ID NO:103, SEQ ID NO:104 или SEQ ID NO:105.

[0245] Вариант осуществления 51. Способ в соответствии с Вариантом осуществления 50, в котором слитый белок имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью формулы (A).

[0246] Вариант осуществления 52. Способ в соответствии с Вариантом осуществления 51, в котором слитый белок имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью формулы (A).

[0247] Вариант осуществления 53. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий: (i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок в соответствии с любым из Вариантов осуществления 1-14; и (ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у субъекта.

[0248] Вариант осуществления 54. Способ в соответствии с любым из Вариантов осуществления 26-53, в котором инфекционное заболевание представляет собой вирусное инфекционное заболевание.

[0249] Вариант осуществления 55. Способ в соответствии с Вариантом осуществления 54, в котором инфекционное заболевание представляет собой флавивирусное инфекционное заболевание.

[0250] Вариант осуществления 56. Слитый белок, содержащий, от N-конца к C-

концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN и домен K \ddot{u} rrel-ассоциированного бокса.

[0251] Вариант осуществления 57. Слитый белок в соответствии с Вариантом осуществления 56, в котором первый линкер XTEN содержит от около 5 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до около 864 аминокислотных остатков.

[0252] Вариант осуществления 58. Слитый белок в соответствии с Вариантом осуществления 57, в котором первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков.

[0253] Вариант осуществления 59. Слитый белок в соответствии с Вариантом осуществления 58, в котором первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 40 аминокислотных остатков.

[0254] Вариант осуществления 60. Слитый белок в соответствии с Вариантом осуществления 59, в котором первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков.

[0255] Вариант осуществления 61. Слитый белок в соответствии с любым из Вариантов осуществления 56-60, в котором дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или эффектор, подобный активатору транскрипции.

[0256] Вариант осуществления 62. Слитый белок в соответствии с Вариантом осуществления 62, в котором дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца.

[0257] Вариант осуществления 63. Слитый белок в соответствии с Вариантом осуществления 62, в котором дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой эффектор, подобный активатору транскрипции.

[0258] Вариант осуществления 64. Слитый белок в соответствии с любым из Вариантов осуществления 56-63, в котором домен ДНК-метилтрансферазы содержит домен Dnmt3A.

[0259] Вариант осуществления 65. Слитый белок в соответствии с Вариантом осуществления 64, в котором домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

[0260] Вариант осуществления 66. Слитый белок в соответствии с любым из Вариантов осуществления 56-65, дополнительно содержащий эпитопную метку, 2A-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

[0261] Вариант осуществления 67. Слитый белок в соответствии с любым из Вариантов осуществления 56-65, дополнительно содержащий сигнальный пептид ядерной

локализации.

[0262] Вариант осуществления 68. Слитый белок в соответствии с Вариантом осуществления 56, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, домен K \ddot{u} rrel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку.

[0263] Вариант осуществления 69. Клетка, содержащая слитый белок в соответствии с любым из Вариантов осуществления 56-68.

[0264] Вариант осуществления 70. Клетка в соответствии с Вариантом осуществления 69, представляющая собой эукариотическую клетку, клетку млекопитающего или стволовую клетку.

[0265] Вариант осуществления 71. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий доставку первого полинуклеотида, кодирующего слитый белок в соответствии с любым из Вариантов осуществления 56-68, в клетку, содержащую нуклеиновую кислоту-мишень; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени.

[0266] Вариант осуществления 72. Способ в соответствии с Вариантом осуществления 71, в котором нуклеиновая кислота-мишень содержит CpG-островок.

[0267] Вариант осуществления 73. Способ в соответствии с Вариантом осуществления 71, в котором нуклеиновая кислота-мишень не содержит CpG-островка.

[0268] Вариант осуществления 74. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок в соответствии с любым из Вариантов осуществления 56-68; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания.

[0269] Вариант осуществления 75. Способ в соответствии с Вариантом осуществления 74, в котором инфекционное заболевание представляет собой вирусное инфекционное заболевание.

[0270] Вариант осуществления 76. Способ в соответствии с Вариантом осуществления 75, в котором вирусное инфекционное заболевание представляет собой флавивирусное инфекционное заболевание.

ПРИМЕРЫ

[0271] Варианты осуществления и аспекты в данном документе дополнительно иллюстрируются следующими примерами. Указанные примеры предназначены только для иллюстрации вариантов осуществления и аспектов и не должны толковаться как ограничивающие объем данного документа.

[0272] Описанная здесь технология позволяет, в частности, обеспечивать постоянный сайленсинг генов в клетках млекопитающих без образования двухцепочечных

разрывов ДНК в геноме хозяина. В некоторых вариантах осуществления, центральный компонент представляет собой одну полипептидную цепь, состоящую из каталитически неактивного Cas9 (dCas9), слитого с Dnmt3A, Dnmt3L и доменом KRAB (в данном документе называется «белок все-в-одном» (all-in-one)). В некоторых вариантах осуществления, центральный компонент представляет собой одну полипептидную цепь, состоящую из домена цинкового пальца, слитого с Dnmt3A, Dnmt3L и доменом KRAB (в данном документе называется «белок все-в-одном» (all-in-one)). Такие слитые белки, предусматриваемые в данном документе, могут быть нацелены на конкретный конкретный сайт в геноме млекопитающего с помощью одиночной направляющей РНК (sgРНК) и могут добавлять ДНК-метиляцию и/или репрессивные хроматиновые метки в сайте. Результатом этого является сайленсинг генов, наследуемый при последующих клеточных делениях. Слитые белки, предложенные в данном документе (и sgРНК), экспрессируются только транзистентно, что позволяет обойтись без использования вирусных методов доставки для индукции постоянного сайленсинга.

[0273] Слитые белки, предложенные в данном документе, обеспечивают надежный долговременный или постоянный сайленсинг экспрессии эндогенных генов путем эпигеномного редактирования, а не редактирования генома. Мишенями могут быть оба аллеля гена или может быть селективно выбран один патогенный аллель. Преимущество слитого белка, предложенного в данном документе, заключается в том, что эпигенетическое редактирование является обратимым и поэтому по своей сути более безопасным, чем редактирование генома. Таким образом, слитые белки, предложенные в данном документе, пригодны для профилактического применения. Например, сайленсинг генов может обеспечить экстренную защиту от инфекции/биологического токсина, а затем быть отменен после исчезновения риска инфекции или интоксикации. Таким образом, слитые белки, предложенные в данном документе, полезны для защиты от вирусного агента или токсина, который попадает в клетку в результате взаимодействия с белком, необходимым для длительного функционирования органа или гомеостаза. Слитые белки, предложенные в данном документе, пригодны для использования в терапевтических препаратах, основанных на редактировании генома.

[0274] Постоянный сайленсинг генов в клетках млекопитающих может быть осуществлен с помощью двух компонентов: одной полипептидной цепи, состоящей из dCas9, слитого с тремя эпигенетическими модуляторами, и одиночной направляющей РНК, которая направляет белок на определенный сайт в геноме хозяина. В некоторых вариантах осуществления, постоянный сайленсинг генов в клетках млекопитающих может быть осуществлен с помощью двух компонентов: одной полипептидной цепи, состоящей из домена цинкового пальца, слитого с тремя эпигенетическими модуляторами, и одиночной направляющей РНК, которая направляет белок на определенный сайт в геноме хозяина. В некоторых вариантах осуществления, компоненты экспрессируются в клетке-хозяине только транзистентно, что снижает токсичность и нецелевые события. В некоторых вариантах осуществления, слитый белок, предложенный в данном документе, не

индуцирует разрывы ДНК в клетке-хозяине для постоянного сайленсинга генов. В некоторых вариантах осуществления, эпигенетические метки, добавленные в геномный сайт, представляющий интерес, являются обратимыми, что позволяет устранять любые возможные нецелевые события.

[0275] Пример 1

[0276] Достижения в области редактирования генов преобразили нашу способность модифицировать геном человека. В частности, CRISPR (короткие палиндромные повторы, регулярно расположенные группами)-Cas9 (CRISPR-ассоциированный белок 9) и другие системы CRISPR могут быть запрограммированы с помощью одиночной направляющей (sg)РНК для внесения разрывов ДНК в определенном сайте для инактивации функции гена или для стимулирования точного редактирования ДНК путем гомологически-направленной репарации (Knott and Doudna, 2018). Кроме того, стратегии редактирования оснований и праймеров позволяют точно изменять последовательность ДНК, но обычно основаны на одном или нескольких одноцепочечных надрезах ДНК (Anzalone et al., 2020). Эти технологии были оптимизированы для целенаправленного изменения базовой последовательности ДНК и потому идеально подходят для репарации или введения патогенных мутаций. Однако зависимость от эндогенного механизма репарации ДНК создает проблемы, поскольку сложность этих путей может затруднить ограничение результата одним желаемым изменением (Yeh et al., 2019).

[0277] Альтернативным способом модуляции функции гена является переписывание эпигенетического ландшафта для контроля экспрессии гена без изменения базовой последовательности ДНК. Слияние белкового каркаса или ферментных доменов с каталитически неактивным dCas9 может усиливать (CRISPRa) или подавлять (CRISPRi) транскрипцию в клетках млекопитающих (Holtzman and Gersbach, 2018; Xu and Qi, 2019). Программируемое эпигеномное редактирование является настраиваемым, обратимым и не требует разрывов ДНК, позволяя эффективно обойти клеточную токсичность, связанную с редактированием генов (Jost et al., 2020). Однако существующие технологии программируемого редактирования эпигенома обычно полагаются на конститутивную экспрессию слитых белков dCas9 для поддержания транскрипционного контроля. Таким образом, эти методики остаются менее пригодными для терапевтической инженерии клеток и организмов.

[0278] Недавние работы продемонстрировали возможность редактирования эпигенома для записи стабильной транскрипционной программы, которая запоминается и распространяется клетками человека без конститутивной экспрессии программируемых эпигенетических модуляторов (Amabile et al., 2016; Bintu et al., 2016; Park et al., 2019; Van et al., 2021). В частности, Amabile et al. показали, что можно наследственно выключать человеческие гены путем рекрутмента коктейля из доменов ДНК-метилтрансферазы и КРАВ. Однако на сегодняшний день лишь небольшое количество эндогенных локусов человека было протестировано на предмет выключения посредством записи в эпигенетическую память (Amabile et al., 2016; O'Geen et al., 2019; Tarjan et al., 2019). Более

того, предыдущие разработки программируемых эпигенетических сайленсеров используют два или три слитых белка для каждого гена-мишени, что экспериментально обременительно - особенно при мультиплексном нацеливании генов - и усложняет оптимальную стратегию нацеливания генов. Кроме того, слияние доменов KRAB с DNMT3A и DNMT3L на основе TALE приводило к низкой эффективности долговременного сайленсинга генов (Mlambo et al., 2018). Таким образом, остается неясным, насколько универсальными являются эти подходы для установления наследственного заглушения генов, и существуют ли геномные особенности, необходимые для записи и поддержания наследуемого эпигенетического сайленсинга. Мы предположили, что эпигенетический редактор, состоящий из одного «мертвого» продукта слияния Cas9, позволит нам более широко исследовать биологию и полезность наследуемого эпигенетического сайленсинга генов.

[0279] Здесь мы представляем дизайн, разработку и характеризацию CRISPRoff - программируемого белка для записи в эпигенетическую память, который может обеспечивать долговременный сайленсинг экспрессии генов. Транзиентная экспрессия CRISPRoff записывает эпигенетическую программу, которую клетки человека поддерживают в течение более 450 клеточных делений, что свидетельствует о стабильности и наследуемости этой формы сайленсинга генов. Используя полногеномные скрининги CRISPRoff, мы показали, что этот подход может долговременно и специфически выключать большую часть генов, кодирующих белки, и имеет широкое окно нацеливания на промоторы генов. Неожиданно, канонические аннотации CpG-островков не являются необходимыми для стабильного сайленсинга генов с помощью CRISPRoff. Наконец, мы продемонстрировали, что CRISPRoff можно использовать для выключения энхансеров и инженерии программ сайленсинга генов в стволовых клетках человека, сохраняющихся до дифференцировки в нейроны. В целом, эта система позволяет нам более широко исследовать биологические правила, лежащие в основе эпигенетического сайленсинга, и обеспечивает надежный инструмент для контроля экспрессии генов, нацеливания энхансеров и изучения принципов эпигенетического наследования.

[0280] Результаты

[0281] Рациональный дизайн единого слитого редактора эпигеномной памяти

[0282] Мы разработали программируемый белок-редактор эпигенома на основе CRISPR, названный CRISPRoff-V1, состоящий из белковых доменов ZNF10 KRAB, Dnmt3A (D3A) и DNMT3L (D3L), слитых с каталитически неактивным dCas9 *S. pyogenes* (Фиг. 1A). Чтобы проверить, может ли транзиентный импульс эпигенетического редактирования CRISPRoff долговременно выключать экспрессию генов, мы транзиентно котрансфицировали клетки HEK293T, стабильно экспрессирующие чувствительный к ДНК-метилированию GFP-репортер GAPDH-Snrpn с использованием CRISPRoff-V1, dCas9-KRAB (CRISPRi) или dCas9-D3A-3L, вместе с sgPHK, нацеленными на синтетический промотор GAPDH-Snrpn (Liu et al., 2016; Stelzer et al., 2015) (Фиг. 1B). Все

три белка эпигенетического редактора транзientно выключали GFP-репортер (**Фиг. 1C**). Как и ожидалось при транзientной трансфекции, экспрессия каждого белка эпигенетического редактора со временем была утрачена, что для dCas9-KRAB и dCas9-D3A-3L привело к восстановлению экспрессии GFP. В отличие от них, в случае CRISPRoff-V1 память о сайленсинге генов и метилирование CpG-островков (CGI) сохранялись в течение длительного времени после утраты экспрессии CRISPRoff (**Фиг. 1C**).

[0283] Сайленсинг с помощью CRISPRoff-V1 оказался метастабильным, поскольку экспрессия гена-репортера постепенно увеличивалась со временем (**Фиг. 1C**). Для стабилизации памяти сайленсинга генов мы закодировали CRISPRoff с устойчивыми к протеолизу линкерами, чтобы минимизировать протеолиз, который может привести к несвязанному D3A-D3L и нецелевому метилированию ДНК, как сообщалось ранее (Galonska et al., 2018; Hofacker et al., 2020). Варианты CRISPRoff программируют переменный долговременный сайленсинг генов (**Фиг. 8A-8B**). Во-вторых, мы предположили, что размещение D3A-3L на N-конце dCas9 обеспечит Dnmt3A оптимальный доступ к сайтам CpG для метилирования ДНК (Zhang et al., 2018) (**Фиг. 8C**). Все сконструированные нами эпигенетические редакторы CRISPRoff-V2 имели схожую стабильность сайленсинга генов, поэтому во всех последующих экспериментах мы использовали CRISPRoff-V2.1 (**Фиг. 8D-8F**).

[0284] Транзientная экспрессия CRISPRoff-V2 программировала долговременную память сайленсинга генов в течение по меньшей мере 50 дней после трансфекции, при этом более 80% трансфицированных клеток выключали репортер Snrpn-GFP и более 90% выключали эндогенно меченый GFP ген HIST2H2BE (H2B) (**Фиг. 1E-1F** и **8G**). Сайленсинг H2B сопровождался метилированием CGI (**Фиг. 1G**). Примечательно, что начиная с 10 дня после трансфекции, белок CRISPR-off не обнаруживался (**Фиг. 1E**). Трансфекция мРНК CRISPRoff-V2 также выключала экспрессию эндогенного GFP-меченого гена CLTA, подтверждая, что транзientная экспрессия CRISPRoff приводит к эффективному сайленсингу генов (**Фиг. 8H**). Эти результаты показывают, что эпигенетическая память CRISPRoff не зависит от устойчивой экспрессии трансгена.

[0285] Для дальнейшего сравнения CRISPRoff-V1 и V2 мы выключили три белка, локализованных на поверхности клетки (ITGB1, CD81, CD151), которые не требуются для пролиферации или выживания клеток. Трансфекция CRISPRoff-V1 с одной sgРНК выключала каждый из генов-мишеней в определенной части клеток, а пул из трех sgРНК улучшал сайленсинг ITGB1 и CD81 (**Фиг. 1H**). CRISPRoff-V2 улучшал сайленсинг каждого гена со степенью сайленсинга по меньшей мере 80% через 3 недели после трансфекции (**Фиг. 1H**).

[0286] Долговременный и мультиплексный сайленсинг эндогенных генов

[0287] Мы продемонстрировали эффективность CRISPRoff-V2 в различных типах клеток, а именно в индуцированных плюрипотентных стволовых клетках (iPSC), HeLa, U2OS и K562 (в качестве индуцируемой доксициклином системы) (**Фиг. 9A-9D**). Далее

мы показали, что CRISPRoff может быть запрограммирован ортогональными ДНК-связывающими белками: dCas9 *S. aureus* (dSauCas9) и dCas12a *Lachnospiracea* (dLbCas12a) (Фиг. 9E-9F). Сайленсинг с помощью dLbCas12a улучшался, когда три sgPНК (PНК CRISPR) были закодированы как один транскрипт, который может процессироваться dLbCas12a в отдельные sgPНК, что указывает на возможность мультиплексного сайленсинга генов.

[0288] Для исследования мультиплексного сайленсинга эндогенных генов человека с помощью CRISPRoff на основе *S. pyogenes*, мы выбрали в качестве мишеней ITGB1, CD81 и CD151 в комбинациях по два, три и четыре гена (Фиг. 1I-1K и 8I-8K). Через 30 дней после трансфекции мы наблюдали надежный мультиплексный сайленсинг каждой комбинации генов (Фиг. 1I). Мы увидели, что клетки с одним выключенным геном имеют более высокую вероятность сайленсинга других генов-мишеней (Фиг. 1I-1K, 8I, и 8K). Например, при совместном нацеливании на ITGB1 и CD81, клетки с успешно выключенным ITGB1 имели в 25 раз больший процент клеток, в которых также был выключен CD81, по сравнению с клетками, в которых не удалось выключить ITGB1 (Фиг. 1I).

[0289] Для измерения долгосрочного поддержания эпигенетической памяти мы нацеливались на эндогенный ген CLTA и отслеживали экспрессию CLTA в одноклеточных клонах (Leonetti et al., 2016a). Примечательно, что через 15 месяцев после трансфекции или после ок. 450 делений клеток, 38 из 39 клонов сохраняли сайленсинг CLTA (Фиг. 1L).

[0290] Эпигеномное редактирование является высокоспецифичным

[0291] Для оценки специфичности CRISPRoff мы провели PНК-секвенирование клеток через 33 дня после трансфекции CRISPRoff и sgPНК, нацеленными на ITGB1, CD81 и CD151, или sgPНК отрицательного контроля. Сравнение нетрансфицированных клеток с клетками, трансфицированными CRISPRoff, показало минимальный нокдаун нецелевых генов (Фиг. 2A). Нацеливание на ITGB1, CD81 и CD151 с помощью CRISPRoff было высокоспецифичным и показало почти полную репрессию гена-мишени (Фиг. 2B-2D и 10A). Анализы методом PНК-секвенирования трех других клеточных линий с эндогенно меченым GFP геном, подавляемым CRISPRoff (RAB11A, CLTA и H2B), также показали надежный и высокоспецифичный нокдаун транскриптов (Фиг. 10B-10D). Анализ соседних генов в окне размером в 1 мегабазу от гена-мишени не выявил значительных изменений в экспрессии генов (Фиг. 10E-10F). При анализе наборов данных, полученных в результате CRISPRoff-нацеливания на ITGB1, CD81 и CD151, мы наблюдали 1-3 нецелевых транскрипта с \log_2 (кратность изменений) > 2 и скорректированным p -значением $< 0,5$ в каждом эксперименте по нокдауну гена, хотя и с гораздо меньшей величиной по сравнению с геном-мишенью. Дифференциальная экспрессия нецелевых транскриптов может быть обусловлена косвенными эффектами, связанными с нокдауном гена-мишени или нецелевой активностью CRISPRoff.

[0292] Мы оценили специфичность ДНК-метилирования CRISPRoff методом полногеномного бисульфитного секвенирования (WGBS) через 30 дней после сайленсинга

CLTA (**Фиг. 10G**). Мы обнаружили единичное доминантное усиление ДНК-метилирования в промоторе CLTA в окне размером 1,5 кб, перекрывающем промотор CLTA, что подчеркивает высокую специфичность CRISPRoff (**Фиг. 2E**). Мы не обнаружили распространения метилирования ДНК на ближайшие соседние гены (**Фиг. 2F**).

[0293] В согласии с результатами предыдущих анализов ДНК-метилирования в клетках, обработанных эпигенетическими редакторами на основе DNMT (ДНК-метилтрансфераза), мы наблюдали умеренно повышенное глобальное ДНК-метилирование в CRISPRoff-трансфицированных клетках (<2%; **Фиг. 10H-10I**) (Galonska et al., 2018; O'Geen et al., 2019). Однако глобальное метилирование ДНК также варьировалось между клонами клеток, не подвергнутых воздействию CRISPRoff в аналогичной степени, и различия в паттернах локального метилирования ДНК в нецелевых участках ДНК между клонами клеток были больше, чем любые умеренные неспецифические изменения метилирования, наблюдаемые после экспрессии CRISPRoff (**Фиг. S3I-S3J**). Чтобы выяснить, может ли зависимое или независимое от sgPНК CRISPRoff нецелевое метилирование ДНК изменять экспрессию генов, мы проверили 10 наиболее дифференциально метилированных участков ДНК. Мы не обнаружили никаких различий в транскрипции генов в 10 лучших нецелевых дифференциально метилированных областях или рядом с ними (**Фиг. 10K-10L**). Таким образом, наши данные WGBS показывают, что любые различия в нецелевом метилировании встречаются нечасто и вряд ли могут модулировать экспрессию генов или клеточные фенотипы.

[0294] Наконец, мы использовали метод ChIP-seq (иммунопреципитация хроматина с последующим секвенированием фрагментов ДНК) для определения профиля изменений в репрессивных модификациях H3K9me3 после CRISPRoff-нацеливания на ген HIST2H2BE (H2B). Мы обнаружили сильное увеличение H3K9me3 в области ок. 5 кб в промоторе H2B на 5 день после трансфекции CRISPRoff, которое сохранялось в течение 30 дней, демонстрируя стабильное распространение H3K9me3 и ДНК-метилирования, как описано ниже (**Фиг. 2G**). Сравнение результатов применения H2B-нацеливающих и ненацеливающих sgPНК показало, что наиболее значительный прирост H3K9me3 происходил в локусе H2B и трех соседних генах: HIST2H2AC, HIST2H2AB и BOLA1 (**Фиг. 2H**). Мы обнаружили нокдаун экспрессии HIST2H2AC, в то время как прочтения секвенирования, отображающие HIST2H2AB, не были обнаружены в наших данных RNA-seq. Мы не обнаружили нокдауна транскрипции BOLA1. Эти данные в сочетании с данными полногеномного бисульфитного секвенирования, которые продемонстрировали ограничение метилирования CpG, показывают транскрипционную и эпигеномную специфичность CRISPRoff, а также документируют локальное эпигенетическое распространение и поддержание от сайта-мишени создания.

[0295] Полногеномное нацеливание CRISPRoff

[0296] Простота дизайна CRISPRoff побудила нас провести объединенные полногеномные скрининги, чтобы определить универсальный характер его применения

для сайленсинга генов в геноме человека. Мы разработали библиотеку sgPНК, которая нацелена на более чем 20000 кодирующих белок генов и включает около 1000 ненацеливающих sgPНК (Horlbeck et al., 2016a; Replogle et al., 2020). Мы сконструировали библиотеку sgPНК для кодирования двух уникальных протоспейсеров, нацеленных на один и тот же ген в каждом лентивирусном векторе, поскольку наши эксперименты показали улучшение активности CRISPRoff при использовании нескольких sgPНК, нацеленных на один и тот же ген (Replogle et al., 2020)(Фиг. 4А).

[0297] Мы проводили объединенные скрининги на основе роста, поскольку наборы данных о существенности генов доступны из предыдущих работ по функциональной геномике, что позволило нам сравнить эффективность CRISPRoff с другими полногеномными отсеивающими скринингами. Для проведения объединенного скрининга CRISPRoff мы упаковали библиотеку sgPНК в лентивирусные частицы, затем трансдуцировали и отобрали клетки НЕК293Т таким образом, чтобы в среднем каждая клетка экспрессировала один вектор sgPНК. Затем мы транзистентно трансфицировали этот пул клеток плазидами, кодирующими CRISPRoff, и отсортировали клетки, экспрессирующие белок CRISPRoff. Мы собрали популяцию трансфицированных CRISPRoff клеток в качестве образца нулевого момента времени (T0) и продолжили пассирование популяции трансфицированных CRISPRoff клеток в течение по меньшей мере 10 удвоений клеток (T10), после чего провели глубокое секвенирование геномной ДНК в оба момента времени для считывания и количественного определения последовательностей sgPНК в качестве косвенного показателя количества клеток. Мы пришли к выводу, что sgPНК, обедненные в популяции T10 по сравнению с T0, являются активными, поскольку эти sgPНК эффективно выключают экспрессию существенных генов и выпадают из популяции (Фиг. 4В). В качестве контроля мы параллельно провели идентичный скрининг с вариантом CRISPRoff, кодирующим мутацию Dnmt3A^{E765A}, который каталитически неактивен и поэтому не способен поддерживать долговременный сайленсинг генов и вместо этого отражает транзистентный эффект сайленсинга CRISPRi (Фиг. 4С). Сравнивая эти два скрининга, мы идентифицировали sgPНК в скрининге CRISPRoff, которые выключают экспрессию генов зависимым от метилирования ДНК образом.

[0298] Анализ оценки фенотипа (γ , при этом более отрицательная оценка указывает на более сильный дефект роста) для каждого гена показал, что экспрессия CRISPRoff воспроизводимо приводила к более выраженным фенотипам дефекта роста по сравнению с мутантом CRISPRoff (Фиг. 4D-4E, 11A-11C). Большой набор генов выявил резкие дефекты роста, которые были специфичны для CRISPRoff-опосредованного нокдауна, что указывает на стойкий эффект сайленсинга генов CRISPRoff. Мы также обнаружили подмножество генов с сопоставимыми фенотипами между двумя скринингами, вероятно, из-за их существенности при транзистентном нокдауне с помощью мутанта CRISPRoff (Фиг. 4D). Мы оценили специфичность сайленсинга во всех скринингах, проанализировав оценки фенотипа контрольных sgPНК. Почти все отрицательные контрольные sgPНК или

sgРНК, нацеленные на неэкспрессируемые гены (ольфакторные гены или гены Y-хромосомы), имели фенотип, измерить который было практически невозможно (1% с $\gamma < -0,2$) (Фиг. 4D).

[0299] С целью оценки универсальности CRISPRoff для глушения генов, мы проанализировали оценки фенотипа генов, которые, как мы ожидали, будут иметь фенотипы роста при нокдауне. Анализ генов, связанных с репликацией ДНК и рибосомой, которые, как предполагается, крайне необходимы для пролиферации клеток, относился к наиболее тяжелым фенотипам роста (Фиг. 11D-11E). Затем мы проанализировали фенотипы обычных существенных генов, которые необходимы для клеточной пролиферации или выживания в большинстве линий раковых клеток (Meyers et al., 2017). Дефекты роста этих генов были гораздо более сильно выражены у CRISPRoff (медиана $\gamma = -0,2$) по сравнению с мутантом CRISPRoff (медиана $\gamma = -0,05$), тогда как большинство несущественных генов не создавали фенотипов роста (Фиг. 4E). Мутант CRISPRoff приводил к слабым оценкам фенотипа из-за отсутствия зависимого от метилирования ДНК стойкого сайленсинга генов. В целом, скрининг CRISPRoff дал высокий процент положительных результатов определения существенных генов с низким количеством ложноположительных результатов, что позволяет предположить, что CRISPRoff способен выключить большинство генов в геноме человека (Фиг. 4F).

[0300] Программируемые редакторы эпигенома могут инициировать эпигенетические метки, которые распространяются от места образования (Hathaway et al., 2012; Stepper et al., 2017). Мы задались вопросом, не обусловлены ли некоторые определения генов методом CRISPRoff «эффектом соседнего гена», вызванным распространением метилирования ДНК от близлежащего существенного гена. Мы каталогизировали генные «совпадения» (определяемые как гены с оценкой фенотипа $-0,2$ или ниже) и определяли их линейное расстояние по геному от ближайшего генного совпадения. Хотя часть совпадений генов находилась в пределах расстояний в окне размером 1 кб, большинство совпадений CRISPRoff находились на расстоянии более 10 кб от ближайшего гена (Фиг. 4G). Поскольку CpG-островки в основном ограничены областью в окне размером 1-2 кб (Deaton and Bird, 2011), мы предположили, что большинство наблюдаемых нами генных совпадений специфичны для промотора гена-мишени, что согласуется со специфичностью, продемонстрированной в наших экспериментах с CRISPRoff методами RNA-seq, WGBS и ChIP-seq.

[0301] CRISPRoff-сайленсинг генов, не имеющих CGI-аннотаций

[0302] По оценкам, около 30% генов человека не связаны с промоторным CpG-островком (CGI) (Deaton and Bird, 2011). Учитывая наблюдаемую универсальность CRISPRoff для сайленсинга генов, мы задались вопросом, могут ли гены, не имеющие CGI-аннотаций, быть долговременно выключены с помощью CRISPRoff. Неожиданно, мы обнаружили с помощью CRISPRoff более 300 генов без CGI-аннотаций с дефектами роста при нокдауне ($\gamma \leq -0,1$), причем 160 генов продуцировали фенотипы роста $\gamma \leq -0,2$ (Фиг. 5A). Большинство этих генов имели слабый фенотип или не имели фенотипа в скрининге

мутантов CRISPRoff, что указывает на то, что их нокдаун является зависимым от ДНК-метилирования, несмотря на отсутствие аннотированного CGI.

[0303] Чтобы подтвердить наше наблюдение о том, что CRISPRoff может выключать гены без аннотированных CGI, мы эндогенно пометили пять генов без аннотированных CGI - CALD1, DYNC2LI1, LAMP2, MYL6 и VPS25 - в HEK293T с помощью mNeonGreen (mNG) и провели оценку долговременного сайленсинга с помощью CRISPRoff (**Фиг. 5B**). Через 14 дней после трансфекции CRISPRoff, мы обнаружили большой процент клеток с выключенными DYNC2LI1, LAMP2, MYL6 и VPS25 (**Фиг. 5C**). Мы не обнаружили стабильного сайленсинга CALD1 на 14-й день, возможно, из-за того, что его промотор почти полностью лишен CpG-динуклеотидов или неоптимальных sgPHK, использованных в эксперименте (**Фиг. 5C**). Трансфекция мутанта CRISPRoff не приводила к долговременному выключению экспрессии генов. Обработка клеток с выключенными DYNC2LI1 и LAMP2 с помощью TETv4 привела к реактивации экспрессии генов около в 70% клеток (**Фиг. 5D**).

[0304] Мы выделили клетки с выключенными LAMP2, DYNC2LI1 и MYL6 и проанализировали статус метилирования ДНК в промоторе гена методом бисульфитного секвенирования. Цитозины в контексте CG были сильно метилированы в подвергнутых сайленсингу клетках (**Фиг. 5E**). Мы пассировали клетки с выключенными DYNC2LI1, LAMP2 и MYL6 в течение 30 дней и наблюдали стабильный сайленсинг DYNC2LI1 и LAMP2 (**Фиг. 5F-5I**). Мы также проследили за 33 одноклеточными клонами с выключенным DYNC2LI1, и все клоны репрессировали ген через 50 дней после трансфекции CRISPRoff (**Фиг. 5I**). Хотя MYL6 подвергся сайленсингу, связанному с метилированием ДНК, в ранние сроки, экспрессия гена реактивировалась до уровня, близкого к уровню до CRISPRoff, к 30-му дню.

[0305] Наконец, чтобы определить степень метилирования ДНК в гене с неаннотированным CGI, мы провели WGBS клеток с выключенным DYNC2LI1 через 30 дней. Мы обнаружили единичное доминантное усиление метилирования ДНК в промоторе DYNC2LI1 (**Фиг. 5J**), состоящее из области промотора размером около 1,2 кб (**Фиг. 5K**). Вместе эти данные показывают, что эпигенетическое редактирование с помощью CRISPRoff не ограничивается генами с каноническими CGI-аннотациями и может быть нацелено на большинство генов, закодированных в геноме человека. Более того, основываясь на этих результатах, мы предположили, что теоретическое обоснование аннотации генов CGI не всегда позволяет прогнозировать наличие функциональных CpG-сайтов, что увеличивает возможности CRISPRoff и CRISPRon для функционального тестирования метилирования CpG в модуляции генной экспрессии.

[0306] Правила нацеливания для платформы CRISPRoff

[0307] Далее мы изучили ландшафт нацеливания CRISPRoff в промоторах генов. Ранее мы и другие авторы использовали мозаичный скрининг sgPHK и методы машинного обучения, чтобы показать, что активные sgPHK для CRISPRi локализованы в узком окне промоторов генов, особенно в области, обедненной нуклеосомами, следующей

непосредственно за сайтом начала транскрипции (TSS) (Gilbert et al., 2014). Несмотря на успешное использование этих правил CRISPRi для разработки полногеномных скринингов существенности CRISPRoff, остается непроверенным, ограничивается ли расположение эффективных направляющих для CRISPRoff таким же узким окном.

[0308] Для эмпирического определения окна нацеливания CRISPRoff, мы разработали объединенную мозаичную библиотеку промоторов sgPHK против подмножества генов, необходимых для роста клеток, на основе предшествующих скринингов CRISPRi в клетках K562 и нашего скрининга CRISPRoff в масштабе генома (**Фиг. 12A**). Библиотека включает PAM-содержащие последовательности на расстоянии +/- 1 кб от TSS для 520 генов (425 с одним аннотированным CGI, 56 с несколькими CGI и 39 без аннотированных CGI; определяются по присутствию CGI в пределах 2,5 кб от TSS), что составляет в общем ок. 116000 уникальных sgPHK (**Фиг. 6A-6B**). Мы провели скрининги CRISPRoff в клетках HEK293T с использованием той же экспериментальной процедуры, которая применялась для полногеномных скринингов, параллельно с CRISPRoff мутанта метилтрансферазы D3A^{E765A}. Мы также трансдуцировали мозаичную библиотеку sgPHK в клетки K562, стабильно экспрессирующие dCas9-KRAB, и провели скрининг CRISPRi для сравнения с сайленсингом генов, опосредованным только dCas9-KRAB.

[0309] Для оценки скринингов мы рассчитали оценку фенотипа (γ) трех наиболее активных sgPHK для каждого гена и сравнили фенотипы в трех скринингах. Сначала мы сфокусировались на 425 генах с одним аннотированным CGI, поскольку было предсказано, что они являются каноническими мишенями для опосредованного CRISPRoff сайленсинга. Фенотипы для скрининга CRISPRoff (медиана $\gamma = -0,33$) были более выраженными по сравнению с фенотипами скрининга мутантов CRISPRoff (медиана $\gamma = -0,15$), что свидетельствует о том, что сильные фенотипы, наблюдаемые в скрининге CRISPRoff, являются зависимыми от ДНК-метилирования (**Фиг. 6C**).

