

(19)



Евразийское
патентное
ведомство

(21) 202293533 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.04.05

(51) Int. Cl. C12N 15/63 (2006.01)
C12N 15/67 (2006.01)
C12N 9/02 (2006.01)
C12N 9/22 (2006.01)

(22) Дата подачи заявки
2021.06.04

(54) КОМПОЗИЦИИ ДЛЯ РЕДАКТИРОВАНИЯ ЭПИГЕНОМА И СПОСОБЫ НА ИХ ОСНОВЕ

(31) 63/035,431; 63/118,832

(32) 2020.06.05; 2020.11.27

(33) US

(86) PCT/US2021/035937

(87) WO 2021/248023 2021.12.09

(88) 2022.01.27

(71) Заявитель:

ДЗЕ РИДЖЕНТС ОФ ДЗЕ
ЮНИВЕРСИТИ ОФ КАЛИФОРНИЯ
(US)

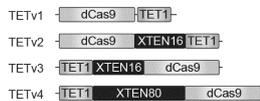
(72) Изобретатель:

Гилберт Люк, Вейссман Джонатан,
Нунез Джеймс, Помье Грег (US)

(74) Представитель:

Медведев В.Н. (RU)

(57) В изобретении представлены, среди прочего, композиции для модулирования экспрессии генов и способы на их основе.



202293533

A1

A1

202293533

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-576647EA/055

КОМПОЗИЦИИ ДЛЯ РЕДАКТИРОВАНИЯ ЭПИГЕНОМА И СПОСОБЫ НА ИХ ОСНОВЕ

Перекрестная ссылка на родственные заявки

[0001] Настоящая заявка претендует на приоритет по заявке США № 63/118832, поданной 27 ноября 2020 г., и заявке США № 63/035431, поданной 5 июня 2020 г., описание которых включено в данный документ посредством ссылки во всей своей полноте.

ЗАЯВЛЕНИЕ О ПРАВАХ НА ИЗОБРЕТЕНИЯ, СОЗДАННЫЕ В СООТВЕТСТВИИ С ПОДДЕРЖИВАЕМЫМИ НА ФЕДЕРАЛЬНОМ УРОВНЕ ИССЛЕДОВАНИЯМИ И РАЗРАБОТКАМИ

[0002] Настоящее изобретение было сделано при поддержке правительства в рамках гранта DARPA-ВAA-16-59, предоставленного Агентством перспективных исследовательских проектов Министерства обороны США. Государство обладает определенными правами на настоящее изобретение.

ССЫЛКА НА «ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ», ТАБЛИЦУ ИЛИ ПРИЛОЖЕНИЕ ПЕРЕЧНЯ В КОМПЬЮТЕРНОЙ ПРОГРАММЕ, ПРЕДСТАВЛЕННОЙ В ФАЙЛЕ ASCII

[0003] Перечень последовательностей, записанный в файле 048536-690001WO_SequenceListing_ST25.txt, созданном в 2021 г., x байт, машинный формат IBM-PC, операционная система MS Windows, настоящим включен посредством ссылки.

УРОВЕНЬ ТЕХНИКИ

[0004] Хотя редактирование генов с использованием технологий на основе CRISPR является перспективным подходом к лечению заболеваний, особенно генетически обусловленных заболеваний, редактирование генов на основе CRISPR основано на разрывах ДНК или редактировании оснований, что может привести к нецелевым модификациям, клеточной токсичности или непредсказуемым исходам в отношении репарации ДНК. Кроме того, большинство технологий, основанных на CRISPR, ограничены редактированием генома и могут вызывать необратимые и вредные изменения. В отличие от этого, модификации, выполненные с помощью эпигенетического редактирования, могут быть долгосрочными и обратимыми, что обеспечивает более безопасный подход к модулированию экспрессии генов. Эпигенетическое редактирование также предоставляет возможности для преобразования как эпигенетического кода ДНК, так и кода гистонов, что позволяет редактировать с использованием различных модальностей и в различных клеточных и генетических контекстах. В данном документе, среди прочего, представлены решения этих и других проблем в данной области техники.

КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0005] В одном аспекте представлен слитый белок, содержащий от N-конца к С-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, или фермент эндонуклеазу с дефицитом нуклеазы. В

аспектах слитый белок дополнительно содержит активатор транскрипции. В аспектах активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0006] В одном аспекте представлен слитый белок, содержащий от N-конца до C-конца, РНК-связывающую последовательность, линкер XTEN и активатор транскрипции. В аспектах активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В аспектах слитый белок дополнительно содержит домен деметилирования, фермент ДНК-эндонуклеазу с дефицитом нуклеазы или фермент эндонуклеазу с дефицитом нуклеазы, последовательность ядерной локализации или комбинацию двух или более из них. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0007] В одном аспекте представлен слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, или фермент эндонуклеазу с дефицитом нуклеазы и активатор транскрипции. В аспектах активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0008] В одном аспекте представлен слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, или фермент эндонуклеазу с дефицитом нуклеазы и последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит активатор транскрипции. В аспектах активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0009] В одном аспекте представлен способ активации последовательности нуклеиновой кислоты-мишени в клетке, включающий: (i) доставку первого полинуклеотида, кодирующего слитый белок, описанный в данном документе, включая его варианты осуществления, в клетку, содержащую подвергнутую сайленсингу нуклеиновую кислоту-мишень; и (ii) доставку в клетку второго полинуклеотида, содержащего: (a) sgRNA или (b) cr:tracrRNA; за счет чего происходит реактивация подвергнутой сайленсингу последовательности нуклеиновой кислоты-мишени в клетке. В аспектах sgRNA содержит

по меньшей мере одну петлю стебля MS2. В аспектах второй полинуклеотид содержит активатор транскрипции. В аспектах второй полинуклеотид содержит две или более sgRNA. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок. В вариантах осуществления слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления способ не включает стадию (ii), когда слитый белок содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0010] В одном аспекте представлен способ активации последовательности нуклеиновой кислоты-мишени или реактивации последовательности подвергнутой сайленсингу нуклеиновой кислоты-мишени в клетке, включающий доставку полинуклеотида, кодирующего слитый белок, описанный в данном документе, включая его варианты осуществления, в клетку, содержащую подвергнутую сайленсингу нуклеиновую кислоту-мишень; за счет чего реактивируется подвергнутая сайленсингу последовательность нуклеиновой кислоты-мишени в клетке. В вариантах осуществления слитый белок содержит домен деметилирования, линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, sgRNA и активатор транскрипции. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0011] Эти и другие варианты осуществления и аспекты раскрытия подробно описаны в данном документе.

Краткое описание графических материалов

[0012] На **фиг. 1** представлена гистограмма клеток HEK293T, реактивирующих подвергнутые сайленсингу с помощью CRISPRoff H2B, Snrpn-GFP или CLTA через 9 дней после Cas9-опосредованного нокаута DNMT1. Столбики погрешностей представляют собой SD из трех независимых экспериментов.

[0013] На **фиг. 2** представлены временные измерения реактивации CLTA после повышения доз 5-aza-dC в клетках HEK293T с подвергнутой сайленсингу CLTA с помощью CRISPRoff. Показан процент клеток с реактивированным CLTA. Этот график показывает, что клетки могут реактивировать экспрессию CLTA посредством деметилирования ДНК.

[0014] На **фиг. 3** представлена медиана флуоресценции CLTA-GFP реактивации CLTA после повышения доз 5-aza-dC в клетках HEK293T с подвергнутой сайленсингу CLTA с помощью CRISPRoff.

[0015] На **фиг. 4** представлена схема эксперимента по реактивации генов. Клетки, кодирующие подвергнутую сайленсингу CLTA-GFP с помощью CRISPRoff, трансфицировали плазмидами, кодирующими dCas9-TET1 и sgRNA.

[0016] На **фиг. 5** представлена схема четырех гибридов TET1 с dCas9 (v1-v4), которые тестировали в отношении реактивации гена CRISPRon.

[0017] На **фиг. 6** представлен график, показывающий временную зависимость

реактивации CLTA после трансфекции четырех гибридов TET, показанных на фиг. 5 с пулом sgRNA, нацеленным на *CLTA*. Ген CLTA имеет островок CpG.

[0018] На **фиг. 7** представлена гистограмма, показывающая сравнение реактивации CLTA с использованием четырех слияний TET на фиг. 5, котрансфицированных одной последовательностью sgRNA или пулом из трех sgRNA. Столбики погрешностей представляют собой диапазон двух технических повторов.

[0019] На **фиг. 8** представлен репрезентативный график FACS реактивации CLTA, измеренный через 28 дней после трансфекции TETv4 и нацеливания на sgRNA.

[0020] На **фиг. 9A** представляет собой бисульфитный ПЦР-анализ CGI *CLTA* после реактивации TET1, показывающий высокие уровни деметилирования цитозина (белые кружки) по сравнению с подвергнутой сайленсингу CLTA с помощью CRISPRoff (черные кружки). Каждая строка представляет одно чтение секвенирования. Процент метилирования локуса представлен на горизонтальной гистограмме.

[0021] На **фиг. 9B** представлена схема CGI *CLTA* (зеленый) с аннотированными сайтами связывания sgRNA (a, b, c). Затенение леденцового графика представляет собой процент каждого динуклеотида CpG с метилированным цитозином, измеренный с помощью бисульфитной ПЦР. Промотор, сплайсинг и аннотации CGI получали из UCSC Genome Browser.

[0022] На **фиг. 10** представлена схема комплекса TETv4 и трансаактиваторного рибонуклеопротеина, опосредованного sgRNA, кодирующей два РНК-аптамера MS2. Домены трансаактиватора включают однокомпонентную, двухкомпонентную и трехкомпонентную архитектуры тетрамера VP16 VP64, домен активации *RELA* (p65) и вирусный активатор транскрипции Rta.

[0023] На **фиг. 11** представлена схема векторов, которые экспрессируют sgRNA, нацеленную на *CLTA*, и слитый продукт на основе белка оболочки MS2 (MCP) с различными активаторами транскрипции.

[0024] На **фиг. 12** представляет собой скрипичный график, который представляет собой медиану флуоресценции CLTA-GFP через 2 дня после трансфекции sgRNA, нацеленных на CLTA, и трансаактиваторы, слитые с dCas9 или dCas9 и MCP, в клетках с эндогенно экспрессированным CLTA-GFP.

[0025] На **фиг. 13** представлена гистограмма, показывающая сравнение кратности изменения доли реактивированных клеток CLTA-GFP, измеренное через два дня после трансфекции трансаактиваторов, слитых с TETv4 и MCP. Данные отображаются как кратное изменение по сравнению с одним только TETv4, рассчитанное на основе среднего значения двух технических повторов.

[0026] На **фиг. 14** показаны гистограммы, иллюстрирующие, что TET1 в комбинации с трансаактиваторами реактивирует экспрессию генов. Уровни экспрессии генов и плазмид измеряли в несколько временных точек после трансфекции.

[0027] На **фиг. 15A-15B** представлены скрипичные графики, иллюстрирующие, что временная экспрессия трансаактиваторов Rta, p65-Rta и VP64-p65 приводила к значимому

повышению экспрессии генов одиночной клетки в реактивированных клетках. На фиг. 15B представлено сравнение медианы флуоресценции отдельных клеток с реактивированным CLTA-GFP, измеренной через 28 дней после трансфекции. Данные являются репрезентативными для двух технических повторностей. * *p*-значение < 0,05, ** *p*-значение < 0,0005, *** *p*-значение, $1e-15$ относительно GFP-положительной популяции в состоянии TETv4 по критерию суммы рангов Уилкоксона.

[0028] На **фиг. 16** представлена гистограмма, показывающая реактивацию гена слитым белком TET1 в клетках с ранее подвергнутыми сайленсингу генами. DYNC2LI1 и LAMP2 не имеют канонического островка CpG.

[0029] На **фиг. 17** показана динамика клеток HEK293T с реактивацией CLTA-GFP после трансфекции sgRNA, нацеленных на CLTA, и либо TETv4 в отдельности, либо TETv4 вместе с различными доменами трансаактиватора, слитыми с MCP, в клетках с подвергнутой сайленсингу CLTA с помощью CRISPRoff. Необработанные клетки представлены белыми кружками. Столбики погрешностей представляют собой SD из трех независимых экспериментов.

[0030] На **фиг. 18** представлена динамика клеток HEK293T с реактивацией CLTA-GFP после трансфекции sgRNA, нацеленных на CLTA, и либо dCas9-VPR, либо dCas9 вместе с различными доменами трансаактиватора, слитыми с MCP, или нетрансфицированными клетками. Трансфекции проводили в отсутствие TETv4 для измерения стойкой активации генов в отсутствие деметилирования ДНК. Столбики погрешностей представляют собой SD из трех независимых экспериментов.

[0031] На **фиг. 19A-19D** показаны слитые белки и реактивация их генов. На **фиг. 19D** представлен график, показывающий слитые белки, описанные в данном документе, включая GCP21 (SEQ ID NO:102), JKNp146 (SEQ ID NO:99) и JKNp147 (SEQ ID NO:101). На **фиг. 19B-19D** показываю реактивацию гена CLTA, гена DYNC2LI1 и гена гистона H2B (соответственно) после трансфекции слитыми белками, измеренную через 13 дней после трансфекции.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0032] Определения

[0033] Если не указано иное, все технические и научные термины, используемые в данном документе, имеют значение, обычно понятное специалисту в области, к которой относится настоящее изобретение. Следующие ссылки предоставляют специалисту в данной области техники общее определение многих из терминов, используемых в настоящем изобретении: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); и Hale & Marham, The Harper Collins Dictionary of Biology (1991). Как используется в данном документе, следующие термины имеют значения, приписываемые им, если не указано иное.

[0034] Использование неопределенной или определенной формы в единственном числе в настоящем изобретении и в следующих пунктах формулы изобретения

соответствует традиционному подходу в патентах со значением «по меньшей мере один», если только в конкретном случае из контекста ясно, что этот термин в данном конкретном случае означает конкретно один и только один. Аналогичным образом, термин «содержащий» является открытым, не исключая дополнительные элементы, признаки, компоненты и т.д. Ссылки, указанные в данном документе, прямо включены в данный документ посредством ссылки во всей своей полноте, если не указано иное.

[0035] Термины «содержать», «включать» и «иметь», а также их производные используются в данном документе взаимозаменяемо как всеобъемлющие, неограниченные термины. Например, использование слов «содержащий», «включающий» или «имеющий» означает, что любой элемент, содержащийся, имеющийся или включенный, не является единственным элементом, охватываемым подлежащим в предложении, содержащем глагол.

[0036] «Нуклеиновая кислота» относится к нуклеотидам (например, дезоксирибонуклеотидам или рибонуклеотидам) и их полимерам в одно-, двух- или многоцепочечной форме или их комплементарным формам. Термины «полинуклеотид», «олигонуклеотид», «олиго» 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, т.е. мономеру. Нуклеотиды могут представлять собой рибонуклеотиды, дезоксирибонуклеотиды или их модифицированные версии. Примеры рассматриваемых в данном документе полинуклеотидов включают одно- и двухцепочечную ДНК, одно- и двухцепочечную РНК и гибридные молекулы, содержащие смеси одно- и двухцепочечной ДНК и РНК. Примеры нуклеиновых кислот, например, полинуклеотиды, рассматриваемые в данном документе, включают без ограничения любой тип РНК, например, mRNA, siRNA, miRNA, sgRNA и направляющую РНК, а также любой тип ДНК, геномную ДНК, плазмидную ДНК и миникольцевую ДНК и любые их фрагменты. В аспектах нуклеиновая кислота представляет собой информационную РНК. В аспектах матричная РНК представляет собой матричный рибонуклеопротеин (RNP). Термин «дуплекс» в контексте полинуклеотидов относится в обычном и общепринятом смысле к двухцепочечному. Нуклеиновые кислоты могут быть линейными или разветвленными. Например, нуклеиновые кислоты могут представлять собой линейную цепь нуклеотидов, или нуклеиновые кислоты могут быть разветвленными, например, так, что нуклеиновые кислоты содержат одно или несколько плеч или ответвлений нуклеотидов. Необязательно разветвленные нуклеиновые кислоты многократно разветвляются с образованием структур более высокого порядка, таких как дендримеры и т.п.

[0037] Используемые в данном документе термины «нуклеиновая кислота», «молекула нуклеиновой кислоты», «олигомер нуклеиновой кислоты», «олигонуклеотид», «последовательность нуклеиновой кислоты», «фрагмент нуклеиновой кислоты» и «полинуклеотид» используются взаимозаменяемо и предназначены для включения без ограничения полимерной формы нуклеотидов, ковалентно связанных друг с другом,

которые могут иметь различную длину, либо дезоксирибонуклеотидов, либо рибонуклеотидов, либо их аналогов, производных или модификаций. Разные полинуклеотиды могут иметь разную трехмерную структуру и выполнять различные функции, известные или неизвестные. Неограничивающие примеры полинуклеотидов включают ген, фрагмент гена, экзон, интрон, межгенную ДНК (включая, помимо прочего, гетерохроматическую ДНК), информационную РНК (mRNA), транспортную РНК, рибосомную РНК, рибозим, cDNA, рекомбинантный полинуклеотид, разветвленный полинуклеотид, плаزمид, вектор, выделенная ДНК последовательности, выделенная РНК последовательности, sgRNA, направляющая РНК, зонд нуклеиновой кислоты и праймер. Полинуклеотиды, применимые в способах настоящего изобретения, могут включать последовательности природных нуклеиновых кислот и их варианты, последовательности искусственных нуклеиновых кислот или комбинацию таких последовательностей.

[0038] Полинуклеотид обычно состоит из определенной последовательности четырех нуклеотидных оснований: аденина (A); цитозина (C); гуанина (G); и тимина (T) (урацила (U) вместо тимина (T), когда полинуклеотид представляет собой РНК). Таким образом, термин «полинуклеотидная последовательность» представляет собой алфавитное представление молекулы полинуклеотида; альтернативно, этот термин может быть применен к самой молекуле полинуклеотида. Это алфавитное представление можно вводить в базы данных на компьютере, имеющем центральный процессор, и использовать для приложений биоинформатики, таких как функциональная геномика и поиск гомологии. Полинуклеотиды могут необязательно включать один или несколько нестандартных нуклеотидов, аналогов нуклеотидов и/или модифицированных нуклеотидов.

[0039] Нуклеиновые кислоты, включая, например, нуклеиновые кислоты с фосфотиоатным 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. Например, нуклеиновая кислота может включать аминокислотный реакционноспособный фрагмент, который реагирует с аминокислотой на белке или полипептиде посредством ковалентного, нековалентного или другого взаимодействия.

[0040] Термины охватывают нуклеиновые кислоты, содержащие известные аналоги нуклеотидов или модифицированные остатки или связи в остове, которые являются синтетическими, встречающимися в природе и не встречающимися в природе, которые характеризуются свойствами связывания, подобными референтной нуклеиновой кислоте, и которые метаболизируются подобно референтным нуклеотидам. Примеры таких аналогов включают без ограничения производные фосфодиэфира, включая, например, фосфорамидат, фосфординидат, фосфоротиоат (также известный как фосфоротиоат, содержащий серу с двойной связью, замещающую кислород в фосфате), фосфородитиоат, фосфонокарбоновые кислоты, фосфонокарбоксилаты, фосфоноуксусную кислоту, фосфономуравьиную кислоту, метилфосфонатные, борфосфонатные или О-

метилфосфорамидитные связи (см. Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press), а также модификации нуклеотидных оснований, таких как 5-метилцитидин или псевдоуридин; и остовы и связи пептидных нуклеиновых кислот. Другие аналоговые нуклеиновые кислоты включают кислоты с положительным остовом; неионогенные остовы, модифицированные сахара и нерибозные остовы (например, фосфородиамидатные морфолиноолигонуклеотиды или закрытые нуклеиновые кислоты (LNA), известные в данной области техники), в том числе описанные в патентах США №№ 5235033 и 5034506, и Chapters 6 and 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 определение нуклеиновых кислот. Модификации рибозо-фосфатного остова могут быть выполнены по разным причинам, например, для повышения стабильности и периода полужизни таких молекул в физиологических условиях или в качестве зондов на биочипе. Могут быть получены смеси встречающихся в природе нуклеиновых кислот и аналогов; альтернативно, могут быть получены смеси различных аналогов нуклеиновых кислот и смеси встречающихся в природе нуклеиновых кислот и аналогов. В аспектах межнуклеотидные связи в ДНК представляют собой фосфодиэфирные, фосфодиэфирные производные или их комбинацию.

[0041] Нуклеиновые кислоты могут включать неспецифические последовательности. Используемый в данном документе термин «неспецифическая последовательность» относится к последовательности нуклеиновой кислоты, которая содержит ряд остатков, не предназначенных для комплементарности любой другой последовательности нуклеиновой кислоты или комплементарных только частично. Например, неспецифическая последовательность нуклеиновой кислоты представляет собой последовательность остатков нуклеиновой кислоты, которая не действует как ингибирующая нуклеиновая кислота при контакте с клеткой или организмом.