[0310] Для сравнения оптимальных sgPHK между CRISPRoff и CRISPRi, мы нормировали фенотипы по всем скринингам и построили агрегированный график активности sgPHK относительно TSS (**Фиг. 6D**). В соответствии с нашими предыдущими работами, высокоактивные sgPHK для CRISPRi были сосредоточены в узком окне (около 75 п.о.) непосредственно следом за TSS. Аналогично, активные sgPHK для мутанта CRISPRoff зеркально отражали CRISPRi, как мы и ожидали, поскольку домен KRAB остается функциональным в этом слитом белке, несмотря на отсутствие активности метилирования ДНК. В отличие от этого, активные sgPHK CRISPRoff были широко распределены по TSS, в частности, в пределах окна размером 1 кб с центром на TSS. Анализ репрезентативных генов DKC1, GPN2 и ZCCHC9 показывает, что даже в пределах одного промотора активные sgPHK для CRISPRoff распределены от -500 до +500 п.о. от TSS - в значительно расширенном окне нацеливания для сайленсинга по сравнению с CRISPRi (**Фиг. 6E-6G**). Мы также наблюдали эффективные sgPHK за пределами CGI, которые были зависимы от метилирования ДНК (**Фиг. 6G**), что указывает на то, что

функциональные CpG не обязательно ограничены каноническими CGI, как это наблюдалось в наших данных WGBS. Наш анализ агрегированного графика активных sgPНК, нацеленных на 56 генов с несколькими аннотированными CGI, также показывает широкое окно нацеливания, аналогично генам с одним аннотированным CGI, с центром в TSS (Фиг. 12В-12С).

[0311] Анализ 39 генов без промоторных CGI показал большое количество высокоактивных sgPНК, сопоставимых по силе фенотипа с генами с аннотированными CGI (Фиг. 6С, окрашены в красный цвет), и фенотипы были сильно понижены в скрининге мутантов CRISPRoff. Репрезентативный график для гена *ORC5* показывает, что подобно генам с аннотированными CGI, активные sgPНК CRISPRoff распределены от -500 до +500 п.о. от TSS (Фиг. 6Н). Более того, мы заметили, что CRISPRoff имеет расширенное окно нацеливания, несмотря на отсутствие аннотированного CGI для этих 39 генов (Фиг. 6I). Наши эксперименты показывают, что оптимальное окно для сайленсинга генов CRISPRoff одинаково широко для генов с аннотированными CGI и без них, вероятно, из-за низкой плотности сайтов CpG, являющихся функциональными для зависимого от метилирования сайленсинга генов, как мы показали для *DYNC2LI1*, *LAMP2*, *MYL6* и *VPS25* (Фиг. 5С).

[0312] Мы обнаружили, что активные sgPНК не распределены равномерно, а расположены с периодическим паттерном в пределах окна от -500 до +500 п.о., как показано для *DKC1*, *GPN2* и *ZCCHC9* (Фиг. 6Е-6G). Наложение заселенности нуклеосом на оценку активности sgPНК CRISPRoff для всех генов показало, что наиболее активные sgPНК расположены в областях промоторов генов, обедненных нуклеосомами, как было показано ранее нами и другими исследователями для инструментов на основе Cas9 и dCas9 (Фиг. 6J-6K, 12D и 12E) (Horlbeck et al., 2016b; Isaac et al., 2016).

[0313] Чтобы подтвердить, что окно нацеливания CRISPRoff является аналогичным для генов, не имеющих фенотипа роста при нокдауне, мы разработали мозаичные библиотеки sgPНК, охватывающие +/- 2,5 кб от TSS для нацеливания на четыре эндогенных гена: *CLTA*, *H2B*, *RAB11A* и *VIM*. Для каждого скрининга пользовательской библиотеки sgPНК мы использовали соответствующую линию клеток HEK293Т, экспрессирующую эндогенно меченый GFP ген (Leonetti et al., 2016b). Каждая клеточная линия была трансдуцирована соответствующей библиотекой sgPНК, трансфицирована CRISPRoff, и клетки пассировали в течение 4 недель, чтобы убедиться, что сайленсинг генов был долговременным (Фиг. 12F). Затем мы сортировали GFP-позитивные и GFP-негативные популяции клеток для каждого скрининга и обрабатывали образцы, как описано выше. Мы рассчитывали эффективность sgPНК путем выявления sgPНК в популяции с выключенным геном (GFP-) по сравнению с популяцией с включенным геном (GFP+).

[0314] Аналогично скринингам роста, активные sgPНК для *CLTA*, *H2B* и *VIM* охватывали большое окно, включающее TSS (Фиг. 6L и 12G-12I). Активные sgPНК CRISPRoff для *CLTA* находились в двух разных областях, причем одна область была

расположена перед TSS за пределами аннотированного CGI (**Фиг. 12G**). Неожиданно, sgPНК, нацеленные на участок ок. 2 кб перед TSS H2B, были высокоактивными (**Фиг. 6L**). Аналогично, для VIM активные sgPНК охватывали окно размером 2 кб, расположенное +/- 1 кб от TSS. В отличие от этого, активные sgPНК для RAB11A были ограничены узким окном в области TSS. Наложение занятости нуклеосом на активность sgPНК показало, что промотор RAB11A является нуклеосомно плотным (**Фиг. 12H**). На основании этих данных мы предположили, что доступность CRISPRoff ограничена нуклеосомами; однако после связывания CRISPRoff может выключать экспрессию генов даже вдали от TSS.

[0315] Долговременный сайленсинг генов зависит от H3K9me3 и поддержания ДНК-метилирования

[0316] Для изучения механизма, лежащего в основе опосредованной CRISPRoff наследуемой памяти, мы использовали широкое окно нацеливания в промоторе H2B, чтобы исследовать создание, распространение и поддержание модификаций гистонов H3K9me3 и меток метилирования CpG. Мы нацелили CRISPRoff на TSS (sgPНК-A) и на дистальный участок на расстоянии ок. 2 кб перед TSS (sgPНК-B) (**Фиг. 13A**). Через 30 дней после трансфекции CRISPRoff, 89% клеток сохраняли сайленсинг H2B при доставке sgPНК-A по сравнению с 76% для sgPНК-B (**Фиг. 13B**). Используя ChIP-seq, мы показали, что обе sgPНК индуцировали отложение H3K9me3 на 5-й день и его поддержание до 30-го дня по всему локусу, несмотря на расстояние ок. 2 кб между сайтами связывания sgPНК (**Фиг. 13C**). Приобретенные модификации H3K9me3 в обработанных CRISPRoff клетках перекрывались с неметилированной областью CpG в необработанных родительских клетках (нижняя дорожка на **Фиг. 13C**). В отличие от этого, отложение и поддержание H3K9me3 было намного слабее при использовании CRISPRoff с мутацией D3A, что согласуется с неспособностью поддерживать репрессию генов с помощью мутанта (**Фиг. 13D**).

[0317] Далее мы определили профиль CRISPRoff-опосредованного метилирования CpG в целевых дистальных и TSS-областях. На 5-й день мы обнаружили формирование и распространение метилирования CpG от сайта инициации до сайта на расстоянии около 2 кб от сайта связывания sgPНК (меченый сайт 1 и сайт 3, **Фиг. 13E-13F**). На 30-й день мы обнаружили стабильное поддержание метилирования ДНК в обоих сайтах связывания sgPНК (**Фиг. 13G-13H**). Мы также обнаружили высокую степень метилирования CpG между сайтами связывания sgPНК, что позволяет предположить линейное продвижение распространения от сайта инициации (сайт 2, **Фиг. 13I**). Эти данные, вместе с нашими данными WGBS, подчеркивают координацию отложения, распространения и поддержания метилирования гистонов и ДНК и указывают на существование основополагающих принципов регуляции, которые, вероятно, зависят от геномного контекста.

[0318] CRISPRoff-сайленсинг генов в iPSC и iPSC-производных нейронах

[0319] Вследствие пригодности стволовых клеток для изучения развития и функционирования конкретных типов клеток, мы использовали CRISPRoff в

индуцированных плюрипотентных стволовых клетках. Мы трансфицировали iPSC с помощью CRISPRoff и sgPНК, нацеленных на CD81, или ненацеливающих контролей, и обнаружили, что через 30 дней после трансфекции многие iPSC имели стабильно выключенный CD81 (Фиг. 7А-7В). Таким образом, кодируемая CRISPRoff память сайленсинга стабильно поддерживается в стволовых клетках.

[0320] Мы выделили iPSC с выключенным CD81 и дифференцировали клетки в нейроны, как было описано ранее (Tian et al., 2019)(Фиг. 7А). Затем мы измерили уровни белка CD81 на стадии предшественников нервных клеток (день 0 дифференцировки) и после дифференцировки клеток в нейроны (день 8 после дифференцировки). Мы обнаружили, что CD81 оставался выключенным во время и после дифференцировки нейронов в 90% клеток (Фиг. 7С-7D). Аналогичная доля недифференцированных iPSC поддерживала сайленсинг CD81 в течение такого же периода времени, что позволяет предположить, что реактивация CD81 около в 10% клеток не была обусловлена процессом дифференцировки. Мы собрали геномную ДНК из нейронов с выключенным CD81 и обнаружили сильно метилированные CpG-динуклеотиды промотора по сравнению с нейронами, обработанными CRISPRoff и ненацеливающей sgPНК (Фиг. 7Е).

[0321] Затем мы применили CRISPRoff-опосредованное редактирование нейронов, полученных из iPSC, чтобы выключить MАРТ - ген, кодирующий тау-белок и вовлеченный в различные неврологические заболевания (Iqbal et al., 2016). MАРТ транскрипционно подавляется в iPSC посредством H3K27me3, а не метилирования ДНК и H3K9me3, и его экспрессия существенно возрастает во время дифференцировки нейронов (Guenther et al., 2010). Мы предположили, что CRISPRoff может записать эпигенетическую память о сайленсинге в локусе MАРТ, которая будет сохраняться после дифференцировки нейронов, чтобы выключить MАРТ в нейронах. Мы транзитивно трансфицировали CRISPRoff в iPSC вместе с sgPНК, нацеленными на MАРТ, или ненацеливающим контролем (Фиг. 7F). На 10-й день протокола дифференцировки мы измерили уровни тау-белка и обнаружили ок. 30% клеток со сниженной экспрессией тау по сравнению с ненацеливающим контролем (Фиг. 7G-7H). В совокупности эти данные подтверждают, что CRISPRoff-опосредованное редактирование эпигенома является пригодной технологией для переписывания программ экспрессии генов в клетках, производных от iPSC, особенно для модуляции экспрессии генов в клетках, для которых доставка платформ редактирования генов остается сложной задачей.

[0322] CRISPRoff-нацеливание на энхансерные элементы

[0323] Наконец, мы исследовали потенциальную полезность CRISPRoff для сайленсинга дистальных от промотора элементов, нацеливая его на энхансеры, контролирующую экспрессию длинной некодирующей PНК PVT1 (Cho et al., 2018; Fulco et al., 2016). Мы транзитивно экспрессировали CRISPRoff в клеточной линии рака молочной железы MDA-MB-231 с sgPНК, нацеленными либо на промотор PVT1, либо на четыре ранее идентифицированных энхансерных элемента после промотора PVT1: E1 (+15,5 кб), E2 (+60 кб), E3 (+105 кб) и E4 (+113 кб) (Фиг. 7I). Мы обнаружили значительное

снижение уровня транскрипта PVT1 при использовании sgРНК, нацеленных на E1, E3 и E4 (ок. 40-60%) по сравнению с 80% нокдауном при использовании sgРНК, нацеленных на промотор (Фиг. 7J, слева). В отличие от этого, параллельные эксперименты с использованием мутанта CRISPRoff-Dnmt3A^{E765A} привели к менее надежному нокдауну (Фиг. 7J, справа). Эти результаты демонстрируют потенциальную возможность использования CRISPRoff для картирования и разделения функций энхансерных элементов и некодирующих регуляторных элементов в геноме человека.

[0324] Обсуждение

[0325] В данной работе мы представляем CRISPRoff и CRISPRon - две технологии программируемой записи и стирания эпигенетической памяти для управления программами экспрессии генов. Транзиентная экспрессия CRISPRoff записывает надежную, специфическую и мультиплексируемую программу сайленсинга генов, которая сохраняется в памяти клеток в процессе клеточного деления и дифференцировки и может быть быстро реверсирована с помощью CRISPRon. Мы показали, что CRISPRoff может специфически и надежно выключать подавляющее большинство генов человека. Наши эксперименты показывают, что CRISPRoff может возмущать энхансеры, открывая потенциал для воздействия на элементы генома, контролирующую тканеспецифическую экспрессию генов (Fulco et al., 2016; Tarjan et al., 2019).

[0326] Наш вывод о том, что целенаправленное метилирование ДНК за пределами аннотированных CGI может приводить к надежному запоминаемому сайленсингу генов, расширяет каноническую модель сайленсинга генов на основе метилирования, согласно которой высокая плотность метилирования CpG является необходимым условием для стабильного распространения сайленсинга (Boyes and Bird, 1992). Хотя опосредованная CRISPRoff запись метилирования ДНК и модификации гистонов программируется искусственно, наши результаты мотивируют необходимость определения функционального метилирования ДНК и разделения регуляторных элементов ДНК и других факторов хозяина, которые опосредуют инициацию, распространение и поддержание меток гистонов и метилирования ДНК. Например, нацеливание CRISPRoff на участок в 2 кб перед TSS H2B приводит к приобретению и поддержанию меток H3K9me3 и метилирования ДНК в тех же позициях генома, что и нацеливание CRISPRoff в непосредственной близости от TSS, что указывает на наличие предустановленных границ, сдерживающих эпигенетическое распространение. Позволяя инициировать сайленсинг в заданное время и в определенном положении генома, CRISPRoff предоставляет уникальный инструмент для решения этих и других фундаментальных задач, касающихся механизма и биологической роли наследственного сайленсинга генов в клетках млекопитающих (Audergon et al., 2015; Iglesias et al., 2018; Ragunathan et al., 2015; Yu et al., 2018).

[0327] CRISPRoff является ценным дополнением к существующим подходам CRISPRi и нуклеазы CRISPR (Doench, 2018; Hanna and Doench, 2020; Shalem et al., 2015). Сайленсинг генов с помощью CRISPRoff может привести к фактически полному

обнулению без индуцирования ответа на повреждение ДНК, что облегчает проведение скринингов с нацеливанием на несколько генов или терапевтическую клеточную инженерию (Ihry et al., 2018). Возможность нацеливания CRISPRoff на большое окно перед TSS позволяет получить доступ к промоторным SNP (однонуклеотидные полиморфизмы), которые могут быть использованы для аллель-специфического нацеливания на связанные с болезнью мутации. Это обеспечит более широкие возможности применения подходов для выключения доминантно-негативных аллелей. Аналогичным образом, сайленсинг длинных некодирующих РНК и регуляторных РНК открывает новые возможности для стабильного перепрограммирования экспрессии генов. Сайленсинг ингибирующих элементов, таких как антисмысловые транскрипты, приведет к наследуемому повышению экспрессии некоторых генов, что позволит предпринимать терапевтические попытки смягчения гаплонедостаточности или нарушений импринтинга (Buiting et al., 2016). В более широком смысле, наследуемый эпигенетический сайленсинг предоставляет универсальный инструмент для редактирования программ экспрессии генов человека.

[0328] Пример 2

[0329] Блокирование флавивирусной инфекции путем эпигеномного редактирования

[0330] Медицинские вмешательства, способствующие повышению иммунитета к вирусам, такие как вакцины, являются одним из наиболее перспективных классов противовирусной терапии. Однако на разработку и внедрение вакцин могут уйти годы. Мы полагаем, что редактирование эпигенома является привлекательной платформой для блокирования вирусных инфекций благодаря транзистному импульсу экспрессии редактора, долговечности и обратимости сайленсинга и, в отличие от редактирования генома, отсутствию необратимых изменений в основной последовательности ДНК (**Фиг. 3А**).

[0331] В качестве первоначального теста для оценки эффективности редактирования эпигенома при блокировании вирусных инфекций мы воздействовали на обработанные CRISPRoff клетки HEK293Т, которые поддерживали сайленсинг CLTA в течение >1 года, живой аттенуированной вакциной против вируса желтой лихорадки (YFV) 17D. YFV представляет собой *флавивирус*, который проникает в клетки хозяина посредством клатрин-опосредованного эндоцитоза. Мы собрали клетки через два дня после инфицирования YFV и обнаружили 50% снижение инфицирования клеток CLTA-off по сравнению с HEK293Т ДТ (дикого типа) (**Фиг. 3В**). CLTA представляет собой одну из двух легких цепей клатрина, которые могут образовывать функциональный клатриновый трискелион, и мы предположили, что инфицирование клеток происходит из-за того, что вирус использует для проникновения молекулу клатринового трискелиона, содержащие легкую цепь CLTB.

[0332] Далее мы применили редактирование эпигенома для блокирования инфекции вируса денге (DENV-2) путем выключения SPCS1 и STT3A, которые являются

ключевыми факторами в патогенезе DENV. Мы трансфицировали CRISPRoff в HEK293T и заразили клетки DENV-2 через 14 дней после трансфекции. Инфицирование DENV-2 значительно снижается в клетках-мишенях SPCS1 и STT3A (Фиг. 3С). Сайленсинг генов SPCS1 и STT3A был измерен на уровне нокдауна около 60%, что соответствует инфекции DENV-2 в отредактированных клетках (Фиг. 3D). Клетки, выключенные с помощью CLTA, инфицировались на том же уровне, что и неотредактированные клетки ДТ, что позволяет предположить, что проникновение DENV-2 не зависит от CLTA.

[0333] Пример 3

[0334] Эксперименты по нокдауну MAPT проводились путем трансфекции CRISPRoff MAPT-нацеливающими sgРНК, как описано ниже. Для усиления сигнала iPSC были котрансфицированы MAPT-нацеливающими sgРНК вместе с CD81-нацеливающими sgРНК. В качестве контроля iPSC были трансфицированы CRISPRoff и ненацеливающими sgРНК. Через неделю после трансфекции iPSC были окрашены антителом CD81, а клетки с нокаутированным CD81 были отсортированы методом FACS и пассировались в течение 1 недели до их дифференцировки в нейроны. Дифференцировку в нейроны проводили, как описано ниже. Через 10 дней после дифференцировки нейроны собирали для анализа методом кПЦР и окрашивания тел антител на тау-белок, и анализировали методом проточной цитометрии. Количественные показатели нокдаунов MAPT/Tau нормировали по отношению к нетрансфицированным клеткам ДТ. Результаты приведены на Фиг. 8А-8С.

[0335] Материалы и способы

[0336] Проектирование и конструирование плазмид

[0337] Последовательности dCas9 и KRAB были получены из ранее созданного конструктора CRISPRi (Gilbert et al., 2013). Последовательности D3A и D3L, включая слитый продукт D3A-D3L, были заимствованы из (Stepper et al., 2017) и собраны с ДНК-последовательностями dCas9 и KRAB в CAG-экспрессионную плазмиду с помощью NEBuilder[®] HiFi DNA Assembly (NEB). Все слитые белки CRISPRoff включают BFP либо в виде прямого продукта слияния, либо с P2A-расщепленной последовательностью для измерения эффективности трансфекции методом проточной цитометрии. Последовательность dSaCas9 (D10A, N508A) была амплифицирована методом ПЦР из pX603 (Addgene, № кат. 61594), а последовательность dLbCas12a была амплифицирована методом ПЦР из (Tak et al., 2017). Лентивирусный репортер GAPDH-Snrpn-GFP был получен от фирмы Addgene, № кат. 70148 (Liu et al., 2016; Stelzer et al., 2015).

[0338] Плазмиды sgРНК были сконструированы путем рестрикционного клонирования протоспейсеров следом за промотором U6 с использованием сайтов разреза BstXI и BspI, как описано ранее. Плазмиды для экспрессии sgРНК также экспрессируют маркер T2A-mCherry для измерения эффективности трансфекции. Последовательности sgРНК, использованные для экспериментов с CRISPRoff, перечислены в Таблице 5. Последовательности sgРНК были выбраны на основе нашего ранее разработанного алгоритма для предсказания активных sgРНК CRISPRi (Horlbeck et al., 2016a).

[0339] Все конструкты мРНК были синтезированы с использованием набора для транскрипции mMESSAGE mMachine™ T7 Ultra Transcription Kit (Thermo Fisher Scientific). Последовательность промотора T7 (SEQ ID NO:111) была сначала клонирована перед последовательностью CRISPRoff. Последовательность T7-CRISPRoff амплифицировали методом ПЦР и использовали в качестве матрицы для реакций синтеза *in vitro*. В соответствии с протоколом производителя для синтеза, реакции очищали экстракцией хлороформом и осаждением изопропанолом.

[0340] Культура клеток, трансфекции ДНК и проточная цитометрия

[0341] Все клеточные линии культивировали при 37°C в инкубаторах для тканевых культур с 5% CO₂. Клетки HEK293T (женские), HeLa (женские) и U2OS (женские) культивировали в модифицированной Дульбекко среде Игла (DMEM) в 10% ФСБ (HyClone), 100 ед/мл стрептомицина, 100 мкг/мл пенициллина и 2 mM глутамина. Клетки K562 (женские) поддерживали в RPMI-1640 с 25 mM HEPES и 2,0 г/л NaHCO₃ в 10% ФСБ, 2 mM глутамина, 100 ед/мл стрептомицина и 100 мг/мл пенициллина. iPSC WTC Gen1c (мужские) культивировали в среде mTESR (STEMCELL Technologies) в условиях отсутствия фидера на матрикеле с пониженным содержанием фактора роста (BD Biosciences). Клетки пассировали с помощью Accutase (STEMCELL Technologies) и высевали на покрытые матрикелем планшеты со средой mTESR, дополненной ингибитором киназы p16-ро-ассоциированная «скрученная спираль» (ROCK) Y-27632 (10 мкМ; Selleckchem).

[0342] Lentivirusные частицы были получены путем трансфекции стандартных упаковочных векторов в HEK293T с использованием реагента для трансфекции TransIT-LT1 (Mirus, MIR2306). Через 24 часа после трансфекции среду меняли на полную DMEM, дополненную 15 mM HEPES. Вирусные супернатанты собирали через 48-60 часов после трансфекции и фильтровали через ПВДФ шприц-фильтр 0,45 мкм. Lentivirusные инфекции содержали полибрен (8 мкг/мл).

[0343] Трансфекции CRISPRoff и анализ

[0344] Эксперименты по транзientной трансфекции в HEK293T проводились в 24-луночных планшетах с использованием реагента для трансфекции TransIT-LT1 (Mirus) и восстановленной сывороточной среды Opti-MEM™ I Reduced Serum Medium (Thermo Fisher Scientific). Клетки при 70-80% конfluenceности трансфицировали 300 нг плазмиды, кодирующей CRISPRoff, dCas9-KRAB или dCas9-D3A-3L, и 150 нг плазмиды, кодирующей sgРНК. Эксперименты с CRISPRoff в клетках HeLa и U2OS проводили путем нуклеофекции плазмид с использованием набора SE Cell Line 96-well Nucleofector Kit (Lonza) и 96-луночного устройства Shuttle™ Device (Lonza), в соответствии с протоколом производителя. Трансфицированные клетки сортировали через 2 дня после трансфекции с помощью BD FACSAria II или FACSAria Fusion и отсортированные клетки пассировали каждые 2-3 дня для измерения продолжительности сайленсинга генов. Эксперименты по сравнению активности сайленсинга различных слитых CRISPRoff (Фиг. 1E, 4C, 8B и 8F) были проведены в клетках, стабильно экспрессирующих нацеливающую sgРНК, для

нормализации экспрессии sgРНК. Для создания клеточных линий, стабильно экспрессирующих sgРНК, клетки трансдуцировали лентивирусными частицами, экспрессирующими sgРНК, и сортировали на sgРНК-положительные клетки через 2-3 дня после трансдукции.

[0345] Количественная оценка уровней белков ITGB1, CD81 и CD151 измерялась путем окрашивания антителами клеточной поверхности живых клеток. Клетки инкубировали с APC- или PE-меченым антителом (BioLegend) в течение ок. 30 мин в темноте при комнатной температуре, дважды промывали ФСБ, содержащим 10% ФБС, и измеряли экспрессию белка на проточном цитометре BD LSR II.

[0346] Все данные проточной цитометрии анализировали с помощью FlowJo, а необработанные графики FACS, представленные на рисунках, приведены в шкале log10.

[0347] Нацеливание на энхансер PVT1

[0348] Количественная оценка экспрессии PVT1 методом ОТ-ПЦР проводилась, как описано в (Cho et al., 2018). Вкратце, клетки МВ-MDA-231 трансфицировали ДНК CRISPR вместе с вектором sgРНК с помощью Neon (1400 вольт, 10 мс, 4 импульса). Двойные положительные клетки сортировали через 2 дня и продолжали культивировать в течение 3 дней. РНК выделяли на центрифужной колонке Zymo, и экспрессию генов определяли количественно с помощью смеси для кПЦР SYBR (LightCycler 480), используя 45 нг РНК. Экспрессию PVT1 нормировали по гену GAPDH методом ddCt. Для расчета статистической значимости использовался Т-критерий с использованием 3-5 биологических реплик на условие. Использовались следующие последовательности праймеров: PVT1 прямой (SEQ ID NO:112), PVT1 обратный (SEQ ID NO:113), GAPDH прямой (SEQ ID NO:114), GAPDH обратный (SEQ ID NO:115). Были использованы sgРНК, нацеленные на энхансеры 1-4 PVT1, промотор и лямбда (контроль), описанные в (Cho et al., 2018) и перечисленные в Таблице 5.

[0349] РНК-секвенирование

[0350] Клетки НЕК293Т, в которых поддерживался стабильный сайленсинг генов-мишеней, собирали через 33 дня (ITGB1, CD81 и CD151) или 28 дней (CLTA, HIST2H2BE, RAB11A и VIM) после трансфекции CRISPRoff. Клетки удаляли из планшетов с помощью ФСБ, центрифугировали при 500×g в течение 5 мин и снова промывали ФСБ. Тотальную РНК экстрагировали с помощью Direct-zol RNA MiniPrep (Zymo R2051). Подготовку библиотек проводили с использованием набора TruSeq Stranded mRNA Library Preparation Kit (Illumina RS-111-2101), начиная с 1000 нг тотальной РНК. Окончательные библиотеки оценивали на биоанализаторе 2100 (Agilent), проводили количественные определения с использованием набора Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) и секвенировали в виде одноконцевых чтений 50 пар оснований на HiSeq 4000 (Illumina). Для обработки результатов чтений секвенирования линкерные последовательности (SEQ ID NO:116) удаляли с помощью FASTX-clipper (FASTX-Toolkit). Затем чтения выравнивали по геному человека (GRCh37) с помощью программы выравнивания STAR (Spliced Transcripts Alignment to a Reference, версия 2.5) против аннотации транскриптома

Gencode Gene V24lift37. Количественную оценку прочтений проводили с помощью featureCounts (Liao et al., 2014). Все последующие анализы были выполнены в Python (версия 2.7) с использованием комбинации библиотек Numpy (v1.12.1), Pandas (v0.17.1) и Scipy (v0.17.0). Эффективность нокдауна рассчитывали путем нормирования генных транскриптов на миллион (TPM) для экспериментальных образцов по среднему TPM контрольных (ненацеливающих) образцов. Дифференциальный анализ экспрессии проводили с помощью DESeq2 (Love et al., 2014). Мы отметили, что нецелевые дифференциально экспрессированные транскрипты представляют собой низкоэкспрессируемые гены.

[0351] Иммунопреципитация и анализ хроматина

[0352] Через 30 дней после трансфекции, 10×10^6 клеток сшивали 1% формальдегидом в течение 10 мин при комнатной температуре и гасили 1,25 М глицином. Сшитые клетки дважды промывали холодным ФСБ, содержащим 1% ингибиторов протеаз HaltTM (Thermo Fisher Scientific), и клеточные осадки мгновенно замораживали при -80°C до подготовки образцов. Клетки лизировали в буфере для лизиса (5 мМ PIPES pH 8, 85 мМ KCl, 1% Igepal, 1% ингибиторов протеаз) в течение 10 мин на льду. Ядра выделяли после центрифугирования суспензии на 2000 об/мин при 4°C в течение 5 мин. Ядра лизировали при 4°C в течение 10 мин в буфере для лизиса ядер (50 мМ Трис pH 8, 10 мМ ЭДТА, 1% ДСН, 1% ингибиторов протеаз). Расщепление хроматина проводили при 4°C с помощью устройства для ультразвуковой обработки Diagenode Bioruptor[®] Pico в микропробирках Bioruptor[®] Pico объемом 1,5 мл. Программа расщепления была оптимизирована для получения фрагментов размером 200-700 п.о. (30 секунд включения, 30 секунд выключения в течение 10 циклов). Озвученные образцы центрифугировали при 13000 об/мин при 4°C в течение 10 мин, супернатант собирали и разводили в 5 раз в буфере для разведения при иммунопреципитации (IP) (50 мМ Трис pH 7,5, 150 мМ NaCl, 1 мМ ЭДТА, 1% игепала (Igepal), 0,25% дезоксихолата, 1% ингибиторов протеаз). Часть исходного материала была сохранена и заморожена (4% от общего количества) перед тем, как приступить к иммунопреципитации. Иммунопреципитацию проводили в течение ночи при 4°C с использованием 5 мкг анти-НЗК9me3 антитела (abcam ab8898). Стадии промывки проводили с использованием магнитных шариков PierceTM Protein A/G (Thermo Fisher Scientific) по следующему протоколу: один раз с буфером для промывки IP 1 (20 мМ трис pH 8, 2 мМ ЭДТА, 50 мМ KCl, 1% тритон X-100, 0,1% ДСН), дважды с высокосолевым буфером (20 мМ трис pH 8, 2 мМ ЭДТА, 500 мМ NaCl, 1% тритон X-100, 0,01% ДСН), один раз с буфером для промывки IP 2 (10 мМ трис pH 8, 1 мМ ЭДТА, 0,25 хлорида лития, 1% игепал, 1% дезоксихолата) и дважды с буфером TE (10 мМ трис pH 8, 1 мМ ЭДТА). Образцы элюировали в буфере для элюции (50 мМ бикарбоната натрия, 1% ДСН) при 65°C в течение 1 часа и подвергали обратному сшиванию сначала в 300 мМ NaCl и РНКазе А в течение 1 часа при 37°C , затем при 63°C в течение ночи с протеиназой К (Thermo Fisher). Образцы ДНК очищали с помощью набора Zymo Clean & Concentrator, а библиотеки готовили с помощью набора для подготовки библиотек ДНК NEBNext[®]

Ultra™ II DNA Library Prep (NEB).

[0353] Чтения выравнивали по геному человека (hg19) с помощью программы bowtie v2.3.2 (Langmead and Salzberg, 2012). Выравнивания обрабатывали с помощью deepTools2 bamCoverage (Ramírez et al., 2016), нормируя чтения к 1x среднему покрытию. Полученные файлы bigWig были визуализированы в программе Integrative Genomics Viewer (IGV). Анализ пиков проводили с использованием MACS (<https://github.com/macs3-project/MACS>). Последующий анализ и обогащение промоторных областей, которые были определены как +/- 2 кб от каждого TSS (с TSS, основанными на ранее опубликованных аннотациях (Horlbeck et al., 2016)), были выполнены с помощью Python 3.6 с пакетом deepTools2. Дифференциальное обогащение H3K9me3 анализировали с помощью DESeq2 (Love et al., 2014), рассматривая разные sgPHK по отношению к одному и тому же TSS как реплики образцов. Мы отметили, что BOLA1 содержит две аннотации TSS, и наши данные ChIP-seq показывают обогащение H3K9me3 вблизи TSS1, ближайшего к промотору H2B, и отсутствие обогащения для TSS2, удаленного на 15 кб (Фиг. 2G-2H). Мы не обнаружили транскрипционного нокдауна BOLA1 в наших данных RNA-seq.

[0354] Вестерн-блоттинг

[0355] Вестерн-блоттинг конструкторов CRISPRoff проводили, собирая клетки HEK293T через 2 дня после трансфекции конструкторов CRISPRoff. Клетки промывали холодным ФСБ и лизировали буфером RIPA (Thermo Fisher Scientific), дополненным коктейлем ингибиторов протеаз Halt™ (Thermo Fisher Scientific). После 30 минут лизиса при 4°C образцы центрифугировали при 20000×g в течение 20 минут при 4 °C. Растворимые фракции собирали, а концентрацию белка определяли количественно с помощью набора Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific). 40 мкг общего белка смешивали и нагревали с буфером для загрузки с ДСН, разделяли на гелях Bolt™ 4-12% Bis-Tris Plus Gels (Thermo Fisher Scientific) и переносили во влажном состоянии на ПВДФ-мембрану в буфере, содержащем 1× MOPS и 10% метанола. Мембраны блокировали с помощью блокирующего буфера Odyssey® (LI-COR), инкубировали с антителами против Cas9 *S. pyogenes* (Active Motif 61577) и калнексина (Abcam ab22595) при 4°C в течение ночи. Мембраны промывали не менее 3 раз блокирующим буфером перед инкубацией с вторичными антителами IRDye® против Cas9 и калнексина. Блоты визуализировали с помощью Odyssey® CLx (LI-COR).

[0356] Бисульфитное секвенирование ПЦР

[0357] Геномную ДНК экстрагировали из ок. $1-2 \times 10^6$ клеток в соответствии с инструкциями производителя с помощью набора PureLink Genomic DNA Mini Kit (Invitrogen). Для каждого условия, 1 мкг геномной ДНК подвергали бисульфитной конверсии и очистке в соответствии с инструкциями производителя с использованием набора EpiTect Bisulfite (Qiagen). Очищенную бисульфит-конвертированную ДНК амплифицировали с помощью EpiMark Hot Start Taq (NEB). Ампликоны очищали в геле с помощью набора Gel DNA Recovery Kit (Zymo) и повторно амплифицировали методом ПЦР с использованием EpiMark Hot Start Taq. Ампликоны клонировали в вектор pCR2.1

ТОРО в соответствии с инструкциями производителя с использованием набора для клонирования ТОРО TA Cloning Kit (Invitrogen). Продукты клонирования трансформировали в клетки Stellar E. coli (Takara) и высевали на карбенициллиновые планшеты с X-gal для сине-белого скрининга. Были отобраны колонии для каждого условия и секвенированы по методу Сэнгера. Последовательности праймеров для бисульфитной ПЦР-амплификации приведены в Таблице 6. Последовательности праймеров для амплификации фрагмента GAPDH-Snrpn были взяты из (Liu et al., 2016).