[0042] Термин «комплементарный» или «комплементарность» относится к способности нуклеиновой кислоты образовывать водородные связи с другой последовательностью нуклеиновой кислоты либо по традиционному методу Уотсона-Крика, либо по другим нетрадиционным типам. Например, последовательность А-Г-Т комплементарна последовательности Т-С-А. Комплементарность в процентах указывает процент остатков в молекуле нуклеиновой кислоты, которые могут образовывать водородные связи (например, спаривание оснований Уотсона-Крика) со второй последовательностью нуклеиновой кислоты (например, 5, 6, 7, 8, 9, 10 из 10 на 50%, 60%, 70%, 80%, 90% и 100% соответственно). «Совершенно комплементарный» означает, что все смежные остатки последовательности нуклеиновой кислоты будут водородными связями с таким же количеством смежных остатков во второй последовательности нуклеиновой кислоты. Термин «по сути комплементарный», используемый в данном документе, относится к степени комплементарности, которая составляет по меньшей мере 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% или 100% в области 8, 9, 10, 11, 12, 13, 14,

15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 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).

[0043] Фраза «жесткие условия гибридизации» относится к условиям, при которых зонд будет гибридизоваться со своей целевой последовательностью, обычно в сложной смеси нуклеиновых кислот, но не с другими последовательностями. Жесткие условия являются зависимыми от последовательностей и будут различными при различных обстоятельствах. Более длинные последовательности гибридизируются специфическим образом при более высоких температурах. Подробное руководство по гибридизации нуклеиновых кислот можно найти 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). Как правило, жесткие условия выбирают так, чтобы они были примерно на 5-10°C ниже температуры плавления (T_m) для конкретной последовательности при определенной ионной силе рН. T_m представляет собой температуру (при определенной ионной силе, рН и концентрации нуклеиновых кислот), при которой 50% зондов, комплементарных мишени, гибридизуются с последовательностью-мишенью в равновесии (поскольку последовательности-мишени присутствуют в избытке, при T_m , 50% зондов заняты в состоянии равновесия). Жесткие условия также могут быть достигнуты с помощью добавления дестабилизирующих средств, таких как формамид. Для селективной или специфической гибридизации положительный сигнал составляет по меньшей мере в два раза больше фона, предпочтительно в 10 раз больше фона гибридизации. Примеры жестких условий гибридизации могут быть следующими: 50% формамид, 5x SSC и 1% SDS, инкубация при 42 °C, или 5x SSC, 1% SDS, инкубация при 65 °C, с промывкой в 0,2x SSC и 0,1% SDS при 65 °C.

[0044] Нуклеиновые кислоты, которые не гибридизуются друг с другом в жестких условиях, по-прежнему являются по сути идентичными, если полипептиды, которые они кодируют, являются по сути идентичными. Это происходит, например, когда копия нуклеиновой кислоты создается с использованием максимальной вырожденности кодона, допускаемого генетическим кодом. В таких случаях нуклеиновые кислоты обычно гибридизуются в умеренно жестких условиях гибридизации. Примеры «умеренно жестких условий гибридизации» включают гибридизацию в буфере из 40% формамида, 1 M NaCl, 1% SDS при 37°C и промывку в 1X SSC при 45 °C. Положительная гибридизация по меньшей мере в два раза превышает фоновую. Специалисту с обычной квалификацией будет легко понять, что можно использовать альтернативные условия гибридизации и промывки для обеспечения условий аналогичной жесткости. Дополнительные рекомендации по определению параметров гибридизации приведены в многочисленных источниках, например, *Current Protocols in Molecular Biology*, ed. Ausubel, et al., см. выше.

[0045] Термин «ген» означает сегмент ДНК, участвующий в образовании белка; он включает участки, предшествующие кодирующему участку («лидерной и трейлерной последовательности») и следующие за таковым, а также промежуточные

последовательности (интроны) между отдельными кодирующими сегментами (экзонами). Лидерная последовательность, трейлерная последовательность, а также интроны включают регуляторные элементы, необходимые при транскрипции и трансляции гена. Кроме того, «белковый генный продукт» представляет собой белок, экспрессируемый конкретным геном.

[0046] Слово «экспрессия» или «экспрессированный», используемое в данном документе в отношении гена, означает транскрипционный и/или трансляционный продукт этого гена. Уровень экспрессии молекулы ДНК в клетке можно определить на основании либо количества соответствующей mRNA, присутствующей в клетке, либо количества белка, кодируемого этой ДНК, продуцируемой клеткой. Уровень экспрессии некодирующих молекул нуклеиновой кислоты (например, sgRNA) можно определить с помощью стандартных методов ПЦР или нозерн-блоттинга, хорошо известных в данной области техники. См. Sambrook et al., 1989 *Molecular Cloning: A Laboratory Manual*, 18.1-18.88.

[0047] Термин «регуляторная последовательность транскрипции», как представлено в данном документе, относится к сегменту ДНК, который способен повышать или уменьшать транскрипцию (например, экспрессию) определенного гена в организме. Неограничивающие примеры регуляторных последовательностей транскрипции включают промоторы, энхансеры и сайленсеры.

[0048] Термины «сайт начала транскрипции» и «сайт инициации транскрипции» могут использоваться взаимозаменяемо для обозначения 5'-конца последовательности гена (например, последовательности ДНК), где РНК-полимераза (например, РНК-полимераза, направляемая ДНК) начинает синтезировать РНК-транскрипт. Сайт начала транскрипции может представлять собой первый нуклеотид транскрибируемой последовательности ДНК, где РНК-полимераза начинает синтез РНК-транскрипта. Квалифицированный специалист может определить сайт инициации транскрипции с помощью рутинных экспериментов и анализа, например, путем проведения анализа прерванной транскрипции или с помощью определений в соответствии с базой данных FANTOM5.

[0049] Используемый в данном документе термин «промотор» относится к области ДНК, которая иницирует транскрипцию конкретного гена. Промоторы обычно расположены рядом с сайтом начала транскрипции гена, выше гена и на той же цепи (т.е. 5' на смысловой цепи) ДНК. Промоторы могут иметь длину от примерно 100 до примерно 1000 пар оснований.

[0050] «Направляющая РНК» или «gRNA», как представлено в данном документе, относится к любой полинуклеотидной последовательности, имеющей достаточную комплементарность с полинуклеотидной последовательностью-мишенью для гибридизации с последовательностью-мишенью и прямое специфичное в отношении последовательности связывание комплекса CRISPR с последовательностью-мишенью. В аспектах степень комплементарности между направляющей последовательностью и соответствующей ей последовательностью-мишенью при оптимальном выравнивании с

использованием подходящего алгоритма выравнивания составляет приблизительно или более приблизительно 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99% и больше.

[0051] В вариантах осуществления полинуклеотид (например, gRNA) представляет собой одноцепочечную рибонуклеиновую кислоту. В аспектах полинуклеотид (например, gRNA) имеет длину от приблизительно 10 до приблизительно 200 остатков нуклеиновой кислоты. В аспектах полинуклеотид (например, gRNA) имеет длину от приблизительно 50 до приблизительно 150 остатков нуклеиновой кислоты. В аспектах полинуклеотид (например, gRNA) имеет длину от приблизительно 80 до приблизительно 140 остатков нуклеиновой кислоты. В аспектах полинуклеотид (например, gRNA) имеет длину от приблизительно 90 до приблизительно 130 остатков нуклеиновой кислоты. В аспектах полинуклеотид (например, gRNA) имеет длину от приблизительно 100 до приблизительно 120 остатков нуклеиновой кислоты. В аспектах длина полинуклеотида (например, gRNA) составляет приблизительно 113 остатков нуклеиновой кислоты.

[0052] В общем, направляющая последовательность (т.е. последовательность, нацеленная на ДНК) представляет собой любую полинуклеотидную последовательность, имеющую достаточную комплементарность с целевой полинуклеотидной последовательностью для гибридизации с последовательностью-мишенью (например, последовательностью-мишенью геномной или митохондриальной ДНК) и прямое специфическое в отношении последовательности связывание комплекса (например, комплекса CRISPR) с последовательностью-мишенью. В аспектах степень комплементарности между направляющей последовательностью и соответствующей ей последовательностью-мишенью при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет приблизительно или более приблизительно 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99% и больше. В аспектах степень комплементарности между направляющей последовательностью и соответствующей ей последовательностью-мишенью при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет по меньшей мере приблизительно 80%, 85%, 90%, 95% или 100%. В аспектах степень комплементарности составляет по меньшей мере 90%. Оптимальное выравнивание можно определить с использованием любого подходящего алгоритма выравнивания последовательностей, неограничивающий пример которого включает алгоритм Смита-Уотермана, алгоритм Нидлмана-Вунша, алгоритмы, основанные на преобразовании Берроуза-Уилера (например, выравниватель Берроуза-Уилера), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, Сан-Диего, Калифорния), SOAP (доступно на soap.genomics.org.cn) и Maq (доступно на maq.sourceforge.net). В аспектах длина направляющей последовательности составляет приблизительно или более приблизительно 10, 20, 30, 35, 40, 45, 50, 75 или больше нуклеотидов. В аспектах длина направляющей последовательности составляет от приблизительно 10 до приблизительно 150, от приблизительно 15 до приблизительно 100 нуклеотидов. В аспектах длина направляющей последовательности составляет менее приблизительно 75, 50, 45, 40, 35, 30, 25, 20, 15, 12

или менее нуклеотидов. В аспектах длина направляющей последовательности составляет приблизительно или более приблизительно 20 нуклеотидов. Способность направляющей последовательности направлять специфичное в отношении последовательности связывание комплекса (например, комплекса CRISPR) с последовательностью-мишенью можно оценить с помощью любого подходящего анализа. Например, компоненты системы CRISPR, достаточные для образования комплекса (например, комплекса CRISPR), включая тестируемую направляющую последовательность, могут быть введены в клетку-хозяина, имеющую соответствующую последовательность-мишень, например, путем трансфекции векторами, кодирующими компоненты последовательности CRISPR с последующей оценкой предпочтительного расщепления в последовательности-мишени, например, с помощью анализа Surveyor, известного в данной области техники. Аналогично, расщепление полинуклеотидной последовательности-мишени можно оценить в пробирке, предоставив последовательность-мишень, компоненты комплекса (например, комплекса CRISPR), включая тестируемую направляющую последовательность и контрольную направляющую последовательность, отличную от тестируемой направляющей последовательности и сравнив связывание или скорость расщепления последовательности-мишени между реакциями тестовой и контрольной направляющей последовательности. Возможны и другие анализы, и это будет понятно специалистам в данной области техники.

[0053] Термины «sgRNA», «одиночная направляющая РНК» и «последовательность одной направляющей РНК» используются взаимозаменяемо и относятся к полинуклеотидной последовательности, включая последовательность crРНК и необязательно последовательность tracrRNA. Последовательность crРНК включает направляющую последовательность (т.е. «направляющую» или «спейсерную») и сопряженную последовательность tracr (т.е. прямой(ые) повтор(ы)). Термин «направляющая последовательность» относится к последовательности, которая указывает сайт-мишень. В аспектах две РНК могут кодироваться отдельно crРНК и tracrРНК в виде 2 молекул РНК, которые затем образуют комплекс РНК/РНК благодаря комплементарному спариванию оснований между crРНК и tracrРНК (т.е. до того, как он станет компетентным для связывания с ферментом ДНК-эндонуклеазой с дефицитом нуклеазы, направляемой РНК). В аспектах первая нуклеиновая кислота включает последовательность tracrРНК, а отдельная вторая нуклеиновая кислота включает последовательность gRNA, в которой отсутствует последовательность tracrRNA. В аспектах первая нуклеиновая кислота, содержащая последовательность tracrРНК, и вторая нуклеиновая кислота, содержащая последовательность gRNA, взаимодействуют друг с другом и необязательно включаются в комплекс (например, комплекс CRISPR). Иллюстративные sgRNA и их последовательности-мишени показаны в таблицах 2, 3 и 4.

[0054] Таблица 2

Название	Последовательность-мишень (от 5' до 3')	последовательность sgRNA (от 5' до 3')
----------	---	--

A (JKNg156)	ACTGCGGAAATTTGAGCGT (SEQ ID NO: 37)	ACGCUCAAAUUUCCGCAGU (SEQ ID NO: 38)
B (JKNg158)	AGGCAATGGCTGCACATGC (SEQ ID NO: 39)	GCAUGUGCAGCCAUUGCCU (SEQ ID NO: 40)
C (JKNg160)	GACGCTTGGTTCTGAGGAG (SEQ ID NO: 41)	CUCCUCAGAACCAAGCGUC (SEQ ID NO: 42)

[0055] Таблица 3

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

[0056] Таблица 4

Название	Последовательность-мишень (от 5' до 3')	последовательность sgRNA (от 5' до 3')
Pcsk9 sgRNA-1	TCCGGAAACGCATTCCTCT (SEQ ID NO: 43)	AGAGGAAUGCGUUUCCGGA (SEQ ID NO: 44)
Pcsk9 sgRNA-2	ACCGGCAGCCTGCGCGTCC (SEQ ID NO: 61)	GGACGCGCAGGCUGCCGGU (SEQ ID NO: 62)

Pcsk9 sgRNA-3	CGATGGGCACCCACTGCTC (SEQ ID NO: 63)	GAGCAGUGGGUGCCCAUCG (SEQ ID NO: 64)
Pcsk9 sgRNA-4	CCTTCACGTGGACGCGCAG (SEQ ID NO: 65)	CUGCGCGUCCACGUGAAGG (SEQ ID NO: 66)
Pcsk9 sgRNA-5	CGTGAAGGTGGAAGCCTTC (SEQ ID NO: 67)	GAAGGCUUCCACCUUCACG (SEQ ID NO: 68)
Npc1 sgRNA-1	CTCCTTGGTCAGGCGCCGG (SEQ ID NO: 69)	CCGGCGCCUGACCAAGGAG (SEQ ID NO: 70)
Npc1 sgRNA-2	TGGTCAGGCGCCGGTTCCG (SEQ ID NO: 71)	CGGAACCGGCGCCUGACCA (SEQ ID NO: 72)
Npc1 sgRNA-3	TAGAGGTCGCCTTCTCCTC (SEQ ID NO: 73)	GAGGAGAAGGCGACCUCUA (SEQ ID NO: 74)
Npc1 sgRNA-4	CGACGCTCGGGTCGCGGTG (SEQ ID NO:75)	CACCGCGACCCGAGCGUCG (SEQ ID NO:76)
Npc1 sgRNA-5	ATGCTGTGCGCCGCGCGGGG (SEQ ID NO:77)	CCCCGCGCGGCGACAGCAU (SEQ ID NO:78)
Spes1 sgRNA-1	CTCACCTCACCGGAGCCA (SEQ ID NO:79)	UGGCUCCGGUGAGGGUGAG (SEQ ID NO:80)
Spes1 sgRNA-2	CCGCAAAC TTTACTCCTTA (SEQ ID NO:81)	UAAGGAGUAAAGUUUGCGG (SEQ ID NO:82)
Spes1 sgRNA-3	CTCGGAGACATCCGCTTCC (SEQ ID NO: 60)	GGAAGCGGAUGUCUCCGAG (SEQ ID NO: 60)
Spes1 sgRNA-4	CTCCTAAGATTGGCTTAC (SEQ ID NO:83)	GUGAAGCCAAUCUUAGGAG (SEQ ID NO:84)
Spes1 sgRNA-5	CCGGAGCCACTCCTAAGAT (SEQ ID NO:85)	AUCUUAGGAGUGGCUCGG (SEQ ID NO:86)
Cd81 sgRNA-1	TTCTCTACCCTACGTCTCA (SEQ ID NO:87)	UGAGACGUAGGGUAGAGAA (SEQ ID NO:88)
Cd81 sgRNA-2	TACGTCTCATTCTCCGCAA (SEQ ID NO:89)	UUGCGGAGAAUGAGACGUA (SEQ ID NO:90)
Cd81 sgRNA-3	GCTAGGCCTCCAGCCCTTC (SEQ ID NO:91)	GAAGGGCUGGAGGCCUAGC (SEQ ID NO:92)
Cd81 sgRNA-4	ACAGGTGGCGCCGCAACTT (SEQ ID NO:93)	AAGUUGCGGCGCCACCUGU (SEQ ID NO:94)

Cd81 sgRNA-5	AGCCGGAGGCGCGAGAGTC (SEQ ID NO:95)	GACUCUCGCGCCUCCGGCU (SEQ ID NO:96)
--------------	---------------------------------------	---------------------------------------

[0057] Последовательности в таблицах 2, 3 и 4 представляют собой нацеливающие последовательности crRNA. Например, полная одиночная направляющая РНК (sgRNA) для SEQ ID NO: 38 представляет собой: GACGCUCAAAUUUCCGCAGUGUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGU UUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUU UUUUU (SEQ ID NO:114). Обычная последовательность tracr каждой отдельной направляющей для Sp Cas9 представляет собой GUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:115). Квалифицированному специалисту понятно, что последовательности sgRNA в таблицах 2, 3 и 4 представляют собой 19 пар оснований и не отражают того, что каждая sgRNA начинается с G, который требуется, если экспрессируется с промотора pol-III для инициации транскрипции. Таким образом, для SEQ ID NO:38 последовательность будет представлять собой GACGCUCAAAUUUCCGCAGU (SEQ ID NO:116), а не ACGCUCAAUUUCCGCAGU (SEQ ID NO:38). В вариантах осуществления каждая 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 содержит G в качестве первого нуклеотида.

[0058] Как правило, сопряженная последовательность tracr включает любую последовательность, которая имеет достаточную комплементарность с последовательностью tracrRNA для стимуляции одного или нескольких из: (1) вырезание направляющей последовательности, фланкированной парными последовательностями tracr, в клетке, содержащей соответствующую последовательность tracr; и (2) образование комплекса (например, комплекса CRISPR) с последовательностью-мишенью, где комплекс (например, комплекс CRISPR) содержит сопряженную последовательность tracr, гибризованную с последовательностью tracr. В общем, степень комплементарности относится к оптимальному выравниванию сопряженной последовательности tracr и последовательности tracrRNA по длине более короткой из двух последовательностей. Оптимальное выравнивание может быть определено с помощью любого подходящего алгоритма выравнивания и может дополнительно учитывать вторичные структуры, такие как самокомплементарность либо в последовательности tracrRNA, либо в сопряженной последовательности tracr. В аспектах степень комплементарности между последовательностью tracrRNA и сопряженной последовательностью tracr по длине более короткой из двух при оптимальном выравнивании составляет приблизительно или более приблизительно 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99% или выше. В аспектах степень комплементарности составляет приблизительно или по меньшей мере приблизительно 80%, 90%, 95% или 100%. В аспектах длина последовательности tracrRNA составляет приблизительно или более приблизительно 5, 10, 15, 20, 30, 40, 50 или больше нуклеотидов. В аспектах последовательность tracrRNA и сопряженная последовательность

tracr содержатся в одном транскрипте, так что гибридизация между ними приводит к образованию транскрипта, имеющего вторичную структуру, такую как шпилька.

[0059] Термин «аминокислота» относится к встречающимся в природе и синтетическим аминокислотам, а также аналогам аминокислот и миметикам аминокислот, которые функционируют аналогично встречающимся в природе аминокислотам. Встречающиеся в природе аминокислоты представляют собой аминокислоты, кодируемые с помощью генетического кода, а также такие аминокислоты, которые позже модифицируются, например, гидроксипролин, γ -карбоксиглутамат и O-фосфосерин. Аналоги аминокислот относятся к соединениям, которые характеризуются той же самой основной химической структурой, что и встречающаяся в природе аминокислота, т.е., α -углеродом, который связан с водородом, карбоксильной группой, аминогруппой и R-группой, например, гомосерин, норлейцин, метионинсульфоксид, метионинметилсульфоний. Такие аналоги имеют модифицированные R-группы (например, норлейцин) или модифицированные пептидные остовы, однако сохраняют ту же самую основную химическую структуру, что и встречающаяся в природе аминокислота. Миметики аминокислот относятся к химическим соединениям, которые имеют структуру, которая отличается от общей химической структуры аминокислоты, однако которые функционируют аналогично встречающейся в природе аминокислоте. Термины «не встречающаяся в природе аминокислота» и «неприродная аминокислота» относятся к аналогам аминокислот, синтетическим аминокислотам и миметикам аминокислот, которые не встречаются в природе.

[0060] Аминокислоты могут называться в данном документе или по их общеизвестным трехбуквенным символам, или по однобуквенным символам, рекомендованным комиссией по биохимической номенклатуре IUPAC-IUB. Аналогичным образом, нуклеотиды могут быть обозначены по их общепринятым однобуквенным кодам.

[0061] Термины «полипептид», «пептид» и «белок» используются в данном документе взаимозаменяемо для обозначения полимера аминокислотных остатков, где полимер может, в некоторых аспектах, быть конъюгирован с фрагментом, который не состоит из аминокислот. Термины применимы к полимерам из аминокислот, в которых один или несколько аминокислотных остатков представляют собой искусственный химический миметик соответствующей встречающейся в природе аминокислоты, а также к полимерам из встречающихся в природе аминокислот и полимерам из не встречающихся в природе аминокислот. «Слитый белок» относится к химерному белку, кодирующему две или более отдельных белковых последовательностей, которые рекомбинантно экспрессируются в виде одного фрагмента.

[0062] «Консервативно модифицированные варианты» относятся как к аминокислотным, так и к нуклеиновым последовательностям. Что касается конкретных последовательностей нуклеиновых кислот, «консервативно модифицированные варианты» относятся к тем нуклеиновым кислотам, которые кодируют идентичные или по сути идентичные аминокислотные последовательности. Из-за вырожденности генетического

кода ряд последовательностей нуклеиновых кислот будет кодировать любой заданный белок. Например, все кодоны GCA, GCC, GCG и GCU кодируют аминокислоту аланин. Таким образом, в каждом положении, в котором аланин определен кодоном, кодон можно изменить на любой из описанных соответствующих кодонов без изменения кодируемого полипептида. Такие варианты нуклеиновой кислоты представляют собой «молчащие варианты», являющиеся одной из разновидностей вариантов с консервативными модификациями. Любая последовательность аминокислот в данном документе, которая кодирует полипептид, также описывает каждую возможную молчащую вариацию нуклеиновой кислоты. Специалисту в данной области техники будет 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 с получением функционально идентичной молекулы. Соответственно, каждая молчащая вариация нуклеиновой кислоты, которая кодирует полипептид, неявно определена в каждой описанной последовательности.

[0063] Что касается аминокислотных последовательностей, специалисту в данной области техники будет понятно, что отдельные замены, делеции или добавления к последовательности нуклеиновой кислоты, пептида, полипептида или белка, которые изменяют, добавляют или удаляют одну аминокислоту или небольшой процент аминокислот в кодируемой последовательности, представляют собой «консервативно модифицированный вариант», если изменение приводит к замене аминокислоты на химически аналогичную аминокислоту. Таблицы консервативных замен, обеспечивающие функционально аналогичные аминокислоты, хорошо известны в данной области техники. Такие консервативно модифицированные варианты являются дополнением 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)).

[0064] «Процент идентичности последовательности» определяют путем сравнения двух оптимально выровненных последовательностей в пределах окна сравнения, при этом для оптимального выравнивания этих двух последовательностей фрагмент полинуклеотидной или полипептидной последовательности в окне сравнения может содержать добавления или делеции (т. е., гэпы) по сравнению с эталонной последовательностью (которая не содержит добавлений или делеций). Процент рассчитывают с помощью определения числа положений, в которых встречается идентичное основание нуклеиновой кислоты и остаток аминокислоты в обеих последовательностях, для получения числа совпадающих положений, деля число совпадающих положений на общее число положений в окне сравнения и умножая результат

на 100 для получения процента идентичности последовательности.

Термины «идентичный» или «процент идентичности» в контексте двух или более последовательностей нуклеиновых кислот или полипептидов относятся к двум или более последовательностям или подпоследовательностям, которые являются одинаковыми или характеризуются определенным процентом аминокислотных остатков или нуклеотидов, которые являются одинаковыми (т.е. Приблизительно 60% идентичность, предпочтительно 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более высокая идентичность по указанной области при сравнении и выравнивании для максимального соответствия по окну сравнения или обозначенной области) при измерении с применением алгоритмов сравнения последовательностей BLAST или BLAST 2.0 с параметрами по умолчанию, описанными ниже, или посредством выравнивания вручную и визуального осмотра (см., например, веб-сайт NCBI по адресу ncbi.nlm.nih.gov/BLAST/ или аналогичный). Такие последовательности затем называют «по сути идентичными». Это определение также может относиться или может быть применимо к комплементу исследуемой последовательности. Указанное определение также включает в себя последовательности, которые имеют делеции и/или добавления, а также последовательности, которые имеют замены. Описанные ниже предпочтительные алгоритмы могут учитывать гэпы и т.п. Предпочтительно идентичность существует в области, которая составляет по меньшей мере приблизительно 25 аминокислот или нуклеотидов в длину, или более предпочтительно в области, которая составляет 50-100 аминокислот или нуклеотидов в длину.

[0065] «Положение» аминокислоты или нуклеотидного основания обозначается числом, которое последовательно идентифицирует каждую аминокислоту (или нуклеотидное основание) в эталонной последовательности на основе ее положения относительно N-конца (или 5'-конца). Из-за делеций, вставок, усечений, слияний и т.п., которые необходимо учитывать при определении оптимального выравнивания, обычно количество аминокислотных остатков в тестируемой последовательности, определяемое простым подсчетом от N-конца, не обязательно будет совпадает с номером соответствующего положения в эталонной последовательности. Например, в случае, когда вариант имеет делецию относительно выровненной эталонной последовательности, в варианте будет отсутствовать аминокислота, которая соответствует положению в эталонной последовательности в сайте делеции. При наличии вставки в выровненной эталонной последовательности эта вставка не будет соответствовать пронумерованному положению аминокислоты в эталонной последовательности. В случае усечения или слияния могут присутствовать участки аминокислот либо в эталонной, либо в выровненной последовательности, которые не соответствуют ни одной аминокислоте в соответствующей последовательности.

[0066] Термины «пронумерованы относительно» или «соответствуют» при использовании в контексте нумерации данной аминокислотной или полинуклеотидной последовательности относятся к нумерации остатков указанной эталонной

последовательности, когда данная аминокислота или полинуклеотид последовательность сравнивается с эталонной последовательностью.

[0067] Для конкретных белков, описанных в данном документе (например, TET1, dCas9), названный белок включает любую из встречающихся в природе форм белка, или его 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 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, части из 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с встречающейся в природе формой. В аспектах белок представляет собой белок, как идентифицировано с помощью его ссылки на последовательность NCBI. В аспектах белок представляет собой белок, идентифицированный с помощью его эталонной последовательности NCBI, или его функциональный фрагмент, или гомолог.

[0068] Термин «ДНК-эндонуклеаза, направляемая РНК» и т.п. относится в обычном и общепринятом смысле к ферменту, который расщепляет фосфодиэфирную связь в полинуклеотидной цепи ДНК, где распознавание фосфодиэфирной связи облегчается отдельной последовательностью РНК (например, одной направляющей РНК).

[0069] Термин «эндонуклеаза CRISPR класса II» относится к эндонуклеазам, которые обладают такой же эндонуклеазной активностью, что и Cas9, и участвуют в системе CRISPR класса II. Примером системы CRISPR класса II является локус CRISPR типа II из *Streptococcus pyogenes* SF370, который содержит кластер из четырех генов Cas9, Cas1, Cas2 и Csn1, а также два некодирующих элемента РНК, tracrRNA и характерный массив повторяющихся последовательностей (прямые повторы), разделенных короткими участками неповторяющихся последовательностей (спейсерами, около 30 п.н. каждый). Фермент Cpf1 принадлежит к предполагаемой системе CRISPR-Cas типа V. Системы типа II и типа V включены в класс II системы CRISPR-Cas.

[0070] «Последовательность ядерной локализации», или «сигнал ядерной локализации», или «NLS» представляет собой пептид, который направляет белки в ядро. В аспектах NLS включает пять основных положительно заряженных аминокислот. NLS может быть расположен в любом месте пептидной цепи. В аспектах NLS представляет собой NLS, происходящую из SV40. В аспектах NLS включает последовательность, представленную SEQ ID NO:4. В аспектах NLS представляет собой последовательность, представленную SEQ ID NO:4. В 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:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности

последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность SEQ ID NO:4.

[0071] Используемый в данном документе термин «клетка» относится к клетке, выполняющей метаболическую или другую функцию, достаточную для сохранения или репликации ее геномной ДНК. Клетка может быть идентифицирована с помощью хорошо известных в данной области техники методов, включая, например, наличие интактной мембраны, окрашивание определенным красителем, способность производить потомство или, в случае гамет, способность объединяться со второй гаметой с получением жизнеспособного потомства. Клетки могут включать прокариотические и эукариотические клетки. Прокариотические клетки включают без ограничения бактерии. Эукариотические клетки включают без ограничения дрожжевые клетки и клетки, полученные из растений и животных, например, клетки млекопитающих, насекомых (например, Spodoptera) и клетки человека. Клетки могут быть пригодны, если они естественным образом не прикрепляются или были обработаны, чтобы не прикрепляться к поверхностям, с помощью трипсинизации.

[0072] Используемый в данном документе термин «вектор» относится к молекуле нуклеиновой кислоты, способной переносить другую нуклеиновую кислоту, с которой она была связана. Одним типом вектора является «плаزمид», которая относится к линейной или кольцевой двухцепочечной петле ДНК, в которую могут быть лигированы дополнительные сегменты ДНК. Другим типом вектора является вирусный вектор, при этом в вирусный геном могут быть лигированы дополнительные сегменты ДНК. Определенные векторы способны к автономной репликации в клетке-хозяине, в которую их вводят (например, бактериальные векторы, имеющие бактериальную точку начала репликации, и эписомные векторы млекопитающих). Другие векторы (например, неэписомные векторы млекопитающих) могут быть интегрированы в геном клетки-хозяина при введении в клетку-хозяина, и, тем самым, реплицироваться вместе с геномом хозяина. Более того, определенные векторы способны направлять экспрессию генов, с которыми они функционально связаны. Такие векторы в данном документе обозначены как «векторы экспрессии». В целом, векторы экспрессии, используемые в методиках рекомбинантной ДНК, часто находятся в форме плазмид. В настоящем описании термины «плаزمид» и «вектор» можно использовать взаимозаменяемо, поскольку плазмид является наиболее распространенной формой вектора. Однако, предполагается, что настоящее изобретение включает и другие такие формы векторов экспрессии, такие как вирусные векторы (например, ретровирусы, дефектные по репликации, аденовирусы и аденоассоциированные вирусы), выполняющие эквивалентные функции. Кроме того, некоторые вирусные векторы

способны нацеливаться на определенный тип клеток либо специфически, либо неспецифически. Некомпетентные к репликации вирусные векторы или вирусные векторы с дефектом репликации относятся к вирусным векторам, которые способны инфицировать свои клетки-мишени и доставлять их вирусную нагрузку, но затем не могут продолжить типичный литический путь, который приводит к лизису и гибели клеток.

[0073] Термины «трансфекция», «трансдукция», «трансфекция» или «трансдукция» могут использоваться взаимозаменяемо и определяются как процесс введения молекулы нуклеиновой кислоты и/или белка в клетку. Нуклеиновые кислоты могут быть введены в клетку с использованием невирусных или вирусных методов. Молекула нуклеиновой кислоты может представлять собой последовательность, кодирующую полные белки или их функциональные части. Как правило, вектор нуклеиновой кислоты содержит элементы, необходимые для экспрессии белка (например, промотор, сайт начала транскрипции и т.д.). Невирусные способы трансфекции включают любой подходящий способ, в котором не используется вирусная ДНК или вирусные частицы в качестве системы доставки для введения молекулы нуклеиновой кислоты в клетку. Иллюстративные способы невирусной 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 посредством теплового шока, магнетифекцию и электропорацию. Для методов, основанных на использовании вирусов, в способах, описанных в данном документе, можно использовать любой пригодный вирусный вектор. Примеры вирусных векторов включают без ограничения ретровирусные, аденовирусные, лентивирусные и аденоассоциированные вирусные векторы. В аспектах молекулы нуклеиновой кислоты вводят в клетку с использованием ретровирусного вектора в соответствии со стандартными процедурами, хорошо известными в данной области техники. Термины «трансфекция» или «трансдукция» также относятся к введению белков в клетку из внешней среды. Как правило, трансдукция или трансфекция белка зависит от присоединения 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.

[0074] «Пептидный линкер», представленный в данном документе, представляет собой линкер, содержащий пептидный фрагмент. В вариантах осуществления пептидный линкер представляет собой двухвалентный пептид, такой как аминокислотная последовательность, присоединенная на N-конце и C-конце к остальной части соединения, например, слитому белку, представленному в данном документе. Пептидный линкер может представлять собой пептидный фрагмент (двухвалентный пептидный фрагмент), способный к расщеплению (например, расщепляемый полипептид P2A). Пептидный линкер, представленный в данном документе, также может взаимозаменяемо называться аминокислотным линкером. В аспектах пептидный линкер содержит от 1 до приблизительно 80 аминокислотных остатков. В аспектах пептидный линкер содержит от

аминокислотные остатки. Разработку и использование XTEN можно найти, например, в Schellenberger et al., *Nature Biotechnology* 27, 1186-1190 (2009). В аспектах линкер XTEN содержит последовательность, представленную SEQ ID NO: 5, 6 или 98.

[0076] «Метка эпитопа» относится к биологическому фрагменту, такому как пептид, который генетически сконструирован в виде рекомбинантного белка и который функционирует как универсальный эпитоп, который легко выявляется с помощью имеющихся в продаже анализов или антител и который обычно не нарушает нативную структуру или функцию белка.

[0077] «Выявляемое средство» или «выявляемый фрагмент» представляет собой композицию, выявляемую соответствующими средствами, такими как спектроскопические, фотохимические, биохимические, иммунохимические, химические, магнитно-резонансные томографические или другие физические средства. Например, пригодные обнаруживаемые средства 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 , $^{154-158}\text{Gd}$, ^{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, обычно используемые в ELISA), биотин, дигоксигенин, парамагнитные молекулы, парамагнитные наночастицы, сверхмалые суперпарамагнитные наночастицы оксида железа («USPIO»), агрегаты наночастиц USPIO, суперпарамагнитные наночастицы оксида железа («SPIO»), агрегаты наночастиц SPIO, наночастицы монокристаллического оксида железа, монокристаллический оксид железа, контрастные вещества в виде наночастиц, липосомы или другие носители доставки, содержащие молекулы хелата гадолиния («хелат Gd»), гадолиний, радиоизотопы, радионуклиды (например, углерод-11, азот-13, кислород-15, фтор-18, рубидий-82), фтордезоксиглюкозу (например, меченную фтором-18), любые радионуклиды, излучающие гамма-лучи, радионуклиды, излучающие позитроны, радиоактивно меченую глюкозу, радиоактивно меченую воду, радиоактивно меченый аммиак, биокolloиды, микропузырьки (например, включая оболочки микропузырьков, включая альбумин, галактозу, липид и/или полимеры; газовое ядро микропузырьков, включая воздух, тяжелые газы, перфторуглерод, азот, октафторпропан, перфлексановые липидные микросферы, перфлутрен и т.д.), йодсодержащие контрастные вещества (например, йогексол, йодиксанол, йоверсол, йопамидол, йоксилан, йопромид, диатризоат, метризоат, иоксагат), сульфат бария, диоксид тория, золото, наночастицы золота, агрегаты наночастиц золота, флуорофоры, двухфотонные флуорофоры или гаптены и белки или другие объекты, которые можно обнаружить, например, путем включения радиоактивной метки в пептид или антитело, специфически реагирующее с пептидом-мишенью.

[0078] Поддающийся обнаружению фрагмент представляет собой одновалентное поддающееся обнаружению средство или поддающееся обнаружению средство, способное образовывать связь с другой композицией. В аспектах подлежащее обнаружению средство

представляет собой метку эпитопа. В аспектах метка эпитопа представляет собой метку HA. В аспектах метка HA содержит последовательность, представленную SEQ ID NO:7. В аспектах метка HA представляет собой последовательность, представленную SEQ ID NO:7. В аспектах метка HA имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:7. В аспектах метка HA имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:7. В аспектах метка HA имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:7. В аспектах метка HA имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:7.

[0079] В аспектах подлежащее обнаружению средство представляет собой флуоресцентный белок. В аспектах флуоресцентный белок представляет собой синий флуоресцентный белок (BFP). В аспектах BFP включает последовательность, представленную SEQ ID NO:8. В аспектах BFP представляет собой последовательность, представленную SEQ ID NO:8. В аспектах BFP имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:8. В аспектах BFP имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:8. В аспектах BFP имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:8. В аспектах BFP имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:8.

[0080] Радиоактивные вещества (например, радиоизотопы), которые можно использовать в качестве средств визуализации и/или маркировки в соответствии с аспектами настоящего изобретения, включают без ограничения ^{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 и ^{225}Ac . Парамагнитные ионы, которые можно использовать в качестве дополнительных визуализирующих средств в соответствии с аспектами настоящего изобретения, включают без ограничения ионы переходных и лантаноидных металлов (например, металлов с атомными номерами 21-29, 42, 43, 44, или 57-71). К таким металлам относятся ионы Cr, V, Mn, Fe, Co, Ni, Cu, La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb и Lu.

[0081] «Приведение в контакт» используется в соответствии с его простым обычным значением и относится к процессу, позволяющему по меньшей мере двум различным соединениям расположиться достаточно близко, чтобы реагировать, взаимодействовать или физически касаться друг друга. Однако следует понимать, что конечный продукт реакции может быть получен непосредственно в результате реакции между добавленными реагентами или из промежуточного продукта из одного или нескольких добавленных

реагентов, которые могут быть получены в реакционной смеси.

[0082] Термин «приведение в контакт» может включать в себя возможность реакции, взаимодействия или физического соприкосновения двух соединений, при этом два соединения могут представлять собой, например, слитый белок, представленный в данном документе, и последовательность нуклеиновой кислоты (например, последовательность ДНК-мишени).

[0083] Как определено в данном документе, термин «активация», «активировать», «осуществление активации», «усиливать», «реактивация», «реактивировать», «осуществление реактивации» и т.п. при использовании в отношении композиции, представленной в данном документе (например, слитого белка, комплекса, нуклеиновой кислоты, вектора) относятся к положительному влиянию (например, повышению) активности (например, транскрипции) последовательности нуклеиновой кислоты (например, повышению транскрипции гена) по сравнению с активностью последовательности нуклеиновой кислоты (например, транскрипции гена) в отсутствие композиции (например, слитого белка, комплекса, нуклеиновой кислоты, вектора). Таким образом, активация или реактивация включает, по меньшей мере, частичное повышение или активацию (например, транскрипцию) экспрессии или предупреждение или обращение вспять снижения или задержки экспрессии (например, транскрипции) последовательности нуклеиновой кислоты. Активированная или реактивированная активность (например, транскрипция) может составлять 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% или более, чем в контроле. В аспектах активация или реактивация является 1,5-кратной, 2-кратной, 3-кратной, 4-кратной, 5-кратной, 10-кратной или больше по сравнению с контролем. В вариантах осуществления активация может представлять собой активацию гена, который ранее был подвергнут сайленсингу. В вариантах осуществления реактивация может представлять собой реактивацию гена, который ранее был подвергнут сайленсингу.

[0084] Используемый в данном документе термин «энхансер» или «активатор» относится к области ДНК, которая может быть связана с белками (например, активаторами транскрипции) и/или полинуклеотидами для повышения вероятности того, что произойдет транскрипция гена. Энхансеры могут иметь длину от примерно 50 до приблизительно 35000 пар оснований. В вариантах осуществления энхансеры могут иметь длину от приблизительно 50 до приблизительно 1500 пар оснований. Энхансеры могут быть расположены ниже или выше сайта инициации транскрипции, который они регулируют, и могут быть удалены от сайта инициации транскрипции на сотни или по меньшей мере миллион пар оснований. В вариантах осуществления энхансер может находиться в нескольких сотнях пар оснований от сайта инициации транскрипции. В вариантах осуществления энхансер может быть связан по меньшей мере с одним активатором транскрипции (например, VP64, p65, Rta). В вариантах осуществления энхансер может представлять собой полинуклеотидную последовательность-мишень, подходящую для редактирования эпигенома. В вариантах осуществления энхансер может быть нацелен на один или несколько белков и/или полинуклеотидов, которые могут активировать или

реактивировать транскрипцию гена.

[0085] Как определено в данном документе, термины «ингибирование», «ингибировать», «осуществление ингибирования», «репрессия», «осуществление репрессии», «сайленсинг», «подвергать сайленсингу» и т.п. при использовании в отношении композиции, представленной в данном документе (например, слитого белка, комплекса, нуклеиновой кислоты, вектора) относятся к отрицательному влиянию (например, снижению) активности (например, транскрипции) последовательности нуклеиновой кислоты (например, снижению транскрипции гена) по сравнению с активностью последовательности нуклеиновой кислоты (например, транскрипции гена) в отсутствие композиции (например, слитого белка, комплекса, нуклеиновой кислоты, вектора). В аспектах ингибирование относится к уменьшению заболевания или симптомов заболевания (например, рака). Таким образом, ингибирование включает, по меньшей мере, отчасти, частично или полное блокирование активации (например, транскрипции) или уменьшение, предупреждение или задержку активации (например, транскрипции) последовательности нуклеиновой кислоты. Ингибированная активность (например, транскрипция) может составлять 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% или меньше от активности в контроле. В аспектах ингибирование является 1,5-кратным, 2-кратным, 3-кратным, 4-кратным, 5-кратным, 10-кратным или больше по сравнению с контролем.