[0358] Таблица 6

Название	Последовательность (от 5' к 3')	SEQ ID NO:
CLTA, прямой праймер	TATTTGTTGATTGGGTAGTTTTTGAATT	SEQ ID NO:184
CLTA, обратный праймер	ACTCCСТАACCTCСТААТТCAACAAAA	SEQ ID NO:185
LAMP2, прямой праймер	АTAGGAAGGGTТGTGAАТТАAAAAGTTAGG	SEQ ID NO:186
LAMP2, обратный праймер	ACAАCTCACCCAAAАCTAAACAAACCAAA	SEQ ID NO:187
DYNC2L1, прямой праймер	GAGATGTAAAAGGATTCCTGAAAATTATATTT	SEQ ID NO:188
DYNC2L1, обратный праймер	CTCTCTCCAAAАТАCTCTCAATAACCTTTCC	SEQ ID NO:189
MYL6, прямой праймер, фрагмент 1	ATTTTAAGCGTTTGAGTGTGCAGGTAGGG	SEQ ID NO:190
MYL6, обратный праймер, фрагмент 1	TAAAACAATAAATAACCTCСТААТААСТААА A	SEQ ID NO:191
MYL6, прямой праймер, фрагмент 2	TTTGATTTTTTTGCTGTTGGGAGGTGTAGAT	SEQ ID NO:192
MYL6, обратный праймер, фрагмент 2	CAACTCTCCGAAACSTTTTCCTACAATAAT	SEQ ID NO:193
CD81, прямой праймер	AAGTTGTGGGGtAttTGTGGGtTttAGGAG	SEQ ID NO:194
CD81, обратный праймер	CCCaCaACCACCaCACCCATCACCACCACA	SEQ ID NO:195
МАРТ, прямой праймер	AtAAAGAtTttAAtTAtAGGAGGTGGAGAAAG	SEQ ID NO:196
МАРТ, обратный праймер	aaAaCCTTCTCCTCCaaCCTAaT	SEQ ID NO:197
H2B, сайт 1, прямой праймер	tAGTGGtATGGGGGAGtAGGATAAGAAAAG	SEQ ID NO:198
H2B, сайт 1, обратный праймер	aATCCCCTTaAaCTCTAAACTaCCAAaTC	SEQ ID NO:199
H2B, сайт 2,	TCCTTGTAААAGAAGtTGTAAtTGtAATTTTAGt	SEQ ID NO:200

прямой праймер		
H2B, сайт 2, обратный праймер	TaaTaACACCCCaAaTAACTTaTTaAaCTCTT	SEQ ID NO:201
H2B, сайт 3, прямой праймер	AttAAGAGGAAAGttTTTATAAGttTTTttTAGGGG	SEQ ID NO:202
H2B, сайт 3, обратный праймер	CCTaATTCTTTTATAACCACCTTATaCAAATTAa aaCTC	SEQ ID NO:203

[0359] Полногеномное бисульфитное секвенирование и анализ

[0360] Мы сгенерировали библиотеки полногеномного бисульфитного секвенирования (WGBS) для 12 образцов, соответствующих ДТ (необработанные), NT (ненацеливающие) и Т (нацеливающие) для экспериментов с CLTA и DYNC2L1, и определили их профили в двух репликах. После выделения ДНК и обработки РНКазой А с помощью набора PureLink Genomic DNA Mini Kit (Invitrogen), 1 мкг ДНК разводили до 7,7 нг/мкл в 130 мкл и обрезали до длины 450-550 п.о. с помощью Covaris E220 evolution с интенсификатором в течение 50 с при 140 В, 7 °С, амплитуде 10%, 200 циклов. Озвученную ДНК выделяли и концентрировали с помощью шариков Ampure XP, а размеры фрагментированной ДНК проверяли на приборе Agilent TapeStation с D1000 HS DNA ScreenTape. Затем фрагментированную ДНК подвергали бисульфитному конвертированию с использованием набора EZ DNA Lightning (Zymo, № кат. D5030) в соответствии с инструкциями производителя и продолжительностью стадии десульфонирования 16 минут. Затем 500 нг фрагментированной и конвертированной ДНК использовали для приготовления библиотеки с помощью набора Swift Accel®-NGS Methyl-Seq DNA Library Kit (№ кат. 30024) и праймеров Methyl-Seq Unique Dual Indexing Primers (№ кат. 39096). Приготовленные библиотеки количественно оценивали с помощью набора KAPA Library Quantification (Roche, № кат. KK4873) и секвенировали с использованием чтений парных концов по 150 п.о. (300 циклов) на приборе Illumina NovaSeq6000 с проточной кюветой S4 и добавлением 35% из другой библиотеки, не относящейся к WGBS, для диверсификации пулов образцов. Мы получили для всех 12 библиотек в среднем 707М чтений парных концов (диапазон 629-835М).

[0361] Перед выравниванием, чтения секвенирования были обрезаны с помощью fastp (версия 0.21.0 (Chen et al., 2018)) со следующими параметрами: --adapter_sequence=(SEQ ID NO:117) --adapter_sequence_r2=(SEQ ID NO:118) --trim_front1=0 --trim_tail1=20 --trim_front2=20 --trim_tail2=0. Обрезка необходима для удаления оснований, добавленных во время реакции с адаптазой, которые могут повлиять на выравнивание и распознавание метилирования ДНК. Обработанные чтения выравнивали по референсному геному hg38 с помощью программ methylTools (версия 1.0.0, <https://github.com/hovestadt/methylTools> (Hovestadt et al., 2014)) и bwa mem (версия 0.7.17, arXiv: 1303.3997v1) с использованием параметров по умолчанию. Более 98% чтений были выровнены как правильные пары в разных образцах. После маркировки дубликатов ПЦР с помощью sambamba (версия 0.8.0, (Tarasov et al., 2015)), значения полногеномного

метилирования CpG определяли с помощью methylTools с использованием параметра --trimPE. Среднее покрытие CpG в образцах было около 25-кратным. Эффективность бисульфитной конверсии была оценена в более чем 99,5% на основе не-CpG метилирования.

[0362] Последующие анализы проводились в R (версия 4.0.2, <https://www.r-project.org/>) с использованием пакетов bsseq (версия 1.24.4, (Hansen et al., 2012)) и DSS (версия 2.36.0, (Park and Wu, 2016)). В частности, мы применили функцию DMLtest для того, чтобы сначала выявить дифференциально метилированные локусы (используя окна сглаживания 500 п.о.) между методами обработки, а затем функцию callDMR для определения дифференциально метилированных областей. Результаты визуализировались путем построения \log_{10} -трансформированных *p*-значений, связанных с отдельными локусами (положительные значения для локусов, которые приобрели метилирование, отрицательные значения для локусов, которые потеряли метилирование). Локусы окрашивались по их разности бета-значений (-1: синий, 0: белый, +1: красный). Крупные планы геномных областей были получены путем визуализации бета-значений отдельных локусов в программе IGV (Robinson et al., 2011). Данные отображались в виде гистограмм (мин./0: синий, ср./0,5, макс./1: красный).

[0363] Полногеномный скрининг CRISPRoff и анализ

[0364] Для полногеномных скринингов CRISPRoff мы создали компактную библиотеку, позволяющую максимизировать нокдаун мишени при минимизации общего размера библиотеки. Мы нацелили на каждый ген в геноме человека две уникальные sgPНК, экспрессируемые из tandemных кассет экспрессии U6 в одном векторе. Для выбора оптимальных sgPНК, нацеленных на каждый ген, мы использовали ранее опубликованную нами библиотеку hCRISPRi v2.1 (Horlbeck et al., 2016a). Для балансировки эмпирических данных с вычислительными прогнозами и выбора наиболее активной пары sgPНК для каждого гена был использован трехуровневый подход. Сначала для сильно существенных генов (*p*-значение < 0,001 и $\gamma < -0,2$ в скрининге роста hCRISPRi v2), sgPНК ранжировали по росту. Далее, для генов, которые были идентифицированы как значимое совпадение с предшествующими скринингами CRISPRi, sgPНК ранжировали по сумме *Z*-оценок фенотипов во всех скринингах. Наконец, для всех остальных генов sgPНК ранжировали по оценкам регрессии в hCRISPRi v2.1. Используя эту схему ранжирования, мы разработали полногеномную библиотеку, состоящую всего из 21378 элементов (20360 нацеливающих элементов плюс 1018 ненацеливающих контролей).

[0365] Для клонирования наших библиотек мы начали с создания модифицированного лентивирусного вектора экспрессии одиночной sgPНК, pJR104, из исходного pJR85 (Addgene 140095) путем: (i) замены флуоресцентного маркера BFP на BsmBI-отрицательный GFP, (ii) замены константной области sgPНК на немодифицированную константную область (т.е. без последовательности захвата Perturb-seq), и (iii) включения элемента UCOE перед промотором EF1alpha для предотвращения

сайленсинга. Были синтезированы олигомеры двойных sgРНК в виде пула олигонуклеотидов (Twist Biosciences) со структурой: 5' - адаптер ПЦР - SEQ ID NO:126 - протоспейсер А - SEQ ID NO:119 - протоспейсер В - SEQ ID NO:120 - адаптер ПЦР - 3'. Пулы олигомеров были ПЦР-амплифицированы, гидролизованы с помощью BstXI/BspI, экстрагированы в геле, лигированы в лентивирусный вектор sgРНК рJR104 и трансформированы для создания промежуточной библиотеки, как описано ранее (Replogle et al., 2020). Вставку рJR98, состоящую из варианта 3 константной области sgРНК (Adamson et al., 2016) и промотора hU6, синтезировали (IDT), гидролизовали BsmBI и лигировали в гидролизованную BsmBI промежуточную библиотеку. Затем итоговую библиотеку с двумя направляющими трансформировали для амплификации и секвенировали методом секвенирования следующего поколения для обеспечения репрезентативности и однородности библиотеки.

[0366] Мозаичный скрининг пула sgРНК в клетках HEK293T проводили путем трансдукции клеток лентивирусными частицами, кодирующими библиотеку sgРНК. Эффективность инфекции измеряли через 2 дня после заражения методом проточной цитометрии, стремясь получить 20-30% sgРНК-позитивных клеток. Скрининги проводились с двумя техническими репликами, и каждая sgРНК была представлена не менее чем 1000 клеток в течение всего времени скрининга. Через два дня после трансдукции клетки обрабатывали пурамицином до тех пор, пока популяция клеток не становилась на 90% sgРНК-позитивной, что подтверждалось mCherry, кодируемым лентивирусным вектором. Для транзientной трансфекции CRISPRoff сначала высевали ок. 8×10^6 клеток на планшеты площадью 15 см^2 . Примерно через 20-24 часа (70-80% конфлюентности) каждый планшет с клетками площадью 15 см^2 трансфицировали 20 мкг плазмид, кодирующих CRISPRoff или каталитический мутант CRISPRoff-Dnmt3A^{E765A}. Через два дня после трансфекции клетки сортировали по экспрессии CRISPRoff (BFP) и высевали на планшеты площадью 15 см^2 . Через четыре дня после сортировки клетки трипсинизировали и аликвоту клеток (ок. 110×10^6) отбирали в качестве исходного момента времени T(0), а остальную часть клеточной популяции пассировали в течение еще как минимум 10 удвоений клеток. Затем клетки были собраны в качестве конечного момента времени (T10).

[0367] Библиотеки ДНК T(0) и T(10) были подготовлены для глубокого секвенирования, по существу, как описано ранее (Jost et al., 2020). Вкратце, геномную ДНК выделяли с помощью набора NucleoSpin Blood XL (Macherey-Nagel). Затем выделенную гДНК непосредственно амплифицировали в 23 циклах ПЦР с помощью NEBNext Ultra II Q5 PCR MasterMix (NEB), присоединяя адаптеры Illumina и уникальные индексы образцов (прямой праймер оJR234: SEQ ID NO:121; индексные праймеры SEQ ID NO:122). Секвенирование проводили на приборе NovaSeq 6000 (Illumina) с использованием считывания 1 длиной 19 п.о., считывания 2 длиной 19 п.о. и индексного считывания 1 длиной 5 п.о., с применением пользовательских праймеров для секвенирования оJR326 (пользовательское считывание 1, SEQ ID NO:123), оJR328

(пользовательское считывание 2, SEQ ID NO:124) и oJR327 (пользовательское индексное считывание 1, SEQ ID NO:125).

[0368] Результаты секвенирования, полученные в ходе CRISPR-скрининга, обрабатывались для расчета фенотипов генов с помощью пользовательского скрипта Python, аналогичного описанному ранее (Horlbeck et al., 2016a), за исключением того, что в данном случае сопоставляли две последовательности протоспейсеров, а не одну. В этом случае генный фенотип совпадает с фенотипом sgPHK и определяется как \log_2 (обогащение sgPHK) / (число удвоений клеток). Все дополнительные анализы данных скрининга CRISPR и построение графиков выполнялись в Python 3.6 с использованием комбинации Numpy (v1.16.2), Pandas (v0.23.4), Scipy (v1.4.1) и sklearn (v0.22.2). Существенные и несущественные гены DepMap были загружены из DepMap Public 20Q2 по веб-адресу <https://depmap.org/portal/download/> (Blomen et al., 2015; Hart et al., 2014). Анализ обогащения набора генов (GSEA) проводили с помощью GSEAPY (v0.9.19) в Python с использованием базы данных 2019 Human KEGG Pathway.

[0369] Дизайн библиотеки мозаичного скрининга, экспериментальные спецификации и анализ

[0370] Для скрининга на основе роста была разработана мозаичная библиотека sgPHK, нацеленная на существенные гены, на основе ранее опубликованного нами полногеномного скрининга CRISPRi в K562. Для генов без канонических CGI (определяемых по отсутствию CGI в пределах 2,5 кб от TSS, причем TSS основывались на ранее опубликованных аннотациях (Horlbeck et al., 2016a) и аннотациях CGI из UCSC Genome Browser), были отобраны все гены со средней оценкой фенотипа роста менее -0,2. Для генов с одним или несколькими CGI (также определяемыми как находящиеся в пределах 2,5 кб от TSS) были отобраны гены со средней оценкой фенотипа роста между -0,2 и -0,4. Всего было отобрано 39 генов без канонических CGI, 425 генов с одним аннотированным CGI и 56 генов с несколькими CGI.

[0371] Для каждого гена, все последовательности +/-2,5 кб (или +/- 1 кб, в зависимости от положения и длины CGI) от TSS, содержащие 19 п.о., за которыми следует PAM (примыкающий к протоспейсеру мотив) NGG, были выделены в качестве потенциальных sgPHK. В начало всех последовательностей был добавлен 5' G для обеспечения надежной транскрипции с промотора U6, независимо от того, присутствовало ли это основание в геномной последовательности. sgPHK оценивали на наличие нецелевых сайтов с помощью взвешенного метода «галстук-бабочка» (Bowtie), как описано ранее (Horlbeck et al., 2016a). Вкратце, sgPHK оценивались по уникальности в геноме, определяемой эмпирически выведенной и экспериментально проверенной метрикой оценки: PAM (примыкающий к протоспейсеру мотив) G1=40, PAM G2=19, PAM N=0, следующие 7 оснований от PAM=28, следующие 5 оснований=19 и последние 7 оснований=10. Затем рассчитывали оценку несовпадения по сумме несовпадений с метрикой оценки. Эта оценка несовпадения была реализована с помощью пороговой функции оценки Phred в Bowtie с использованием флагов --nomaground, -n 3, -l 15, -a и --

best. Для наиболее строгого порога, sgPНК должны были иметь не более 1 выравнивания (сам сайт-мишень sgPНК) в геноме с оценкой несовпадения 39. Контрольные ненацеливающие sgPНК были выделены из ранее протестированного списка контрольных sgPНК (Horlbeck et al., 2016a).

[0372] Мозаичные библиотеки для эндогенно GFP-меченных CLTA, HIST2H2BE (H2B), RAB11A и VIM были сконструированы аналогично, с отбором sgPНК +/- 2,5 кб от TSS и получением ок. 600 sgPНК на ген. Последовательности протоспейсеров для каждой GFP-меченной библиотеки генов представлены в Таблице 5.

[0373] Пулы олигонуклеотидов были сконструированы с фланкирующими ПЦР и сайтами рестрикции (BstXI и BspI), синтезированы фирмой Agilent Technologies и клонированы в вектор экспрессии sgPНК pCRISPRi-v2 (Addgene, № кат. 84832), как описано ранее (Horlbeck et al., 2016a). Экспрессионный вектор содержит промотор U6, направляющий экспрессию sgPНК, а также промотор EF1 α , направляющий экспрессию пурамицина T2A-mCherry.

[0374] Мозаичные скрининги в HEK293Ts проводились по той же схеме, что и полногеномные скрининги CRISPRoff. Для проведения мозаичного скрининга в K562, клетки, стабильно экспрессирующие dCas9-KRAB, сначала трансдуцировали лентивирусными частицами библиотеки мозаичных sgPНК. Через два дня после трансдукции (20-30% инфекции) клетки обрабатывали пурамицином до тех пор, пока популяция не состояла на 90% из клеток, экспрессирующих sgPНК. Затем отбирали образец момента времени T(0) и клетки продолжали пассировать в течение еще 10 удвоений клеток, чтобы получить образец момента времени T(10).

[0375] Результаты секвенирования, полученные при проведении скрининга CRISPR, обрабатывали с помощью конвейера ScreenProcessing на базе Python (<https://github.com/mhorlbeck/ScreenProcessing>), как описано ранее (Horlbeck et al., 2016a) для расчета фенотипов sgPНК. Оценка фенотипа sgPНК определяется как \log_2 (обогащение sgPНК) / (число удвоений клеток). Все дополнительные анализы данных скрининга CRISPR и построение графиков были выполнены в Python 2.7 с использованием комбинации Numpy (v1.12.1), Pandas (v0.17.1) и Scipy (v0.17.0). Данные MNase-seq (секвенирование сайтов, доступных для микрококковой нуклеазы) для K562 и GM12878 были получены от консорциума ENCODE в виде обработанных данных непрерывного сигнала (в файловом формате BigWig; лаборатория Майкла Снайдера (Michael Snyder), Стэнфордский университет). Использовались средние значения данных MNase-seq для K562 и GM12878. Мы отметили, что фенотипы в CRISPRi K562 (медиана $\gamma = -0,46$) являются более выраженными по сравнению со скринингом CRISPRoff HEK293T (медиана $\gamma = -0,33$). Однако, как мы уже показали ранее, поскольку гены были выбраны на основе существенности в K562, эта разница, вероятно, может быть объяснена вариабельностью типов клеток между K562 и HEK293T.

[0376] Мозаичный скрининг GFP-меченных sgPНК

[0377] Мозаичные скрининги sgPНК в клеточных линиях HEK293T с GFP-метками

проводились по той же схеме, что и описанные выше скрининги на основе роста. Ранее опубликованные эндогенно GFP-меченые клеточные линии (CLTA, HIST2H2BE, RAB11A, VIM) были дополнительно отсортированы FACS для получения >99% GFP-положительных клеток, чтобы минимизировать фоновые GFP-отрицательные клетки. После создания клеточных линий, стабильно экспрессирующих соответствующую библиотеку sgPHK, была произведена трансфекция плазмидами, экспрессирующими CRISPRoff. Через два дня трансфицированные клетки сортировали и затем пассировали в течение 4 недель путем трипсинизации каждые 2-3 дня. В момент времени 4 недели каждая клеточная линия имела следующую обнаруживаемую популяцию с выключенным GFP: 21,8% CLTA, 22,7% HIST2H2BE, 3,05% RAB11A и 24,7% VIM. Популяции с включенным и выключенным GFP (GFP-on и GFP-off) сортировали FACS в отдельные контейнеры, собирая ок. 2×10^6 клеток на популяцию для каждой клеточной линии. Изменение кратности \log_2 представленности sgPHK количественно оценивали по присутствию каждой sgPHK в популяции GFP-off по сравнению с общей популяцией. Анализ проводили с использованием Python 2.7, аналогично другим мозаичным скринингам, описанным ранее.

[0378] Манипуляции с iPSC и дифференцировка нейронов

[0379] Транзистентные трансфекции iPSC проводили в 6-луночных планшетах с использованием реагента для трансфекции TransIT-LT1 Transfection Reagent (Mirus). Сначала смесь 0,5 мл mTeSR1 и 2 мкл 10 мМ ингибитора ROCK Y-27632 добавляли в каждую лунку 6-луночного планшета, покрытого матригелем. Затем смесь плазмид, кодирующих dCas9-KRAB или CRISPRoff (1 мкг), 1 мкг плазмид sgPHK и 200 нг плазмиды, кодирующей BCL-XL (Li et al., 2018), добавляли к 0,4 мл Opti-MEM™. TransIT-LT1 (12 мкл) добавляли к смеси ДНК-Opti-MEM™ и вносили в каждую лунку 6-луночного планшета. Клетки при 70-80% конфлюентности отделяли от поверхности с помощью Accutase, промывали DPBS и подсчитывали с помощью Countess (Thermo Fisher AMQAX1000). Около $1,5 \times 10^6$ клеток ресуспендировали в 1 мл mTeSR1 и добавляли в каждую лунку, содержащую смеси для трансфекции. Трансфицированные клетки сортировали через 3 дня после трансфекции с помощью BD FACS Fusion и высевали в среду mTeSR, дополненную 10 мкМ ингибитора ROCK Y-27632.

[0380] Дифференцировку нейронов проводили на iPSC при пассаже номер 46 с использованием доксициклин-индуцибельного NGN2 (Tian et al., 2019). В день -3, клетки при 70-80% конфлюентности отделяли от поверхности с помощью Accutase и промывали DPBS. Около $7,5 \times 10^5$ клеток ресуспендировали в 2 мл среды для предифференцировки N2 Pre-differentiation Media в каждой лунке 6-луночного планшета, покрытого матригелем. В день 0 клетки отделяли от поверхности с помощью Accutase и промывали DPBS. Около 5×10^5 клеток ресуспендировали в 2 мл классической среды для дифференцировки N2/B27 и высевали на планшеты, покрытые поли-D-лизином (Corning 354413). На 3-й день среду в каждой лунке аспирировали и заменяли на 2 мл свежей среды для дифференцировки N2/B27. На 7-й день удаляли из каждой лунки 1 мл среды и заменяли ее на 1 мл свежей

среды для дифференцировки N2/B27. Среда для предифференцировки N2 была приготовлена из 1X Knockout DMEM/F12 (Thermo Fisher 11320-033), 1X NEAA (Thermo Fisher 11140-050), 1X добавки N2 (Thermo Fisher 17502-048), 10 нг/мл NT-3 (PreproTech 450-03), 10 нг/мл BDNF (PreproTech 450-02), 1 мкг/мл мышинового ламинина (Thermo Fisher 23017-015), 10 нМ ингибитора ROCK Y-27632 и 2 мкг/мл доксициклина (Sigma-Aldrich D3447-500MG). Классическая среда для дифференцировки N2/B27 была приготовлена из 0,5X DMEM/F12 (Thermo Fisher 10888-033), 0,5X Neurobasal-A (10888-022), 1X NEAA, 0,5X GlutaMAX (Thermo Fisher 35050-061), 0,5X добавки N2, 0,5X добавки B27-VA (Thermo Fisher 12587010), 10 нг/мл NT-3, 10 нг/мл BDNF, 1 мкг/мл мышинового ламинина и 2 мкг/мл доксициклина.

[0381] Мы использовали iNeuron RNA-Seq (<https://kampmannlab.ucsf.edu/ineuron-rna-seq>) для подтверждения активации экспрессии гена MAPT посредством нейрональной дифференцировки iPS-клеток.

[0382] Неформальный перечень последовательностей

[0383] Специалистам будет понятно, что в приведенных в данном документе последовательностях метионин (M) может присутствовать на N-конце белка для того, чтобы инициировать трансляцию. Таким образом, SEQ ID NOS:1-15 могут дополнительно содержать метионин на N-конце.

[0384] SEQ ID NO:1 - CRISPRoff-V1 или V1 или P76 (описаны в WO 2019/204766)): KRAB (выделен жирным шрифтом; из Gilbert et al., Cell, 2013, 2014), линкеры (выделены подчеркиванием), dCas9 (курсив), метка HA (строчные буквы), сигнал ядерной локализации (NLS) SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив; остатки 612-912; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (синий флуоресцентный белок) (строчные буквы, подчеркивание))

**DAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEEPGGSGGSMDDKYSIGLAIGTNSVGWAVITDEYKVPSSKF
KVLGNTDRHSIKKNLIGALLFDSGETAATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS
FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
MIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLSKDQTYDDDLNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIH
LGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE
EVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSG
EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWRLSRKLING
IRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA**

IKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG IKELGSQI
 LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRO
 LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHA
 HDAYLNAVVG TALIKKYPKLESEFVYGDYKVVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
 KRNSDKLIARKKDWDPKKYGGFDSPTVAVSVLVAKVEK GKSKKLKSVKELLGITIMERSFFE
 KNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIR
 EQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
SRADypdyvdpdyaSGSpkkkrkvEASGSGRASPGIPGSTRNHDQEFDPKPKVYPPVPAEKRPPIRV
LSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQ
EWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL
FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDK
LELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCTEMERVF GFP
VHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPFS SGLVP
LSLRGSHMGPMIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLK
YVEDVTNVVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILOYALPR
QESORPFFWIFMDNLLTEDDOETTTRFLOTEAVTLQDVRGRDYQONAMRVWSNIP
GLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSONSLP
LSRADpkkkrkvGSGatnfsllkqagdveenpgpselikenmhmklymegtdnhhfktsegegkpyegtqtmrikvve
ggplpfadilatsflygsktfinhqtgipdffkqsfepegftwervttyedggvltatqdtlqdgeliynvkirgvnftsngpvmqkktlg
weaftetlypadgglegrndmalklvvgshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavarycdlpsk
lghkln*

[0385] SEQ ID NO:2 (p90 (KRAB-dCas9-XTEN16-Dnmt3A-Dnmt3L-P2A-BFP):
 KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014); линкеры (подчеркивание), dCas9
 (курсив); метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN16
 (прописные буквы, последовательность из 16 аминокислот), Dnmt3A (жирный шрифт,
 курсив; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот
 (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L
 (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016),
 последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP
 (строчные буквы, подчеркивание))

DAKSLTAWSR TLVTFKDV FVDF TREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGE EPPGGSGGGSMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF
 KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDS
 FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYALAH
 MIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
 LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLSKD TYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF

DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH
 LGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFE
 EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSG
 EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
 LDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLING
 IRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQI
 LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRO
 LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHA
 HDAYLNAVVGTAIIKKYKLESEFVYGDYKVVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIKKTEVQTGGFSKESILP
 KRNSDKLIARKKDWDPKKYGGFDSPTVAVSVLVAKVEK GKSKKLKSVKELLGITIMERSFFE
 KNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHDKPIR
 EQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
SRADypdyvdpdyaSGSpkkkrkvSPGSGSETPGTSESATPESNHDQEFDPKVPYPPVPAEKRKPI
RVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQK
HIQEWGPFDLVIGGSPCNDLSIVN PARKGLYE GTGRLFFEFYRLLHDARPKEGDDRPFF
 WLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVN
 DKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVF
 FPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSESSGL
VPLSLRGSHMGPMIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTL
KYVEDVTNVVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILOYALP
ROESORPFFWIFMDNLLL TEDDQETTTRFLOTEAVTLODVRGRDYQONAMRVWSNI
PGLKSKHAPLTPKEEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSL
PLSRADpkkkrkvGSGatnfsllkqagdveenpgpselikenmhmkl ymegtdnhhfck tsegegkpyegtqtmrikvv
eggplpfadilatsflygsktfinhtqgipdffkqspegftwervttyedggvltatqdt slqdgcli ynvkirgvnftsngpvmqkktl
gweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavarycdlps
klghkln*

[0386] SEQ ID NO:3 (p91 (KRAB-dCas9-Dnmt3A-Dnmt3L-P2A-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

**DAKSLTAWSR TLVTFKDV FVDF TREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGE EPPGGSGGGSMDDKYSIGLAIGTNSVGWAVITDEYKVP SKKF**

KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS
 FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
 MIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
 LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLQSKDQYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
 DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIH
 LGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFE
 EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG
 EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDF
 LDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLING
 IRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGKELGSQI
 LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELDKAGFIKRQ
 LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHA
 HDAYLNAVVGTAIIKKYKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKT
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVVKKTEVQTGGFSKESILP
 KRNSDKLIARKKDWP KKYGGFDSPTVAYSVLVAKVEK GKSKKLKSVKELLGITIMERSFFE
 KNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIR
 EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
SRADypdyvdpdySGSpkkkrkvEASGSGRASPGIPGSTRNHDQEFDPKVPYPPVPAEKRRKPIRV
LSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQ
EWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPF FWL
FENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDK
LELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHF PVFMNEKEDILWCTEMERVF GFP
VHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPFSSSLVP
LSLRGSHMGPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGGGTLK
YVEDVTNVVRRDVEKVGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILOYALPR
QESORPFFWIFMDNLLLTEDDQETTTRFLOTEAVTLODVRGRDYQONAMRVWSNIP
GLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLP
LSRADpkkkrkvGSGatnfsllkqagdveenppGSGatnfsllkqagdveenppselikenmhmklymegtvdnhh
fktsegegekpyegtqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqsfpegftwerttyvedggvltatqdtslqdgcli
ynvkirgvnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerik
eannetyveqhevavarycdlpsklghkln*

[0387] SEQ ID NO:4 (p92 (KRAB-dCas9-XTEN16-Dnmt3A-Dnmt3L-P2A-P2A-BFP):
 KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линеры (подчеркивание), dCas9
 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN16
 (последовательность из 16 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et
 al., JMB, 2013; Stepper et al., NAR, 2016), линер из 27 аминокислот (курсив,

подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A-пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

**DAKSLTAWSRTLVTFKDVFDFTREEWKLDDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEEPGGSGGGSMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF
 KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS
 FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
 MIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
 LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLQSKD TYDDDLDNLLAQIGDQYA
 DLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
 DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIH
 LGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFE
 EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG
 EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDF
 LDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLING
 IRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGKELGSQI
 LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELDKAGFIKRQ
 LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHA
 HDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKT
 EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVK KTEVQTGGFSKESILP
 KRNSDKLIARKKDWP KKYGGFDSPTVAYSVLVAKVEK GKSKKLKSVKELLGITIMERS SFE
 KNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLKGS PEDNEQKQLFVEQH KHYLDEIIEQISEFSKR VILADANLDKVL SAYNKHRDKPIR
 EQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
SRADypdyvdpdySGSpkkrkvSPGSGSETPGTSESATPESNHDQEFDPPKVYPPVPAEKRKPI
RVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQK
HIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPF
WLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVN
DKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHF PVMNEKEDILWCTEMERVF
FPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSESSGL
VPLSLRGSHMGPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGGGTL
KYVEDVTNVVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILOYALP
ROESORPFFWIFMDNLLLTEDDQETTTRFLOTEAVTLQDVRGRDYQNAMRVWSNI
PGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSL
PLSRADpkkrkvGSGatnfsllkqagdveenppgGSGatnfsllkqagdveenppgselikenmhmkl ymegtvdn
 hhfktsegegkpyegtqtmrikvvegplpfafdilatsflygsktfinhtqgipdffkqsfpegftwerttyedggvltatqdtslqdg
 cliynvki rgnvftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianiktyrskk paknlkmpgvyyvdyrler**

ikeannetyveqhevavarycdlpsklghkln*

[0388] SEQ ID NO:5 (p93 (KRAB-dCas9-XTEN80-Dnmt3A-Dnmt3L-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN80 (последовательность из 80 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

**DAKSLTAWSRTLVTFKDVVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEEPGGSGGGSMDDKYSIGLAIGTNSVGWAVITDEYKVP SKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSD
FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
MIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLSKDQYDDDLNLLAQIGDQYA
DLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIH
LGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFE
EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSG
EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLING
IRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
IKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIIEGKELGSQI
LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLT
RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELKAGFIKRQ
LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHA
HDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKT
EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTVEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGGKSKKLKSVKELLGITIMERSSE
KNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIR
EQAENIIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
SRADypdyvdpdyaSGSpkkrkvSPGGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST
EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSENHDQEFDPPKVYPPVPAEKRK
PIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQ
KHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLHDARPKEGDDRPF
FWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTV
NDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKQHF PVFMNEKEDILWCTEMERVF
GF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGP SFSS**

GLVPLSLRGSHMGPMIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGSGGG
TLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYA
LPROESORPFFWIFMDNLLLTEDDOETTTRFLOTEAVTLQDVRGRDYQONAMRVWS
NIPGLKSKHAPLTPKEEEYLQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQN
SLPLSRADpkkkrkvGSGatnfsllkqagdveenpgpselikenmhmklymegtvdnhhfktsegegkpyegtqtmri
kvveggplpfafdilatsflygsktfinhtqgipdffkqsfpegtwerwttyedggvltatqdtslqdgeliynvkirgvnftsngpvmq
kktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavaryc
dlpsklghkln*

[0389] SEQ ID NO:6 (p94 (KRAB-dCas9-XTEN80-Dnmt3A-Dnmt3L-P2A-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN80 (последовательность из 80 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A-пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

DAKSLTAWSRTLVTFKDVVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEEPGGSGGGSMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS
FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYALAH
MIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN
LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLSKDQTYDDDLNLLAQIGDQYA
DLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH
LGELHAILRRQEDFYFLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFE
EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSG
EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDF
LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWRLSRKLING
IRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
IKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGKELGSQI
LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYVD AIVPQSFLKDDSIDNKVLT
RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ
LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHA
HDAYLNAVVG TALIKKYPKLESEFVYGDYKVVYDVRKMIKAKSEQEIGKATAKYFFYSNIMNFFKT
EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGGKSKKLKSVKELLGITIMERSFFE
KNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHDRDKPIR
EQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD

SRADypdyvdpdyaSGSpkkkrkvSPGGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST
 EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSENHDQEFDPKVPYPPVPAEKRK
 PIRVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQ
 KHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPF
 FWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTV
 NDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCTEMERVF
 GPFVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPFSSS
GLVPLSLRGSHMGPMIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGG
TLKYVEDVTNVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYA
LPROESORPFFWIFMDNLLTEDDQETTTRFLOTEAVTLQDVRGRDYQONAMRVWS
NIPGLKSKHAPLTPKEEEYLQAOVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQN
 SLPLSRADpkkkrkvGSGatnfsllkqagdveenpgpGSGatnfsllkqagdveenpgpselikenmhmklymegtv
 dnhhfketsegegkpyegtqtmrikvvegplpfadilatsflygsktfinhtqgipdffkqsfpegftwerttyedggvltatqdtlq
 dgcliynvkirgvnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianiktyrskkpaknlkmpgvyyvdyr
 lerikeannetyveqhevavarycdlpsklghkln*

[0390] SEQ ID NO:7 (p95 (KRAB-XTEN16-dCas9-Dnmt3A-Dnmt3L-P2A-BFP):
 KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN16 (последовательность из 16 аминокислот), dCas9 (курсив), метка HA (строчные
 буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из
 Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив,
 подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный
 шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016),
 последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP
 (строчные буквы, подчеркивание))

DAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEEPSGSETPGTSESATPESMDKKYSIGLAIGTNSVGWAVITDE
 YKVP SKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
 MAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADLR
 LIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLL
 AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRF AWMTRKSE
 ETITPWNFEVVDKGAASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD
 LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGW
 GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHE
 HIANLAGSPAIKKGILQTVKVVDLKVVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE
 EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE

LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK
 VREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VVKTEVQ
 TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELL
 GITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
 PSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL
 AYNKHRDKPIREQAENIHLFTLNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYE
 TRIDLSQLGGDSRADypdyvpdyasGSspkkkrkvEASGSGRASPGIPGSTRNHDQEFDPKVPYPP
 VPAEKRPVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVG
 DVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK
 EGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGM
 NRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKQDQHFVFMNEKEDILW
 CTEMERVF GF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNA
NSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLES
GSGSGGGTLKYVEDVTNVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQF
HRILOYALPROESQRPFWFIFMDNLLTEDDOETTTRFLOTEAVTLQDVRGRDYQ
NAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLRE
YFKYFSQNSLPLSRADpkkkrkvGSGatnfsllkqagdveenpgpselikenmhmklmegtvdnhhfktsege
gkpyegtqtmrikvveggplpafdilatsflygsktfinhtqgipdffkqsfpegtwerwttyedggvltatqdtlsldgcliynvking
vnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannety
veqhevavarycdlpsklghkln*

[0391] SEQ ID NO:8 (p96 (KRAB-XTEN16-dCas9-Dnmt3A-Dnmt3L-P2A-P2A-BFP):
 KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN16 (последовательность из 16 аминокислот), dCas9 (курсив), метка HA (строчные
 буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из
 Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив,
 подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный
 шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016),
 последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт); BFP
 (строчные буквы, подчеркивание))

DAKSLTAWSR TLVTFKDV FVDF TREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEEPSGSETPGTSESATPESMDKKYSIGLAIGTNSVGWAVITDE
 YKVP SKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
 MAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR
 LIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNLFGNLIASLGLTPNFKSNFDLAEDAQLQSKDQYDDDLNLL
 AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYFKIPILEKMDGTEELLVKNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRFAWMTRKSE
 ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG

MRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD
 LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW
 GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHE
 HIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE
 EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
 LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK
 VREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVVYDVRKMKIAKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VVKTEVQ
 TGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
 GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
 PSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKKHYLDEIEQISEFSKRVLADANLDKVL
 AYNKHRDKPIREQAENIIHLFTLNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYE
 TRIDLSQLGGDSRADypdyvpydyaSGSpkkkrkvEASGSGRASPGIPGSTRNHDQEFDPPKVYPP
 VPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVG
 DVRSVTQKHIQE WGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK
 EGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGM
 NRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHF PVMNEKEDILW
 CTEMERVF GF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLF APLKEYF ACVSSGNSNA
 NSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLES
GSSGGGTLKYVEDVTNVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQF
HRILOYALPROESORPFFWIFMDNLLTEDDOETTTRFLOTEAVTLODVRGRDYQ
NAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLRE
YFKYFSQNSLPLSRADpkkkrkvGSGatnfsllkqagdveenpgpGSGatnfsllkqagdveenpgpseliken
mhmklymegetvdnhhfktsegegkpyegtqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqsfpegtwerwttye
dggvltatqdtslqdgcliynvkirgvnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknl
kmpgvyyvdyrlerikeannetyveqhevavarycdlpsklghkln*