[0086] Термин «сайленсер», используемый в данном документе, относится к последовательности ДНК, способной связывать факторы регуляции транскрипции, известные как репрессоры, за счет чего происходит отрицательное влияние на транскрипцию гена. Последовательности ДНК-сайленсера могут быть обнаружены во многих различных положениях по всей ДНК, включая без ограничения расположенные выше от гена-мишени, для которого он действует, репрессируя транскрипцию гена (например, осуществлять сайленсинг экспрессию гена).

[0087] «Контрольный» образец или значение относится к образцу, который служит эталоном, обычно известным эталоном, для сравнения с тестируемым образцом. Например, тестируемый образец может быть взят в тестируемых условиях, например, в присутствии тестируемого соединения, и сравнен с образцами в известных условиях, например, в отсутствие тестируемого соединения (отрицательный контроль) или в присутствии известного соединения (положительный контроль). Контроль также может представлять собой среднее значение, полученное из ряда тестов или результатов. Специалист в данной области техники поймет, что контроли могут быть разработаны для оценки любого количества параметров. Например, контроль может быть разработан для сравнения терапевтического эффекта на основе фармакологических данных (например, периода полужизни) или терапевтических мер (например, сравнения побочных эффектов). Специалист в данной области техники поймет, какие контроли являются ценными в данной ситуации, и сможет анализировать данные на основе сравнений с контрольными значениями. Контроли также пригодны для определения значимости данных. Например,

если значения данного параметра широко различаются в контроле, вариации в тестируемых образцах не будут считаться значимыми.

[0088] Термин «домен деметилирования» относится к части белковой последовательности или структуры, способной к деметилированию ДНК. Например, домен деметилирования может удалять метильную группу из азотистого основания (т.е. превращение 5-метилцитозина в цитозин). В вариантах осуществления домены деметилирования включают ферменты транслокации Ten-Eleven (TET) или функциональные домены ферментов TET. В вариантах осуществления домен деметилирования представляет собой бактериальную ДНК-деметилазу.

[0089] Термин «транслокация Ten-Eleven» или «TET» относится к семейству ферментов, включающему TET1, TET2 и TET3. Не желая быть связанными какой-либо теорией, ферменты TET могут удалять репрессивные метки 5mC и/или катализировать окисление метильной группы 5-метилцитозина (5mC) с образованием 5-гидроксиметилцитозина (5hmC) и других окисленных метилцитозинов, облегчая деметилирование.

[0090] Термин «TET1» или «белок TET1», как представлено в данном документе, включает любую из рекомбинантных или встречающихся в природе форм транслокационной метилцитозиндиоксигеназы 1 Ten-eleven (TET1), также известной как **метилцитозиндиоксигеназа TET1**, белок 6 цинковых пальцев CXXC-типа, ассоциированный с лейкозом белок с доменом CXXC или его варианты или гомологи, которые сохраняют активность белка TET1 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком TET1). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка TET1. В вариантах осуществления белок TET1 представляет собой белок, идентифицированный с помощью ссылочного номера UniProt Q8NFU7, или его вариант, гомолог или функциональный фрагмент. В аспектах TET1 содержит аминокислотную последовательность SEQ ID NO:1. В аспектах TET1 имеет аминокислотную последовательность SEQ ID NO:1. В аспектах TET1 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:1. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:1. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:1. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:1. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности

последовательности с SEQ ID NO:1. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:1. В аспектах TET1 содержит аминокислотную последовательность SEQ ID NO:86. В аспектах TET1 имеет аминокислотную последовательность SEQ ID NO:86. В аспектах TET1 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:86. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:86. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:86. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:86. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:86. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:86. В аспектах TET1 содержит аминокислотную последовательность SEQ ID NO:97. В аспектах TET1 имеет аминокислотную последовательность SEQ ID NO:97. В аспектах TET1 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:97. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:97. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:97. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:97. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:97. В аспектах TET1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:97.

[0091] Термин «TET2» или «белок TET2», как представлено в данном документе, включает любую из рекомбинантных или встречающихся в природе форм транслокационной метилцитозиндиоксигеназы 2 Ten-eleven (TET2), также известной как **метилцитозиндиоксигеназа TET2**, или его варианты или гомологи, которые сохраняют активность белка TET2 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком TET2). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка TET2. В вариантах осуществления белок TET2 представляет собой белок, идентифицированный с помощью ссылочного номера UniProt Q6N021, или его вариант, гомолог или функциональный фрагмент. В аспектах TET2 содержит аминокислотную последовательность SEQ ID NO:2. В аспектах TET2 имеет аминокислотную последовательность SEQ ID NO:2. В аспектах TET2 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:2. В аспектах TET2 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:2. В аспектах TET2 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:2. В аспектах TET2 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:2. В аспектах TET2 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:2. В аспектах TET2 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:2.

[0092] Термин «TET3» или «белок TET3», как представлено в данном документе, включает любую из рекомбинантных или встречающихся в природе форм транслокационной метилцитозиндиоксигеназы 3 Ten-eleven (TET3), также известной как **метилцитозиндиоксигеназа TET3**, или его варианты или гомологи, которые сохраняют активность белка TET3 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком TET3). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка TET3. В вариантах осуществления белок TET3 представляет собой белок, идентифицированный с помощью ссылочного номера UniProt O43151, или его вариант, гомолог или функциональный фрагмент. В аспектах TET3 содержит аминокислотную последовательность SEQ ID NO:3. В аспектах TET3 имеет аминокислотную последовательность SEQ ID NO:3. В аспектах TET3 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:3. В аспектах TET3 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:3. В аспектах TET3 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности

последовательности с SEQ ID NO:3. В аспектах TET3 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:3. В аспектах TET3 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:3. В аспектах TET3 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:3.

[0093] Термин «активатор транскрипции», «активатор» и т.п. относится в обычном и общепринятом смысле к белку (т.е. фактору транскрипции), который повышает транскрипцию гена или набора генов. Например, активаторы транскрипции могут представлять собой ДНК-связывающие белки, которые связываются с энхансерами или проксимальными элементами промотора. В вариантах осуществления активатор транскрипции представляет собой VP64, p65 или Rta. В вариантах осуществления активатор транскрипции может усиливать транскрипцию гена или набора генов, которые ранее были подвергнуты сайленсингу. Активаторы транскрипции и их применение можно найти, например, в Tanenbaum et al., A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell. 2014 Oct 23;159(3):635-46, и Zalatan et al., Engineering Complex Synthetic Transcriptional Programs With CRISPR RNA Scaffolds. Cell. 2015 Jan 15;160(1-2):339-50, которые включены в данный документ посредством ссылки во всей своей полноте и для всех целей.

[0094] Термин «p65» или «белок p65», как представлено в данном документе, включает любую из рекомбинантных или встречающихся в природе форм фактора транскрипции p65 (p65), также известного как субъединица p65 ядерного фактора NF-каппа-B, или его варианты или гомологи, которые поддерживают активность белка p65 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком p65). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка p65. В вариантах осуществления белок p65 представляет собой белок, идентифицированный с помощью ссылочного номера UniProt Q04206, или его вариант, гомолог или функциональный фрагмент. В аспектах p65 содержит аминокислотную последовательность SEQ ID NO:13. В аспектах p65 имеет аминокислотную последовательность SEQ ID NO:13. В аспектах p65 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:13. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:13. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности

последовательности с SEQ ID NO:13. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:13. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:13. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:13. В аспектах p65 содержит аминокислотную последовательность SEQ ID NO:14. В аспектах p65 имеет аминокислотную последовательность SEQ ID NO:14. В аспектах p65 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:14. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:14. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:14. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:14. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:14. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:14. В аспектах p65 содержит аминокислотную последовательность SEQ ID NO:100. В аспектах p65 имеет аминокислотную последовательность SEQ ID NO:100. В аспектах p65 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:100. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:100. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:100. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:100. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:100. В аспектах p65 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:100.

[0095] Термин «Rta» или «белок Rta», как представлено в данном документе, включает любую из рекомбинантных или встречающихся в природе форм **активатора репликации и транскрипции (Rta)**, также известного как R-трансактиватор,

немедленно-ранний белок Rta, или его варианты или гомологи, которые поддерживать активность белка Rta (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком Rta). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка Rta. В вариантах осуществления белок Rta представляет собой белок, идентифицированный с помощью ссылочного номера UniProt P03209, или его вариант, гомолог или функциональный фрагмент. В аспектах Rta содержит аминокислотную последовательность SEQ ID NO:15. В аспектах Rta имеет аминокислотную последовательность SEQ ID NO:15. В аспектах Rta 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:15. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:15. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:15. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:15. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:15. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:15. В аспектах Rta содержит аминокислотную последовательность SEQ ID NO:16. В аспектах Rta имеет аминокислотную последовательность SEQ ID NO:16. В аспектах Rta 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:16. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:16. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:16. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:16. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:16. В аспектах Rta имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:16.

[0096] Термин «VP64» или «белок VP64», как представлено в данном документе,

включает любую из рекомбинантных или встречающихся в природе форм белка тегумента VP16 (VP64), также известного как альфа-транс-индуцирующий белок, Alpha-TIF, или его варианты или гомологи, которые поддерживают активность белка VP64 (например, в пределах по меньшей мере 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком VP64). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка VP64. В вариантах осуществления белок VP64 представляет собой белок, идентифицированный с помощью ссылочного номера UniProt P06492, или его вариант, гомолог или функциональный фрагмент. В аспектах VP64 содержит аминокислотную последовательность SEQ ID NO:17. В аспектах VP64 имеет аминокислотную последовательность SEQ ID NO:17. В аспектах VP64 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:17. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:17. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:17. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:17. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:17. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:17. В аспектах VP64 содержит аминокислотную последовательность SEQ ID NO:18. В аспектах VP64 имеет аминокислотную последовательность SEQ ID NO:18. В аспектах VP64 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:18. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:18. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:18. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:18. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:18. В аспектах VP64 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности

последовательности с SEQ ID NO:18.

[0097] Термин «МСР» или «белок МСР», как представлено в данном документе, включает любые рекомбинантные или встречающиеся в природе формы капсидного белка (МСР), также известные как СР, белок оболочки, или их варианты или гомологи, которые сохраняют активность белка МСР (например, по меньшей мере в пределах 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% или 100% активности по сравнению с белком МСР). В 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным полипептидом белка МСР. В вариантах осуществления белок МСР представляет собой белок, идентифицированный с помощью ссылочного номера UniProt P03612, или его вариант, гомолог или функциональный фрагмент. В аспектах МСР содержит аминокислотную последовательность SEQ ID NO:21. В аспектах МСР имеет аминокислотную последовательность SEQ ID NO:21. В аспектах МСР 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:21. В аспектах МСР имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:21. В аспектах МСР имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:21. В аспектах МСР имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:21. В аспектах МСР имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:21. В аспектах МСР имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:21.

[0098] Термин «ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК» и т.п. относится в обычном и общепринятом смысле к ДНК-эндонуклеазе, направляемой РНК (например, мутантной форме встречающейся в природе ДНК-эндонуклеазы, направляемой РНК), которая нацелена на специфическую фосфодиэфирную связь в полинуклеотиде ДНК, где распознавание фосфодиэфирной связи облегчается отдельной полинуклеотидной последовательностью (например, последовательностью РНК (например, одиночной направляющей РНК (sgRNA)), но неспособной расщепить фосфодиэфирную связь-мишень до значительной степени (например, в физиологических условиях не происходит измеримого расщепления фосфодиэфирной связи). Таким образом, ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, сохраняет способность связываться с ДНК (например, специфическое связывание с последовательностью-мишенью) при образовании комплекса с полинуклеотидом (например, sgRNA), но не обладает значительной эндонуклеазной активностью (например, любое количество обнаруживаемой

эндонуклеазной активности). В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой белок, ассоциированный с CRISPR. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas9, dCas12a, dCpf1, ddCpf1, Cas-phi, вариант Cas9 с дефицитом нуклеазы, эндонуклеазу CRISPR класса II с дефицитом нуклеазы, домен лейциновой молнии, домен «крылатой спирали», мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокс, домен Wog3, домен ОВ-складка, домен иммуноглобулина или домен В3. В аспектах фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен лейциновой молнии, домен «крылатой спирали», мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокс, домен Wog3, домен ОВ-складка, домен иммуноглобулина или домен В3. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен лейциновой молнии. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен «крылатой спирали». В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой мотив спираль-поворот-спираль. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен спираль-петля-спираль. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен НМВ-бокс. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен Wog3. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен ОВ-складка. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен иммуноглобулина. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой домен В3. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, ddCpf1, Cas-phi, вариант Cas9 с дефицитом нуклеазы или эндонуклеазу CRISPR класса II с дефицитом нуклеазы. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas9. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas9 из *S. pyogenes*. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9 из *S. aureus*. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas12a. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas12a из *Lachnospiraceae bacterium*. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas12. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент ddCas12a. В аспектах ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент Cas-phi.

[0099] Термин «белок, ассоциированный с CRISPR» или «белок CRISPR» относится к любому белку CRISPR, который функционирует как фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, т.е. белок CRISPR, в котором каталитические сайты эндонуклеазной активности дефектны или лишены активности. Иллюстративные

белки CRISPR включают dCas9, dCpf1, ddCpf1, dCas12, ddCas12, dCas12a Cas-phi, вариант Cas9 с дефицитом нуклеазы, эндонуклеазу CRISPR класса II с дефицитом нуклеазы и т.п.

[0100] Термин «ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК» относится к ДНК-эндонуклеазе (например, мутантной форме встречающейся в природе ДНК-эндонуклеазы), которая нацелена на конкретную фосфодиэфирную связь в полинуклеотиде ДНК, но не требует направляющей РНК. В вариантах осуществления «ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК» представляет собой домен цинковых пальцев или эффектор, подобный активатору транскрипции (TALE).

[0101] В вариантах осуществления фермент ДНК-эндонуклеаза с дефицитом нуклеазы представляет собой «домен цинковых пальцев». Термины «домен цинковых пальцев» или «связывающий домен цинковых пальцев» или «связывающий домен ДНК цинковых пальцев» используются взаимозаменяемо и относятся к белку или домену в более крупном белке, который связывает ДНК специфичным для последовательности образом посредством одного или несколько цинковых пальцев, которые представляют собой области аминокислотной последовательности внутри связывающего домена, структура которых стабилизируется за счет координации иона цинка. В вариантах осуществления домен цинковых пальцев не встречается в природе, поскольку он сконструирован так, чтобы связываться с выбранным сайтом-мишенью. В аспектах домен, связывающий цинковые пальцы, относится к белку, домену в более крупном белке или ферменту ДНК-эндонуклеазе с дефицитом нуклеазы, направляемой РНК, который способен связываться с любыми белком цинковых пальцев, известным в данной области техники, таким как типа C2H2, типа CCHC, типа PHD или типа RING цинковыми пальцами.

[0102] Используемый в данном документе термин «цинковый палец» представляет собой структурный мотив полипептида, свернутый вокруг связанного катиона цинка. В вариантах осуществления полипептид цинкового пальца имеет последовательность в форме X_3 -Cys- X_{2-4} -Cys- X_{12} -His- X_{3-5} -His- X_4 , где X представляет собой любую аминокислоту (например, X_{2-4} означает олигопептид длиной 2-4 аминокислоты). Как правило, существует широкий диапазон вариаций последовательностей 28-31 аминокислот известных полипептидов цинковых пальцев. Только два консенсусных остатка гистидина и два консенсусных остатка цистеина, связанные с центральным атомом цинка, являются инвариантными. Из оставшихся остатков от трех до пяти являются высококонсервативными, в то время как среди других остатков могут быть значительные различия. Несмотря на широкий диапазон вариаций последовательности полипептида, цинковые пальцы этого типа имеют сходную трехмерную структуру. Однако существует широкий диапазон специфичностей связывания различных цинковых пальцев, т.е. разные цинковые пальцы связывают двухцепочечные полинуклеотиды, имеющие широкий диапазон последовательностей нуклеотидов. В некоторых аспектах цинковый палец относится к типу C2H2. В аспектах цинковый палец относится к типу CCHC. В аспектах цинковый палец относится к типу PHD. В аспектах цинковый палец относится к типу RING.

[0103] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы,

направляемая РНК, представляет собой фермент TALE. «TALE» или «подобный активатору транскрипции эффектор» относится к искусственным ферментам рестрикции, полученным путем слияния эффекторного ДНК-связывающего домена 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 активаторам транскрипции (TALE), могут быть быстро сконструированы для связывания практически любой последовательности ДНК. Используемый в данном документе термин TALE является широким 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 для расщепления ДНК в одном и том же сайте. TALE, которые работают вместе, могут называться левым TALE и правым TALE, что указывает на хиральность ДНК. TALE представляет собой белки, секретируемые бактериями *Xanthomonas*. ДНК-связывающий домен содержит высококонсервативную последовательность из 33-34 аминокислот, за исключением аминокислот 12 и 13. Эти два положения очень переменны (повторяющийся переменный остаток (RVD)) и демонстрируют сильную корреляцию с распознаванием специфических нуклеотидов. Эта простая взаимосвязь между аминокислотной последовательностью и распознаванием ДНК позволила разработать специфические ДНК-связывающие домены путем выбора комбинации повторяющихся сегментов, содержащих соответствующие RVD.

[0104] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas9. Термины «dCas9» или «белок dCas9», используемые в данном документе, представляют собой белок Cas9, в котором оба каталитических сайта для эндонуклеазной активности дефектны или лишены активности. В аспектах белок dCas9 имеет мутации в положениях, соответствующих D10A и H840A Cas9 *S. pyogenes*. В некоторых аспектах белку dCas9 не обладает эндонуклеазной активностью из-за точечных мутаций в обоих эндонуклеазных каталитических сайтах (RuvC и HNH) Cas9 дикого типа. Точечные мутации могут представлять собой D10A и H840A. В некоторых аспектах dCas9 практически не обладает обнаруживаемой эндонуклеазной (например, эндодезоксирибонуклеазной) активностью. В аспектах dCas9 содержит аминокислотную последовательность SEQ ID NO:9. В аспектах dCas9 имеет аминокислотную последовательность SEQ ID NO:9. В аспектах dCas9 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:9. В аспектах dCas9 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:9. В аспектах dCas9 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:9. В аспектах dCas9 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности

последовательности с SEQ ID NO:9. В аспектах dCas9 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:9. В аспектах dCas9 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:9.

[0105] «Ассоциированный с CRISPR белок 9», «Cas9», «Csn1» или «белок Cas9», как упоминается 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 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным белком Cas9. В аспектах белок Cas9 по сути идентичен белку, идентифицированному UniProt с помощью ссылочного номера Q99ZW2, или варианту или гомологу, имеющему по сути идентичность с ним. В аспектах белок Cas9 имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt Q99ZW2. В аспектах белок Cas9 имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt Q99ZW2. В аспектах белок Cas9 имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt Q99ZW2. В аспектах белок Cas9 имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt Q99ZW2. В аспектах белок Cas9 имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt Q99ZW2.

[0106] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент «ddCpf1» или «ddCas12a». Термины «ДНКазы-dead Cpf1» или «ddCpf1» относятся к мутировавшим Cpf1 (*AsCpf1*) *Acidaminococcus sp.*, что приводит к инактивации активности ДНКазы Cpf1. В аспектах ddCpf1 включает мутацию E993A в домене RuvC *AsCpf1*. В некоторых аспектах ddCpf1 практически не обладает обнаруживаемой эндонуклеазной (например, эндодезоксирибонуклеазной) активностью. В аспектах ddCpf1 содержит аминокислотную последовательность SEQ ID NO:10. В аспектах ddCpf1 имеет аминокислотную последовательность SEQ ID NO:10. В аспектах ddCpf1 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:10. В аспектах ddCpf1 имеет аминокислотную

последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:10. В аспектах ddCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:10. В аспектах ddCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:10. В аспектах ddCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:10. В аспектах ddCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:10.

[0107] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dLbCpf1. Термин «dLbCpf1» относится к мутантному Cpf1 из ND2006 (LbCpf1) бактерии *Lachnospiraceae*, у которого отсутствует активность ДНКазы. В аспектах dLbCpf1 включает мутацию D832A. В аспектах dLbCpf1 по сути не обладает обнаруживаемой эндонуклеазной (например, эндодезоксирибонуклеазной) активностью. В аспектах dLbCpf1 содержит аминокислотную последовательность SEQ ID NO:11. В аспектах dLbCpf1 имеет аминокислотную последовательность SEQ ID NO:11. В аспектах dLbCpf1 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:11. В аспектах dLbCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:11. В аспектах dLbCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:11. В аспектах dLbCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:11. В аспектах dLbCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:11. В аспектах dLbCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:11.