[0392] SEQ ID NO:9 (p97 (KRAB-XTEN80-dCas9-Dnmt3A-Dnmt3L-P2A-BFP):
 KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN80 (последовательность из 80 аминокислот), dCas9 (курсив); метка HA (строчные
 буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из
 Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив,
 подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный
 шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016),
 последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP
 (строчные буквы, подчеркивание))

DAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEPPGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTS
 TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVIT
 DEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS

NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA
DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILS
ARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDL
NLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ
QLPEKYKEIFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRK
SEETITPWNFEVVDKGSASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
DYLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK
RIEIGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQS
FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQ
FYKREINNYHHAHDAYLNAVVGTAIIKKYKLESEFVYGDYKVVYDVRKMIKSEQEIGKATA
KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKK
EVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGGKSKLKS
VKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQK
GNE LALPSKYVNFYLAHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADAN
LDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLI
HQSITGLYETRIDLSQLGGDSRADypdyvpyaSGSpkkkrkvEASGSGRASPGIPGSTRNHDQEFDP
PKV YPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDA
RPKEGDDRPFWFLENVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLP
GMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKQDQHFVFMNEKEDI
LWCTEMERVF GF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNS
NANSRGPSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRVLSLERNIDKVLKSLGF
LESGSGSGGGTLKYVEDVTNVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYM
FOFHRILOYALPROESQRPFWFIFMDNLLTEDDQETTTRFLOTEAVTLQDVRGRD
YQNAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPL
REYFKYFSQNSLPLSRADpkkkrkvGSGatnfsllkqagdeenpgpselikenmhmklmegtvdn
hhfkcts egegkpyegtqtmrikvveggplpfadilatsflygsktfinhtqgipdffkqspegftwerv
ttiedggvltatqdtslqdgcliynvk irgvnftsnqpvmqkktlgweaftetlypadgglegrndmalklv
ggshlianikttyrskkpaknlkmpgvyyvdyrlerikeanne
tyveqhevavarycdlpsklghkln*

[0393] SEQ ID NO:10 (p98 (KRAB-XTEN80-dCas9-Dnmt3A-Dnmt3L-P2A-P2A-BFP):
KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линеры (подчеркивание),
XTEN80 (последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные
буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из
Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив,
подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный

шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

**DAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEPPGGPSSGAPPPSSGGSPAGSPTSTEEGTSESATPESGPGTS
TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVIT
DEYKVP SKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA
DLRLIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLF EENPINASGVDAKAILS
ARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFLAEDAQLQLSKD TYDDDL
NLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ
QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRK
SEETITPWNFEVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH
DLLKIIKDKDFLDNEENEDILEDIVLTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMK
RIEIGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQS
FLKDDSIDNKVLRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQ
FYKREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATA
KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKT
EVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVK
ELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNE
LALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVLADANLDK
VLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITG
LYETRIDLSQLGGDSRADypdyvpdyasGSspkkkrkvEASGSGRASPGIPGSTRNHDQEFDPK V
YPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
YVGDVRSVTQKHIQE WGPFDLVIGGSPCNDSLIVNPARKGLYEGTGRLFFEFYRLLHDA
RPKEGDDRPF WLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLP
GMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHF PVMNEKEDI
LWCTEMERVF GF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLF APLKEYFACVSSGNS
NANSRGP SFSSGLVPLSLRGSHMGPMIYKTVSA WKROPVRVLSLERNIDKVLKSLGF
LESGSGSGGGTLKYVEDVTNVRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYM
FOFHRILOYALPROESQRPFWFIFMDNLLTEDDQETTTRFLOTEAVTLODVRGRD
YQNAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPL
REYFKYFSQNSLPLSRADpkkkrkvGS GatnfsllkqagdeenpgpGS Gatnfsllkqagdeenpgpselik
enmhmklymegtvdnhhfctsegegkpyegtqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqspegftwervtt
yedggvltatqdtslqdgcliynvkirgvnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpa**

knlkmpgvyyvdyrlerikeannetyveqhevavarycdlpsklghkln*

[0394] SEQ ID NO:11 (p99 (KRAB-XTEN16-dCas9-XTEN80-Dnmt3A-Dnmt3L-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), XTEN16 (последовательность из 16 аминокислот), dCas9 (курсив), метка HA (строчные буквы), линкеры (подчеркивание), NLS SV40 (строчные буквы, курсив), XTEN80 (строчные буквы, курсив, жирный шрифт, последовательность из 80 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт (old), подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

**DAKSLTAWSRTLVTFKDVVFVDFREEWKLLDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEEPSGSETPGTSESATPESMDKKYSIGLAIGTNSVGWAVITDE
YKVP SKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
MAKVDDSSFHRLVESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR
LIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
SKSRLENLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLL
AQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSE
ETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD
LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGW
GRLSRKLINGIRDKQSGKTILDFLKSDFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHE
HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE
EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLK
DDSIDNKVLRSDKNRKGSDNVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK
VREINNYHHAHDAYLNAVVG TALIKKYPKLESEF VYGDYKVVYDVRKMIKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVAKVEK GKSKKLKSVKELL
GITIMERSSEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNF LYLASHYEKLGSPEDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVLADANLDKVL
AYNKHRDKPIREQAENIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYE
TRIDLSQLGGDSRADypdydpdySGSpkkkrkvSPGggpssgappssgspagsptsteegtsesatpespggststep
segsapgsapgsptsteegtstepsegsapgtstepseNHDQEF DPPKVYPVPAEK RKPIRVLSLFDGIATG
LLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVI
GGSPCNDLSIVNPARKGLYE GTGRLFFEFYRLLHDARPKEGDDRPF F WLFENVVAMGV
SDKRDISRFLSNPVMIDAKEVSA AHRARYF WGNLPGMNRPLASTVNDKLELQECLEH
GRIAKFSKVRTITTRSNSIKQGKDQHF PVFMNEKEDILWCTEMERVF GF PVHYTDVSN**

MSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSSFSSGLVPLSLRGSHM
GPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLYVEDVTNV
VRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILLOYALPROESORPFFW
IFMDNLLLTEDDQETTTRFLOTEAVTLQDVRGRDYQONAMRVWSNIPGLKSKHAPL
TPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLSRAD_{pkkkrkv}
GSGatnfsllkqagdveenppgselikenmhmklmegtvdnhhfktsegegkpyegtqtmrikvveggplpfafdilatsfly
gsktfinhtqgipdffkqsfpegftwervttyedggvltatqdtslqdgcliynvkiqrgvnftsnpgvmqkktlgweaftetlypadggle
grndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavarycdlpsklghkln*

[0395] SEQ ID NO:12 (p100 (KRAB-XTEN16-dCas9-XTEN80-Dnmt3A-Dnmt3L-P2A-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), XTEN16 (последовательность из 16 аминокислот), dCas9 (курсив), метка HA (строчные буквы), линкеры (подчеркивание), NLS SV40 (строчные буквы, курсив), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

DAKSLTAWSRTLVTFKDVVDFDTREEWKLDDTAQQIVYRNVMLENYKNLV
SLGYQLTKPDVILRLEKGEEPSGSETPGTSESATPESMDKKYSIGLAIGTNSVGWAVITDE
 YKVP SKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
 MAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR
 LIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLL
 AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRF AWMTRKSE
 ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD
 LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGW
 GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHE
 HIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE
 EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
 LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK
 VREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VVKTEVQ
 TGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
 GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
 PSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVL S

AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYE
 TRIDLSQLGGDSRADypdyvpdyasGSppkkkrkvSPGggpssgapppsggspagsptsteegtsesatpesgpgtstep
 segsapgsapgsptsteegtstepsegsapgtstepseNHDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATG
 LLVLKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVI
 GGSPCNDLSIVNPARKGLYE GTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGV
 SDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEH
 GRIAKFSKVRTITTRSNSIKQGDQHF PVFMNEKEDILWCTEMERVF GPFVHYTDVSN
 MSRLARQRLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHM
GPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGGGTLYVEDVTNV
VRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILLOYALPROESQRPFFW
IFMDNLLLTEDDQETTTRFLOTEAVTLQDVRGRDYQONAMRVWSNIPGLKSKHAPL
TPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLSRADpkkkrkv
GSGatnfsllkqagdveenppgGSGatnfsllkqagdveenppgselikenmhmklymegtvdnhhfktsegegkpyeg
tqtmrikvveggplpfadilatsflygsktfinhtqgipdffkqspegftwervttyedggvltatqdtslqdgeliynvkirgvnftsg
pvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqheva
varycdlpsklghkln*

[0396] SEQ ID NO:13 (p101 (KRAB-XTEN80-dCas9-XTEN16-Dnmt3A-Dnmt3L-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN16 (последовательность из 16 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

DAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEEPggpssgapppsggspagsptsteegtsesatpesgpgtstepsegsapgsapgspt
 steegtstepsegsapgtstepseMDKKYSIGLAIGTNSVGVAVITDEYKVP SKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEEDK
 KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN
 PDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFG
 NLIALLSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS
 DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFDQSKNGYAGYIDGGA
 SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYF
 FLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIERM
 TNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK
 VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT
 LTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
 DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDL

VKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNE
 KLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNRGKSDNVPS
 EEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVETRQITKHVAQIL
 DSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF YK VREINNYHHAHDAYLNAVVG TALI
 KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
 ETNGETGEIVWDKGRDFATVRKVL SMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKEK GKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYK
 EVK KDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPED
 NEQKQLFVEQH KHYLDEPIEQISEFSKR VILADANLDKVL SAYNKH RDKPIREQAENIHLFTLT
 NLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDSRADypydvpdyaS
GSpkkkkrkvSPGSGSETPGTSESATPESNHDQEFDPPKVYPPVPAEK RKPIRVL SFLDGIATG
 LLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQE WGPFDLVI
 GGSPCNDLSIVN PARKGLYE GTGRLFFEFY RLLHDARPKEGDDRPF FWFENVVAMGV
 SDKRDISRFLESNPVMIDAKEVSA AHRARYF WGNLPGMNRPLASTVNDKLELQECLEH
 GRIAKFSKVRTITTRS NSIKQGKDQHF PVMNEKEDILWCTEMERVF GPFVHYTDVSN
 MSRLARQRLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHM
GPMEIYKTVSAWK RQPVRLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNV
VRRDVEKWGPFDLVYGSTOPLGSSCDRC PGWYMFQFHRIL OYALPROESORPFFW
IFMDNLLL TEDDQETTTRFLOTEAVTLQDVRGRDYQONAMRVWSNIPGLKSKHAPL
TPKEEYLOAQVRSR SKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLSRADpkkkkrkv
GSGatnfsllkqagdveenpgpselikenmhmklymegtvdnhhfktsegegkpyegtqtmrikvveggplpfafdilatsfly
 gsktfinhtqgipdfkqsfpegftwervttyedggvltatqdtslqdgcliynv kirgvnftsngpvmqkktlgweaftetlypadggle
 grndmalklv gsshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavarycdlpsklghkln*

[0397] SEQ ID NO:14 (p102 (KRAB-XTEN80-dCas9-XTEN16-Dnmt3A-Dnmt3L-P2A-BFP): KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), XTEN16 (последовательность из 16 аминокислот), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), последовательность расщепляемого P2A пептида (строчные буквы, жирный шрифт), BFP (строчные буквы, подчеркивание))

DAKSLTAWSR TLVTFKDV FVDF TREEWKLLDTAQQIVYRNVMLENYKNLV
 SLGYQLTKPDVILRLEKGEEPggpssgappssgspagsptsteegtsesatpesgpgtstepsegsapgsapspt
 steegtstepsegsapgtstepseMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEEDK
 KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLN
 PDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRENLIAQLPGEKKNGLFG
 NLIALLSLGLTPNFKSNFDLAEDA KLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLS

DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA
 SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYF
 FLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIERM
 TNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK
 VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT
 LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
 DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDEL
 VKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE
 KLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNRKGSDNVPS
 EEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL
 DSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGITALI
 KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
 ETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYK
 EVKCDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPED
 NEQKQLFVEQHKLHYLEPIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLT
 NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQITGLYETRIDLSQLGGDSRADypydvpdyaS
GSpkkkrkvSPGSGSETPGTSESATPESNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATG
LLVLKDLGIQVDRYIASEVCESITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVI
GGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGV
SDKRDISRFLSNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEH
GRIAKFSKVRTITTRSNSIKQGDQHFPVFMNEKEDILWCTEMERVFGFVHYTDVSN
MSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSSGLVPLSLRGSHM
GPMEIYKTVSAWKROPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNV
VRRDVEKWGPFDLVYGSTOPLGSSCDRCPGWYMFQFHRILLOYALPROESORPFFW
IFMDNLLLTEDDQETTTRFLOTEAVTLQDVRGRDYQONAMRVWSNIPGLKSKHAPL
TPKEEYLOAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLSRADpkkkrkv
GSGatnfsllkqagdveenpgpGSGatnfsllkqagdveenpgpselikenmhmklymegtvdnhhfktsegegkpyeg
tqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqspegftwervttiedggvltatqdtlsldgcliynvkirgvnftsng
pvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknllkmpgvyyvdyrlerikeannetyveqheva
varycdlpsklghkln*

[0398] SEQ ID NO:15 (V2.1 или p112 (Dnmt3A-Dnmt3L-XTEN80-dCas9-BFP-KRAB); KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), BFP (строчные буквы, подчеркивание))

NHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE

DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVNPARKGLYEG
 TGRLFF EFYRLLHDARPKEGDDRPF F WLFENVVAMGVSDKRDISRFLESNPVMIDAKE
 VSAHRARYF WGNLPGMNRPLASTVNDKLELQECLEHGRIAKF SKVRTITTRSNSIKQG
 KDQHF PVMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRH
 LFAPLKEYF ACVSSGNSNANSRGPSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTO
PLGSSCDRCPGWYMFQFHRILQYALPROESORPFFWIFMDNLLLTEDDOETTTRFL
QTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDA
PKVDLLVKNCLLPLREYFKYFSQNSLPLggpssgappppsggpagsptsteegtsesatpesgpgtstepseg
 saggspagsptsteegtstepsegsapgtstepseMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF KVLGNTD
 RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSSFFHRLLEE
 SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH
 FLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGE
 KKNGLFGNLIALSLGLTPNF KSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFDQSKNGY
 AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
 RRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGA
 SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
 EDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQS
 GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGIL
 QTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHP
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKN
 RGKSDNVPSEE VVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVETR
 QITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVINNYHHAHDAY
 LNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITL
 ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGF SKESILPKNR
 SDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPI
 DFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
 EKLLKGS PEDNEQQLFVEQH KHYLDEIIEQISEFSKR VILADANLDKVL SAYNKHRDKPIREQ
 AENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDAyp
 ydvpdyaSLGSGSpkkkrkvEDpkkkrkvDGIGSGSNGSSGSselikenmhmkl ymegtvdnhhfctsegegkpy
egtqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqspegftwervttyedggvltatqdtlsldgcliynv kirgvnfts
ngpvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhe
vavarycdlpsklghklnGGGGMDAKSLTAWSR TLVTFKDV FVDFTREEWKLLDTAQQIV
YRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP*

[0399] SEQ ID NO:16 (KRAB; из Gilbert et al., Cell, 2013, 2014)

DAKSLTAWSR TLVTFKDV FVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG
YQLTKPDVILRLEKGEEP

[0400] SEQ ID NO:17 (линкер)

GGSGGGS

[0401] SEQ ID NO:18 (линкер)

SGS

[0402] SEQ ID NO:19 (линкер)

EASGSGRASPGIPGSTR

[0403] SEQ ID NO:20 (линкер)

SRAD

[0404] SEQ ID NO:21 (линкер)

GSG

[0405] SEQ ID NO:22 (линкер)

SPG

[0406] SEQ ID NO:23 (dCas9)

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
 NPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
 NGLFGNLIASLGLTPNFKSNFDLAEDAQLSKDQTYDDDLNLLAQIGDQYADLFLAA
 KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQ
 SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
 HLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVVGPLARGNSRFAWMTRKSEETIP
 WNFEEVVDKGAASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLK
 RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ
 KGQKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDI
 NRLSDYDVDAIVPQSFLKDDSIDNKVLRSDKNRSGSDNPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLES
 EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 GETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEA
 KGYKEVKKDLIIPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
 KLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRDKPIR
 EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQL
 GGD

[0407] SEQ ID NO:24 (метка HA)

YPYDVPDYA

[0408] SEQ ID NO:25 (сигнал ядерной локализации (NLS) SV40)

PKKKRKV

[0409] SEQ ID NO:26 (Dnmt3A; остатки 612-912; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016)

NHDQEFDPKPYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE
 DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEG
 TGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEV
 SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK
 DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGLRSWSVPVIRHLF
 APLKEYFACV

[0410] SEQ ID NO:27 (линкер из 27 аминокислот; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016)

SSGNSNANSRGPFSGLVPLSLRGS

[0411] SEQ ID NO:28 (Dnmt3L; из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016)

MGPMEIYKTVSAWKRQPVRLVLSLFRNIDKVLKSLGFLESGSGGGTLKYVEDV
 TNVVRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPF
 IFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPK
 EEEYLQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL

[0412] SEQ ID NO:29 (последовательность расщепляемого P2A пептида)

ATNFSLLKQAGDVEENPGP

[0413] SEQ ID NO:30 (BFP)

SELIKENMHMCLYMEGTVDNHHFKCTSEGEKPYEGTQTMRIKVVVEGGPLPFAF
 DILATSFLYGSKTFINHTQGIPDFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGCLI
 YNVKIRGVNFTSNGPVMQKKTGWAEFTETLYPADGGLEGRNDMALKLVGGSHLIANI
 KTTYRSKKPAKNLKMGPVYYVDYRLRIKEANNETYVEQHEVAVARYCDLPSKLGHK
 LN*

[0414] SEQ ID NO:31 (XTEN16 (последовательность из 16 аминокислот))

SGSETPGTSESATPES

[0415] SEQ ID NO:32 (XTEN80 (последовательность из 80 аминокислот))

GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST
 EEGTSTEPSEGSAPGTSTEPSE

[0416] SEQ ID NO:33 (домен Dnmt3A-Dnmt3L)

NHDQEFDPKPYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE
 DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEG
 TGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEV
 SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK
 DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGLRSWSVPVIRHLF
 APLKEYFACVSSGNSNANSRGPFSGLVPLSLRGS
 HMGPMMEIYKTVSAWKRQPVRLVLSLFRNIDKVLKSLGFLESGSGGGTLKYVEDV
 TNVVRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPF
 WIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPK
 EEEYLQAQVRSRSLDAPKVDLLVKN

LLPLREYFKYFSQNSLPL

[0417] SEQ ID NO:34 (ddAsCfp1)

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNNDHYKELKPIIDR
 IYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDFYFIGRTD
 NLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYE
 NRKNVFS AEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSI
 EEVFSFPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIISLP
 HRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLNRNENVLETAEALFNELNSIDLT
 HIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEII
 SAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQDLSLLGLYHLLDWF
 AVDESNEVDPEFSARLTGIKLEMPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASG
 WDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEKTSEGFDMYYDYFPD
 AAKMIPKCSTQLKAVTAHFQTHHTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQATAYAK
 KTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHI
 SFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKL
 NGQAELFYRPKSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDL
 SDEARALLPNVITKEVSHEIHKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEH
 PETPIIGIDRGERNLIYITVIDSTGKILEQRSLNITQQFDYQKKLDNREKERVAAARQAWSV
 VGTIKDLKQGYLSQVIHEIVDLMIHYQAVVLANLNFQFKSKRTGIAEKAVYQQFEKML
 IDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTG
 FVDPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAW
 DIVFEKNETQFDAQGTPFIAGKRIVPVIEHRFTGRYRDLYPANELIALLEEKGIVFRDGS
 NILPKLLENDSDHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNP
 WPMADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELRN

[0418] SEQ ID NO:35 (ddLbCfp1)

MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLEVEDEKRAEDYKGVKLL
 DRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGN
 YKSLFKKDIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFR
 CI NENLTRYISNMDIFEKVD AIFDKHEVQEIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVY
 NAIIGGFVTESEGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLSFYGEGYTS
 D EEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNPAISTISKDIFGEWNVIR
 DKWNAEYDDIHLKKA VVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLKEII
 IQKVDEIYKVYGSSEKLFDAADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYKAFFGEG
 KETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWD
 KDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKML
 PKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPKWSNA
 YDFNFSETEKYKDIAGFYREVVEEQYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSD
 KSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNP
 DNPKKTTLSDYVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGI
 ARGERNLLYIVVVDGKGNIVEQYSLNEIINNFNIRIKTDYHSLDKKEKERFEARQNW

SIENIKELKAGYISQVVHKICELVEKYDAVIALADLNSGFKNRSRVKVEKQVYQKFEKMLI
 DKLNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFV
 NLLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNR
 IRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMAL
 MSLMLQMRNSITGRTDVDFLISPVKNSDGFYDSRNYEAQENAILPKNADANGAYNIAR
 KVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKH

[0419] SEQ ID NO:36 (ddFnCfp1)

MYPYDVPDYASGSGMSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDE
 KRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSA
 KDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDNIELFKANSDITDI
 DEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLNKAKEYESLKD
 KAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFLSDEVFEIANFNLYLNQSGITKFNTII
 GGKFNVENTKRKGINEYINLYSQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDD
 SDVVTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQ
 VFDDYSVIGTAVLEYITQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALFEFNKH
 RDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAI
 KDLLDQTNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYIT
 QKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYLLGVMNKKNNKIFDDKA
 IKENKGEQYKKIVYKLLPGANKMLPKVFFSAKSIKFNPSSEDILRIRNHSTHTKNGSPQK
 GYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNIDEFYREVENQGYKLT
 FENISESYIDSVVNQGLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKL
 NGEAELFYRKQSIPKKITHPAKEAIAKNKNDNPKKESVFEYDLIKDKRFTEDKFFFHCPT
 INFKSSGANKFNDEINLLLKEKANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGN
 DRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVV
 FEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAYQLTAPFE
 TFKKMGKQTGIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKG
 YFEFSFDYKNFGDKAAKGGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKELEKLLKD
 YSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNF
 FDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKNLVIKNEEYFEFVQNR
 NN

[0420] SEQ ID NO:97=V2.2 (DNMT3A-DNMT3L-XTEN80-dCas9-HA-NLS-NLS-
 XTEN16-KRAB-P2A-BFP)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80
 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы,
 курсив), DNMT3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al.,
 NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB,
 2013; Stepper et al., NAR, 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et
 al., JMB, 2013; Stepper et al., NAR, 2016), BFP (строчные буквы, подчеркивание)); линкер
 XTEN16 указан как: (**SGSETPGTSESATPES**); последовательность расщепляемого P2A

пептида указана как ((ATNFSLLKQAGDVEENPGP)). Скобки в последовательности использованы только с целью идентификации линкера XTEN16 и последовательности расщепляемого P2A-пептида.

**MNHDQEFDPKPYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEV
 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
 EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDA
 KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
 QGKDQHPVFMNEKEDILWCTEMERVF GFFPVHYTDVSNMSRLARQRLGRSWSVPVI
 RHLFAPLKEYFACVSSGNSNANSRGPSSFSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFLESGSGSGGTLKYVEDVTNVVRRDVEKWGPFDLVGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLLTEDDQETTTRFLOTEAV
TLODVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLOAQVRSRSKLDAPKVDLLVK
NCLLPLREYFKYFSONSLPLgpssgappsgspagsptsteegtseatpesgptsepsesapgtsepse
 gtsepsesapgtsepseMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
 LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH
 ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD
 NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL
 IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQLPEKYKEIFDQSKNGYAGYIDGGAS
 QEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP
 FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM
 TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK
 VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNENEDILEDIVLT
 LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
 DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVDEL
 VKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE
 KLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS
 EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL
 DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHAHDAYLNAVVGTALI
 KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
 ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVASVLVVAKVEKGKSKKLSVKELLGITIMERSSFEKNPIDFLEAKGYK
 EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
 NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT
 NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDApydvpdyaSLGS
GSpkkrkvEDpkkrkvDGIGSGSNGSSGS(SGSETPGTSESATPES)GGGGGMDAKSLTAWS
RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILR
LEKGEEP((ATNFSLLKQAGDVEENPGP))selikenmhmklymegtvdnhhfktsegegkpyegtmri
kvveggplpfadilatsflygsktfinhtgipdffksfpegftwervtyedggvltatqdtslqdgcliynkirgvnftsngpvmq
kktlgweaftetlypadgglegrndmaklvggshlianikttyrskkpaknlkmpgvyvdyrlerikeannetyveqhevaryc**

dlpsklghkln*

[0421] SEQ ID NO:98=V2.3 (DNMT3A-DNMT3L-XTEN80-NLS-dCas9-NLS-XTEN16-KRAB-P2A-BFP)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), DNMT3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), BFP (строчные буквы, подчеркивание)); линкер XTEN16 указан как (**SGSETPGTSESATPES**); последовательность расщепляемого P2A пептида указана как ((**ATNFSLLKQAGDVEENPGP**)). Скобки в последовательности использованы только с целью идентификации линкера XTEN16 и последовательности расщепляемого P2A пептида.

**MNHDQEFDPKVPYPPVPAEKRRKPIRVLSLFDGIATGLLVKDLGIQVDRIASEV
CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
EGTGRLLFFEFYRLLHDARPKEGDDRPFWFLENVAMGVSDKRDISRFLESNPVMIDA
KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
QGKDQHFVFMNEKEDILWCTEMERVFVFPVHYTDVSNMSRLARQRLGRSWSVPVI
RHLFAPLKEYFACVSSGNSNANSRGPSSGLVPLSLRGSHMGPMIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFLGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLLTEDDQETTTRFLQTEAV
TLODVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVK
NCLLPLREYFKYFSQNSLPLggpssgapppsggspagsptsteegtsesatpesgpgtstepsegsapgsapgsptste
gstepsegsapgtstepsepkkkrkvMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF KVLGNTDRHSIKK
NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEE
DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYALAHMIKFRGHFLIED
LNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGL
FGNLIASLGLTPNFKSNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSF
IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILED
IVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILD
FLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV
DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGIKELGSQILKEHPVENTQL
QNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH
VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAV**

GTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR
 KRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIAR
 KKDWDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLSVKELLGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSP
 EDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHDPKIREQAENIHLF
 TLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDpkkkrkv(SGSE
 TPGTSESATPES)RTLVTFKDVFVDFTREEWKL LDTAQQIVYRNVMLENYKNLVSL
 GYQLTKPDVILRLEKGEEP((ATNFSLKQAGDVEENPGP))selikenmhmklymegtvdnhhf
kctsegegkpyegtqtmrikvveggplpfafdilatsflygsktfinhtqgipdffkqsfpegtwerwttyedggvltatqdtlsldgcli
ynvkirgvnftsnpgvmqkktlgweaftetlypadgglegrndmalklvggshlianikttyrskkpaknlkmpgvyyvdyrlerik
eannetyveqhevavarycdlpsklghkln*

[0422] SEQ ID NO:99=V2.4 (DNMT3A-DNMT3L-XTEN80-NLS-dCas9-NLS-XTEN16-BFP-KRAB)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), метка HA (строчные буквы), NLS SV40 (строчные буквы, курсив), DNMT3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), BFP (строчные буквы, подчеркивание)); линкер XTEN16 указан как (SGSETPGTSESATPES). Скобки в последовательности использованы только с целью идентификации последовательности линкера XTEN16.

MNHDQEFDPKVPVPAEKRKPIRVL SLDGIATGLLVKDLGIQVDRIASEV
 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVNPARKGLY
 EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENNVAMGVSDKRDISRFLSNPVMIDA
 KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
 QGKDQHF PVFMNEKEDILWCTEMERVF GPFVHYTDVSNMSRLARQRLGRSWSVPVI
 RHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFL ESGSGGGT LKYVEDVTNVVRRDVEKWGPF DLVYGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLLTEDDQETTTRFLQTEAV
TLODVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVK
NCLLPLREYFKYFSQNSLPLggpssgappsggspagsptsteegtsesatpespgtstepsegsapgsptste
gtstepsegsapgtstepsepkkkkrkvMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF KVLGNTDRHSIKK
 NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FHRLEESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYALAHMIKFRGHFLIEGD
 LNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFLAEDAQLSKDQTYDDLDNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
 GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSF

IERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED
 IVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILD
 FLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVV
 DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
 QNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVETRQITKH
 VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAV
 GTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR
 KRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIAR
 KKDWDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLKSVKELLGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSP
 EDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLF
 TLTNLGAFAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDpkkkrkv(SGSE
TPGTSESA TPES)selikenmhmkl ymegtvdnhhfcktsegegkpyegtqtmrikvveggplpfafdilatsflygsktfin
htqgipdffkqspegftwervttyedggvltatqdtslqdgeliynv kirgvnftsngpvmqkktlgweaftetlypadgglegrndm
alklvggshlianikttyrskkpaknlkmpgvyyvdyrlerikeannetyveqhevavarycdlpsklghkl nGGGGGMDAK
SLTAWSR TLVTFKDV FVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTK
PDVILRLEKGEEP*

[0423] SEQ ID NO:100

GACGCTCAAATTTCCGCAGTGTTTAAGAGCTAAGCTGGAAACAGCATAGCAA
 GTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTT
 TTTT

[0424] SEQ ID NO:101

GTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCG
 TTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

[0425] SEQ ID NO:102

GACGCTCAAATTTCCGCAGT

[0426] SEQ ID NO:103 (KRAB с иницирующим метионином)

MDAKSLTAWSR TLVTFKDV FVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL
 GYQLTKPDVILRLEKGEEP

[0427] SEQ ID NO:104 (KRAB)

RTLVTFKDV FVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVIL
 RLEKGEEP

[0428] SEQ ID NO:105 (KRAB)

MRTLVTFKDV FVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDV
 ILRLEKGEEP

[0429] SEQ ID NO:106 (Dnmt3A с метионином на N-конце)

MNHDQEFDPKVPVPAEKRPVLSLFDGIATGLLVKDLGIQVDRYIASEV
 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY

EGTGRLFFEFYRLLHDARPKEGDDRPFVWFENNVAMGVSDKRDISRFLSNPVMIDAK
 EVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQ
 GKDQHFPVFMNEKEDILWCTEMERVFVGFVHYTDVSNMSRLARQRLGRSWSVPVIRH
 LFAPLKEYFACV

[0430] SEQ ID NO:107 (V2.1 или p112 (Dnmt3A-Dnmt3L-XTEN80-dCas9-BFP-KRAB); KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), NLS SV40 (строчные буквы, курсив), Dnmt3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), Dnmt3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016),

**NHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCE
 DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDSLIVNPARKGLYEG
 TGRLFFEFYRLLHDARPKEGDDRPFVWFENNVAMGVSDKRDISRFLSNPVMIDAKE
 VSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQG
 KDQHFPVFMNEKEDILWCTEMERVFVGFVHYTDVSNMSRLARQRLGRSWSVPVIRH
 LFAPLKEYFACV****S****S****G****N****S****N****A****N****S****R****G****P****S****F****S****S****G****L****V****P****L****S****L****R****G****S****H****M****G****P****M****E****I****Y****K****T****V****S****A****W****K****R****O****P****V****R****V**
L**S****L****F****R****N****I****D****K****V****L****K****S****L****G****F****L****E****S****G****S****G****S****G****G****T****L****K****Y****V****E****D****V****T****N****V****R****R****D****V****E****K****W****G****P****F****D****L****V****Y****G****S****T****Q**
P**L****G****S****S****C****D****R****C****P****G****W****Y****M****F****Q****F****H****R****I****L****O****Y****A****L****P****R****O****E****S****O****R****P****F****F****W****I****F****M****D****N****L****L****L****T****E****D****D****Q****E****T****T****R****F****L**
Q**T****E****A****V****T****L****O****D****V****R****G****R****D****Y****Q****N****A****M****R****V****S****N****I****P****G****L****K****S****K****H****A****P****L****T****P****K****E****E****E****Y****L****Q****A****Q****V****R****S****R****S****K****L****D****A**
P**K****V****D****L****L****V****K****N****C****L****L****P****L****R****E****Y****F****K****Y****F****S****Q****N****S****L****P****L****g****p****s****s****g****a****p****p****p****s****g****s****p****a****g****s****p****t****s****t****e****e****g****t****s****e****s****a****t****p****e****s****g****p****t****s****e****p****e****s****a****p****t****s****e****p****s****e****M****D****K****K****Y****S****I****G****L****A****I****G****T****N****S****V****G****W****A****V****I****T****D****E****Y****K****V****S****K****K****F****K****V****L****G****N****T****D**
 RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEE
 SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH
 FLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGE
 KKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFDQSKNGY
 AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
 RRQEDFYFLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGA
 SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
 EDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQS
 GKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGIL
 QTVKVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHP
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYVDVAIVPQSFLKDDSIDNKVLTRSDKN
 RGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR
 QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVINNYHHAHDAY
 LNAVVGTAIIKKYPKLESEFVYGDYKVDYVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITL
 ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIKKTEVQTGGFSKESILPKRN

SDKLIARKKDWDPPKYGGFDSPTVAVSVLVVAKVEKKGSKKLKSVKELLGITIMERSSSFENPI
DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQ
AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDAS
LGSGSpkkkrkvEDpkkkrkvDGIGSGSNGSSSGSGGGGMDAKSLTAWRSLVTFKDVVDF
TREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP*

[0431] SEQ ID NO:108=V2.2 (DNMT3A-DNMT3L-XTEN80-dCas9-HA-NLS-NLS-XTEN16-KRAB-P2A-BFP)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание), XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80 аминокислот), dCas9 (курсив), NLS SV40 (строчные буквы, курсив), DNMT3A (жирный шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), BFP (строчные буквы, подчеркивание)); линкер XTEN16 указан как: (**SGSETPGTSESATPES**). Скобки в последовательности использованы только с целью идентификации линкера XTEN16.

MNHDQEFDPKVPVPAEKRRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEV
CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDA
KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
QGKDQHF PVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVI
RHLFAPLKEYFACVSSGNSNANSRGPSSFSSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLLTEDDQETTTRFLQTEAV
TLODVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLOAQVRSRSLDAPKVDLLVK
NCLLPLREYFKYFSQNSLPLggpssgappsggspagsptsteegtsesatpesgpgtstepsegsapgsapgsptste
gstepsegsapgtstepseMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKF KVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKH
ERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLA LAHMIKFRGHFLIEGDLNPD
NSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNL
IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFDQSKNGYAGYIDGGAS
QEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYF
FLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDK GASAQSFIERM
TNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK
VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT
LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNE

KLYLYLQNGRDMYVDQELDINRLSDYVDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS
 EEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL
 DSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF YKREINNYHHAHDAYLNAVVGITALI
 KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
 ETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLSVKELLGITIMERSSSFENPIDFLEAKGYK
 EVKCDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPED
 NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIHLFTLT
 NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDASLGSGSpkkkrkv
EDpkkkrkvDGIGSGSNGSSGS(SGSETPGTSESATPES)GGGGGMDAKSLTAWSRTLVTF
KDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGE
P

[0432] SEQ ID NO:109=V2.3 (DNMT3A-DNMT3L-XTEN80-NLS-dCas9-NLS-XTEN16-KRAB-P2A-BFP)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80
 аминокислот), dCas9 (курсив), NLS SV40 (строчные буквы, курсив), DNMT3A (жирный
 шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27
 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR,
 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et
 al., NAR, 2016), линкер XTEN16 указан как (**SGSETPGTSESATPES**). Скобки в
 последовательности использованы только с целью идентификации линкера XTEN16.

MNHDAQEFDPKVPVPAEKRRKPIRVL SLDGIATGLLVKDLGIQVDRIASEV
CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVNPARKGLY
EGTGRLEFFEFYRLLHDARPKEGDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA
KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
QGKDQHF PVFMNEKEDILWCTEMERVF GPFVHYTDVSNMSRLARQRLGRSWSVPVI
RHLFAPLKEYFACVSSGNSNANSRGPSSGLVPLSLRGSHMGPMIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPF DLVYGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLLTEDDQETTTRFLQTEAV
TLODVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVK
NCLLPLREYFKYFSONSLPLggpssgapppsggspagsptsteegtsesatpesgpgtstepsegsapgsptste
gtstepsegsapgtstepsepkkkrkvMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFVVLGNTDRHSIKK
 NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYALAHMIKFRGHFLIEGD
 LNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFLAEDAQLSKDQTYDDLDNLLAQIGDQYADFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
 GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSF

IERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED
 IVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILD
 FLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVV
 DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
 QNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLRSDKNRGKSD
 NVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH
 VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAV
 GTALIKKYPKLESEFVYGDYKVDVVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR
 KRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIAR
 KKDWDPKKYGGFDSPTVAVSVLVAKVEKKGSKKLSVKELGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSP
 EDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIHLF
 TLTNL GAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDpkkkrkv(SGSE
 TPGTSESATPES)RTLVTFKDVFVDFTREEWKL LDTAQQIVYRNVMLENYKNLVSL
 GYQLTKPDVILRLEKGEEP

[0433] SEQ ID NO:110=V2.4 (DNMT3A-DNMT3L-XTEN80-NLS-dCas9-NLS-
 XTEN16-BFP-KRAB)

KRAB (жирный шрифт, из Gilbert et al., Cell, 2013, 2014), линкеры (подчеркивание),
 XTEN80 (строчные буквы, жирный шрифт, курсив, последовательность из 80
 аминокислот), dCas9 (курсив), NLS SV40 (строчные буквы, курсив), DNMT3A (жирный
 шрифт, курсив, из Siddique et al., JMB, 2013; Stepper et al., NAR, 2016), линкер из 27
 аминокислот (курсив, подчеркивание, из Siddique et al., JMB, 2013; Stepper et al., NAR,
 2016), DNMT3L (жирный шрифт, подчеркивание, из Siddique et al., JMB, 2013; Stepper et
 al., NAR, 2016), линкер XTEN16 указан как (SGSETPGTSESATPES). Скобки в
 последовательности использованы только с целью идентификации последовательности
 линкера XTEN16.

MNHDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEV
 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
 EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDA
 KEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIK
 QGKDQHF PVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVI
 RHLFAPLKEYFACVSSGNSNANSRGPSESSGLVPLSLRGSHMGPMEIYKTVSAWKROPVRV
LSLFRNIDKVLKSLGFL ESGSGGGT LKYVEDVTNVVRRDVEKWGPFDLVYGSTOPLG
SSCDRCPGWYMFQFHRILOYALPROESORPFFWIFMDNLLL TEDDQETTTRFLQTEAV
TLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVK
NCLLPLREYFKYFSQNSLPLggpssgappsggspagsptsteegtsesatpesgpgtstepsegsapgsptste
gstepsegsapgtstepsepkkkkrkvMDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKK
 NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEE
 DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD

LNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID
 GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSF
 IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED
 IVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILD
 FLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVV
 DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
 QNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGELSEDKAGFIKRQLVETRQITKH
 VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAV
 GTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR
 KRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIAR
 KKDWDPKKYGGFDSPTVAYSVLVAKVEKKGSKKLKSVKELLGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSP
 EDNEQKQLFVEQHKKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLF
 TLTNLGAFAAFKYFDTTIDRKRYTSTKEVLDATLIHQISITGLYETRIDLSQLGGDpkkkrkv(SGSE
 TPGTSESA TPES)GGGGGMDAKSLTAWSR TLVTFKDV FVDF TREEWKLLDTAQQIV
 YRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP*

[0434] SEQ ID NO:111

5'-TAATACGACTCACTATAGG-3'

[0435] SEQ ID NO:112

CGAGCTGCGAGCAAAGAT

[0436] SEQ ID NO:113

CGTGTCTCCACAGGTCACAG

[0437] SEQ ID NO:114

GGGCTCTCCAGAACATCATC

[0438] SEQ ID NO:115

CCTGCTTACCACCTTCTTG

[0439] SEQ ID NO:116

AGATCGGAAGAGCACACGTCTGAACTC

[0440] SEQ ID NO:117

AGATCGGAAGAGCACACGTCTGAACTCCAGTCA

[0441] SEQ ID NO:118

AGATCGGAAGAGCGTCTGTAGGGAAAGAGTGTCCACCTTGTTG

[0442] SEQ ID NO:119

gtttcagagcgagacgtgctgcaggatacgtctcagaaacatg

[0443] SEQ ID NO:120

GTTTAAGAGCTAAGCTG

[0444] SEQ ID NO:121

AATGATACGGCGACCACCGAGATCTACACCGCGGTCTGTATCCCTTGGAGAA
CCACCT

[0445] SEQ ID NO:122

CAAGCAGAAGACGGCATAACGAGATXXXXXGCGGCCGGCTGTTTCCAGCTTA
GCTCTTAAA

[0446] SEQ ID NO:123

CGCGGTCTGTATCCCTTGGAGAACCACCTTGTTGG

[0447] SEQ ID NO:124

GCGGCCGGCTGTTTCCAGCTTAGCTCTTAAAC

[0448] SEQ ID NO:125

GTTTAAGAGCTAAGCTGGAAACAGCCGGCCGC

[0449] SEQ ID NO:126

CCACCTTGTTG

[0450] Заголовки разделов, используемые в данном документе, предназначены только для использования в организационных целях и не должны рассматриваться как ограничивающие описываемый предмет изобретения. Все документы или части документов, упоминаемых в заявке, включая, без ограничений, патенты, заявки на патенты, статьи, книги, руководства и трактаты, настоящим явным образом включены посредством ссылок во всей их полноте для любых целей.

[0451] Хотя в данном документе показаны и описаны различные варианты осуществления и аспекты настоящего изобретения, специалистам в данной области техники будет очевидно, что такие варианты осуществления и аспекты представлены только в качестве примера. Многочисленные варианты, изменения и замены будут доступны специалистам в данной области без выхода за рамки изобретения. Следует понимать, что при осуществлении изобретения могут быть использованы различные альтернативы вариантам осуществления изобретения, описанным в данном документе.

[0452] Список литературы

Adamson, B., Norman, T.M., Jost, M., Cho, M.Y., Nuñez, J.K., Chen, Y., Villalta, J.E., Gilbert, L.A., Horlbeck, M.A., Hein, M.Y., et al. (2016). A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* 167, 1867-1882.e21. Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., and Lombardo, A. (2016). Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* 167, 219-232.e14. Anzalone, A.V., Koblan, L.W., and Liu, D.R. (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824-844. Audergon, P.N.C.B., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A.L., and Allshire, R.C. (2015). Restricted epigenetic inheritance of H3K9 methylation. *Science* 348, 132-135. Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell

level. *Science* 351, 720-724. Blomen, V.A., Májek, P., Jae, L.T., Bigenzahn, J.W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F.R., Olk, N., Stukalov, A., et al. (2015). Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092-1096. Boyes, J., and Bird, A. (1992). Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.* 11, 327-333. Buiting, K., Williams, C., and Horsthemke, B. (2016). Angelman syndrome - insights into a rare neurogenetic disorder. *Nat. Rev. Neurol.* 12, 584-593. Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly-efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326-328. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinforma. Oxf. Engl.* 34, i884-i890. Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., et al. (2018). Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. *Cell* 173, 1398-1412.e22. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010-1022. Doench, J.G. (2018). Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* 19, 67-80. Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 354, 769-773. Galonska, C., Charlton, J., Mattei, A.L., Donaghey, J., Clement, K., Gu, H., Mohammad, A.W., Stamenova, E.K., Cacchiarelli, D., Klages, S., et al. (2018). Genome-wide tracking of dCas9-methyltransferase footprints. *Nat. Commun.* 9. Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442-451. Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647-661. Guenther, M.G., Frampton, G.M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., and Young, R.A. (2010). Chromatin Structure and Gene Expression Programs of Human Embryonic and Induced Pluripotent Stem Cells. *Cell Stem Cell* 7, 249-257. Hanna, R.E., and Doench, J.G. (2020). Design and analysis of CRISPR-Cas experiments. *Nat. Biotechnol.* 38, 813-823. Hansen, K.D., Langmead, B., and Irizarry, R.A. (2012). BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 13, R83. Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R., and Moffat, J. (2014). Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Mol. Syst. Biol.* 10, 733. Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. *Cell* 149, 1447-1460. Hofacker, D., Broche, J., Laistner, L., Adam, S., Bashtrykov, P., and Jeltsch, A. (2020). Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects. *Int. J. Mol. Sci.* 21. Holtzman, L., and Gersbach, C.A. (2018). Editing the Epigenome: Reshaping the Genomic Landscape. *Annu. Rev. Genomics Hum. Genet.* 19, 43-71. Horlbeck,

M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., et al. (2016a). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *ELife* 5, e19760. Horlbeck, M.A., Witkowsky, L.B., Guglielmi, B., Replogle, J.M., Gilbert, L.A., Villalta, J.E., Torigoe, S.E., Tjian, R., and Weissman, J.S. (2016b). Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *ELife* 5. Hovestadt, V., Jones, D.T.W., Picelli, S., Wang, W., Kool, M., Northcott, P.A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H.-J., et al. (2014). Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* 510, 537-541. Iglesias, N., Currie, M.A., Jih, G., Paulo, J.A., Siuti, N., Kalocsay, M., Gygi, S.P., and Moazed, D. (2018). Automethylation-induced conformational switch in Ctr4 (Suv39h) maintains epigenetic stability. *Nature* 560, 504-508. Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939-946. Iqbal, K., Liu, F., and Gong, C.-X. (2016). Tau and neurodegenerative disease: the story so far. *Nat. Rev. Neurol.* 12, 15-27. Isaac, R.S., Jiang, F., Doudna, J.A., Lim, W.A., Narlikar, G.J., and Almeida, R. (2016). Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *ELife* 5, e13450. Jost, M., Santos, D.A., Saunders, R.A., Horlbeck, M.A., Hawkins, J.S., Scaria, S.M., Norman, T.M., Hussmann, J.A., Liem, C.R., Gross, C.A., et al. (2020). Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nat. Biotechnol.* 38, 355-364. Knott, G.J., and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science* 361, 866-869. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583-588. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357-359. Leonetti, M.D., Sekine, S., Kamiyama, D., Weissman, J.S., and Huang, B. (2016a). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci. U. S. A.* 113, E3501-3508. Leonetti, M.D., Sekine, S., Kamiyama, D., Weissman, J.S., and Huang, B. (2016b). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci.* 113, E3501-E3508. Li, X.-L., Li, G.-H., Fu, J., Fu, Y.-W., Zhang, L., Chen, W., Arakaki, C., Zhang, J.-P., Wen, W., Zhao, M., et al. (2018). Highly efficient genome editing via CRISPR-Cas9 in human pluripotent stem cells is achieved by transient BCL-XL overexpression. *Nucleic Acids Res.* 46, 10195-10215. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930. Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., and Jaenisch, R. (2016). Editing DNA Methylation in the Mammalian Genome. *Cell* 167, 233-247.e17. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., et al. (2013). Targeted DNA demethylation and activation of endogenous genes using programmable

TALE-TET1 fusion proteins. *Nat. Biotechnol.* 31, 1137-1142. Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., Pantel, S., et al. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* 49, 1779-1784.

Mlambo, T., Nitsch, S., Hildenbeutel, M., Romito, M., Müller, M., Bossen, C., Diederichs, S., Cornu, T.I., Cathomen, T., and Mussolino, C. (2018). Designer epigenome modifiers enable robust and sustained gene silencing in clinically relevant human cells. *Nucleic Acids Res.* 46, 4456-4468.

O'Geen, H., Bates, S.L., Carter, S.S., Nisson, K.A., Halmai, J., Fink, K.D., Rhie, S.K., Farnham, P.J., and Segal, D.J. (2019). Ezh2-dCas9 and KRAB-dCas9 enable engineering of epigenetic memory in a context-dependent manner. *Epigenetics Chromatin* 12, 26.

Park, Y., and Wu, H. (2016). Differential methylation analysis for BS-seq data under general experimental design. *Bioinforma. Oxf. Engl.* 32, 1446-1453.

Park, M., Patel, N., Keung, A.J., and Khalil, A.S. (2019). Engineering Epigenetic Regulation Using Synthetic Read-Write Modules. *Cell* 176, 227-238.e20.

Ragunathan, K., Jih, G., and Moazed, D. (2015). Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* 348, 1258699.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160-165.

Replogle, J.M., Norman, T.M., Xu, A., Hussmann, J.A., Chen, J., Cogan, J.Z., Meer, E.J., Terry, J.M., Riordan, D.P., Srinivas, N., et al. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* 38, 954-961.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24-26.

Schellenberger, V., Wang, C.-W., Geething, N.C., Spink, B.J., Campbell, A., To, W., Scholle, M.D., Yin, Y., Yao, Y., Bogin, O., et al. (2009). A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. *Nat. Biotechnol.* 27, 1186-1190.

Shalem, O., Sanjana, N.E., and Zhang, F. (2015). High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* 16, 299-311.

Stelzer, Y., Shivalila, C.S., Soldner, F., Markoulaki, S., and Jaenisch, R. (2015). Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution. *Cell* 163, 218-229.

Stepper, P., Kungulovski, G., Jurkowska, R.Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A., and Jurkowski, T.P. (2017). Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res.* 45, 1703-1713.

Tak, Y.E., Kleinstiver, B.P., Nuñez, J.K., Hsu, J.Y., Horng, J.E., Gong, J., Weissman, J.S., and Joung, J.K. (2017). Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods* 14, 1163-1166.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. *Bioinforma. Oxf. Engl.* 31, 2032-2034.

Tarjan, D.R., Flavahan, W.A., and Bernstein, B.E. (2019). Epigenome editing strategies for the functional annotation of CTCF insulators. *Nat. Commun.* 10, 4258.

Tian, R., Gachechiladze, M.A., Ludwig, C.H., Laurie, M.T., Hong, J.Y., Nathaniel, D., Prabhu, A.V., Fernandopulle, M.S., Patel, R., Abshari, M., et al. (2019). CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-

Derived Neurons. *Neuron* 104, 239-255.e12. Van, M.V., Fujimori, T., and Bintu, L. (2021). Nanobody-mediated control of gene expression and epigenetic memory. *Nat. Commun.* 12, 537.

Xu, X., and Qi, L.S. (2019). A CRISPR-dCas Toolbox for Genetic Engineering and Synthetic Biology. *J. Mol. Biol.* 431, 34-47.

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G., and Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. *Cell Discov.* 2, 1-12.

Yeh, C.D., Richardson, C.D., and Corn, J.E. (2019). Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* 21, 1468-1478.

Yu, R., Wang, X., and Moazed, D. (2018). Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. *Nature* 558, 615-619.

Zhang, Z.-M., Lu, R., Wang, P., Yu, Y., Chen, D., Gao, L., Liu, S., Ji, D., Rothbart, S.B., Wang, Y., et al. (2018). Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* 554, 387-391.

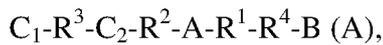
ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Слитый белок, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, второй линкер XTEN и домен K \ddot{u} rrel-ассоциированного бокса.
2. Слитый белок по п. 1, отличающийся тем, что первый линкер XTEN содержит от около 5 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до около 864 аминокислотных остатков.
3. Слитый белок по п. 2, отличающийся тем, что первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков, и домен K \ddot{u} rrel-ассоциированного бокса.
4. Слитый белок по п. 3, отличающийся тем, что первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 40 аминокислотных остатков.
5. Слитый белок по п. 4, отличающийся тем, что первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков.
6. Слитый белок по п. 1, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3.
7. Слитый белок по п. 6, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.
8. Слитый белок по п. 6, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.
9. Слитый белок по п. 1, отличающийся тем, что домен ДНК-метилтрансферазы содержит домен Dnmt3A.
10. Слитый белок по п. 9, отличающийся тем, что домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).
11. Слитый белок по п. 1, дополнительно содержащий эпитопную метку, 2A-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.
12. Слитый белок по п. 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, сигнальный пептид ядерной локализации и домен K \ddot{u} rrel-ассоциированного бокса.
13. Слитый белок по п. 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, сигнальный пептид ядерной локализации, второй линкер

ХТЕН и домен Kgrüppel-ассоциированного бокса.

14. Слитый белок по п. 11, содержащий, от N-конца к С-концу, домен ДНК-метилтрансферазы, первый линкер ХТЕН, дефицитный по нуклеазе фермент РНК-направляемой эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер ХТЕН, домен Kgrüppel-ассоциированного бокса, расщепляемый 2А-пептид и флуоресцентную белковую метку.

15. Слитый белок, имеющий по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью формулы (А); где аминокислотная последовательность формулы (А) представляет собой, от N-конца к С-концу:



где:

C_1 содержит SEQ ID NO:26 или SEQ ID NO:106;

R^3 отсутствует или R^3 содержит SEQ ID NO:27;

C_2 содержит SEQ ID NO:28;

R^2 отсутствует или R^2 содержит SEQ ID NO:32;

А содержит SEQ ID NO:23;

R^1 отсутствует или R^1 содержит SEQ ID NO:25;

R^4 отсутствует или R^4 содержит SEQ ID NO:31; и

В содержит SEQ ID NO:16, SEQ ID NO:103, SEQ ID NO:104 или SEQ ID NO:105.

16. Слитый белок по п. 15, имеющий по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью формулы (А).

17. Слитый белок по п. 16, имеющий по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью формулы (А).

18. Слитый белок, имеющий по меньшей мере 85% идентичности последовательности с SEQ ID NO:97, 98, 99, 107, 108, 109 или 110.

19. Слитый белок по п. 18, имеющий по меньшей мере 90% идентичности последовательности с SEQ ID NO:97, 98, 99, 107, 108, 109 или 110.

20. Клетка, содержащая слитый белок по п. 1.

21. Клетка по п. 20, представляющая собой эукариотическую клетку, клетку млекопитающего или стволовую клетку.

22. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий:

(i) доставку первого полинуклеотида, кодирующего слитый белок по п. 1 в клетку, содержащую нуклеиновую кислоту-мишень; и

(ii) доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени.

23. Способ по п. 22, отличающийся тем, что нуклеиновая кислота-мишень содержит CpG-островок.

24. Способ по п. 22, отличающийся тем, что нуклеиновая кислота-мишень не содержит CpG-островка.

25. Способ по п. 22, отличающийся тем, что второй полинуклеотид содержит sgРНК.

26. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий:

(i) доставку первого полинуклеотида, кодирующего слитый белок, в клетку, содержащую нуклеиновую кислоту-мишень, причем нуклеиновая кислота-мишень не содержит CpG-островка; при этом слитый белок содержит дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы, домен Krüppel-ассоциированного бокса и домен ДНК-метилтрансферазы; и

(ii) доставку в клетку второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени в клетке.

27. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий:

(i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок, содержащий дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы, домен Krüppel-ассоциированного бокса и домен ДНК-метилтрансферазы; и

(ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания.

28. Способ по п. 26, отличающийся тем, что второй полинуклеотид содержит sgРНК.

29. Способ по п. 26, отличающийся тем, что слитый белок содержит, от N-конца к C-концу, домен ДНК-метилтрансферазы, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен Krüppel-ассоциированного бокса.

30. Способ по п. 26, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3.

31. Способ по п. 30, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.

32. Способ по п. 30, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.

33. Способ по п. 26, отличающийся тем, что домен ДНК-метилтрансферазы

содержит домен Dnmt3A.

34. Способ по п. 33, отличающийся тем, что домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

35. Способ по п. 26, отличающийся тем, что dCas9 ковалентно связан с доменом Dnmt3A с помощью пептидного линкера, и домен Dnmt3A ковалентно связан с доменом Krüppel-ассоциированного бокса с помощью пептидного линкера.

36. Способ по п. 35, отличающийся тем, что пептидный линкер представляет собой линкер XTEN.

37. Способ по п. 26, отличающийся тем, что слитый белок содержит, от N-конца к C-концу, Krüppel-ассоциированный бокс, дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы и домен ДНК-метилтрансферазы.

38. Способ по п. 37, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой CRISPR-ассоциированный белок, домен лейциновой молнии, домен крылатой спирали, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокса, домен Wor3, домен OB-складки, домен иммуноглобулина или домен В3.

39. Способ по п. 38, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCas9.

40. Способ по п. 38, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы представляет собой dCpf1 или ddCpf1.

41. Способ по п. 37, отличающийся тем, что домен ДНК-метилтрансферазы содержит домен Dnmt3A.

42. Способ по п. 41, отличающийся тем, что домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

43. Способ по п. 39, отличающийся тем, что dCas9 ковалентно связан с доменом Dnmt3A с помощью пептидного линкера, и домен Krüppel-ассоциированного бокса ковалентно связан с dCas9 с помощью пептидного линкера.

44. Способ по п. 43, отличающийся тем, что пептидный линкер представляет собой линкер XTEN.

45. Способ по п. 37, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы ковалентно связан с доменом Krüppel-ассоциированного бокса с помощью пептидного линкера.

46. Способ по п. 37, отличающийся тем, что дефицитный по нуклеазе фермент РНК-направляемой ДНК-эндонуклеазы ковалентно связан с доменом ДНК-метилтрансферазы с помощью пептидного линкера.

47. Способ по п. 37, отличающийся тем, что домен Krüppel-ассоциированного бокса ковалентно связан с доменом ДНК-метилтрансферазы с помощью пептидного линкера.

48. Способ по п. 26, отличающийся тем, что слитый белок дополнительно содержит эпитопную метку, 2А-пептид, флуоресцентную белковую метку, сигнальный пептид

ядерной локализации, или комбинацию двух или более из них.

49. Способ по п. 26, отличающийся тем, что слитый белок дополнительно содержит сигнальный пептид ядерной локализации.

50. Способ по п. 26, отличающийся тем, что слитый белок имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью формулы (A); при этом аминокислотная последовательность формулы (A) представляет собой, от N-конца к C-концу:



где:

C_1 содержит SEQ ID NO:26 или SEQ ID NO:106;

R^3 отсутствует или R^3 содержит SEQ ID NO:27;

C_2 содержит SEQ ID NO:28;

R^2 отсутствует или R^2 содержит SEQ ID NO:32;

A содержит SEQ ID NO:23;

R^1 отсутствует или R^1 содержит SEQ ID NO:25;

R^4 отсутствует или R^4 содержит SEQ ID NO:31; и

B содержит SEQ ID NO:16, SEQ ID NO:103, SEQ ID NO:104 или SEQ ID NO:105.

51. Способ по п. 50, отличающийся тем, что слитый белок имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью формулы (A).

52. Способ по п. 51, отличающийся тем, что слитый белок имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью формулы (A).

53. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий:

(i) доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок по п. 1; и

(ii) доставку субъекту эффективного количества второго полинуклеотида, содержащего sgРНК или cr:tracrРНК; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у субъекта.

54. Способ по п. 26, отличающийся тем, что инфекционное заболевание представляет собой вирусное инфекционное заболевание.

55. Способ по п. 54, отличающийся тем, что инфекционное заболевание представляет собой флавивирусное инфекционное заболевание.

56. Слитый белок, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, второй линкер XTEN и домен Kgyrrel-ассоциированного бокса.

57. Слитый белок по п. 56, отличающийся тем, что первый линкер XTEN содержит

от около 5 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до около 864 аминокислотных остатков.

58. Слитый белок по п. 57, отличающийся тем, что первый линкер XTEN содержит от более чем 50 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 5 до 50 аминокислотных остатков.

59. Слитый белок по п. 58, отличающийся тем, что первый линкер XTEN содержит от около 60 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 40 аминокислотных остатков.

60. Слитый белок по п. 59, отличающийся тем, что первый линкер XTEN содержит от около 70 до около 864 аминокислотных остатков, и второй линкер XTEN содержит от около 10 до около 30 аминокислотных остатков.

61. Слитый белок по п. 56, отличающийся тем, что дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца или эффектор, подобный активатору транскрипции.

62. Слитый белок по п. 62, отличающийся тем, что дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой домен цинкового пальца.

63. Слитый белок по п. 62, отличающийся тем, что дефицитный по нуклеазе фермент ДНК-эндонуклеазы представляет собой эффектор, подобный активатору транскрипции.

64. Слитый белок по п. 56, отличающийся тем, что домен ДНК-метилтрансферазы содержит домен Dnmt3A.

65. Слитый белок по п. 64, отличающийся тем, что домен Dnmt3A связан с доменом Dnmt3L (домен Dnmt3A-3L).

66. Слитый белок по п. 56, дополнительно содержащий эпитопную метку, 2A-пептид, флуоресцентную белковую метку, сигнальный пептид ядерной локализации, или комбинацию двух или более из них.

67. Слитый белок по п. 56, дополнительно содержащий сигнальный пептид ядерной локализации.

68. Слитый белок по п. 56, содержащий, от N-конца к C-концу, домен ДНК-метилтрансферазы, первый линкер XTEN, дефицитный по нуклеазе фермент эндонуклеазы, эпитопную метку, сигнальный пептид ядерной локализации, второй линкер XTEN, домен K \ddot{u} rpel-ассоциированного бокса, расщепляемый 2A-пептид и флуоресцентную белковую метку.

69. Клетка, содержащая слитый белок по п. 56.

70. Клетка по п. 69, представляющая собой эукариотическую клетку, клетку млекопитающего или стволовую клетку.

71. Способ сайленсинга последовательности нуклеиновой кислоты-мишени в клетке, включающий доставку первого полинуклеотида, кодирующего слитый белок по п. 56, в клетку, содержащую нуклеиновую кислоту-мишень; тем самым обеспечивая сайленсинг последовательности нуклеиновой кислоты-мишени.

72. Способ по п. 71, отличающийся тем, что нуклеиновая кислота-мишень содержит CpG-островок.

73. Способ по п. 71, отличающийся тем, что нуклеиновая кислота-мишень не содержит CpG-островка.

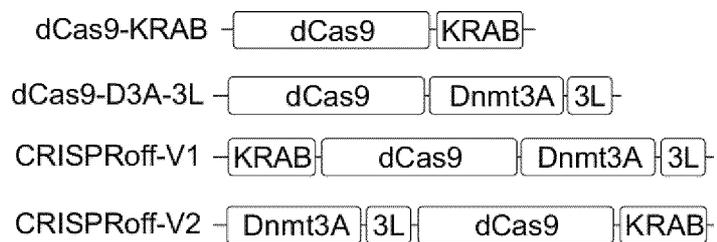
74. Способ лечения синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания у нуждающегося в этом субъекта, включающий доставку субъекту эффективного количества первого полинуклеотида, кодирующего слитый белок по п. 56; тем самым, обеспечивая лечение синдрома Ангельмана, инфекционного заболевания, тау-патологии или нейродегенеративного заболевания.

75. Способ по п. 74, отличающийся тем, что инфекционное заболевание представляет собой вирусное инфекционное заболевание.

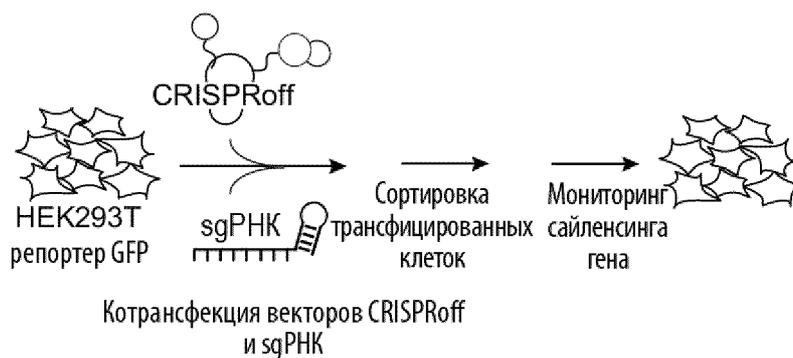
76. Способ по п. 75, отличающийся тем, что вирусное инфекционное заболевание представляет собой флавивирусное инфекционное заболевание.

По доверенности

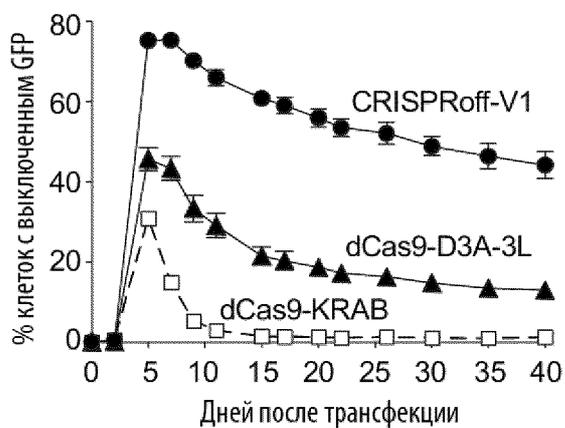
1/32



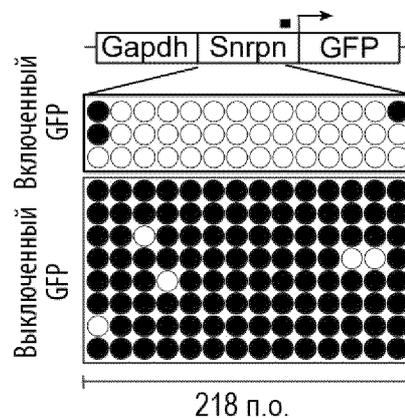
Фиг. 1А



Фиг. 1В

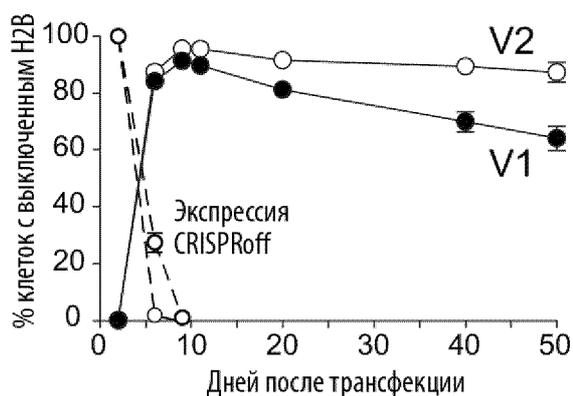


Фиг. 1С

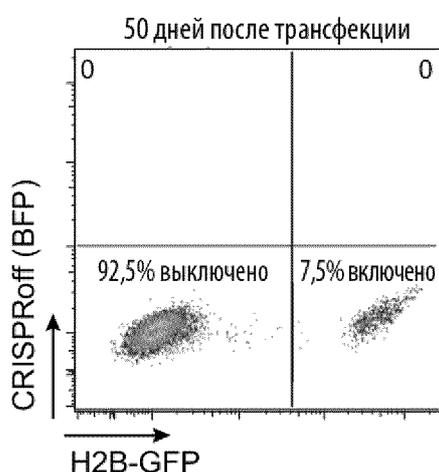


Фиг. 1D

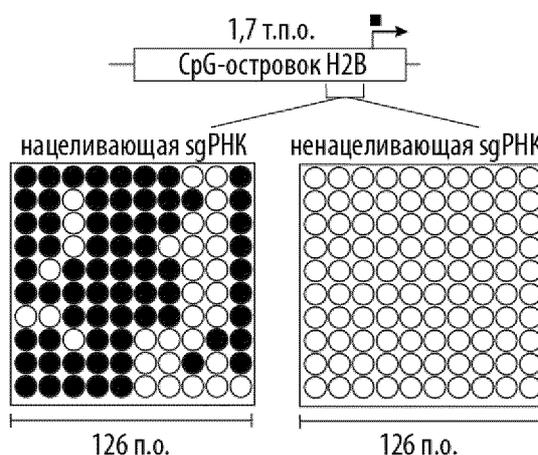
2/32



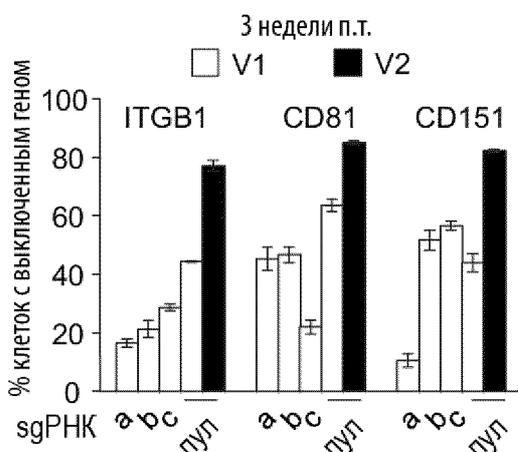
Фиг. 1Е



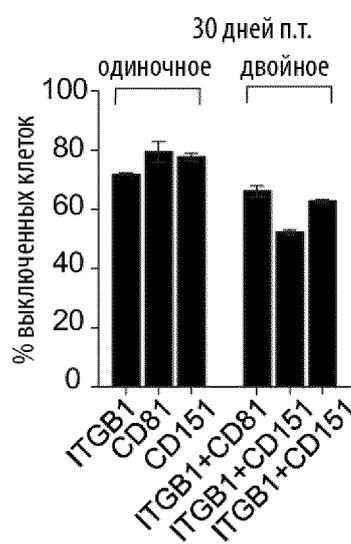
Фиг. 1F



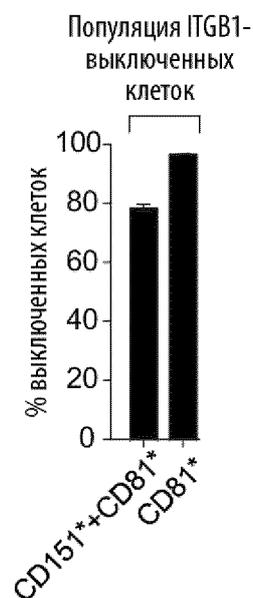
Фиг. 1G



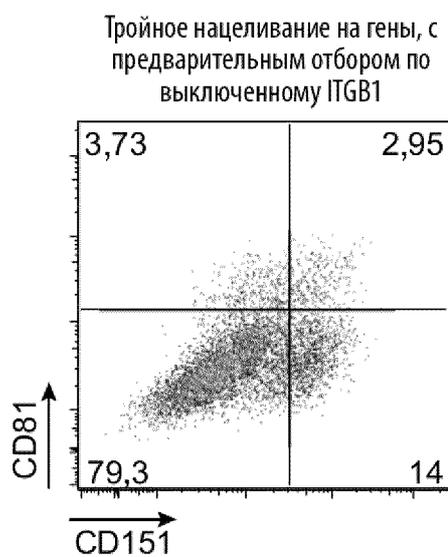
Фиг. 1H



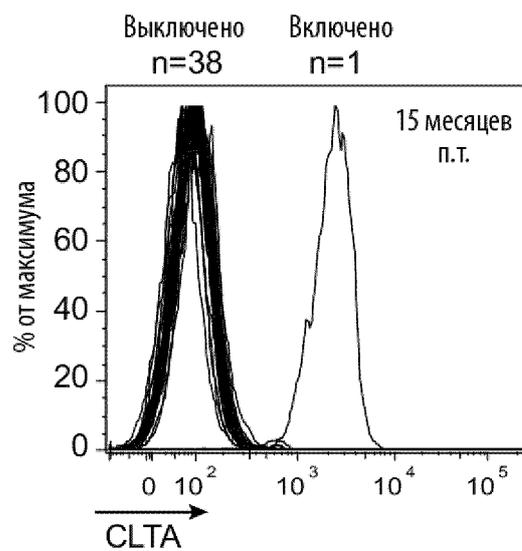
Фиг. 1I



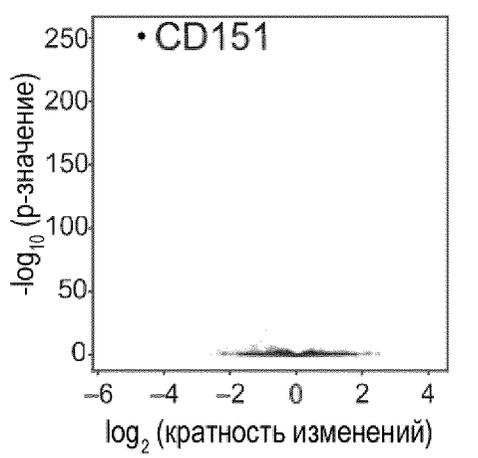
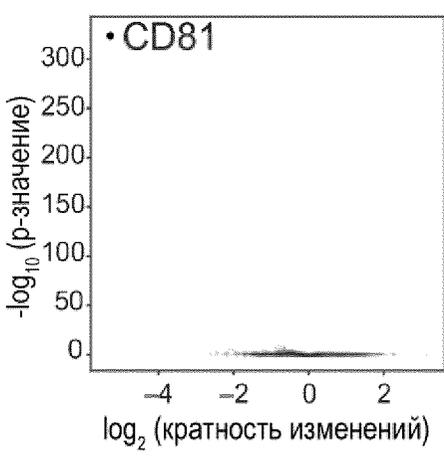
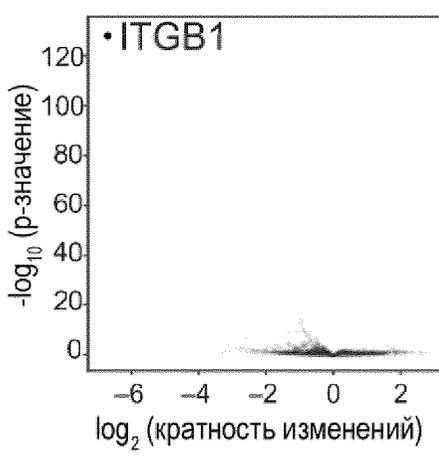
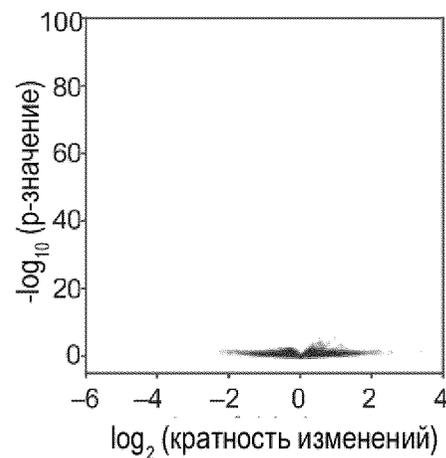
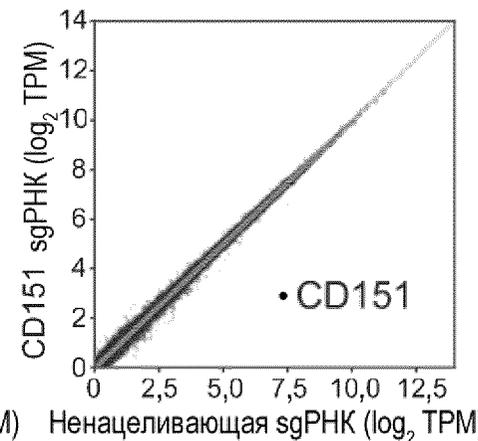
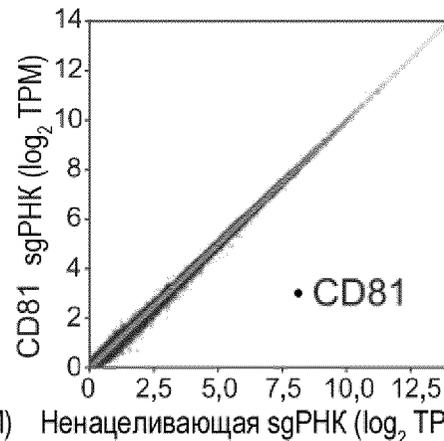
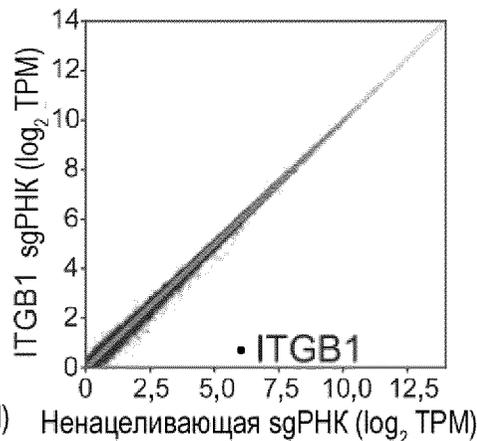
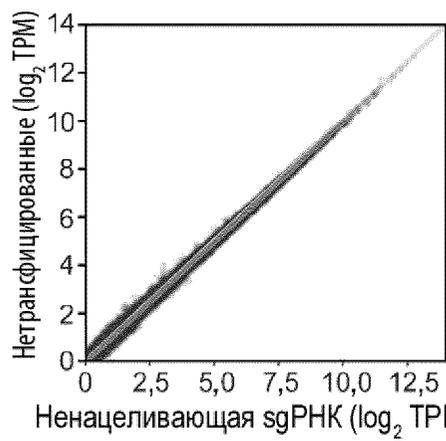
Фиг. 1J