[0108] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dFnCpf1. Термин «dFnCpf1» относится к мутантному Cpf1 из U112 (FnCpf1) *Francisella novicida*, у которого отсутствует активность ДНКазы. В аспектах dFnCpf1 включает мутацию D917A. В аспектах dFnCpf1 по сути не обладает обнаруживаемой эндонуклеазной (например, эндодезоксирибонуклеазной) активностью. В аспектах dFnCpf1 содержит аминокислотную последовательность SEQ ID NO: 12. В аспектах dFnCpf1 имеет аминокислотную последовательность SEQ ID NO:12. В аспектах dFnCpf1 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:12. В аспектах dFnCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:12. В аспектах dFnCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:12. В аспектах dFnCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:12. В аспектах dFnCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:12. В аспектах dFnCpf1 имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:12.

[0109] «Cpf1» или «белок Cpf1», как упоминается в данном документе, включает любую из рекомбинантных или встречающихся в природе форм эндонуклеазы Cpf1 (CRISPR из *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 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, 50, 100, 150 или 200 непрерывных аминокислот) по сравнению с природным белком Cpf1. В аспектах белок Cpf1 по сути идентичен белку, идентифицированному UniProt с помощью ссылочного номера U2UMQ6, или варианту или гомологу, имеющему по сути идентичность с ним. В аспектах белок Cpf1 идентичен белку, идентифицированному UniProt с помощью ссылочного номера U2UMQ6. В аспектах белок Cpf1 имеет по меньшей мере 75% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt U2UMQ6. В аспектах белок Cpf1 имеет по меньшей мере 80% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt U2UMQ6. В аспектах белок Cpf1 идентичен белку, идентифицированному UniProt с помощью ссылочного номера U2UMQ6. В аспектах белок Cpf1 имеет по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt U2UMQ6. В аспектах белок Cpf1 идентичен белку, идентифицированному UniProt с помощью ссылочного номера U2UMQ6. В аспектах белок Cpf1 имеет по меньшей мере 90% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt U2UMQ6. В аспектах белок Cpf1 идентичен белку, идентифицированному UniProt с помощью ссылочного номера U2UMQ6. В аспектах белок Cpf1 имеет по меньшей мере 95% идентичности последовательности с аминокислотной последовательностью белка, идентифицированного с помощью ссылочного номера UniProt U2UMQ6.

[0110] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы,

направляемая РНК, представляет собой вариант Cas9 с дефицитом нуклеазы. Термин «вариант Cas9 с дефицитом нуклеазы» относится к белку Cas9, имеющему одну или несколько мутаций, которые повышают его специфичность связывания с PAM по сравнению с Cas9 дикого типа, и дополнительно включают мутации, которые делают белок неспособным или имеющим значительно нарушенную эндонуклеазную активность. Не желая быть связанными какой-либо теорией, полагают, что последовательность-мишень должна быть ассоциирована с PAM (мотивом, смежным с протоспейсером); т.е. короткой последовательностью, распознаваемой комплексом CRISPR. Требования к точной последовательности и длине PAM различаются в зависимости от используемого фермента CRISPR, но PAM обычно представляют собой последовательности из 2-5 пар оснований, примыкающие к протоспейсеру (то есть к последовательности-мишени). Специфичность связывания вариантов Cas9 с дефицитом нуклеазы с PAM можно определить любым способом, известным в данной области техники. Описание и использование известных вариантов Cas9 можно найти, 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, которые включены в данный документ посредством ссылки во всей своей полноте и для всех целей. Иллюстративные варианты Cas9 перечислены в таблице 1 ниже.

[0111] Таблица 1

Варианты Cas9	Домены PAM	Литературные источники
Cas9 <i>Streptococcus pyogenes</i> (Sp)	NGG	Hsu et al. 2014 Cell
Cas9 <i>Staphylococcus aureus</i> (Sa)	NNGRRT или NNGRR NNGGGT, NNGAAT, NNGAGT (Zetsche)	Ran et al. 2015 Nature
мутант VQR SpCas9 (D1135V, R1335Q, T1337R)	NGAG>NGAT=NGAA>NGAC NGCG	Kleinstiver et al. 2015 Nature
Мутант VRER SpCas9 (D1135V/G1218R/R1335E/T1337R)	NGCG	Kleinstiver et al. 2015 Nature
D1135E SpCas9	NGG, более высокая точность, меньше разрезания в сайтах NAG и NGA	Kleinstiver et al. 2015 Nature
мутант 1.1 eSpCas9 (K848A/K1003A/R1060A)	NGG	Slaymaker et al. Science 2015
HF1 SpCas9 (Q695A, Q926A, N497A,	NGG	Kleinstiver et al. 2016 Nature

R661A)		
AsCpf1	TTTN (5' sgRNA)	Zetsche et al. 2015 Cell
HyraCas9 (N692A, M694A, Q695A, H698A)		Chen et al., Nature volume 550, pages 407-410 (19 October 2017)

[0112] В вариантах осуществления ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой эндонуклеазу CRISPR класса II с дефицитом нуклеазы. Термин «эндонуклеаза CRISPR класса II с дефицитом нуклеазы», используемый в данном документе, относится к любой эндонуклеазе CRISPR класса II, имеющей мутации, приводящие к сниженной, нарушенной или неактивной эндонуклеазной активности.

[0113] В вариантах осуществления пептидный линкер представляет собой линкер XTEN. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 17 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 18 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 19 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 20 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 30 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 40 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 50 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 70 до приблизительно 80 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 70 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 60 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 50 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 40 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 35 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 30 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 25 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 16 до приблизительно 20 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 16 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 17 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 18 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 19 аминокислотных остатков. В аспектах линкер XTEN

содержит приблизительно 20 аминокислотных остатков.

[0114] В аспектах слитый белок содержит по меньшей мере два линкера XTEN, которые являются одинаковыми или разными. В аспектах слитый белок содержит первый линкер XTEN, имеющий большее количество аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 10 до 150 больше аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 20 до 120 больше аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 30 до 110 больше аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 40 до 110 больше аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 50 до 100 больше аминокислотных остатков, чем второй линкер XTEN. В аспектах слитый белок содержит первый линкер XTEN, содержащий от 60 до 100 больше аминокислотных остатков, чем второй линкер XTEN.

[0115] В вариантах осуществления линкер XTEN содержит от приблизительно 50 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 50 до приблизительно 200 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 55 до приблизительно 180 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 150 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 120 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 110 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 100 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 70 до приблизительно 90 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 75 до приблизительно 85 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 80 аминокислотных остатков. В аспектах, когда слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер XTEN, содержащий от приблизительно 50 до приблизительно 200 аминокислотных остатков, называется первым пептидным линкером XTEN.

[0116] В вариантах осуществления линкер XTEN содержит от приблизительно 5 до приблизительно 55 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 5 до приблизительно 50 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 5 до приблизительно 40 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 30 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 25 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 20 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 14 до приблизительно 18 аминокислотных остатков. В

аспектах линкер XTEN содержит приблизительно 16 аминокислотных остатков. В аспектах, когда слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер XTEN, содержащий от приблизительно 5 до приблизительно 55 аминокислотных остатков, называется вторым пептидным линкером XTEN.

[0117] В вариантах осуществления линкер XTEN содержит последовательность, представленную SEQ ID NO:5. В аспектах линкер XTEN представляет собой последовательность, представленную SEQ ID NO:5. В аспектах линкер XTEN 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:5. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:5. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:5. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:5. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:5. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:5.

[0118] В вариантах осуществления линкер XTEN содержит последовательность, представленную SEQ ID NO:6. В аспектах линкер XTEN представляет собой последовательность, представленную SEQ ID NO:6. В аспектах линкер XTEN 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:6. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:6. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:6. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:6. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:6. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:6.

[0119] В вариантах осуществления линкер XTEN содержит последовательность, представленную SEQ ID NO:98. В аспектах линкер XTEN представляет собой последовательность, представленную SEQ ID NO:98. В аспектах линкер XTEN 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:98. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:98. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:98. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:98. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:98. В аспектах линкер XTEN имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:98.

[0120] Слитый белок может содержать аминокислотные последовательности, пригодные для нацеливания слитого белка на определенные области клетки (например, цитоплазму, ядро). Таким образом, в аспектах слитый белок дополнительно содержит пептид сигнала ядерной локализации (NLS). В аспектах NLS включает последовательность, представленную SEQ ID NO:4. В аспектах NLS представляет собой последовательность, представленную SEQ ID NO:4. В 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:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 75% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 80% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 85% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 90% идентичности последовательности с SEQ ID NO:4. В аспектах NLS имеет аминокислотную последовательность, которая имеет по меньшей мере 95% идентичности последовательности с SEQ ID NO:4.

[0121] Слитые белки

[0122] В данном документе представлены, среди прочего, слитые белки, которые могут быть нацелены на любой локус в геноме человека для активации экспрессии генов человека в долгосрочной перспективе (т. е. унаследованы в результате множественных клеточных делений), и которые могут временно доставляться в виде mRNA, ДНК или RNP. Слитые белки обладают возможностями мультиплексного эпигенетического редактирования для активации транскрипции и контроля транскрипции путем удаления эпигенетических меток, включая метильные группы на азотистых основаниях и репрессивные модификации гистонов. Представленные в данном документе слитые белки

дополнительно содержат несколько доменов, действующих согласованно для устойчивой активации транскрипции.

[0123] В вариантах осуществления в настоящем изобретении представлен слитый белок, содержащий от N-конца до С-конца, домен деметилирования и фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления фермент эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой белок, ассоциированный с CRISPR. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилирования представляет собой домен TET2. В вариантах осуществления домен деметилирования представляет собой домен TET3. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации. В вариантах осуществления слитый белок имеет по меньшей мере 85% идентичности последовательности с соединением формулы (I): $R^1-L^1-R^2$; где R^1 содержит SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:86 или SEQ ID NO:97; L^1 отсутствует, SEQ ID NO:5, SEQ ID NO:6 или SEQ ID NO:98; и R^2 содержит SEQ ID NO:9. В вариантах осуществления слитый белок имеет по меньшей мере 90% идентичности последовательности с соединением формулы (I). В вариантах осуществления слитый белок имеет по меньшей мере 92% идентичности последовательности с соединением формулы (I). В вариантах осуществления слитый белок имеет по меньшей мере 94% идентичности последовательности с соединением формулы (I). В вариантах осуществления слитый белок имеет по меньшей мере 95% идентичности последовательности с соединением формулы (I). В вариантах осуществления слитый белок имеет по меньшей мере 96% идентичности последовательности с соединением формулы (I). В вариантах осуществления слитый белок имеет по меньшей мере 98% идентичности последовательности с соединением формулы (I).

[0124] В вариантах осуществления в настоящем изобретении представлен слитый белок, содержащий от N-конца до С-конца, РНК-связывающую последовательность и по меньшей мере один активатор транскрипции. В вариантах осуществления слитый белок содержит от N-конца до С-конца РНК-связывающую последовательность, линкер XTEN и по меньшей мере один активатор транскрипции. В вариантах осуществления слитый белок содержит от N-конца до С-конца РНК-связывающую последовательность, линкер XTEN и по меньшей мере один активатор транскрипции, выбранный из группы, состоящей из VP64, p65, Rta или комбинации двух или более того. В вариантах осуществления активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции представляет собой VP64. В вариантах осуществления активатор транскрипции представляет собой p65. В вариантах

осуществления активатор транскрипции представляет собой Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta. В вариантах осуществления слитый белок имеет по меньшей мере 85% идентичности последовательности с соединением формулы (II): $R^4-L^1-R^3$; где R^4 содержит SEQ ID NO:21; L^1 отсутствует, SEQ ID NO:5, SEQ ID NO:6 или SEQ ID NO:98; и R^3 содержит SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:100 или комбинацию двух или более из них. В вариантах осуществления R^3 содержит SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:100 или комбинацию двух или более из них. В вариантах осуществления слитый белок имеет по меньшей мере 90% идентичности последовательности с соединением формулы (II). В вариантах осуществления слитый белок имеет по меньшей мере 92% идентичности последовательности с соединением формулы (II). В вариантах осуществления слитый белок имеет по меньшей мере 94% идентичности последовательности с соединением формулы (II). В вариантах осуществления слитый белок имеет по меньшей мере 95% идентичности последовательности с соединением формулы (II). В вариантах осуществления слитый белок имеет по меньшей мере 96% идентичности последовательности с соединением формулы (II). В вариантах осуществления слитый белок имеет по меньшей мере 98% идентичности последовательности с соединением формулы (III).

[0125] В вариантах осуществления слитый белок, имеющий от N-конца к C-концу РНК-связывающую последовательность, линкер XTEN и по меньшей мере один активатор транскрипции, содержит SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110. В аспектах слитый белок содержит аминокислотную последовательность, имеющую по меньшей мере 80% идентичности последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110. В аспектах слитый белок содержит аминокислотную последовательность, имеющую по меньшей мере 85% идентичности последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110. В аспектах слитый белок содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110. В аспектах слитый белок содержит аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110.

[0126] В вариантах осуществления в настоящем изобретении представлен слитый белок, содержащий от N-конца до С-конца, домен деметилирования, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, и активатор транскрипции. В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, и активатор транскрипции. В вариантах осуществления фермент эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой белок, ассоциированный с CRISPR. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилирования представляет собой домен TET2. В вариантах осуществления домен деметилирования представляет собой домен TET3. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации. В вариантах осуществления слитый белок имеет по меньшей мере 85% идентичности последовательности с соединением формулы (III): $R^1-L^1-R^2-R^3$; где R^1 содержит SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:86, SEQ ID NO:97; L^1 отсутствует, SEQ ID NO:5, SEQ ID NO:6 или SEQ ID NO:98; R^2 содержит SEQ ID NO:9; и R^3 содержит SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:100 или комбинацию двух или более из них. В вариантах осуществления R^3 содержит SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:100 или комбинацию двух или более из них. В вариантах осуществления слитый белок имеет по меньшей мере 90% идентичности последовательности с соединением формулы (III). В вариантах осуществления слитый белок имеет по меньшей мере 92% идентичности последовательности с соединением формулы (III). В вариантах осуществления слитый белок имеет по меньшей мере 94% идентичности последовательности с соединением формулы (III). В вариантах осуществления слитый белок имеет по меньшей мере 95% идентичности последовательности с соединением формулы (III). В вариантах осуществления слитый белок имеет по меньшей мере 96% идентичности последовательности с соединением формулы (III). В вариантах осуществления слитый белок имеет по меньшей мере 98% идентичности последовательности с соединением

формулы (III).

[0127] В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления фермент эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой белок, ассоциированный с CRISPR. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилирования представляет собой домен TET2. В вариантах осуществления домен деметилирования представляет собой домен TET3. В вариантах осуществления слитый белок дополнительно содержит активатор транскрипции. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации.

[0128] В вариантах осуществления фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, dCpf1, домен цинковых пальцев, домен лейциновой молнии, домен «крылатой спирали», TALE, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокс, домен Wor3, домен OB-складка, домен иммуноглобулина или домен В3. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой белок, ассоциированный с CRISPR. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCas9. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент dCpf1. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент Cas-phi. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен лейциновой молнии. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен «крылатой спирали». В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой мотив спираль-поворот-спираль. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен спираль-петля-спираль. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен НМВ-бокс. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен Wor3. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен OB-складка. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен иммуноглобулина. В вариантах осуществления ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой домен В3.

[0129] В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы. В вариантах осуществления фермент эндонуклеазы с дефицитом нуклеазы представляет собой «домен цинковых пальцев». В вариантах осуществления эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент TALE. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилирования представляет собой домен TET2. В вариантах осуществления домен деметилирования представляет собой домен TET3. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации.

[0130] В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы и активатор транскрипции. В вариантах осуществления фермент эндонуклеазы с дефицитом нуклеазы представляет собой «домен цинковых пальцев». В вариантах осуществления эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент TALE. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилирования представляет собой домен TET2. В вариантах осуществления домен деметилирования представляет собой домен TET3. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации.

[0131] В вариантах осуществления слитый белок содержит от N-конца к С-концу домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы. В вариантах осуществления фермент эндонуклеазы с дефицитом нуклеазы представляет собой «домен цинковых пальцев». В вариантах осуществления эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой фермент TALE. В вариантах осуществления домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них. В вариантах осуществления домен

деметилирования представляет собой домен TET1. В вариантах осуществления домен деметилюрования представляет собой домен TET2. В вариантах осуществления домен деметилюрования представляет собой домен TET3. В вариантах осуществления слитый белок дополнительно содержит активатор транскрипции. В вариантах осуществления активатор транскрипции содержит VP64, р65, Rta или комбинацию двух или более из них. В аспектах слитый белок дополнительно содержит последовательность ядерной локализации. В аспектах слитый белок дополнительно содержит две или три последовательности ядерной локализации.

[0132] В вариантах осуществления линкер XTEN содержит от приблизительно 5 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 20 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 30 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 40 до приблизительно 864 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 50 до приблизительно 200 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 55 до приблизительно 180 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 150 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 120 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 110 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 60 до приблизительно 100 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 70 до приблизительно 90 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 75 до приблизительно 85 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 80 аминокислотных остатков. В аспектах, когда слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер XTEN, содержащий от приблизительно 50 до приблизительно 200 аминокислотных остатков, называется первым пептидным линкером XTEN.

[0133] В вариантах осуществления линкер XTEN содержит от приблизительно 5 до приблизительно 55 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 5 до приблизительно 50 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 5 до приблизительно 40 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 30 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 25 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 10 до приблизительно 20 аминокислотных остатков. В аспектах линкер XTEN содержит от приблизительно 14 до приблизительно 18 аминокислотных остатков. В аспектах линкер XTEN содержит приблизительно 16 аминокислотных остатков. В аспектах, когда слитый белок содержит по меньшей мере два пептидных линкера XTEN, то линкер

XTEN, содержащий от приблизительно 5 до приблизительно 55 аминокислотных остатков, называется вторым пептидным линкером XTEN.

[0134] Касательно слитого белка, представленного в данном документе, в вариантах осуществления слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них. В вариантах осуществления слитый белок дополнительно содержит метку эпитопа. В вариантах осуществления слитый белок дополнительно содержит пептид 2A. В вариантах осуществления слитый белок дополнительно содержит флуоресцентную белковую метку. В вариантах осуществления слитый белок дополнительно содержит сигнальный пептид ядерной локализации.

[0135] Касательно слитого белка, представленного в данном документе, в вариантах осуществления слитый белок дополнительно содержит по меньшей мере один активатор транскрипции. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta.

[0136] В вариантах осуществления РНК-связывающая последовательность представляет собой РНК-связывающую последовательность MS2. В вариантах осуществления РНК-связывающая последовательность MS2 содержит белок MCP.

[0137] Слитый белок может содержать линкер XTEN, как описано в данном документе. В вариантах осуществления линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0138] В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, белок, ассоциированный с CRISPR, линкер XTEN, последовательность ядерной локализации, активатор транскрипции и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, домен цинкового пальца, линкер XTEN, последовательность ядерной локализации, активатор транскрипции и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, TALE, линкер XTEN, последовательность ядерной локализации, Rta и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, dCas9, линкер XTEN, последовательность ядерной локализации, активатор транскрипции и последовательность ядерной локализации. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах

осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит р65. В вариантах осуществления активатор транскрипции содержит VP64 и р65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит р65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, р65 и Rta. В вариантах осуществления слитый белок содержит от N-конца к С-концу SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:9, SEQ ID NO:6, SEQ ID NO:4, SEQ ID NO:15 и SEQ ID NO:4. В вариантах осуществления слитый белок содержит SEQ ID NO:99. В вариантах осуществления слитый белок представляет собой SEQ ID NO:99. В аспектах слитый белок 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:99. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 75% идентична последовательности с SEQ ID NO:99. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности с SEQ ID NO:99. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 85% идентична последовательности с SEQ ID NO:99. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности с SEQ ID NO:99. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности с SEQ ID NO:99.

[0139] В вариантах осуществления слитый белок содержит от N-конца к С-концу домен TET1, линкер XTEN, белок, ассоциированный с CRISPR, линкер XTEN, последовательность ядерной локализации, два активатора транскрипции и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к С-концу домен TET1, линкер XTEN, домен цинкового пальца, линкер XTEN, последовательность ядерной локализации, р65, Rta и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к С-концу домен TET1, линкер XTEN, TALE, линкер XTEN, последовательность ядерной локализации, два активатора транскрипции и последовательность ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к С-концу домен TET1, линкер XTEN, dCas9, линкер XTEN, последовательность ядерной локализации, два активатора транскрипции и последовательность ядерной локализации. В вариантах осуществления активатор транскрипции содержит по меньшей мере два из VP64, р65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64 и р65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит р65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, р65 и Rta. В вариантах осуществления слитый белок содержит от N-конца к С-концу SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:9, SEQ ID NO:6,

SEQ ID NO:4, SEQ ID NO:100, SEQ ID NO:15 и SEQ ID NO:4. В вариантах осуществления слитый белок содержит SEQ ID NO:101. В вариантах осуществления слитый белок представляет собой SEQ ID NO:101. В аспектах слитый белок 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:101. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 75% идентична последовательности с SEQ ID NO:101. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности с SEQ ID NO:101. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 85% идентична последовательности с SEQ ID NO:101. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности с SEQ ID NO:101. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности с SEQ ID NO:101.