Фиг. 1К



Фиг. 1Л

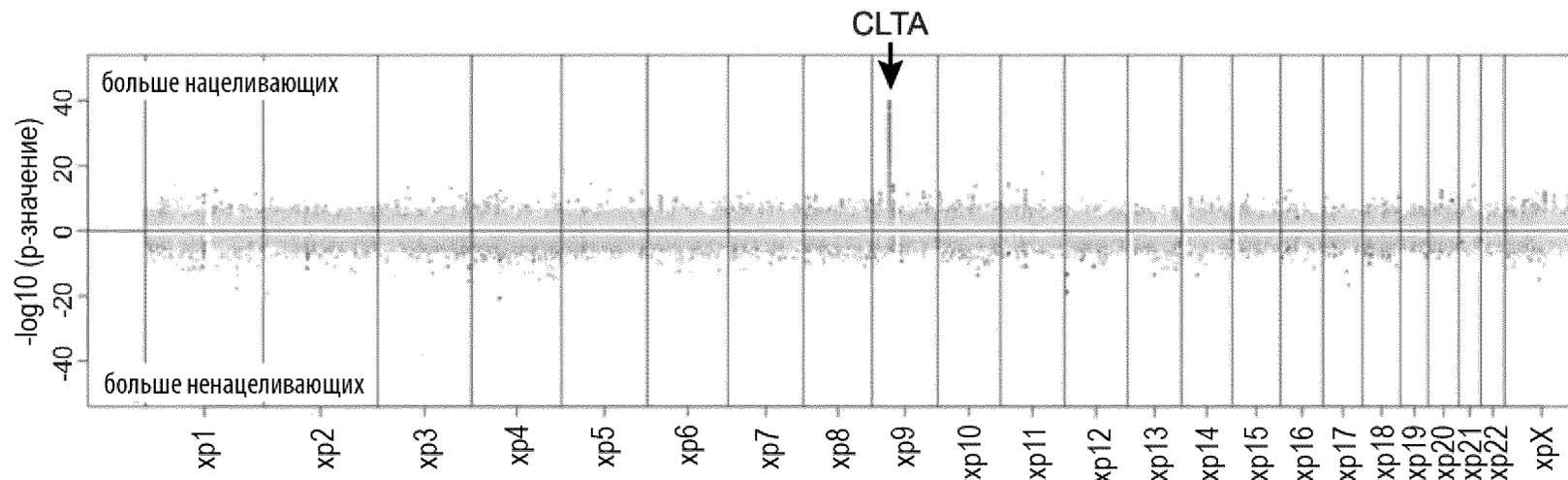


Фиг. 2А

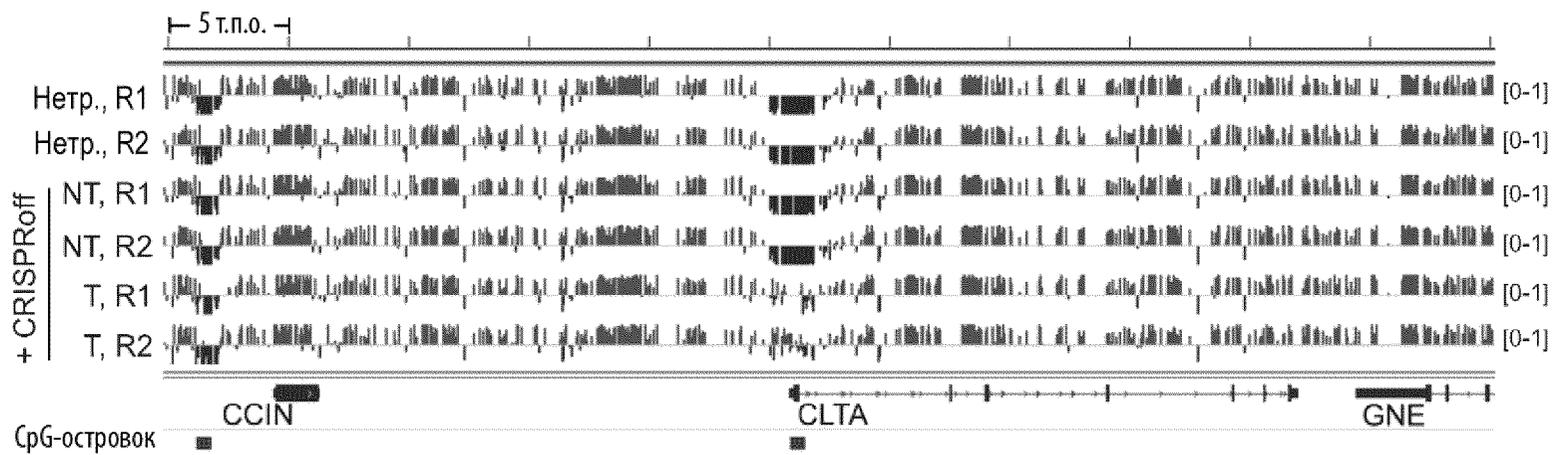
Фиг. 2В

Фиг. 2С

Фиг. 2D

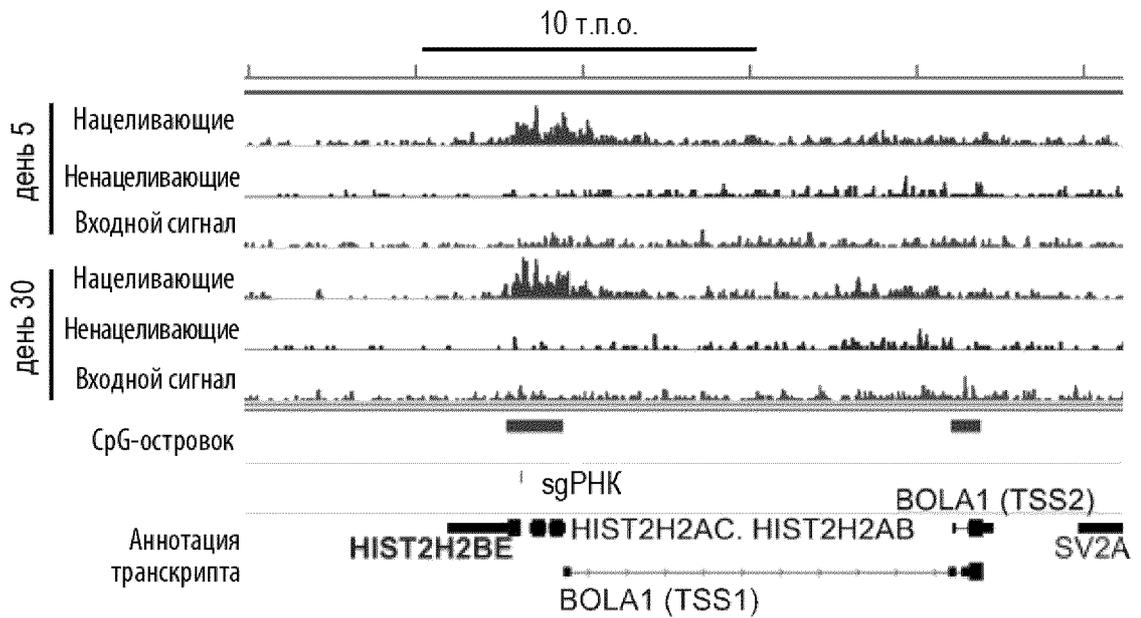


Фиг. 2Е

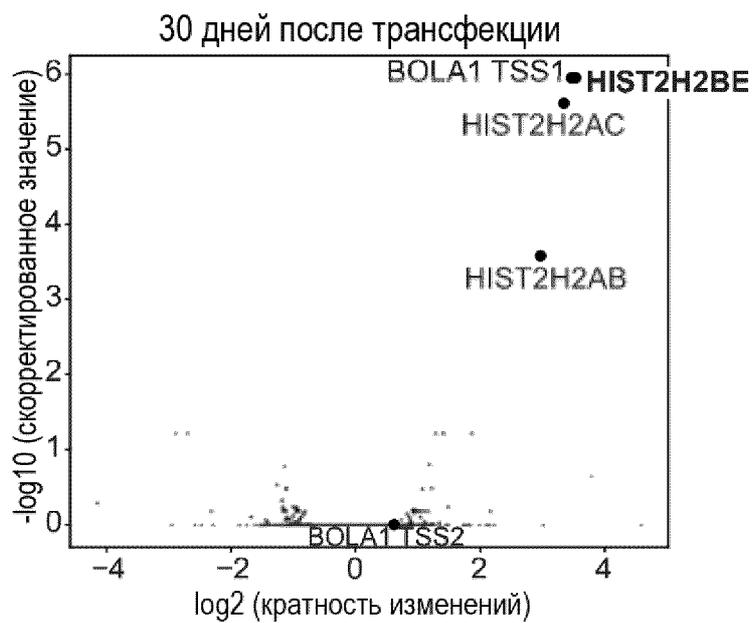


Фиг. 2F

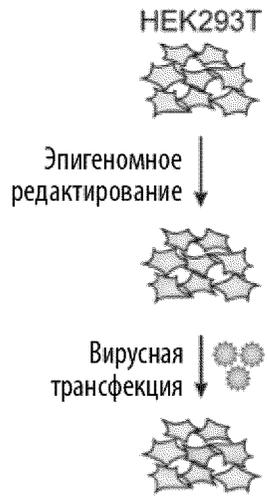
6/32



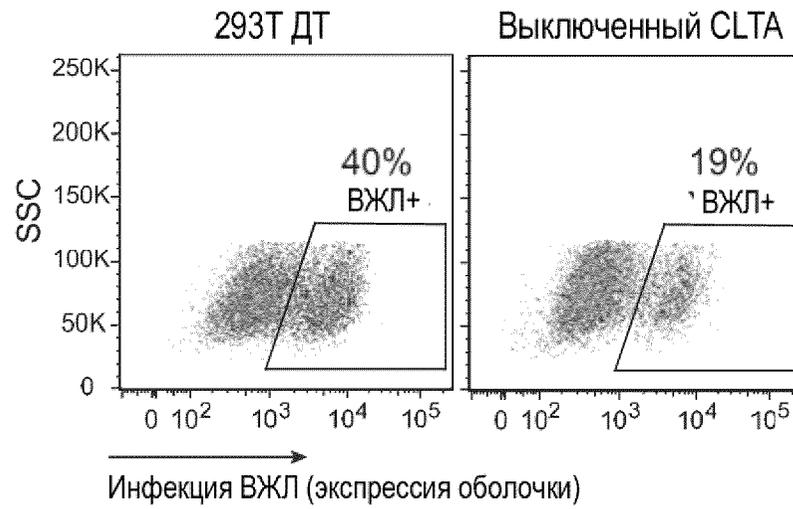
Фиг. 2G



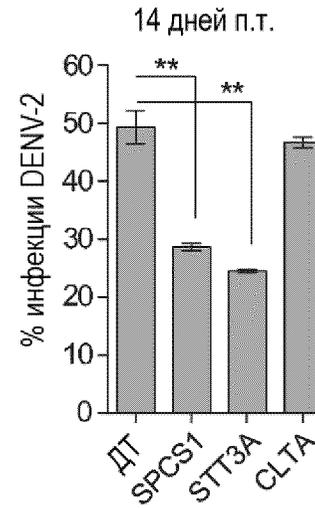
Фиг. 2H



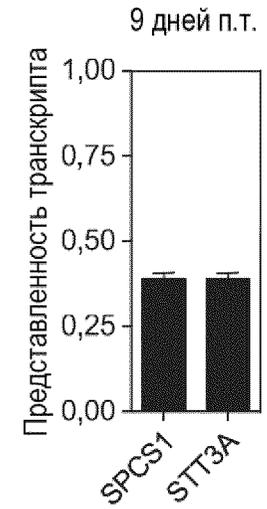
Фиг. 3А



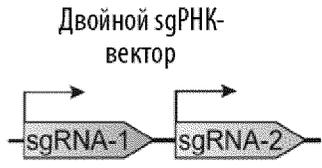
Фиг. 3В



Фиг. 3С



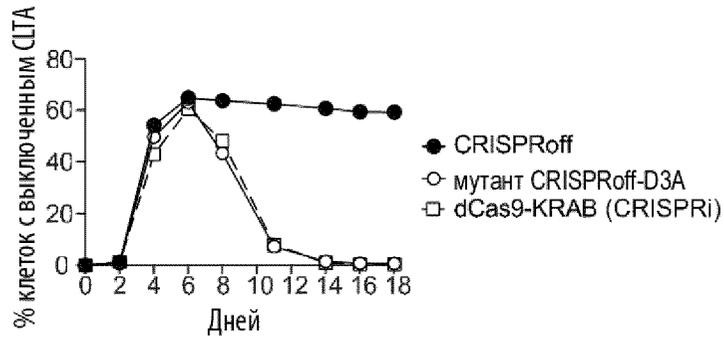
Фиг. 3D



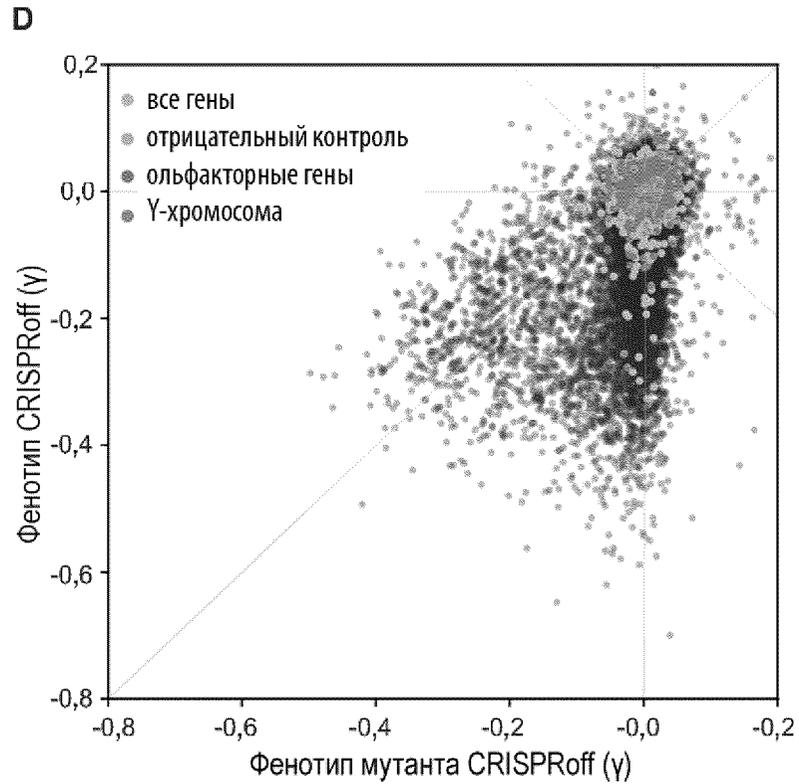
Фиг. 4А



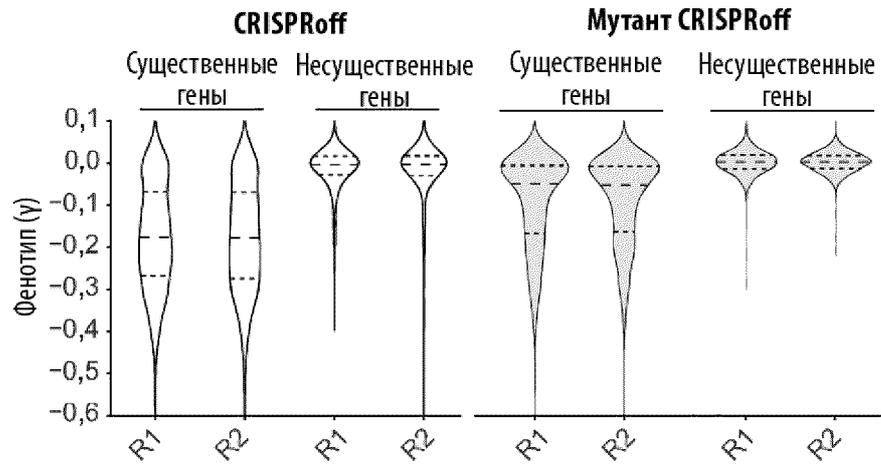
Фиг. 4В



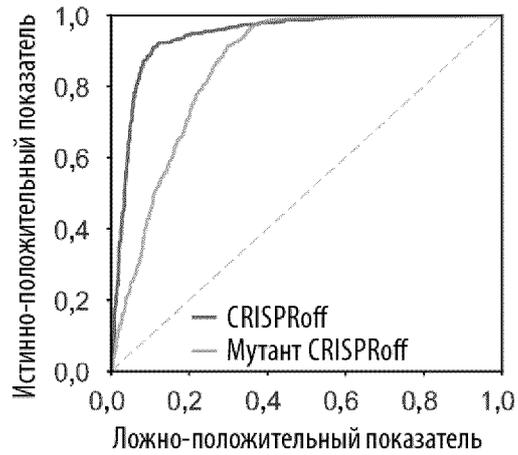
Фиг. 4С



Фиг. 4D



Фиг. 4Е

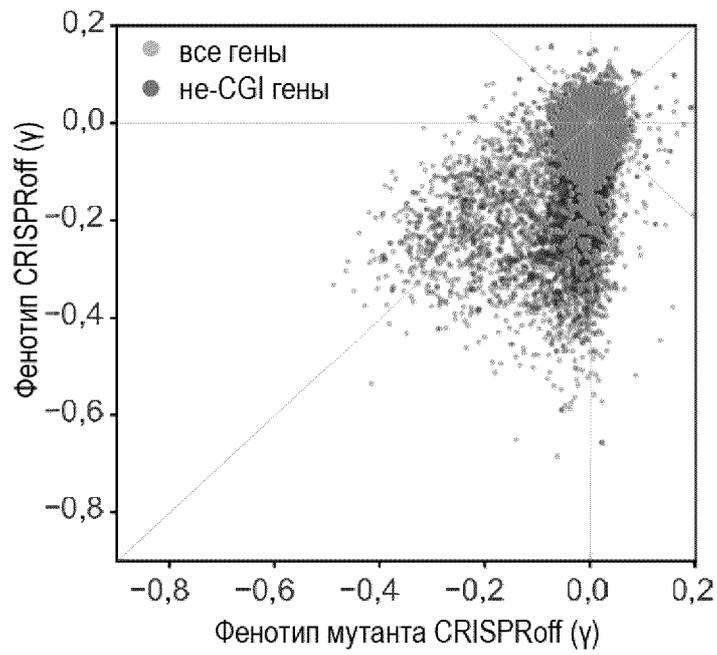


Фиг. 4F

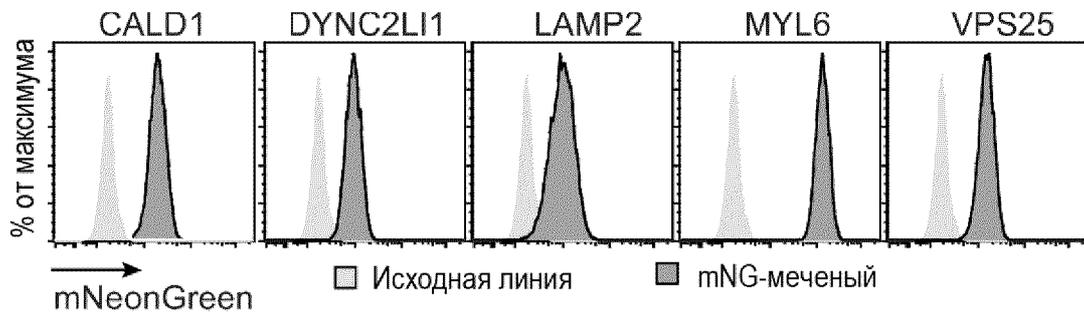


Фиг. 4G

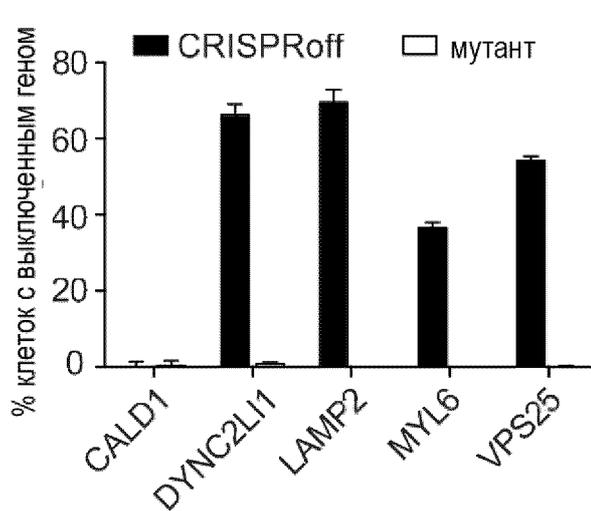
10/32



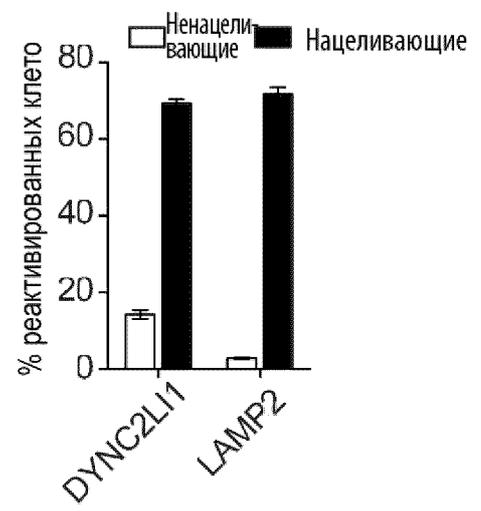
Фиг. 5А



Фиг. 5В



Фиг. 5С



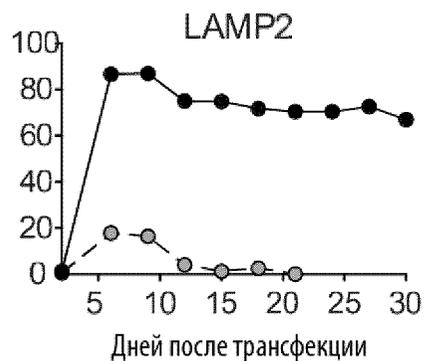
Фиг. 5D



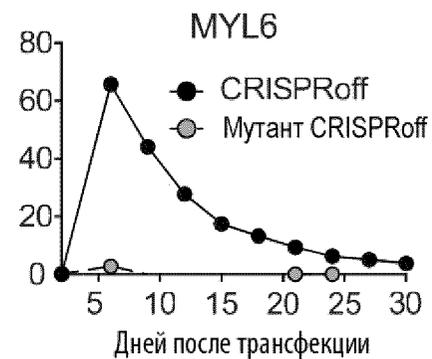
Фиг. 5Е



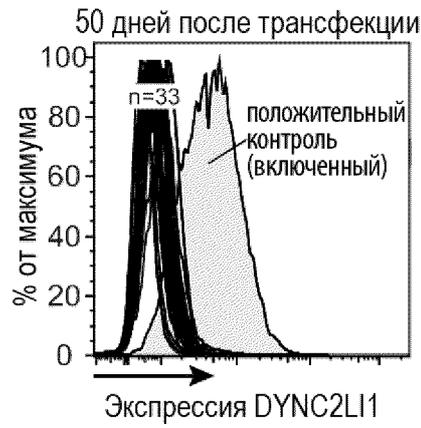
Фиг. 5F



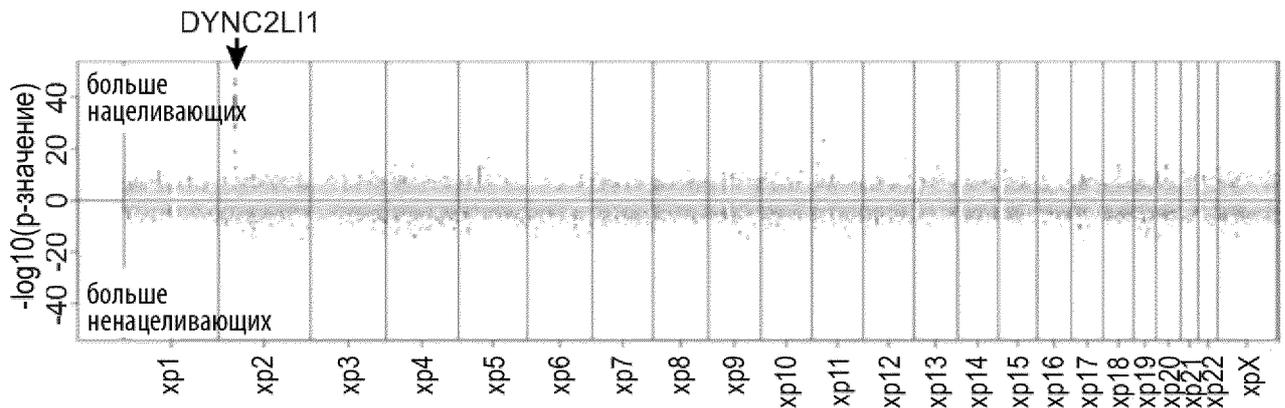
Фиг. 5G



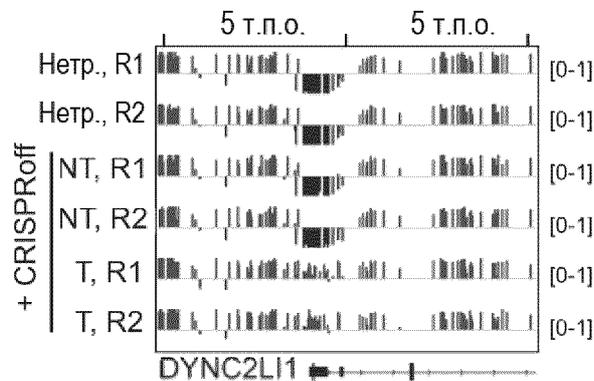
Фиг. 5H



Фиг. 5I



Фиг. 5J



Фиг. 5K

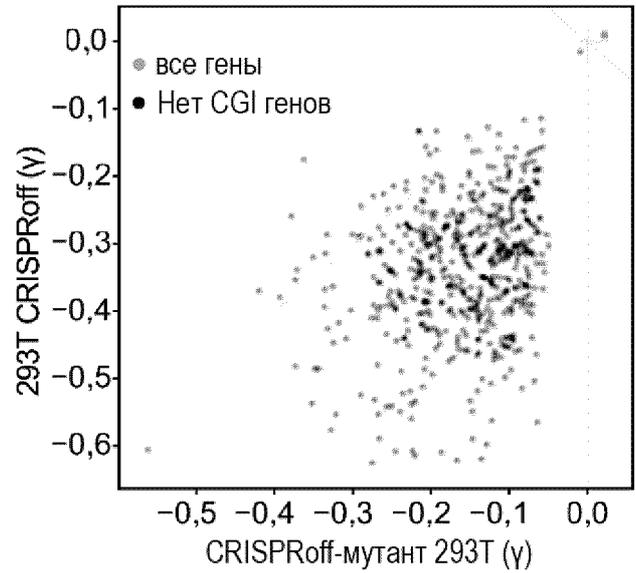
Мозаичный скрининг sgRNA



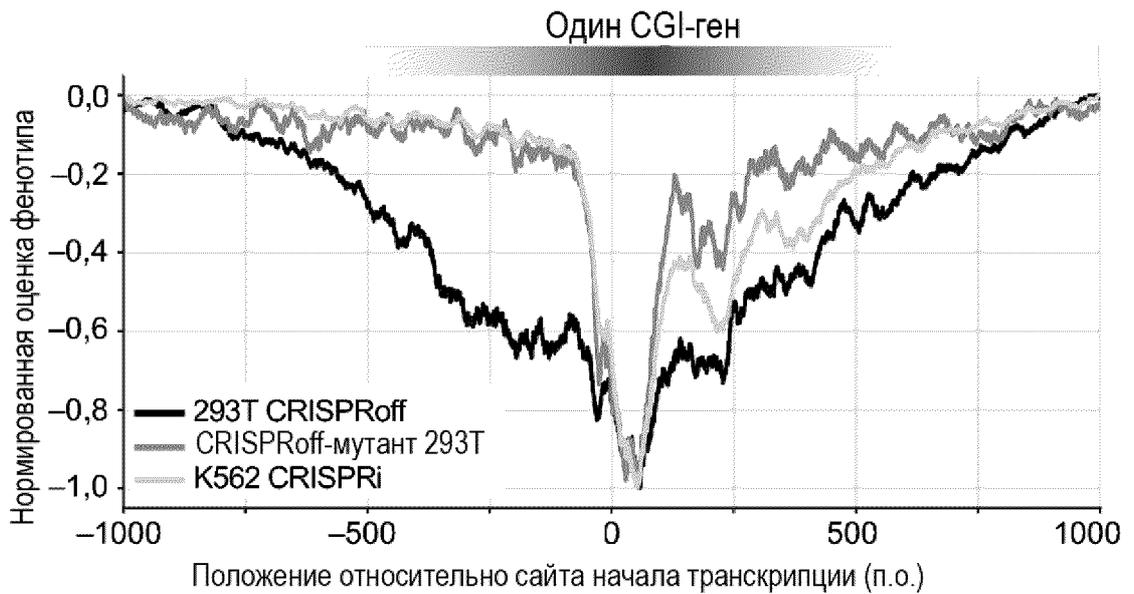
Фиг. 6А

число гидов	116 085
Число НН гидов	3 790
Всего генов	520
Один CGI	425
Несколько CGI	56
Нет CGI	39

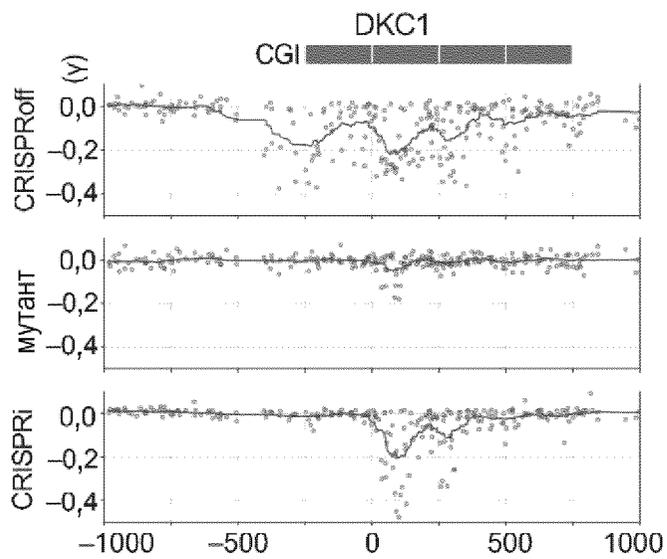
Фиг. 6В



Фиг. 6С

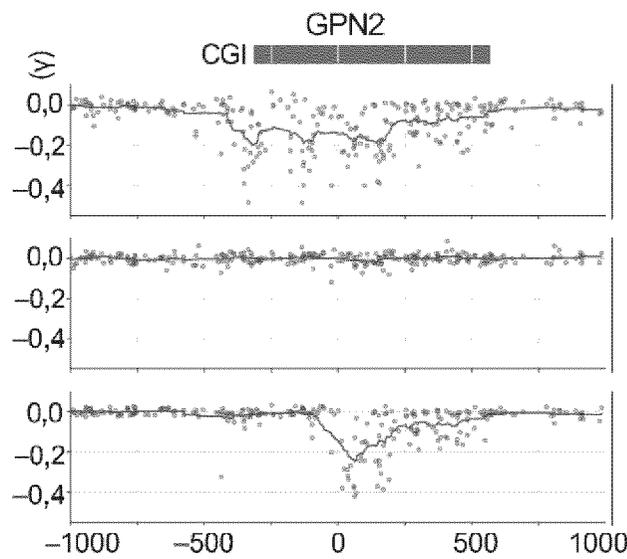


Фиг. 6D



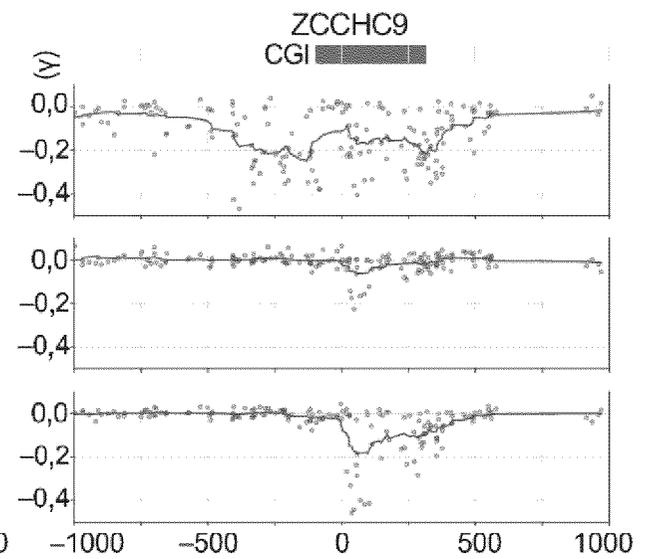
Положение относительно сайта начала транскрипции (п.о.)

Фиг. 6E



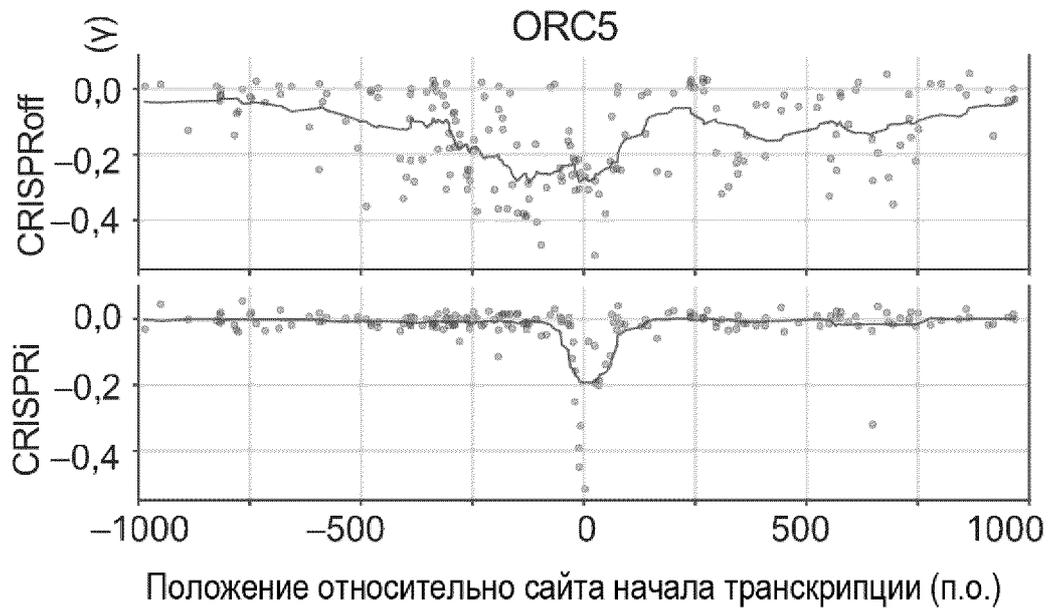
Положение относительно сайта начала транскрипции (п.о.)

Фиг. 6F

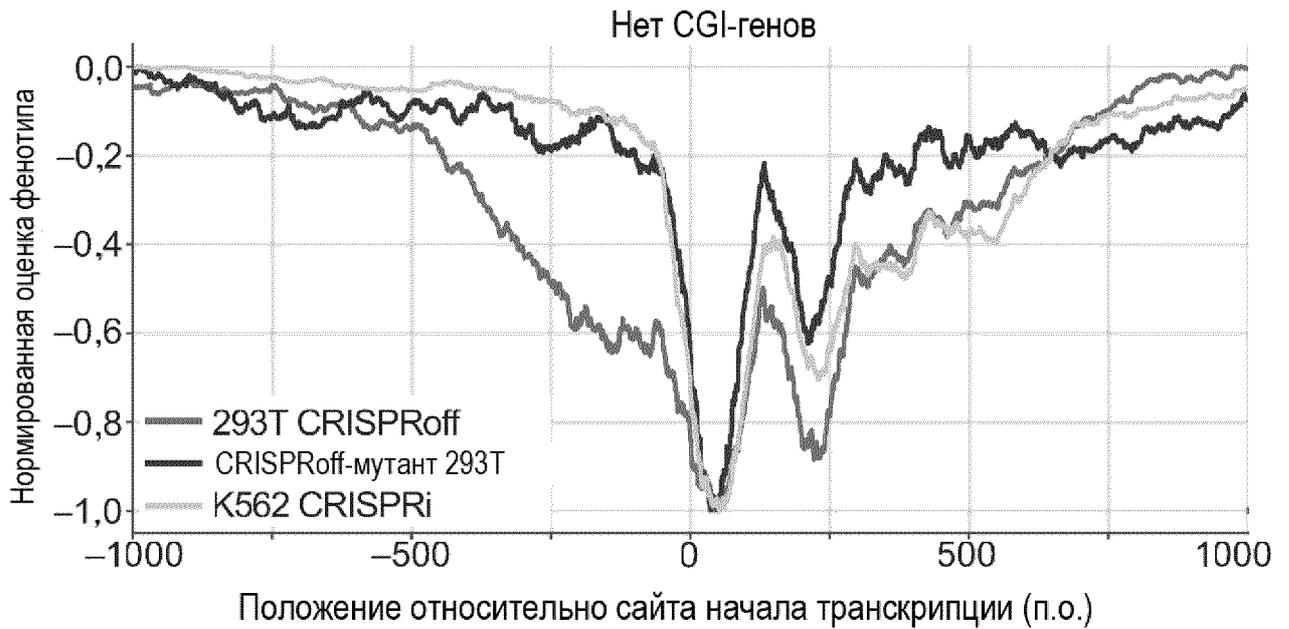


Положение относительно сайта начала транскрипции (п.о.)