[0140] В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, белок, ассоциированный с CAS, и от 1 до 3 последовательностей ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, домен цинкового пальца и от 1 до 3 последовательностей ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, TALE и от 1 до 3 последовательностей ядерной локализации. В вариантах осуществления слитый белок содержит от N-конца к C-концу домен TET1, линкер XTEN, dCas9 и от 1 до 3 последовательностей ядерной локализации. В вариантах осуществления слитый белок дополнительно содержит активатор транскрипции. В вариантах осуществления слитый белок содержит от N-конца к C-концу SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:9 и SEQ ID NO:4. В вариантах осуществления слитый белок содержит SEQ ID NO:102. В вариантах осуществления слитый белок представляет собой SEQ ID NO:102. В аспектах слитый белок 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:102. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 75% идентична последовательности с SEQ ID NO:102. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности с SEQ ID NO:102. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 85% идентична последовательности с SEQ ID NO:102. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности с SEQ ID NO:102. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности с SEQ ID NO:102.

аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности с SEQ ID NO:112.

[0144] В вариантах осуществления слитый белок содержит SEQ ID NO:113. В вариантах осуществления слитый белок представляет собой SEQ ID NO:113. В аспектах слитый белок 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:113. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 75% идентична последовательности с SEQ ID NO:113. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности с SEQ ID NO:113. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 85% идентична последовательности с SEQ ID NO:113. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности с SEQ ID NO:113. В аспектах слитый белок имеет аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности с SEQ ID NO:113.

[0145] В данном документе представлены соединения формулы (III) или соединения, имеющие по меньшей мере 85% идентичности последовательности с соединением формулы (III), где соединение формулы (III) представляет собой $R^{10}-L^1-R^{11}-R^{12}-L^2-L^3-(R^{13}-L^4)_x-R^{14}-X^1-L^5-X^2-L^6-X^3-L^7-R^{15}$. В вариантах осуществления соединения имеют по меньшей мере 90% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют по меньшей мере 92% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют по меньшей мере 94% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют по меньшей мере 95% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют по меньшей мере 96% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют по меньшей мере 98% идентичности последовательности с соединением формулы (III). В вариантах осуществления соединения имеют формулу (III). R^{10} представляет собой домен деметилирования. В вариантах осуществления R^{10} содержит SEQ ID NO:1, 2, 3, 86, 97 (включая их варианты осуществления). В вариантах осуществления R^{10} содержит SEQ ID NO:97 (включая ее варианты). L^1 представляет собой связь или пептидный линкер. В вариантах осуществления L^1 представляет собой связь. R^{11} представляет собой линкер XTEN. В вариантах осуществления R^{11} содержит SEQ ID NO: 5, 6 или 98 (включая их варианты осуществления). В вариантах осуществления R^{11} содержит SEQ ID NO: 5 (включая ее варианты осуществления). В вариантах осуществления R^{11} содержит SEQ ID NO: 6 (включая ее варианты осуществления). В вариантах осуществления R^{11} содержит SEQ ID NO:98 (включая ее варианты). R^{12} содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК, или фермент эндонуклеазу с дефицитом нуклеазы. В вариантах

осуществления R^{12} содержит фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК. В вариантах осуществления R^{12} содержит белок, ассоциированный с CRISPR. В вариантах осуществления R^{12} содержит SEQ ID NO:9 (включая ее варианты). В вариантах осуществления R^{12} содержит фермент эндонуклеазу с дефицитом нуклеазы. В вариантах осуществления R^{12} содержит домен цинковых пальцев или TALE. В вариантах осуществления R^{12} содержит домен цинковых пальцев. В вариантах осуществления R^{12} содержит TALE. L^2 представляет собой связь или линкер XTEN. В вариантах осуществления L^2 представляет собой связь или линкер XTEN. В вариантах осуществления L^2 представляет собой связь. В вариантах осуществления L^2 представляет собой линкер XTEN. В вариантах осуществления L^2 содержит SEQ ID NO: 5, 6 или 98 (включая их варианты). В вариантах осуществления L^2 содержит SEQ ID NO: 5 (включая ее варианты осуществления). В вариантах осуществления L^2 содержит SEQ ID NO: 6 (включая ее варианты осуществления). В вариантах осуществления L^2 содержит SEQ ID NO:98 (включая ее варианты). L^3 представляет собой связь или пептидный линкер. В вариантах осуществления L^3 представляет собой связь. В вариантах осуществления L^3 представляет собой пептидный линкер. В вариантах осуществления L^3 представляет собой пептидный линкер, содержащий от 1 аминокислоты до приблизительно 10 аминокислот. В вариантах осуществления L^3 представляет собой пептидный линкер, содержащий от 3 аминокислот до приблизительно 5 аминокислот. R^{13} содержит последовательность ядерной локализации. В вариантах осуществления R^{13} содержит SEQ ID NO:4 (включая ее варианты). L^4 отсутствует или представляет собой пептидный линкер. В вариантах осуществления L^4 отсутствует. В вариантах осуществления L^4 представляет собой пептидный линкер. В вариантах осуществления L^4 представляет собой пептидный линкер, содержащий от 1 аминокислоты до приблизительно 10 аминокислот. В вариантах осуществления L^4 представляет собой пептидный линкер, содержащий от 1 аминокислоты до приблизительно 5 аминокислот. В вариантах осуществления L^4 представляет собой пептидный линкер, содержащий от 1 аминокислоты до приблизительно 4 аминокислот. x представляет собой целое число от 0 до 4. В вариантах осуществления x равно 0. В вариантах осуществления x равно 1. В вариантах осуществления x равно 2. В вариантах осуществления x равно 3. R^{14} отсутствует или представляет собой последовательность ядерной локализации. В вариантах осуществления R^{14} отсутствует. В вариантах осуществления R^{14} представляет собой последовательность ядерной локализации. В вариантах осуществления R^{14} содержит SEQ ID NO:4 (включая ее варианты). X^1 , X^2 и X^3 независимо отсутствуют или представляют собой активатор транскрипции. В вариантах осуществления X^1 , X^2 и X^3 независимо представляют собой активатор транскрипции. В вариантах осуществления X^1 , X^2 и X^3 независимо представляют собой р65, Rta или VP64. В вариантах осуществления X^1 , X^2 и X^3 независимо представляют собой р65, Rta или VP64, где каждый из X^1 , X^2 и X^3 является различным. В вариантах осуществления X^1 и X^2 независимо представляют собой р65, Rta или VP64, а X^3 отсутствует. В вариантах осуществления X^1 и X^2 независимо представляют собой р65, Rta или VP64; X^3 отсутствует; и X^1 и X^2 являются различными. В вариантах осуществления X^1 представляет

собой р65, Rta или VP64; X² отсутствует; и X³ отсутствует. В вариантах осуществления р65 содержит SEQ ID NO: 13, 14 или 100 (включая их варианты). В вариантах осуществления р65 содержит SEQ ID NO:13 (включая ее варианты). В вариантах осуществления р65 содержит SEQ ID NO:14 (включая ее варианты). В вариантах осуществления р65 содержит SEQ ID NO:100 (включая ее варианты). В вариантах осуществления Rta содержит SEQ ID NO: 15 или 16 (включая их варианты). В вариантах осуществления Rta содержит SEQ ID NO:15 (включая ее варианты). В вариантах осуществления Rta содержит SEQ ID NO:16 (включая ее варианты). В вариантах осуществления VP64 содержит SEQ ID NO: 17 или 18 (включая их варианты). В вариантах осуществления VP64 содержит SEQ ID NO:17 (включая ее варианты). В вариантах осуществления VP64 содержит SEQ ID NO:18 (включая ее варианты). L⁵ отсутствует или представляет собой пептидный линкер. В вариантах осуществления L⁵ отсутствует. В вариантах осуществления L⁵ содержит пептидный линкер. В вариантах осуществления пептидный линкер содержит от 1 до приблизительно 10 аминокислот. В вариантах осуществления пептидный линкер содержит от 3 до приблизительно 5 аминокислот. L⁶ отсутствует или представляет собой пептидный линкер. В вариантах осуществления L⁶ отсутствует. В вариантах осуществления L⁶ содержит пептидный линкер. В вариантах осуществления пептидный линкер содержит от 1 до приблизительно 10 аминокислот. В вариантах осуществления пептидный линкер содержит от 3 до приблизительно 5 аминокислот. L⁷ отсутствует или представляет собой пептидный линкер. В вариантах осуществления L⁷ отсутствует. В вариантах осуществления L⁷ содержит пептидный линкер. В вариантах осуществления пептидный линкер содержит от 1 до приблизительно 10 аминокислот. В вариантах осуществления пептидный линкер содержит от 3 до приблизительно 5 аминокислот. В вариантах осуществления, если X¹ отсутствует, то L⁵ отсутствует. В вариантах осуществления, если X² отсутствует, то L⁶ отсутствует. В вариантах осуществления, если X³ отсутствует, то L⁷ отсутствует. В вариантах осуществления, если X² отсутствует, то X³ отсутствует, и L⁶ и L⁷ отсутствуют. В вариантах осуществления, если X¹ отсутствует, то X² и X³ отсутствуют, и L⁵, L⁶ и L⁷ отсутствуют. R¹⁵ отсутствует или представляет собой последовательность ядерной локализации. В вариантах осуществления R¹⁵ отсутствует. В вариантах осуществления R¹⁵ представляет собой последовательность ядерной локализации. В вариантах осуществления R¹⁵ содержит SEQ ID NO:4 (включая ее варианты).

[0146] В последовательностях, приведенных в данном документе, специалисту в данной области техники будет понятно, что метионин (M) может присутствовать на N-конце белка для инициации трансляции. Таким образом, последовательности, описанные в данном документе, могут необязательно дополнительно содержать метионин на N-конце.

[0147] Комплексы

[0148] Для того, чтобы слитый белок мог осуществлять редактирование эпигенома, слитый белок взаимодействует (например, нековалентно связан) с полинуклеотидом (например, sgRNA), комплементарным полинуклеотидной последовательности-мишени (например, последовательности ДНК-мишени, подлежащей редактированию) и

дополнительно содержит последовательность (т.е. связывающую последовательность), с которой может связываться фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, слитого белка, как описано в данном документе. В аспектах полинуклеотид, комплементарный полинуклеотидной последовательности-мишени (например, последовательности геномной ДНК-мишени, подлежащей редактированию) и дополнительно содержащий последовательность связывания, с которой фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, слитого белка, как описано в данном документе, может связываться с sgRNA. В аспектах полинуклеотид, комплементарный полинуклеотидной последовательности-мишени (например, последовательности ДНК-мишени, подлежащей редактированию) и дополнительно содержащий последовательность связывания, с которой фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, слитого белка, как описано в данном документе, может связываться с cr:tracrRNA. Образуя этот комплекс, слитый белок занимает подходящее положение для выполнения редактирования эпигенома. Термин «комплекс» относится к композиции, которая содержит два или более компонентов, где компоненты связаны друг с другом, образуя функциональную единицу. В аспектах комплекс, описанный в данном документе, включает слитый белок, описанный в данном документе, и полинуклеотид, описанный в данном документе. Таким образом, в аспекте представлен слитый белок, как описано в данном документе, включая варианты его осуществления и аспекты, и sgRNA или cr:tracrRNA (т.е. полинуклеотид, содержащий: (1) последовательность, нацеленная на ДНК, которая комплементарна последовательности полинуклеотида-мишени; и (2) последовательность связывания для фермента ДНК-эндонуклеазы с дефицитом нуклеазы, направляемой РНК, где фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, связывается с полинуклеотидом через связывающую последовательность (например, аминокислотную последовательность, способную связываться с последовательностью, нацеленной на ДНК)). В аспектах полинуклеотид содержит по меньшей мере одну петлю MS2.

[0149] В аспектах комплекс, описанный в данном документе, включает слитый белок, описанный в данном документе, полинуклеотид, описанный в данном документе, и второй слитый белок, описанный в данном документе. В аспектах второй слитый белок содержит активатор транскрипции, описанный в данном документе.

[0150] Последовательность, нацеленная на ДНК, относится к полинуклеотиду, который включает последовательность нуклеотидов, комплементарную последовательности полинуклеотида-мишени (ДНК или РНК). В аспектах последовательность, нацеленная на ДНК, может представлять собой одну молекулу РНК (один полинуклеотид РНК), которая может включать «одиочную направляющую РНК» или «sgRNA». В аспектах последовательность, нацеленная на ДНК, включает две молекулы РНК (например, две sgRNA), называемые направляющей РНК (gRNA) (например, соединенные вместе посредством гибридизации в связывающей последовательности (например, dCas9-связывающей последовательности)). В аспектах последовательность,

нацеленная на ДНК (например, sgRNA), на по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% комплементарна полинуклеотидной последовательности-последовательности. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), на по меньшей мере 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), связывается с последовательностью клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA) на по меньшей мере 75% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA) на по меньшей мере 80% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), связывается с последовательностью клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA) на по меньшей мере 85% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), связывается с последовательностью клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA) на по меньшей мере 90% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), связывается с последовательностью клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA) на по меньшей мере 95% комплементарна последовательности клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), связывается с последовательностью клеточного гена. В аспектах последовательность, нацеленная на ДНК (например, sgRNA), содержит по меньшей мере одну петлю стебля MS2. В вариантах осуществления петля стебля MS2 содержит последовательность SEQ ID NO:19. В вариантах осуществления петля стебля MS2 имеет последовательность SEQ ID NO:19. В аспектах петля стебля MS2 имеет 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:19.

[0151] «Полинуклеотидная последовательность-мишень», как представлено в данном документе, представляет собой последовательность нуклеиновой кислоты, присутствующую в клетке или экспрессируемую ею, для которой направляющая последовательность (или последовательность, нацеленная на ДНК) предназначена для обеспечения комплементарности, где гибридизация между последовательностью-мишенью и направляющая последовательность (или последовательность, нацеленная на ДНК) способствует образованию комплекса (например, комплекса CRISPR). Полная комплементарность не обязательно требуется, при условии, что комплементарности достаточно, чтобы вызвать гибридизацию и способствовать образованию комплекса (например, комплекса CRISPR). В аспектах полинуклеотидная последовательность-последовательность представляет собой экзогенную последовательность нуклеиновой кислоты. В аспектах полинуклеотидная последовательность-мишень представляет собой

последовательность эндогенной нуклеиновой кислоты.

[0152] Последовательность полинуклеотида-мишени может представлять собой любую область полинуклеотида (например, последовательность ДНК), подходящую для редактирования эпигенома. В аспектах полинуклеотидная последовательность-мишень является частью гена. В аспектах полинуклеотидная последовательность-мишень является частью регуляторной последовательности транскрипции. В аспектах полинуклеотидная последовательность-мишень является частью промотора, энхансера или сайленсера. В аспектах полинуклеотидная последовательность-мишень является частью промотора. В аспектах полинуклеотидная последовательность-мишень является частью энхансера. В аспектах полинуклеотидная последовательность-мишень является частью сайленсера.

[0153] В вариантах осуществления полинуклеотидная последовательность-мишень представляет собой гиперметилованную последовательность нуклеиновой кислоты. «Последовательность гиперметилованной нуклеиновой кислоты» используется в данном документе в соответствии со стандартным значением в данной области техники и относится к частому метилированию цитозина до 5-метилцитозина (например, в CpG). Частота или встречаемость метильных групп может быть представлена относительно стандартного контроля. Гиперметилирование может происходить, например, при раке (например, в путях репарации ДНК или апоптоза) по сравнению с нераковой клеткой, соответственно. Таким образом, комплекс может быть пригоден для восстановления нормальных (например, непатологических) уровней метилирования.

[0154] В вариантах осуществления полинуклеотидная последовательность-мишень находится внутри или рядом с сайтом начала транскрипции. В аспектах полинуклеотидная последовательность-мишень находится в пределах примерно 3000, 2500, 2000, 1500, 500, 100, 80, 70, 60, 50, 40, 30, 20, 10 или менее пар оснований (п.н.), фланкирующих сайт начала транскрипции.

[0155] В вариантах осуществления полинуклеотидная последовательность-мишень находится в промоторной последовательности, рядом с ней или внутри нее. В аспектах полинуклеотидная последовательность-мишень находится в островке CpG. В аспектах полинуклеотидная последовательность-мишень находится в отличном от островка CpG островке. В аспектах известно, что полинуклеотидная последовательность ассоциирована с заболеванием или состоянием, характеризующимся гиперметилованием или гипометилованием ДНК.

[0156] В вариантах осуществления комплекс включает dCas9, связанный с полинуклеотидом посредством связывания связывающей последовательности полинуклеотида и, таким образом, образующий рибонуклеопротеиновый комплекс. В аспектах последовательность связывания образует структуру шпильки. В аспектах связывающая последовательность имеет длину 10-200 нуклеотидов, 15-150 нуклеотидов, 20-140 нуклеотидов, 30-100 нуклеотидов.

[0157] В вариантах осуществления последовательность связывания (например, последовательность связывания Cas9) взаимодействует или связывается с белком Cas9

(например, белком dCas9), и вместе они связываются с последовательностью полинуклеотида-мишени, распознаваемой последовательностью, нацеленной на ДНК. Связывающая последовательность (например, Cas9-связывающая последовательность) включает два комплементарных участка нуклеотидов, которые гибридизуются друг с другом с образованием дуплекса двухцепочечной РНК (дуплекса dsRNA). Эти два комплементарных участка нуклеотидов могут быть ковалентно связаны промежуточными нуклеотидами, известными как линкеры или линкерные нуклеотиды (например, в случае одномолекулярного полинуклеотида), и гибридизоваться с образованием двухцепочечного дуплекса РНК (дуплекса dsRNA, или «Cas9-связывающей шпильки») связывающей последовательности (например, Cas9-связывающей последовательности), что приводит к структуре «стебель-петля». Альтернативно, в некоторых аспектах два комплементарных участка нуклеотидов могут быть не связаны ковалентно, а вместо этого удерживаются вместе за счет гибридизации между комплементарными последовательностями (например, двухмолекулярный полинуклеотид).

[0158] Связывающая последовательность (например, Cas9-связывающая последовательность) может иметь длину от 10 нуклеотидов до 200 нуклеотидов, например, от 20 нуклеотидов (нт) до 150 нт. В аспектах связывающая последовательность имеет длину от 80 нуклеотидов (нт) до 100 нт. Дуплекс dsRNA связывающей последовательности (например, Cas9-связывающей последовательности) может иметь длину от 6 пар оснований (п.н.) до 200 п.н. Например, дуплекс dsRNA связывающей последовательности (например, Cas9-связывающей последовательности) может иметь длину от 6 до 200 п.н., от 10 до 180 п.н., от 10 до 150 п.н., от 80 до 100 п.н., и тому подобное.

[0159] Нуклеиновые кислоты и векторы

[0160] Слитый белок, описанный в данном документе, включая варианты его осуществления, может быть доставлен в клетку различными способами, известными в данной области техники. Слитый белок может экспрессироваться временно, минуя необходимость вирусных способов доставки. Слитый белок может быть закодирован на РНК или ДНК, доставленных в клетки в виде модифицированной или немодифицированной РНК или плазмидной ДНК. РНК или ДНК, кодирующие белок, могут быть доставлены трансфекцией, липидными наночастицами, вирусоподобными частицами (VLP) или вирусом. Теоретически белок также может быть доставлен напрямую посредством трансфекции или липидных наночастиц или VLP.

[0161] Слитый белок, описанный в данном документе, включая его варианты осуществления и аспекты, может быть представлен в виде последовательности нуклеиновой кислоты, которая кодирует слитый белок. Таким образом, в аспекте представлена последовательность нуклеиновой кислоты, кодирующая слитый белок, описанный в данном документе, включая его варианты осуществления и аспекты. В одном аспекте представлена последовательность нуклеиновой кислоты, кодирующая слитый белок, описанный в данном документе (включая последовательность, нацеленную на ДНК), включая варианты осуществления и их аспекты. В 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 аспектах нуклеиновая кислота представляет собой РНК. В аспектах нуклеиновая кислота представляет собой информационную РНК. В аспектах слитый белок доставляется в виде ДНК, mRNA, белка или RNP. Для RNP белок будет представлять собой dCas9, а РНК будет кодировать sgRNA. Аналогично sgRNA может быть доставлена в виде ДНК, кодирующей промотор, и sgRNA, РНК, кодирующей промотор, и sgRNA. В аспектах последовательность нуклеиновой кислоты кодирует слитые белки, описанные в данном документе, включая их варианты осуществления и аспекты.