Фиг. 6G



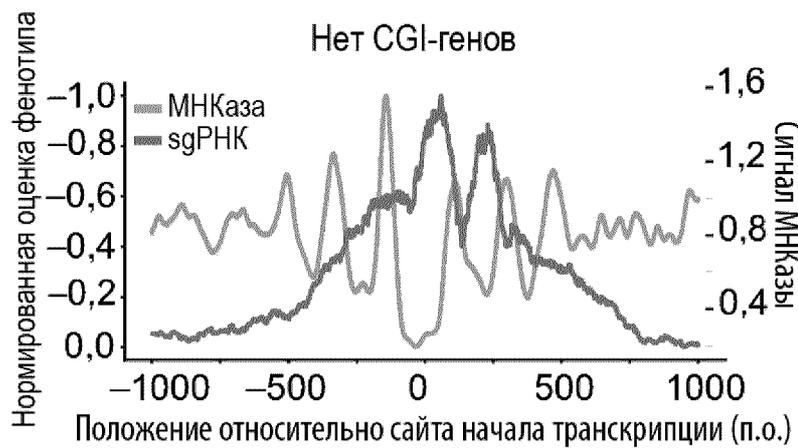
Фиг. 6H



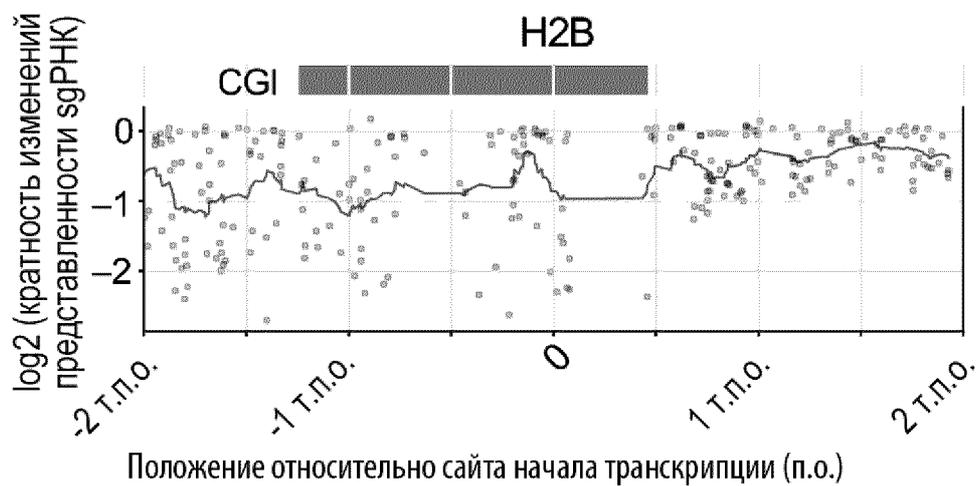
Фиг. 6I



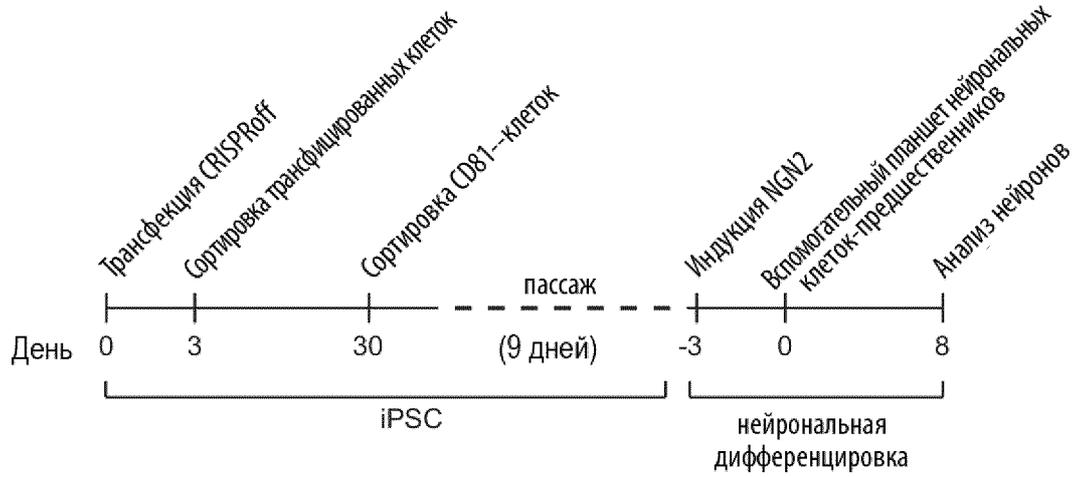
Фиг. 6J



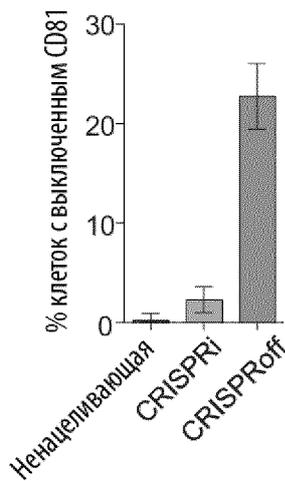
Фиг. 6K



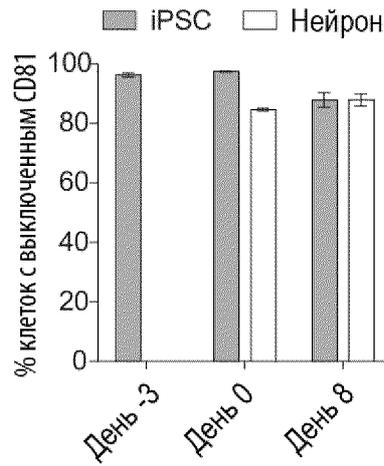
Фиг. 6L



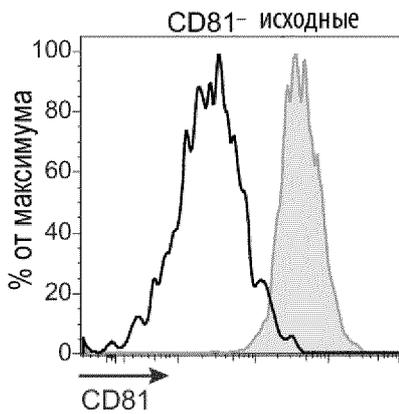
Фиг. 7А



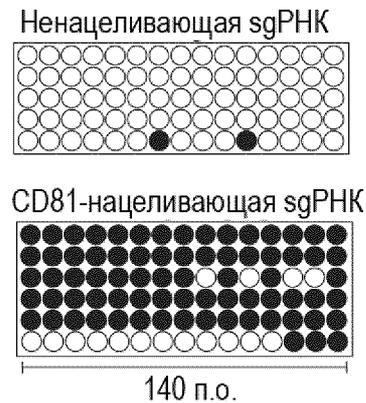
Фиг. 7В



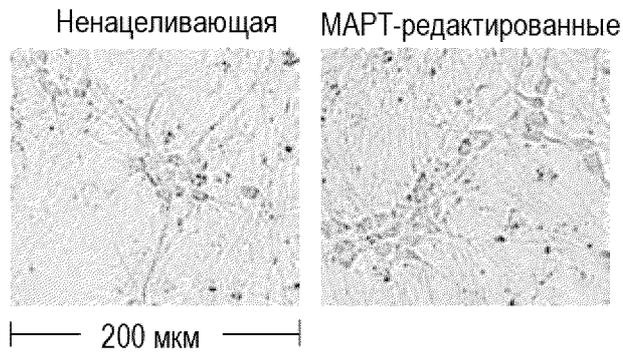
Фиг. 7С



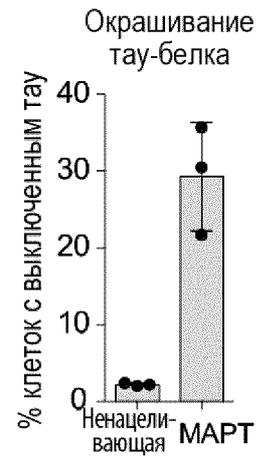
Фиг. 7D



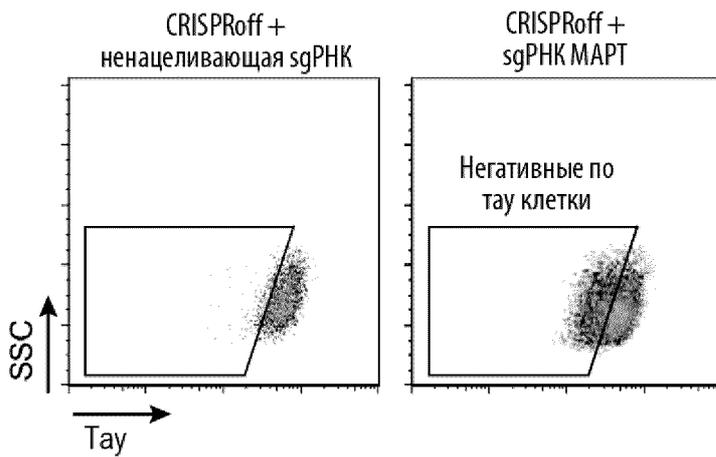
Фиг. 7Е



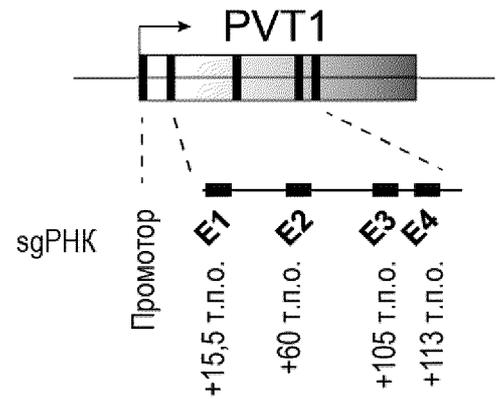
Фиг. 7F



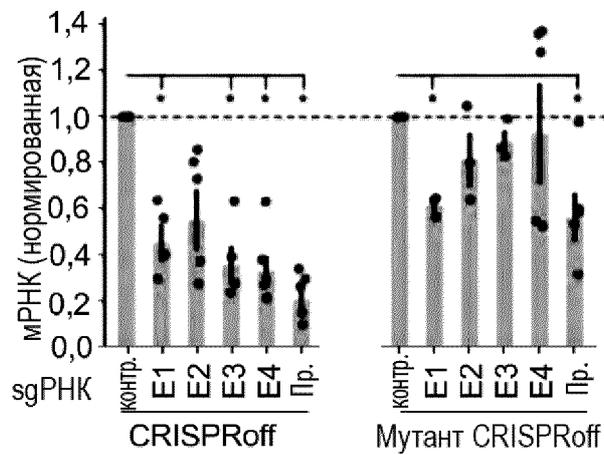
Фиг. 7G



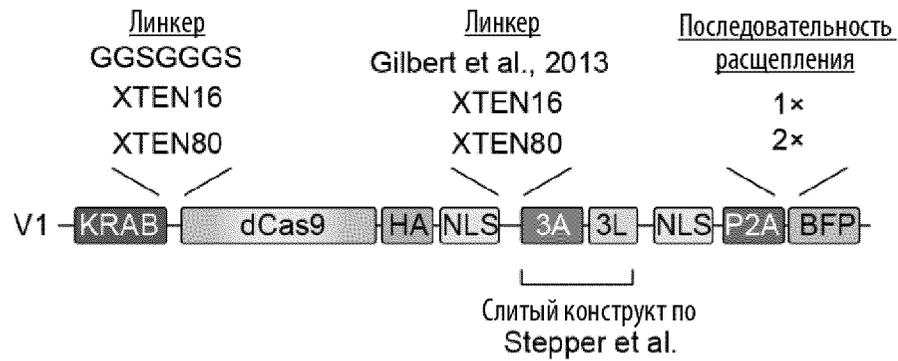
Фиг. 7H



Фиг. 7I



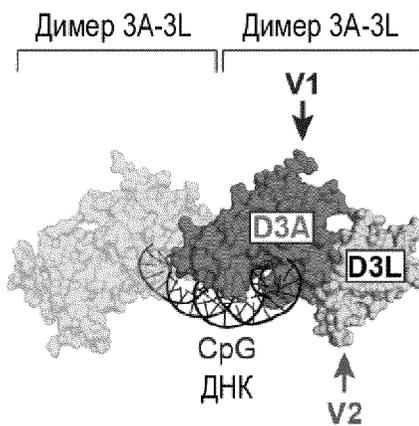
Фиг. 7J



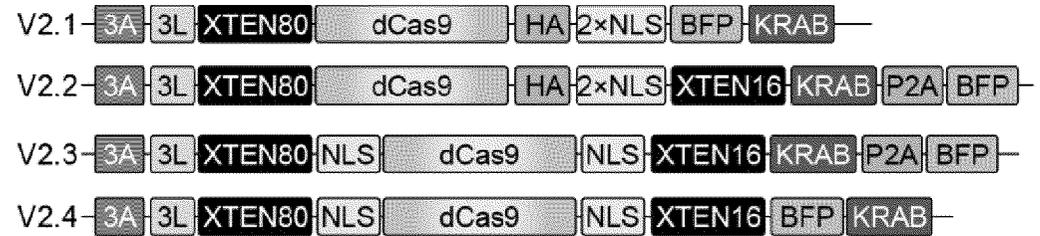
Фиг. 8А



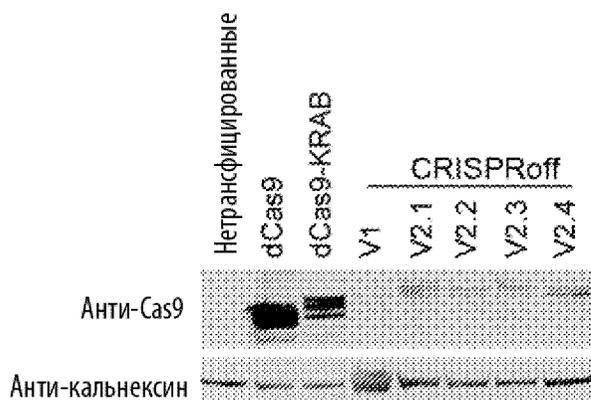
Фиг. 8В



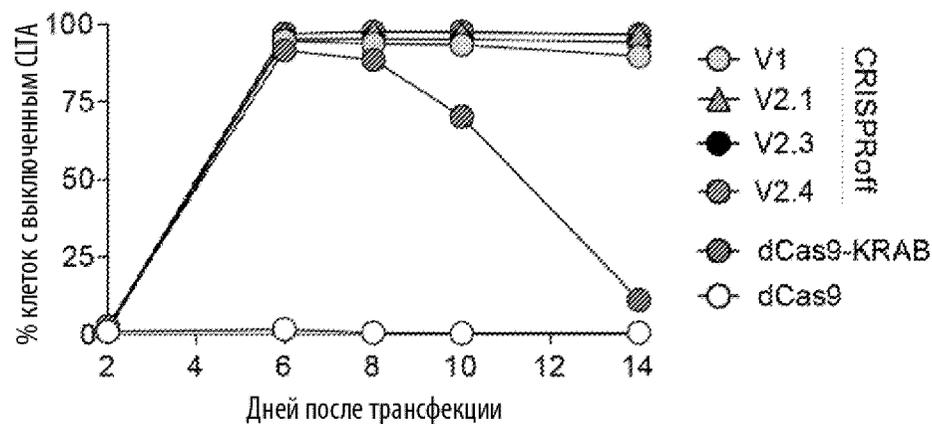
Фиг. 8С



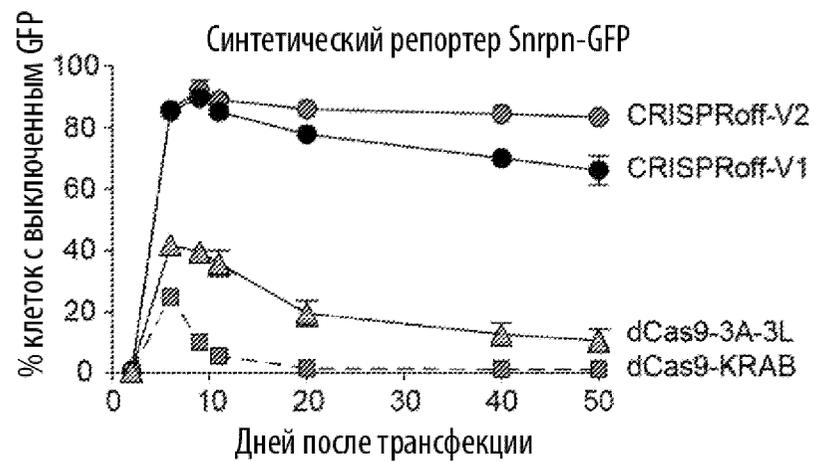
Фиг. 8D



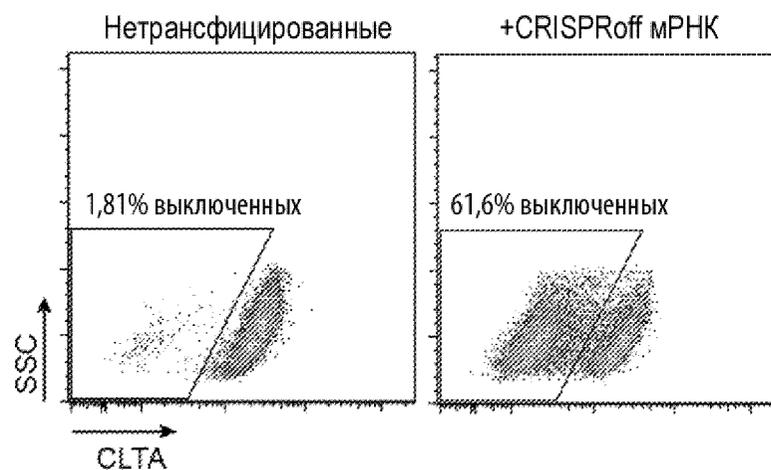
Фиг. 8Е



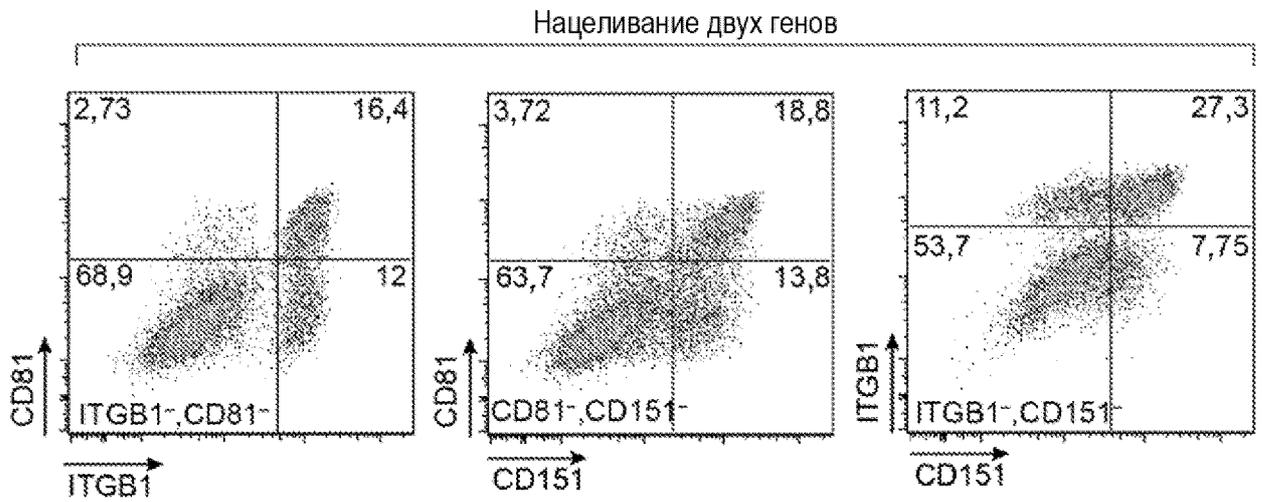
Фиг. 8F



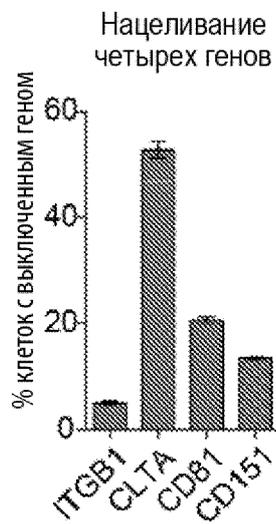
Фиг. 8G



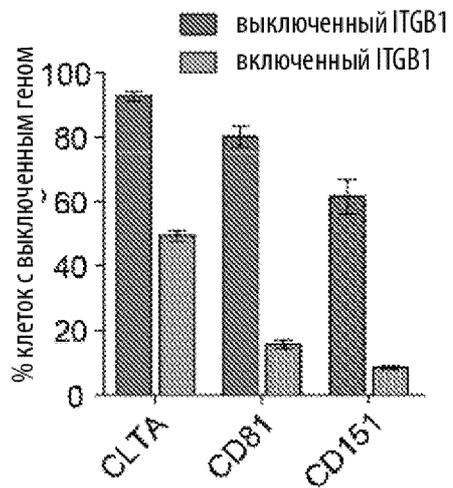
Фиг. 8H



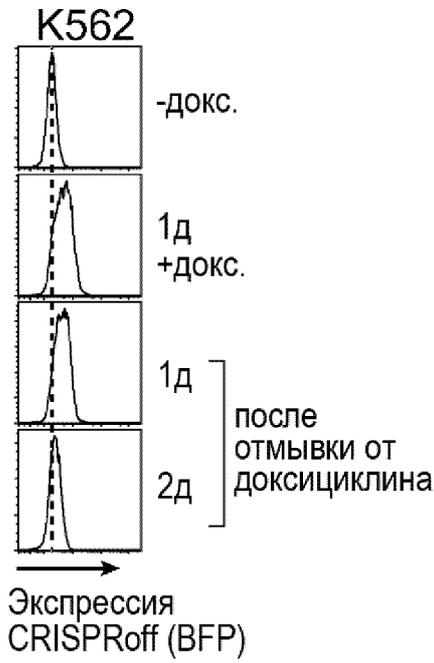
Фиг. 8I



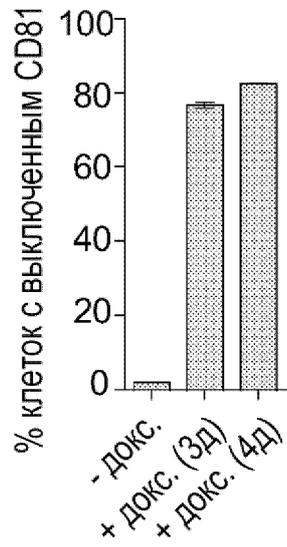
Фиг. 8J



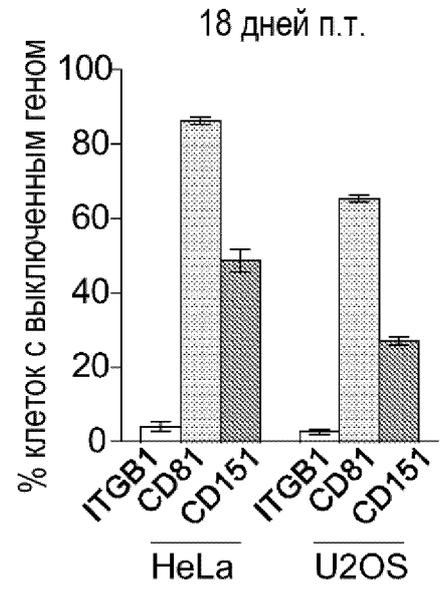
Фиг. 8K



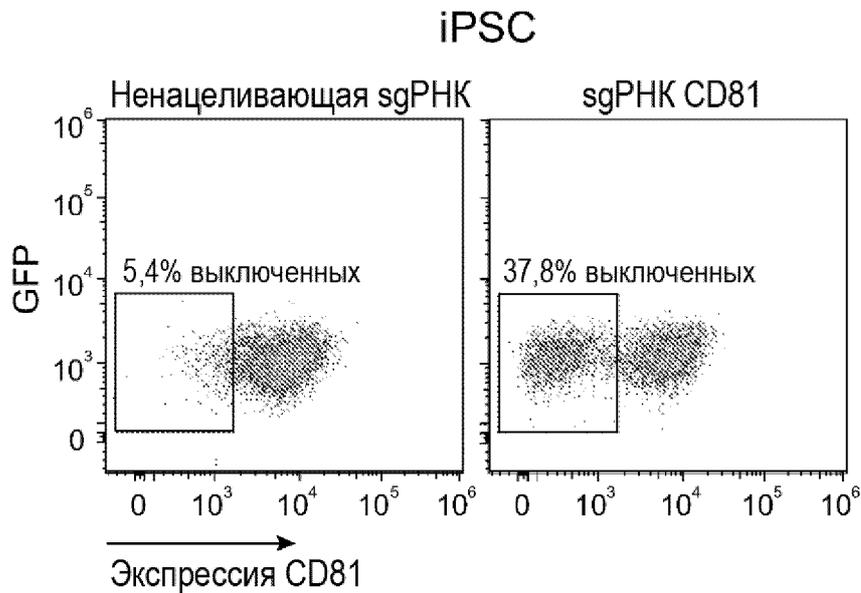
Фиг. 9А



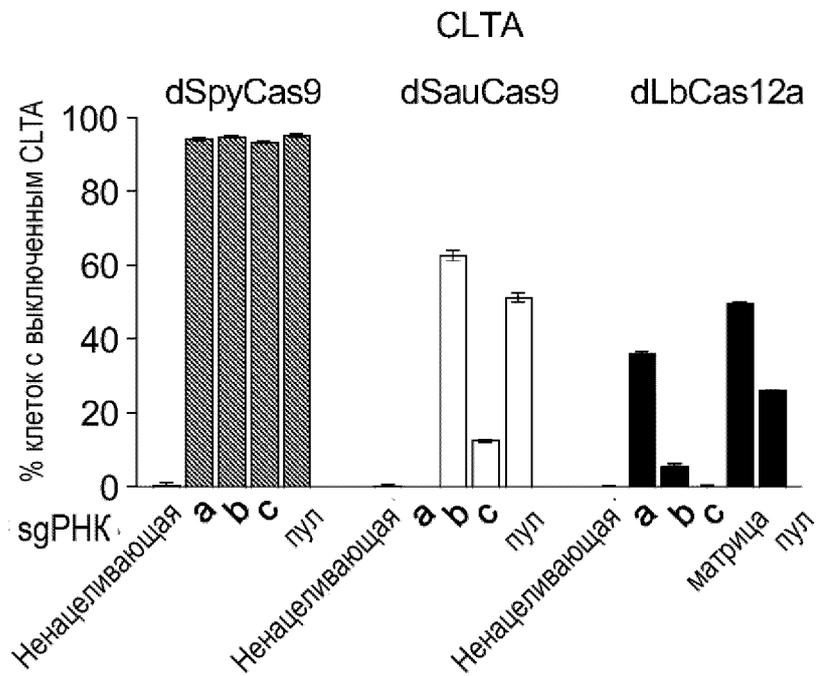
Фиг. 9В



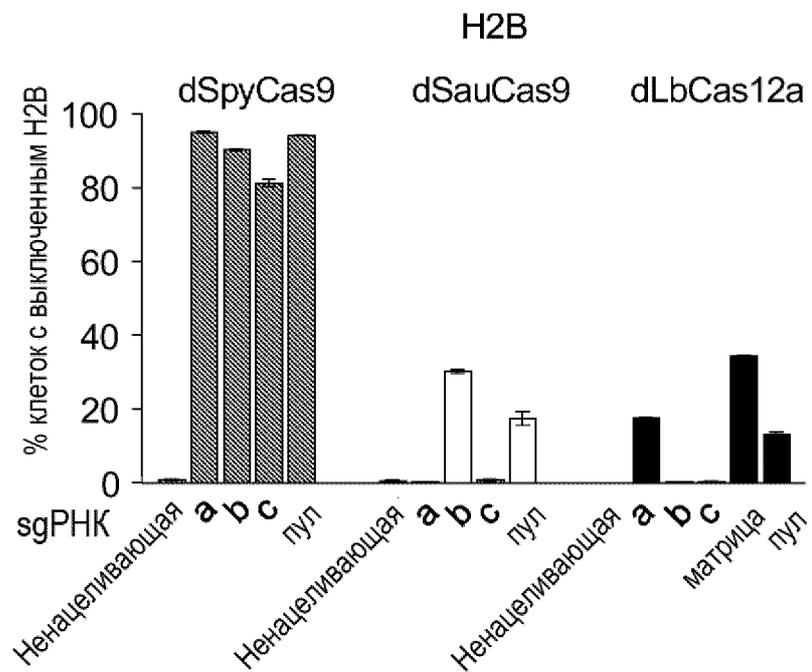
Фиг. 9С



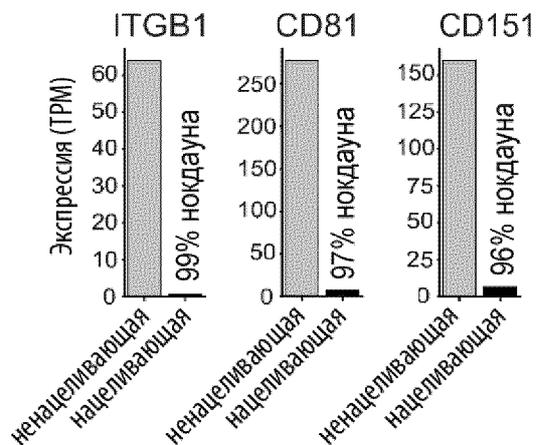
Фиг. 9D



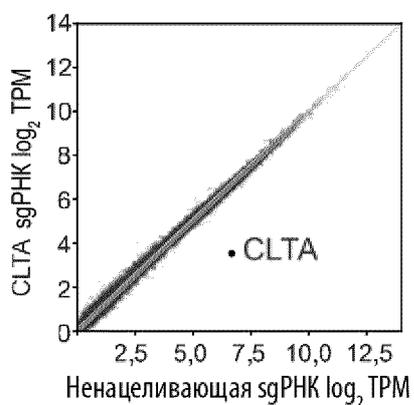
Фиг. 9Е



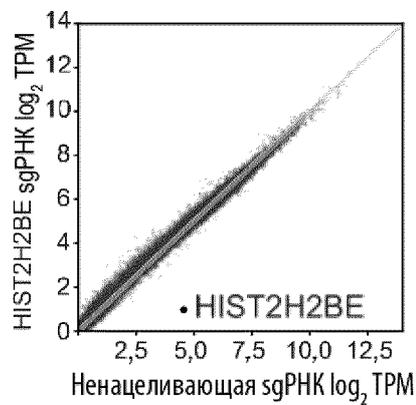
Фиг. 9F



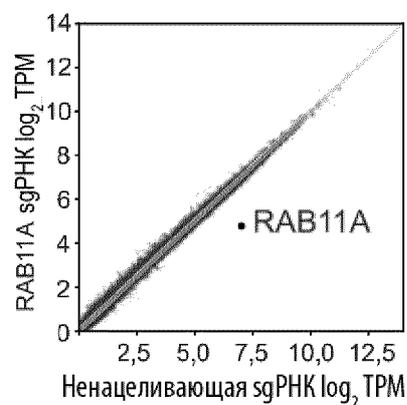
Фиг. 10А



Фиг. 10В



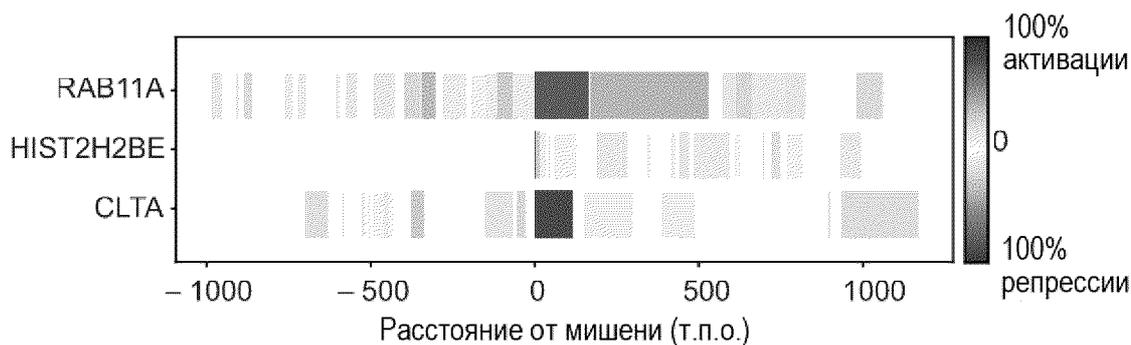
Фиг. 10С



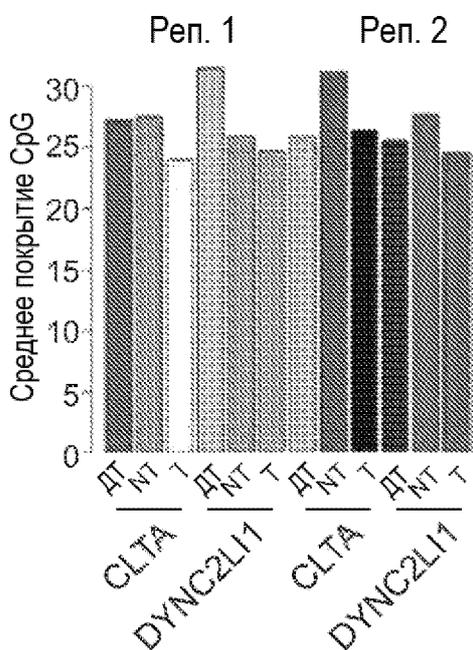
Фиг. 10D



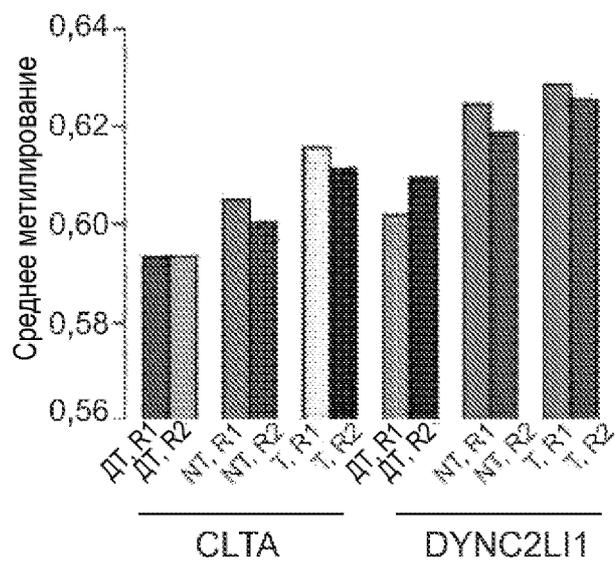
Фиг. 10Е



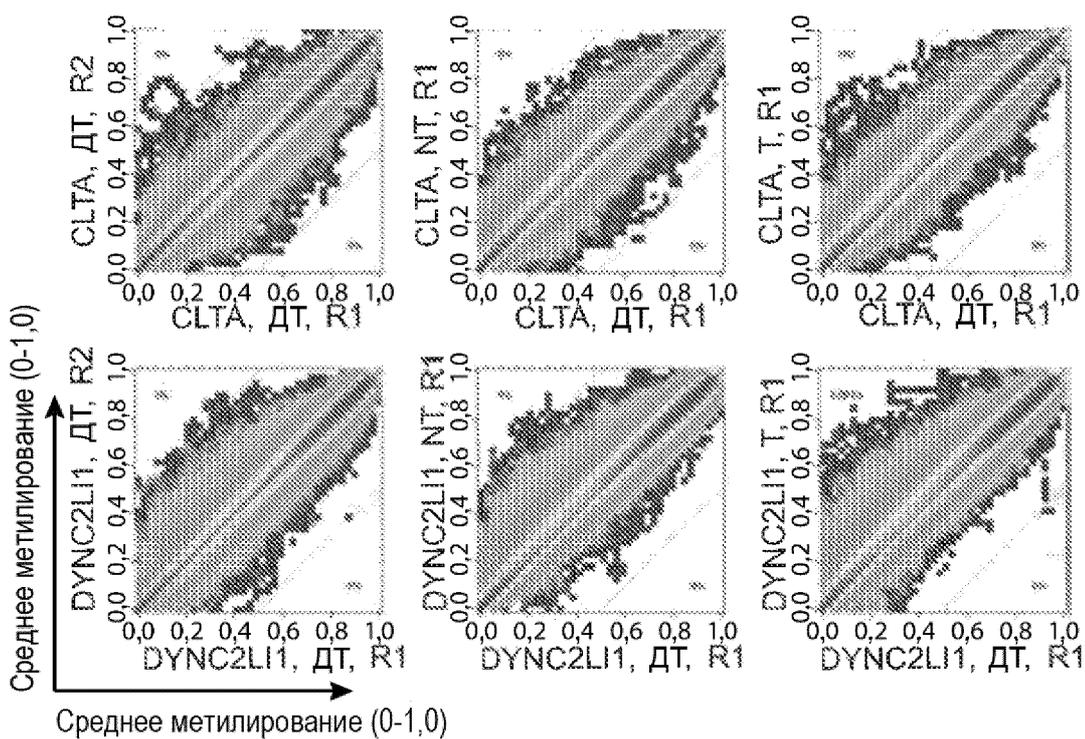
Фиг. 10F



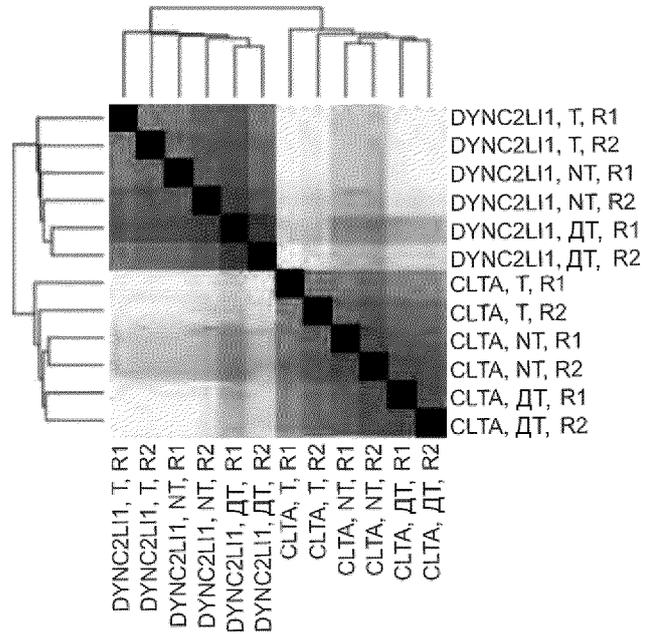
Фиг. 10G



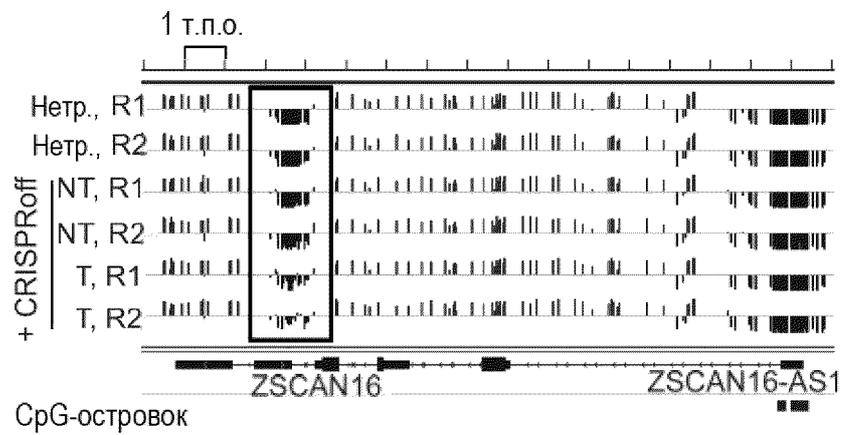
Фиг. 10H



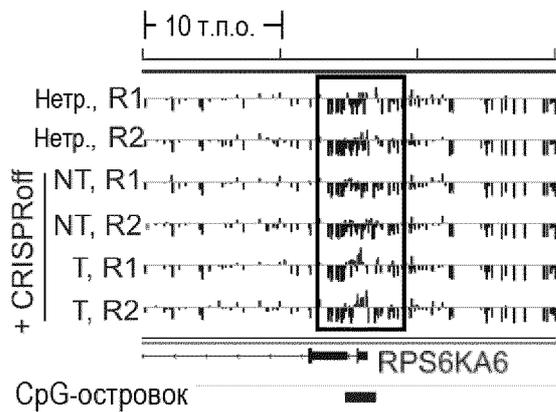
Фиг. 10I



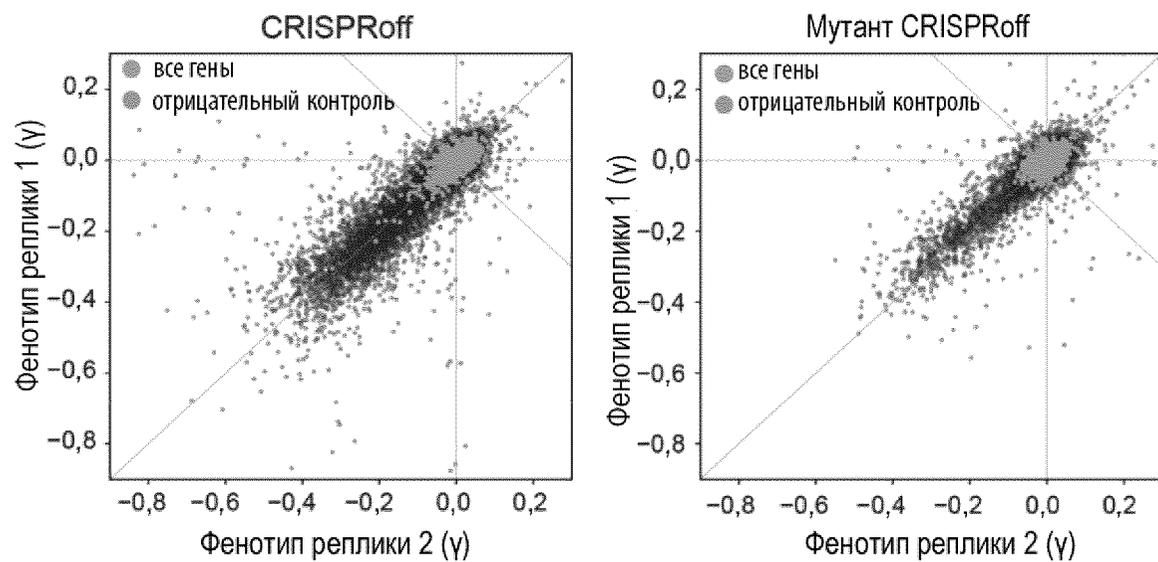
Фиг. 10J



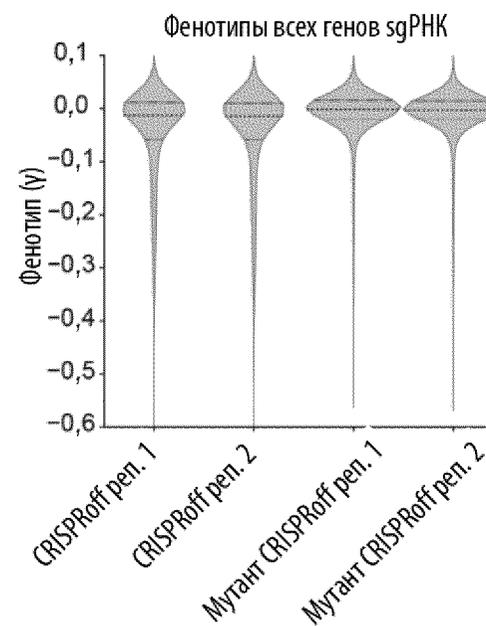
Фиг. 10K



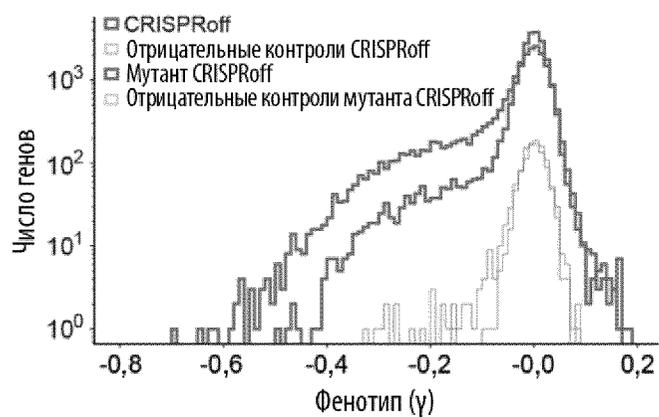
Фиг. 10L



Фиг. 11А



Фиг. 11В



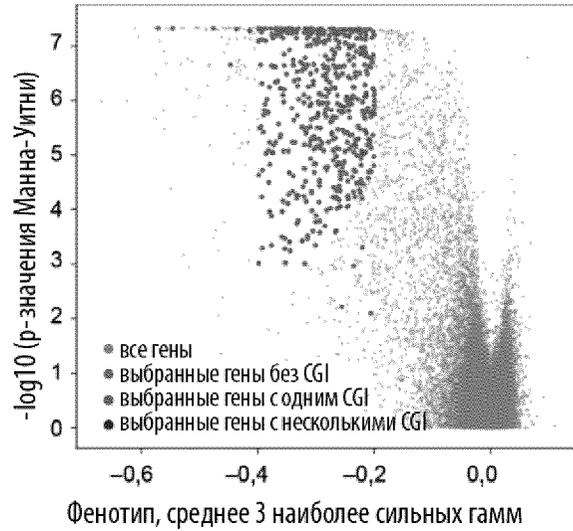
Фиг. 11С



Фиг. 11D



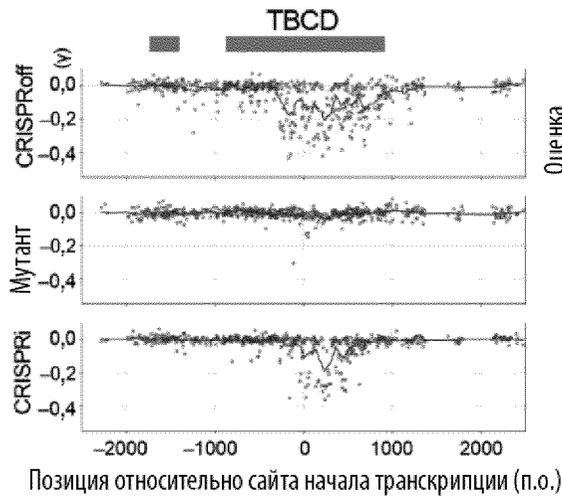
Фиг. 11Е



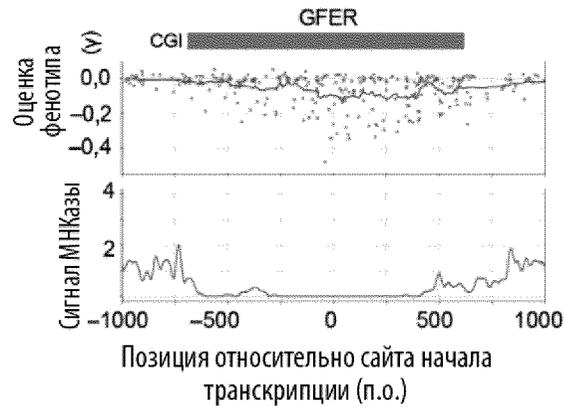
Фиг. 12А



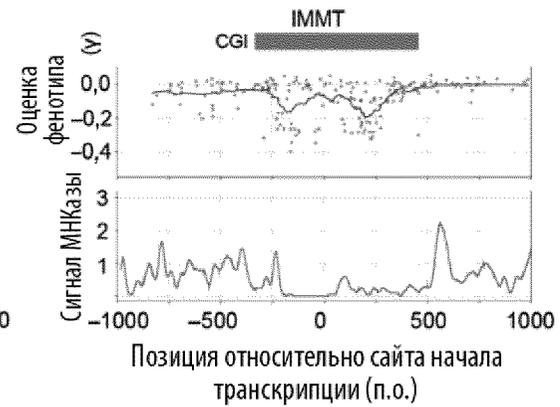
Фиг. 12В



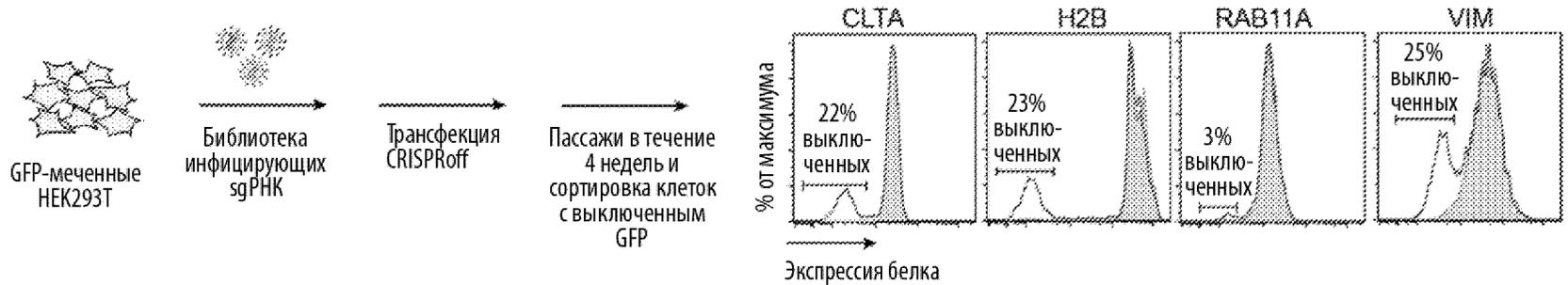
Фиг. 12С



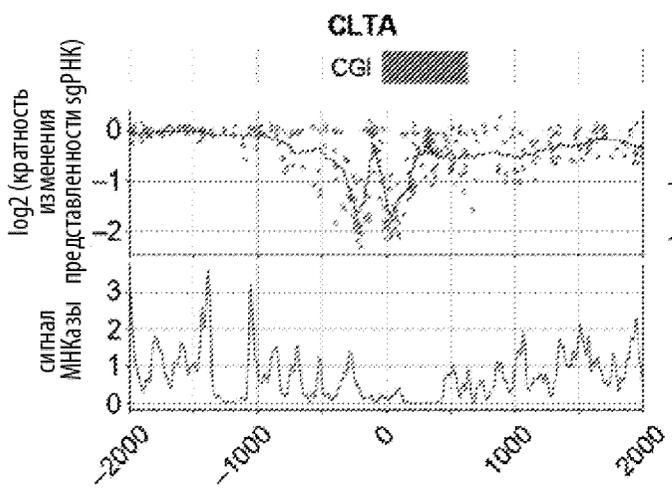
Фиг. 12D



Фиг. 12E

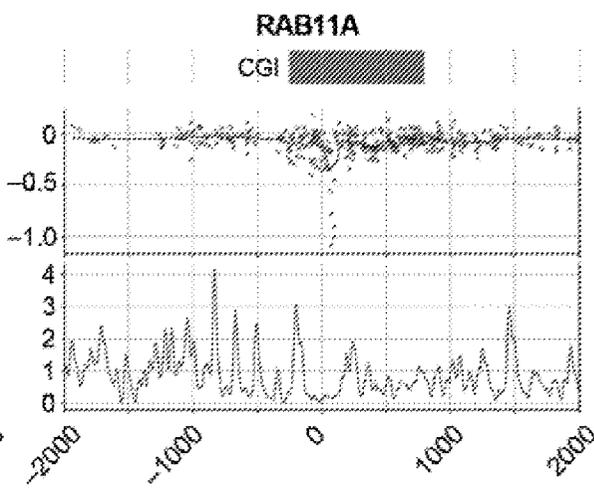


Фиг. 12F



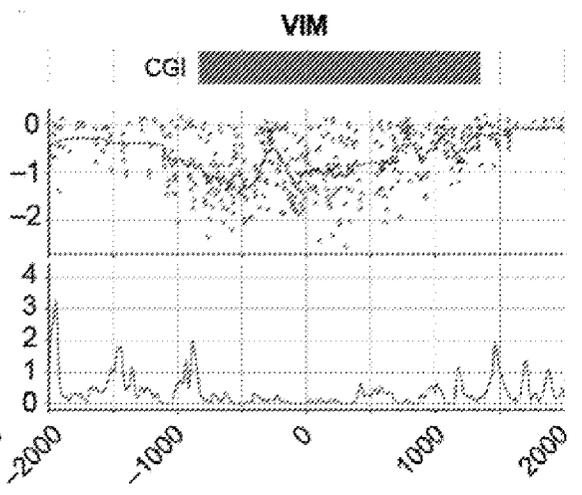
Позиция относительно сайта начала транскрипции (п.о.)

Фиг. 12G



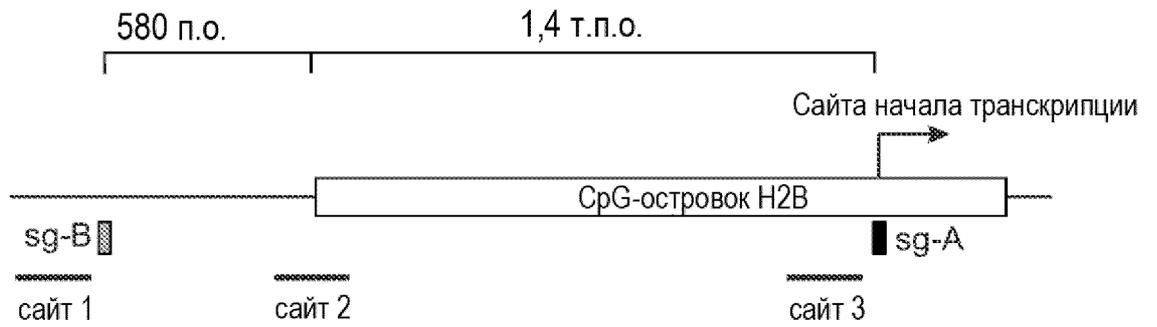
Позиция относительно сайта начала транскрипции (п.о.)

Фиг. 12H

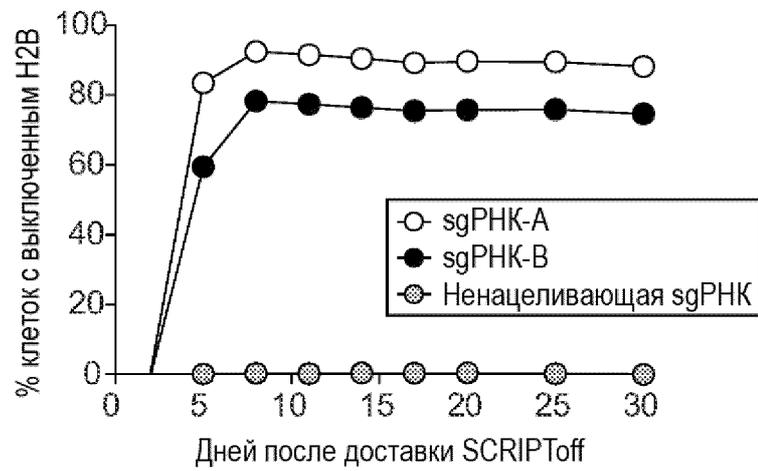


Позиция относительно сайта начала транскрипции (п.о.)

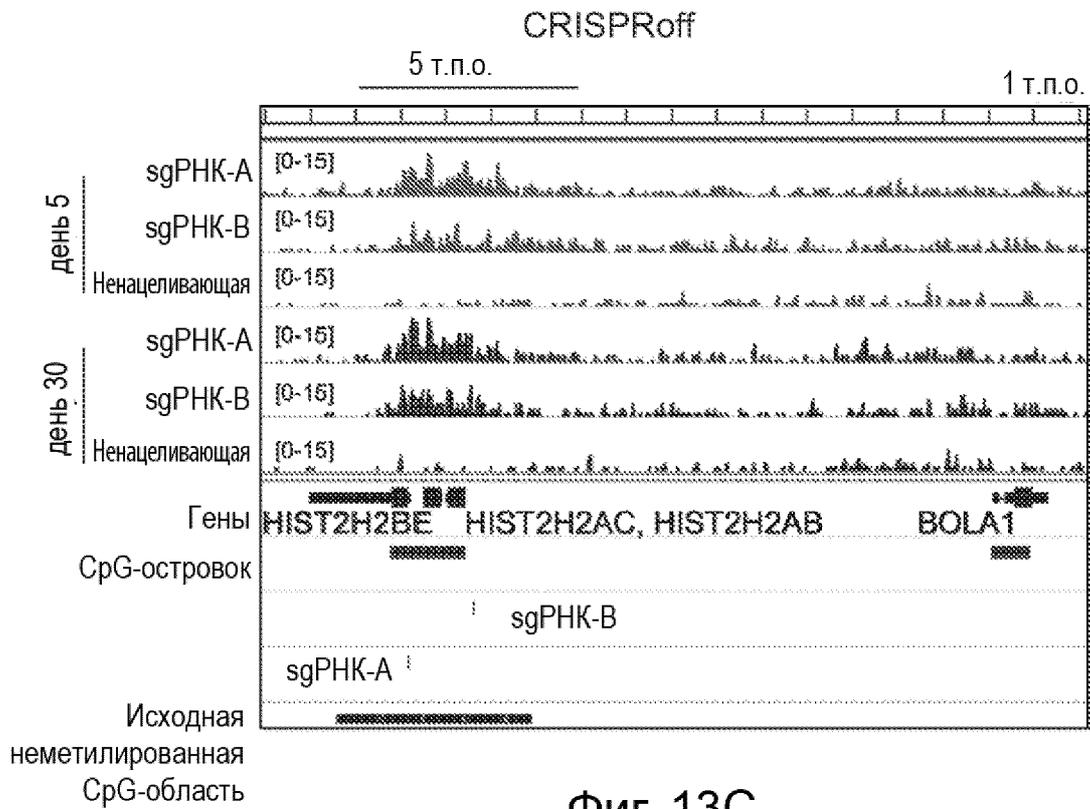
Фиг. 12I



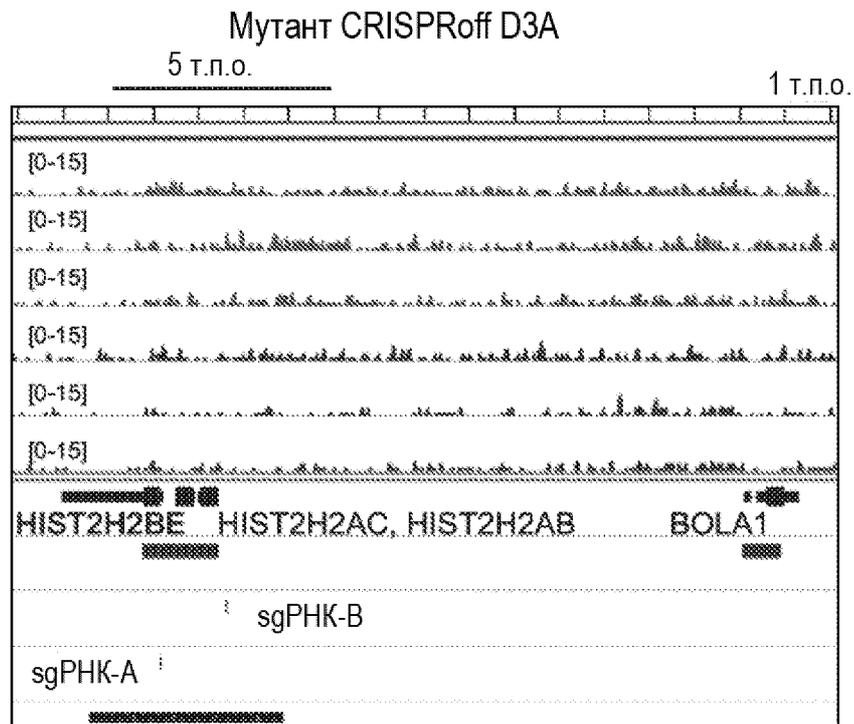
Фиг. 13А



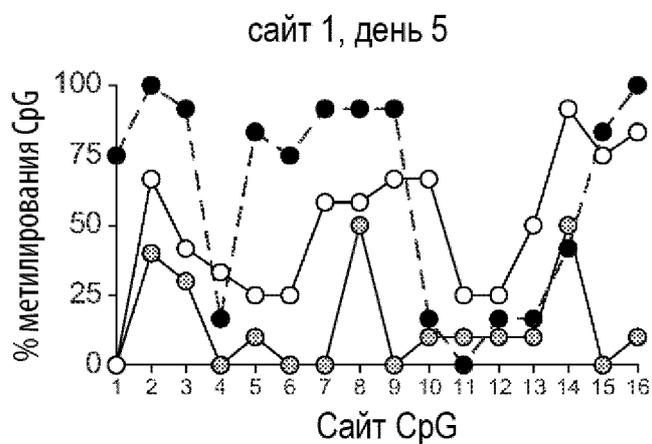
Фиг. 13В



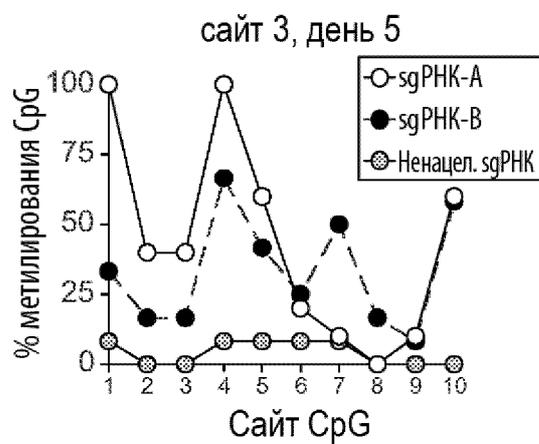
Фиг. 13С



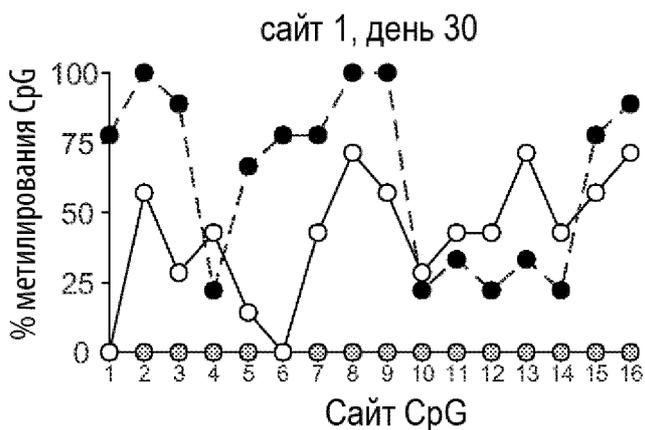
Фиг. 13D



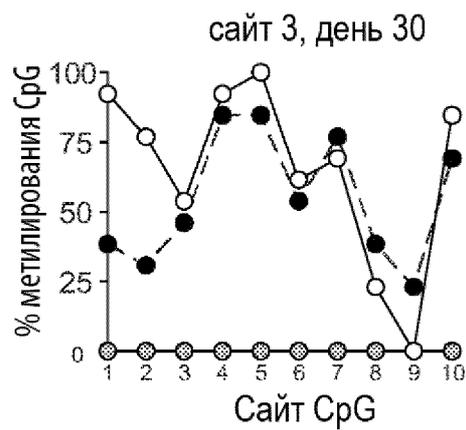
Фиг. 13Е



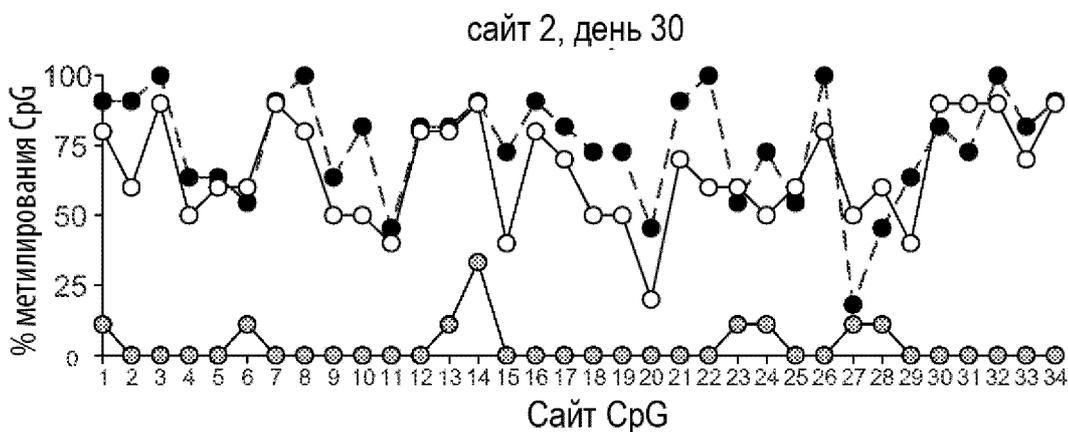
Фиг. 13F



Фиг. 13G



Фиг. 13H



Фиг. 13I