[0162] В аспектах слитые белки и sgRNA или cr:tracrRNA, представленные в данном документе, включая их варианты осуществления, могут быть представлены в виде одной нуклеиновой кислоты, которая кодирует слитый белок и sgRNA или cr:tracrRNA. В аспектах слитые белки и sgRNA или cr:tracrRNA, представленные в данном документе, включая их варианты осуществления, могут быть представлены в виде множества нуклеиновых кислот, которые кодируют слитый белок и sgRNA или cr:tracrRNA. В вариантах осуществления слитый белок и sgRNA или cr:tracrRNA представлены в виде отдельных транскриптов.

[0163] В одном аспекте представлена нуклеиновая кислота, кодирующая слитый белок, содержащий домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК.

[0164] В одном аспекте представляется вторая нуклеиновая кислота, кодирующая sgRNA или cr:tracrRNA. В вариантах осуществления sgRNA содержит по меньшей мере одну последовательность MS2. В вариантах осуществления sgRNA содержит две последовательности MS2. В вариантах осуществления вторая последовательность нуклеиновой кислоты дополнительно кодирует последовательность связывания MS2-РНК и по меньшей мере один активатор транскрипции, представленный в данном документе.

[0165] В одном аспекте представлена третья нуклеиновая кислота, кодирующая активатор транскрипции. В вариантах осуществления третья нуклеиновая кислота дополнительно кодирует РНК-связывающую последовательность и линкер XTEN. В вариантах осуществления РНК-связывающая последовательность представляет собой РНК-связывающую последовательность MS2.

[0166] Кроме того, предполагается, что последовательность нуклеиновой кислоты, кодирующая слитый белок, как описано в данном документе, включая варианты его осуществления и аспекты, может быть включена в вектор. Таким образом, в одном из аспектов представлен вектор, включающий последовательность нуклеиновой кислоты, как описано в данном документе, включая ее варианты осуществления и аспекты. В 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 аспектах нуклеиновая кислота представляет собой информационную

РНК. В аспектах информационная РНК представляет собой информационный RNP.

[0167] В вариантах осуществления вектор дополнительно включает полинуклеотид, при этом полинуклеотид содержит: (1) последовательность, нацеленная на ДНК, которая комплементарна последовательности полинуклеотида-мишени; и (2) последовательность связывания для фермента ДНК-эндонуклеазы с дефицитом нуклеазы, направляемой РНК. В аспектах вектор дополнительно включает полинуклеотид, где полинуклеотид включает sgRNA. В аспектах вектор дополнительно включает полинуклеотид, где полинуклеотид включает cr:tracrRNA. Таким образом, один или несколько векторов могут включать все необходимые компоненты для предварительного редактирования эпигенома.

[0168] Клетки

[0169] Композиции, описанные в данном документе, могут быть включены в клетку. Композиции, описанные в данном документе, включая их варианты и аспекты, внутри клетки могут выполнять редактирование эпигенома. Соответственно, в аспекте представлена клетка, содержащая слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты, нуклеиновую кислоту, как описано в данном документе, включая его варианты осуществления и их аспекты, комплекс, описанный в данном документе, включая его варианты осуществления и их аспекты, или вектор, как описано в данном документе, включая его варианты осуществления и аспекты. В аспектах представлена клетка, содержащая слитый белок, как описано в данном документе, включая его варианты осуществления и аспекты. В аспектах представлена клетка, содержащая нуклеиновую кислоту, как описано в данном документе, включая варианты ее осуществления и аспекты. В аспектах представлена клетка, содержащая комплекс, описанный в данном документе, включая варианты осуществления и их аспекты. В аспектах представлена клетка, содержащая вектор, как описано в данном документе, включая его варианты осуществления и аспекты. В аспектах клетка представляет собой эукариотическую клетку.

[0170] В аспектах клетка представляет собой клетку млекопитающего. В вариантах осуществления клетка млекопитающего представляет собой клетку HEK293T. В вариантах осуществления клетка млекопитающего представляет собой T-клетку. В вариантах осуществления клетка млекопитающего представляет собой гемопоэтическую стволовую клетку. В вариантах осуществления клетка млекопитающего представляет собой индуцированную плюрипотентную стволовую клетку. В вариантах осуществления клетка млекопитающего представляет собой эмбриональную стволовую клетку.

[0171] Способы

[0172] Предполагается, что способы, описанные в данном документе, могут быть использованы для редактирования эпигенома и, более конкретно, редактирования эпигенома, приводящего к активации или реактивации последовательностей нуклеиновых кислот-мишеней (например, генов). Предлагаемые в данном документе способы включают рекрутирование одного или нескольких слитых белков для мультиплексного редактирования эпигенетического кода ДНК и гистонового кода. Эти способы

обеспечивают длительную, но обратимую активацию транскрипции и могут быть использованы для активации ранее подвергнутых сайленсингу генов. Представленные в данном документе способы могут быть использованы в терапевтических целях. Например, рекрутирование одного или нескольких слитых белков, представленных в данном документе, может активировать экспрессию гена путем редактирования негативных регуляторных последовательностей. Этот способ можно использовать для редактирования последовательностей, блокирующих экспрессию генов.

[0173] Слитые белки, описанные в данном документе, программируют долговременную память об активации генов с течением времени. Активация гена (или реактивация) достигается путем трансфекции mRNA, кодирующей описанные в данном документе слитые белки. Таким образом, временная экспрессия слитого белка приводит к эффективной активации гена (или реактивации). Эпигенетическая память CRISPRon с использованием описанных в данном документе слитых белков распространяется клеткой, а не за счет устойчивой экспрессии трансгена.

[0174] В вариантах осуществления настоящего изобретения представлены способы активации последовательности нуклеиновой кислоты-мишени в клетке, при этом способ включает: (i) доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все варианты его осуществления и аспекты (например, содержащий фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК), в клетку, содержащую нуклеиновую кислоту-мишень; и (ii) доставку в клетку второго полинуклеотида, содержащего: (a) sgRNA или (b) cr:tracrRNA; за счет чего активируется последовательность нуклеиновой кислоты-мишени в клетке. В вариантах осуществления второй полинуклеотид содержит sgRNA. В вариантах осуществления sgRNA содержит по меньшей мере петлю стебля MS2. В вариантах осуществления sgRNA содержит две петли стебля MS2. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0175] В вариантах осуществления в настоящем изобретении представлены способы активации последовательности нуклеиновой кислоты-мишени в клетке, где способ включает доставку полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все варианты осуществления и их аспекты (например, включающий фермент ДНК-эндонуклеазу с дефицитом нуклеазы), в клетку, содержащую нуклеиновую кислоту-мишень; за счет чего активируется последовательность нуклеиновой кислоты-мишени в клетке. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0176] В вариантах осуществления настоящего изобретения представлены способы реактивации последовательности нуклеиновой кислоты-мишени, которая была подвергнута сайленсингу, в клетке, при этом способ включает: (i) доставку первого полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая

все варианты его осуществления и аспекты (например, содержащий фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК), в клетку, содержащую нуклеиновую кислоту-мишень, которая была подвергнута сайленсингу; и (ii) доставку в клетку второго полинуклеотида, содержащего: (a) sgRNA или (b) cr:tracrRNA; за счет чего происходит реактивация подвергнутой сайленсингу последовательности нуклеиновой кислоты-мишени в клетке. В вариантах осуществления второй полинуклеотид содержит sgRNA. В вариантах осуществления sgRNA содержит по меньшей мере петлю стебля MS2. В вариантах осуществления sgRNA содержит две петли стебля MS2. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0177] В вариантах осуществления в настоящем изобретении представлены способы реактивации последовательности нуклеиновой кислоты-мишени в клетке, где способ включает доставку полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все варианты осуществления и их аспекты (например, включающий фермент ДНК-эндонуклеазу с дефицитом нуклеазы), в клетку, содержащую нуклеиновую кислоту-мишень; за счет чего активируется последовательность нуклеиновой кислоты-мишени в клетке. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0178] В вариантах осуществления настоящего изобретения представлены способы активации последовательности нуклеиновой кислоты-мишени в клетке, при этом способ включает: (i) доставку полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все варианты его осуществления и аспекты (например, содержащий фермент ДНК-эндонуклеазу с дефицитом нуклеазы, направляемую РНК), в клетку, содержащую нуклеиновую кислоту-мишень; при этом полинуклеотид дополнительно кодирует (a) sgRNA или (b) cr:tracrRNA; за счет чего активируется последовательность нуклеиновой кислоты-мишени в клетке. В вариантах осуществления полинуклеотид содержит sgRNA. В вариантах осуществления sgRNA содержит по меньшей мере петлю стебля MS2. В вариантах осуществления sgRNA содержит две петли стебля MS2. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0179] В вариантах осуществления в настоящем изобретении представлены способы реактивации последовательности подвергнутой сайленсингу нуклеиновой кислоты-мишени в клетке, где способ включает: доставку полинуклеотида, кодирующего слитый белок, как описано в данном документе, включая все варианты осуществления и их аспекты (например, включая фермент ДНК-эндонуклеазу с дефицитом нуклеазы, управляемую РНК) в клетку, содержащую подвергнутую сайленсингу нуклеиновую кислоту-мишень; при этом полинуклеотид дополнительно кодирует (a) sgRNA или (b) cr:tracrRNA; за счет

чего реактивируется подвергнутая сайленсингу последовательность нуклеиновой кислоты-мишени в клетке. В вариантах осуществления полинуклеотид содержит sgRNA. В вариантах осуществления sgRNA содержит по меньшей мере петлю стебля MS2. В вариантах осуществления sgRNA содержит две петли стебля MS2. В аспектах последовательность нуклеиновой кислоты-мишени содержит островок CpG. В аспектах последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0180] В способах активации последовательности нуклеиновой кислоты-мишени или повторной активации последовательности нуклеиновой кислоты-мишени, описанных в данном документе, нуклеиновая кислота-мишень содержит островок CpG и отличный от островка CpG островок. «Содержит островок CpG» или «содержит отличный от островка CpG островок» относится к одному или более островкам CpG или отличным от островков CpG островкам соответственно. В аспектах последовательность нуклеиновой кислоты-мишени содержит множество островков CpG (например, 2, 3, 4, 5 или более островков CpG). В аспектах последовательность нуклеиновой кислоты-мишени содержит множество отличных от островков CpG островков (например, 2, 3, 4, 5 или более отличных от островков CpG островков). В аспектах последовательность нуклеиновой кислоты-мишени не содержит островок CpG и не содержит отличный от островка CpG островок.

[0181] В вариантах осуществления петля стебля MS2 содержит последовательность SEQ ID NO:19. В вариантах осуществления петля стебля MS2 имеет последовательность SEQ ID NO:19. В аспектах петля стебля MS2 имеет 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:19. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 85% идентична последовательности с SEQ ID NO:19. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 90% идентична последовательности с SEQ ID NO:19. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 95% идентична последовательности с SEQ ID NO:19. В аспектах петля стебля MS2 имеет 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:20. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 85% идентична последовательности с SEQ ID NO:20. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 90% идентична последовательности с SEQ ID NO:20. В аспектах петля стебля MS2 имеет последовательность, которая на меньшей мере 95% идентична последовательности с SEQ ID NO:20.

[0182] В вариантах осуществления второй полинуклеотид дополнительно кодирует второй слитый белок, который содержит активатор транскрипции. В вариантах осуществления активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции

представляет собой VP64. В вариантах осуществления активатор транскрипции представляет собой p65. В вариантах осуществления активатор транскрипции представляет собой Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta.

[0183] В вариантах осуществления второй слитый белок содержит РНК-связывающую последовательность MS2. В вариантах осуществления РНК-связывающая последовательность MS2 содержит белок MCP или его функциональный фрагмент.

[0184] В вариантах осуществления способ дополнительно включает доставку в клетку третьего полинуклеотида, кодирующего второй слитый белок, который содержит активатор транскрипции. В вариантах осуществления активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции представляет собой VP64. В вариантах осуществления активатор транскрипции представляет собой p65. В вариантах осуществления активатор транскрипции представляет собой Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65, Rta или комбинацию двух или более из них. В вариантах осуществления активатор транскрипции содержит VP64. В вариантах осуществления активатор транскрипции содержит p65. В вариантах осуществления активатор транскрипции содержит Rta. В вариантах осуществления активатор транскрипции содержит VP64 и p65. В вариантах осуществления активатор транскрипции содержит VP64 и Rta. В вариантах осуществления активатор транскрипции содержит p65 и Rta. В вариантах осуществления активатор транскрипции содержит VP64, p65 и Rta.

[0185] Для способов, представленных в данном документе, в вариантах осуществления второй слитый белок дополнительно содержит линкер XTEN, метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них. В вариантах осуществления второй слитый белок дополнительно содержит линкер XTEN. В вариантах осуществления второй слитый белок дополнительно содержит метку эпитопа. В вариантах осуществления второй слитый белок дополнительно содержит пептид 2A. В вариантах осуществления второй слитый белок дополнительно содержит флуоресцентную белковую метку. В вариантах осуществления второй слитый белок дополнительно содержит сигнальный пептид ядерной локализации.

[0186] Термин «островок CpG» используется в его обычном смысле для обозначения областей в нуклеиновой кислоте, которые имеют высокую частоту нуклеотидов G и C

рядом друг с другом (т.е. динуклеотиды CpG). В аспектах островок CpG относится к области последовательности нуклеиновой кислоты, имеющей по меньшей мере 200 пар оснований и содержание GC более 50%, с соотношением наблюдаемого и ожидаемого CpG более 60%. Процент CpG представляет собой отношение нуклеотидных оснований CpG (удвоенное количество CpG) к длине. Отношение наблюдаемого к ожидаемому CpG рассчитывается по формуле:

$$\text{Obs/Exp CpG} = \text{Количество CpG} * N / (\text{Количество C} * \text{Количество G}),$$

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 островка CpG» или «нуклеиновая кислота-мишень, которая не содержит островок CpG» или «отличный от островка островок CpG» относится к нуклеиновой кислоте-мишени, которая не содержит «островка CpG», поскольку термин определен в данном документе. Эта область может быть любой областью, кодируемой геномом млекопитающего (например, человека). В аспектах фраза «нуклеиновая кислота-мишень не содержит островок CpG» относится к областям в нуклеиновой кислоте-мишени, которые не имеют нуклеотидов G и C рядом друг с другом (т.е. динуклеотидов CpG) или которые имеют низкую частоту нуклеотидов G и C рядом друг с другом. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющей область с содержанием динуклеотидов GC менее 50%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 50%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 50%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 50%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 50%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 45%, с соотношением наблюдаемого и ожидаемого CpG менее 55%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC менее 40%, с соотношением наблюдаемого и ожидаемого CpG менее 50%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 1% до 45%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островок относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 1% до 45%, с соотношением наблюдаемого и ожидаемого CpG менее 55%. В аспектах отличный от

островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 1% до 45%, с соотношением наблюдаемого и ожидаемого CpG менее 50%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 5% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 5% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 55%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 5% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 50%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 10% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 60%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 10% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 55%. В аспектах отличный от островка CpG островков относится к областям нуклеиновой кислоты-мишени, имеющим содержание динуклеотидов GC от 10% до 40%, с соотношением наблюдаемого и ожидаемого CpG менее 50%. В аспектах нуклеиновая кислота-мишень, которая не содержит островков CpG, имеет менее 200 пар оснований.

[0188] Варианты осуществления 1-69.

[0189] Вариант осуществления 1. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемую РНК.

[0190] Вариант осуществления 2. Слитый белок по варианту осуществления 1, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

[0191] Вариант осуществления 3. Слитый белок по варианту осуществления 2, отличающийся тем, что домен деметилирования представляет собой домен TET1.

[0192] Вариант осуществления 4. Слитый белок по варианту осуществления 2, отличающийся тем, что домен TET1 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 1, SEQ ID NO: 86 или SEQ ID NO: 97.

[0193] Вариант осуществления 5. Слитый белок по любому из вариантов осуществления 1-4, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, dCpf1, Cas-phi, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен HMB-бокс, домен Wor3, домен OB-складка, домен иммуноглобулина или домен V3.

[0194] Вариант осуществления 6. Слитый белок по варианту осуществления 5, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9.

[0195] Вариант осуществления 7. Слитый белок по любому из вариантов осуществления 1-6, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0196] Вариант осуществления 8. Слитый белок по варианту осуществления 7, отличающийся тем, что линкер XTEN содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 5, SEQ ID NO: 6 или SEQ ID NO: 98.

[0197] Вариант осуществления 9. Слитый белок по любому из вариантов осуществления 1-8, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

[0198] Вариант осуществления 10. Слитый белок, содержащий от N-конца до C-конца, РНК-связывающую последовательность, линкер XTEN и по меньшей мере один активатор транскрипции.

[0199] Вариант осуществления 11. Слитый белок по варианту осуществления 10, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

[0200] Вариант осуществления 12. Слитый белок по варианту осуществления 11, отличающийся тем, что р65 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 13, SEQ ID NO: 14 или SEQ ID NO: 100.

[0201] Вариант осуществления 13. Слитый белок по варианту осуществления 11 или варианту осуществления 12, отличающийся тем, что Rta содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 15 или SEQ ID NO: 16.

[0202] Вариант осуществления 14. Слитый белок по любому из вариантов осуществления 11-13, отличающийся тем, что VP64 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 17 или SEQ ID NO: 18.

[0203] Вариант осуществления 15. Слитый белок по любому из вариантов осуществления 10-14, отличающийся тем, что РНК-связывающая последовательность представляет собой РНК-связывающую последовательность MS2.

[0204] Вариант осуществления 16. Слитый белок по варианту осуществления 15, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

[0205] Вариант осуществления 17. Слитый белок по любому из вариантов осуществления 10-16, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0206] Вариант осуществления 18. Слитый белок по варианту осуществления 10, имеющий аминокислотную последовательность с по меньшей мере 90% идентичностью

последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110.

[0207] Вариант осуществления 19. Слитый белок по любому из вариантов осуществления 10-18, причем слитый белок дополнительно содержит метку эпитопа, пептид 2А, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

[0208] Вариант осуществления 20. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, первый линкер XTEN, фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемую РНК, второй линкер XTEN и активатор транскрипции.

[0209] Вариант осуществления 21. Слитый белок по варианту осуществления 20, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

[0210] Вариант осуществления 22. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемую РНК.

[0211] Вариант осуществления 23. Слитый белок по любому из вариантов осуществления 20-22, дополнительно содержащий последовательность ядерной локализации.

[0212] Вариант осуществления 24. Слитый белок по любому из вариантов осуществления 20-23, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

[0213] Вариант осуществления 25. Слитый белок по варианту осуществления 24, отличающийся тем, что домен деметилирования представляет собой домен TET1.

[0214] Вариант осуществления 26. Слитый белок по любому из вариантов осуществления 20-25, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, dCpf1, Cas-phi, домен лейциновой молнии, домен «крылатой спирали», мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокс, домен Wor3, домен ОВ-складка, домен иммуноглобулина или домен В3.

[0215] Вариант осуществления 27. Слитый белок по варианту осуществления 26, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9.

[0216] Вариант осуществления 28. Слитый белок по любому из вариантов осуществления 20-27, отличающийся тем, что каждый из первого линкера XTEN и второго линкера XTEN независимо содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0217] Вариант осуществления 29. Слитый белок по любому из вариантов осуществления 20-28, причем слитый белок дополнительно содержит метку эпитопа, пептид 2А, метку флуоресцентного белка или комбинацию двух или более из них.

[0218] Вариант осуществления 30. Слитый белок, содержащий аминокислотную

последовательность, имеющую по меньшей мере 90% идентичность последовательности с SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:111, SEQ ID NO:112 или SEQ ID NO:113.

[0219] Вариант осуществления 31. Слитый белок по варианту осуществления 30, содержащий аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 102, SEQ ID NO: 111, SEQ ID NO: 112, или SEQ ID NO:113.

[0220] Вариант осуществления 32. Слитый белок по варианту осуществления 31, содержащий SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 102, SEQ ID NO: 111, SEQ ID NO: 112 или SEQ ID NO: 113.

[0221] Вариант осуществления 33. Способ активации или реактивации последовательности нуклеиновой кислоты-мишени в клетке, включающий: (i) доставку первого полинуклеотида, кодирующего слитый белок по любому из вариантов осуществления 1-32, в клетку, содержащую подвергнутую сайленсингу нуклеиновую кислоту-мишень; и (ii) доставку в клетку второго полинуклеотида, содержащего: (a) sgRNA или (b) cr:tracrRNA; за счет чего активируется или реактивируется последовательность нуклеиновой кислоты-мишени в клетке.

[0222] Вариант осуществления 34. Способ по варианту осуществления 32, отличающийся тем, что последовательность нуклеиновой кислоты-мишени содержит островок CpG.

[0223] Вариант осуществления 35. Способ по варианту осуществления 32, отличающийся тем, что последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

[0224] Вариант осуществления 36. Способ по любому из вариантов осуществления 32-35, отличающийся тем, что второй полинуклеотид содержит sgRNA.

[0225] Вариант осуществления 37. Способ по любому из вариантов осуществления 32-36, отличающийся тем, что sgRNA содержит по меньшей мере одну петлю стебля MS2.

[0226] Вариант осуществления 38. Способ по варианту осуществления 37, отличающийся тем, что sgRNA содержит две петли стебля MS2.

[0227] Вариант осуществления 39. Способ по любому из вариантов осуществления 32-38, отличающийся тем, что второй полинуклеотид кодирует активатор транскрипции.

[0228] Вариант осуществления 40. Способ по варианту осуществления 39, отличающийся тем, что активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них.

[0229] Вариант осуществления 41. Способ по любому из вариантов осуществления 32-40, отличающийся тем, что второй полинуклеотид дополнительно кодирует РНК-связывающую последовательность MS2.

[0230] Вариант осуществления 42. Способ по варианту осуществления 41, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

[0231] Вариант осуществления 43. Способ по любому из вариантов осуществления 32-42, отличающийся тем, что второй полинуклеотид дополнительно кодирует линкер XTEN, метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

[0232] Вариант осуществления 44. Способ по любому из вариантов осуществления 32-43, дополнительно включающий доставку в клетку третьего полинуклеотида, кодирующего второй слитый белок, который содержит активатор транскрипции.

[0233] Вариант осуществления 45. Способ по варианту осуществления 44, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

[0234] Вариант осуществления 46. Способ по варианту осуществления 44 или варианту осуществления 45, отличающийся тем, что второй слитый белок дополнительно содержит РНК-связывающую последовательность MS2.

[0235] Вариант осуществления 47. Способ по варианту осуществления 46, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

[0236] Вариант осуществления 48. Способ по любому из вариантов осуществления 44-47, отличающийся тем, что слитый белок дополнительно содержит линкер XTEN, метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

[0237] Вариант осуществления 49. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

[0238] Вариант осуществления 50. Слитый белок по варианту осуществления 49, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

[0239] Вариант осуществления 51. Слитый белок по варианту осуществления 49, отличающийся тем, что домен деметилирования представляет собой домен TET1.

[0240] Вариант осуществления 52. Слитый белок по варианту осуществления 51, отличающийся тем, что домен TET1 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 1, SEQ ID NO: 86 или SEQ ID NO: 97.

[0241] Вариант осуществления 53. Слитый белок по любому из вариантов осуществления 49-52, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы представляет собой домен цинкового пальца.

[0242] Вариант осуществления 54. Слитый белок по любому из 49-52, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы представляет собой TALE.

[0243] Вариант осуществления 55. Слитый белок по любому из вариантов осуществления 49-54, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0244] Вариант осуществления 56. Слитый белок по варианту осуществления 55, отличающийся тем, что линкер XTEN содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 5, SEQ ID NO: 6 или SEQ ID NO: 98.

[0245] Вариант осуществления 57. Слитый белок по любому из вариантов осуществления 49-56, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

[0246] Вариант осуществления 58. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, первый линкер XTEN, фермент ДНК-эндонуклеазы с дефицитом нуклеазы, второй линкер XTEN и активатор транскрипции.

[0247] Вариант осуществления 59. Слитый белок по варианту осуществления 58, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

[0248] Вариант осуществления 60. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазы с дефицитом нуклеазы.

[0249] Вариант осуществления 61. Слитый белок по любому из вариантов осуществления 58-60, дополнительно содержащий последовательность ядерной локализации.

[0250] Вариант осуществления 62. Слитый белок по любому из вариантов осуществления 58-61, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

[0251] Вариант осуществления 63. Слитый белок по варианту осуществления 62, отличающийся тем, что домен деметилирования представляет собой домен TET1.

[0252] Вариант осуществления 64. Слитый белок по любому из вариантов осуществления 58-63, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы представляет собой домен цинкового пальца.

[0253] Вариант осуществления 65. Слитый белок по любому из 58-63, отличающийся тем, что фермент ДНК-эндонуклеазы с дефицитом нуклеазы представляет собой TALE.

[0254] Вариант осуществления 66. Слитый белок по любому из вариантов осуществления 58-65, отличающийся тем, что каждый из первого линкера XTEN и второго линкера XTEN независимо содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

[0255] Вариант осуществления 67. Слитый белок по любому из вариантов осуществления 58-66, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка или комбинацию двух или более из них.

[0256] Вариант осуществления 68. Способ активации или реактивации последовательности нуклеиновой кислоты-мишени в клетке, включающий доставку полинуклеотида, кодирующего слитый белок по любому из вариантов осуществления 58-

67, в клетку, содержащую нуклеиновую кислоту-мишень; за счет чего активируется или реактивируется последовательность нуклеиновой кислоты-мишени в клетке.

[0257] Вариант осуществления 69. Способ по варианту осуществления 68, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

ПРИМЕРЫ

[0258] Варианты осуществления и аспекты в данном документе дополнительно проиллюстрированы следующими примерами. Примеры предназначены только для иллюстрации вариантов осуществления и аспектов и не должны толковаться как ограничивающие объем данного документа.

[0259] Пример 1

[0260] Сайленсинг генов обратим путем целенаправленного метилирования ДНК.

[0261] Привлекательным свойством редактирования эпигенома является возможность обратить вспять эпигенетические изменения, вызванные искусственными редакторами. Для проверки обратимости сайленсинга генов, опосредованного CRISPRoff, мы сначала использовали глобальные способы блокировки поддержания метилирования ДНК во время клеточного деления. Мы использовали редактирование гена Cas9 для инактивации DNMT1 - основного фермента поддержания метилирования ДНК в клетках млекопитающих - в клетках НЕК293Т с ранее подвергнутыми сайленсингу H2B, CLTA или Snrpn-GFP. Через 9 дней после нокаута DNMT1 60-80% клеток реактивируют экспрессию гена. Потеря DNMT1, который является важным геном, оказывает заметное цитотоксическое действие и исключает возможность нокаута DNMT1 как возможного метода реактивации генов (фиг. 1). Аналогично, обработка клеток низкомолекулярным ингибитором DNMT1, 5-aza-2'-дезоксцитидином (5-aza-dC), реактивировала экспрессию гена CLTA, хотя и с меньшей эффективностью по сравнению с нокаутом DNMT1 (фиг. 2-3). Эти результаты демонстрируют, что деплеции метилирования ДНК достаточно, чтобы обратить вспять сайленсинг гена CRISPRoff. Поэтому мы стремились разработать геноспецифические и программируемые инструменты для реактивации подвергнутых сайленсингу генов CRISPR.

[0262] Пример 2

[0263] Метилирование ДНК цитозинов в диаде цитозин-гуанин может быть активно устранено ферментами семейства TET (транслокация десять-одиннадцать), которые были переназначены для программируемого деметилирования промоторов генов человека для активации генов. Мы проверили, можем ли мы повторно активировать подвергнутые сайленсингу гены CRISPRoff путем целенаправленного деметилирования ДНК CLTA, гена, который мы подвергали сайленсингу более 1 года. Первоначально мы использовали ранее описанный слитый продукт dCas9 с каталитическим доменом ДНК-деметилязы TET1 (TETv1) (Liu et al., Cell, 167-233-247 (2016)). Мы совместно трансфицировали плазмиды, экспрессирующие TETv1, и sgRNA, нацеленные на промотор CLTA, и измеряли уровни белка CLTA (GFP) с течением времени. (фиг. 4-5). Наши результаты демонстрируют

направленное деметилирование ДНК с помощью реактивированной экспрессии гена TETv1, однако через 28 дней после трансфекции только около 20% трансфицированных клеток сохраняют экспрессию CLTA в соответствии с переменной реактивацией, типичной для предыдущих исследований. (фиг. 6) Для улучшения реактивации, мы оптимизировали слитые белки, кодируя линкеры XTEN между dCas9 и TET1, и перемещали TET1 на N-конце dCas9. Размещение TET1 на N-конце с линкером XTEN16 из 16 аминокислот (TETv3) улучшало реактивацию CLTA до приблизительно 50% клеток. Более того, разделение TET1 и dCas9 с помощью линкера XTEN80 из 80 аминокислот (TETv4) приводило к стабильной реактивации CLTA более чем в 70% клеток. Реактивация CLTA была стабильной в течение по меньшей мере 28 дней после трансфекции (фиг. 6-8). Реактивация гена была достигнута в 60% клеток, трансфицированных TETv4, с помощью одной последовательности sgRNA, но была улучшена за счет объединения трех sgRNA на промоторе гена (фиг. 7).

[0264] Чтобы оценить степень деметилирования ДНК подвергнутого сайленсингу гена, мы проводили бисульфитное секвенирование локуса *CLTA* до и после реактивации, опосредованной dCas9-TET. Мы наблюдали высокие уровни метилирования ДНК на всем протяжении CLTA CGI после CRISPRoff-опосредованного сайленсинга, включая > 400 п.н. ниже сайтов связывания sgRNA. (фиг. 9А-9В) После TET1-опосредованной реактивации гена CGI деметилируется почти до завершения, что коррелирует с полной реактивацией экспрессии CLTA (фиг. 9А).

[0265] Мы наблюдали, что реактивация CLTA неизменно достигает пика и стабилизируется, начиная с дня 9 после обработки TET1. (фиг. 6) Мы предположили, что экспрессия генов может быть реактивирована в более ранние временные точки путем рекрутирования доменов активатора транскрипции в TET1v4. Стремясь модулировать кинетику реактивации генов, мы разработали систему под названием CRISPRon, состоящую из TETv4, ранее описанной модифицированной sgRNA, которая кодирует две последовательности стебля MS2, и белка оболочки MS2 (MCP), слитого с различными комбинациями транскрипционных трансактиваторных доменов VP64, p65 (p65-AD) и Rta (Konermann et al., 2015a) (фиг. 10-11). Сначала мы подтвердили, что совместная экспрессия слитых белков dCas9 и трансактиватора MCP может повышать экспрессию генов в отсутствие TET1, мы сливали домены с белком оболочки MS2 (MCP) и подвергали рекрутингу слитые продукты с dCas9, нацеленные на промотор эндогенно экспрессируемого CLTA с sgRNA, кодирующими петли MS2. Через два дня после трансфекции dCas9, слитых продуктов MCP и sgRNA мы обнаруживали повышенную эндогенную экспрессию гена CLTA с каждой комбинацией трансактиваторов с максимальной реактивацией с использованием VPR и p65-Rta (фиг. 12), указывая на то, что эти белки функциональны для рекрутирования транскрипционного аппарата.

[0266] Затем мы экспрессировали отрицательный контроль (отр. контр.) или sgRNA, нацеленные на *CLTA* (sg-A), вместе с различными комбинациями CRISPRon или TETv4 только в клетках с подвергнутым сайленсингу CLTA и отслеживали экспрессию CLTA с течением времени. Неожиданно мы обнаружили, что некоторые комбинации CRISPRon,

такие как TETv4 с p65-Rta и TETv4 с VPR, надежно реактивировали экспрессию CLTA в течение 2 дней. Между тем, TETv4 показал небольшую реактивацию гена в этой временной точке (**фиг. 13** и **17**). Затем мы совместно привлекли трансаktиваторы и TETv4 к промотору подвергнутого сайленсингу CLTA с помощью CRISPRoff. Через два дня после трансфекции экспрессия CLTA реактивируется только в присутствии TETv4 и трансаktиваторов (**фиг. 13** и **14**). Каждая комбинация трансаktиваторов повышает долю клеток с реактивацией CLTA на разных уровнях, в диапазоне от 2 до 46 раз, по сравнению только с TETv4, при этом VPR и p65-Rta вызывают самые высокие уровни экспрессии CLTA. Через восемь дней после трансфекции рекрутинг однокомпонентных Rta или VP64-p65 приводит к максимальному повышению доли реактивированных клеток по сравнению с другими трансаktиваторами (**FIGS. 14** и **15A**). TETv4 и коактиватор sgRNA присутствуют в клетках при низких уровнях в этой временной точке (<10% клеток), что означает, что повышенная экспрессия реактивированного гена с использованием TETv4 с p65-Rta или VP64-p65 наследуется и запоминается клетками. Через 28 дней после трансфекции медиана флуоресценции реактивированного CLTA-GFP была значительно выше при использовании CRISPRon-комбинаций TETv4 с Rta и TETv4 с p65-Rta по сравнению только с TETv4 (**фиг. 15B**). Мы не обнаруживали экспрессию слитого белка TETv4 или MCP в этой временной точке. В качестве дополнительного контроля коэкспрессия слитых трансаktиваторов MCP с dCas9 (без TET) или одиночного слияния dCas9-VPR показала только временную активацию CLTA, и через 10 дней после трансфекции уровни CLTA возвращаются к состоянию сайленсинга (**фиг. 18**). Вместе эти результаты показывают, что наши оптимизированные слитые белки TET1-dCas9 могут устойчиво реактивировать подвергнутые сайленсингу гены CRISPRoff, как форму транскрипционной памяти, а динамику реактивации можно дополнительно модулировать с помощью наших комбинаций CRISPRon. В совокупности эти данные подчеркивают нашу способность модулировать динамику реактивации подвергнутых сайленсингу генов CRISPR, и кодировать клеточную память об экспрессии генов, аналогичную CRISPRa типа «бей и беги».

[0267] Пример 3

[0268] Сайленсинг и реактивация генов, в которых отсутствует аннотация CpG

[0269] Чтобы подтвердить наше наблюдение, что CRISPRoff может отключать гены без аннотированных CGI, мы эндогенно помечали пять генов без аннотированных CGI в HEK293T с помощью mNeonGreen (mNG) и оценивали устойчивый сайленсинг с помощью CRISPRoff. Через 9 дней после трансфекции мы обнаруживали высокий процент клеток, которые выключили DYNC2LI1, LAMP2, MYL6 и VPS25. Сайленсинг DYNC2LI1 и LAMP2 остается стабильным в течение 14 дней после трансфекции, а MYL6 и VPS25 обнаруживают дефекты клеточного роста после нокдауна. Трансфекция мутанта CRISPRoff Dnmt3A не поддерживает сайленсинг генов, и, таким образом, наблюдаемый устойчивый фенотип зависит от метилирования ДНК. В отличие от этого, трансфекция CRISPRoff в клетки CALD1-mNG не приводила к сайленсингу с помощью CRISPRoff или мутанта CRISPRoff,

что позволяет предположить, что ген не поддается зависимому от метилирования ДНК редактированию эпигенома типа «бей и беги».

[0270] Мы выделяли клетки, которые выключили LAMP2, DYNC2LI1 и MYL6 с помощью CRISPRoff, и определили статус метилирования ДНК промоторов с помощью бисульфитного секвенирования. Анализ цитозинов в контексте CG показал высокую степень метилирования в подвергнутых сайленсингу клетках. Кроме того, мы обработали клетки DYNC2LI1 и LAMP2-off с помощью TETv4, и приблизительно 70% клеток реактивировали подвергнутый сайленсингу ген через 14 дней после трансфекции TETv4 (фиг. 16).

[0271] Материалы и способы для примеров

[0272] Разработка и конструкция плазмид

[0273] Конструкцию TETv1 конструировали путем ПЦР-амплификации последовательности dCas9-TET1CD из Fuw-dCas9-Tet1CD (Addgene № 84475) и собирали в CAG-экспрессирующую плазмиду. Линкерные последовательности XTEN были опубликованы ранее (Schellenberger et al.). Все слитые белки CRISPRoff и TET1 включают BFP либо в виде прямого слитого продукта, либо с последовательностью расщепления P2A для измерения эффективности трансфекции с помощью проточной цитометрии. Последовательность dSaCas9 (D10A, N508A) амплифицировали с помощью ПЦР из pX603 (Addgene № 61594), а последовательность dLbCas12a амплифицировали с помощью ПЦР из Tak et al. VP64, p65 и Rta амплифицировали с помощью ПЦР из SP-dCas9-VPR (Addgene № 63798). Лентивирусный репортер GAPDH-Snrpn-GFP получали от Addgene № 70148 (Liu et al., 2016; Stelzer et al., 2015).

[0274] Плазмиды sgRNA конструировали путем рестрикционного клонирования протоспейсеров ниже промотора U6 с использованием сайтов разрезания BstXI и BspI, как описано ранее. Плазмиды экспрессии sgRNA также экспрессируют маркер T2A-mCherry для измерения эффективности трансфекции. Последовательности sgRNA, используемые для экспериментов CRISPRoff и CRISPRon, перечислены в таблице 1. Последовательности sgRNA выбирали на основе нашего предыдущего алгоритма для прогнозирования активных sgRNA CRISPRi (Horlbeck et al., 2016).

[0275] Плазмиды MS2 конструировали, сначала перенося кассету промотор mU6-sgRNA-EF1a-пурамицин-T2A-mCherry в отличный от лентивирусного вектор путем рестрикционного клонирования. Кассету MCP-XTEN80-NLS-(трансактиваторный домен)-2xP2A заказывали как четыре gBlocks (IDT) и клонирована в вышеупомянутую отличную от лентивирусной плазмиду с помощью сборки Gibson. Последовательность петель sgRNA-MS2 разрабатывали на основе системы SAM (Konermann et al., 2015b) с сайтами рестрикции BstXI и BspI, включенными из нашей предыдущей разработки экспрессии sgRNA mU6 (Addgene № 84832). Последовательность ДНК, кодирующая каркас MS2-sgRNA, представляет собой SEQ ID NO:117. Для конструирования трансактиваторных плазмид каждый домен или комбинацию доменов амплифицировали с помощью ПЦР и клонировали с помощью сборки Gibson в плазмиду, кодирующую sgRNA и белок оболочки MS2 (MCP).

Направляющие последовательности клонировали путем двойного расщепления и лигирования отоженных олигонуклеотидов, как описано ранее.

[0276] Все конструкции mRNA синтезировали с использованием набора mMESSAGE mMachinTM T7 Ultra Transcription (Thermo Fisher Scientific). Последовательность промотора T7 (SEQ ID NO:118) сначала клонировали выше последовательности CRISPRoff. Последовательность T7-CRISPRoff амплифицировали с помощью ПЦР и использовали в качестве матрицы для реакций синтеза *in vitro*. Следуя протоколу синтеза производителя, реакционную смесь очищали экстракцией хлороформом и осаждением изопропанолом.

[0277] Культура клеток, трансфекция ДНК и проточная цитометрия

[0278] Все клеточные линии культивировали при 37°C в инкубаторах для тканевых культур с 5% CO₂. Клетки HEK293T (жен. пол), HeLa (жен. пол) и U2OS (жен. пол) культивировали в модифицированной по Дульбекко среде Игла (DMEM) в 10% FBS (HyClone), 100 единиц/мл стрептомицина, 100 мкг/мл пенициллина и 2 mM глутамина. Клетки K562 (жен. пол) поддерживали в RPMI-1640 с 25 mM HEPES и 2,0 г/л NaHCO₃ в 10% FBS, 2 mM глутамина, 100 единиц/мл стрептомицина и 100 мг/мл пенициллина. iPSCs WTC Gen1c (муж. пол) культивировали в среде mTESR (STEMCELL Technologies) в условиях отсутствия питающих клеток на матрикеле со сниженным содержанием фактора роста (BD Biosciences). Клетки пассировали с использованием Accutase (STEMCELL Technologies) и высевали на планшеты, покрытые матрикелем, в среду mTESR с добавлением ингибитора p16-Rho-ассоциированной спиральной киназы (ROCK) Y-27632 (10 мкМ; Selleckchem).

[0279] Lentivirusные частицы получали путем трансфекции стандартных упаковочных векторов в HEK293T с использованием трансфекционного реагента *TransIT-LT1* (Mirus, MIR2306). Среду меняли через 24 часа после трансфекции полной DMEM с добавлением 15 mM HEPES. Вирусные супернатанты собирали через 48-60 часов после трансфекции и фильтровали через шприцевой фильтр 0,45 мкм PVDF. Lentivirusные инфекции включали полибрен (8 мкг/мл).

[0280] CRISPRon

[0281] Все эксперименты с CRISPRon проводили в 24-луночных планшетах. Вкратце, 1×10^5 клеток подвергнутым CLTA-GFP сайленсингу HEK293T высевали в каждую лунку. Когда на следующий день клетки достигали 60-80% конfluenceности, клетки трансфицировали 500 нг плазмиды dCas9 (dCas9 или TETv1-4) и 300 нг плазмиды sgRNA-трансактиватора (только sgRNA, VP64, p65, Rta, VP64-p65, p65-Rta или VPR). Клетки контролировали в отношении предмет экспрессии BFP (dCas9 или TETv1-4) и mCherry (направляющей трансактиватор) через 24 часа после трансфекции. Через два дня после трансфекции $7,5 \times 10^4$ клеток BFP и mCherry с двойной положительной реакцией сортировали с помощью сортировщика BD FACSAria Fusion. После сортировки клеткам давали возможность восстановиться в течение 4 дней, а затем каждые 2-3 дня анализировали с помощью проточной цитометрии на цитометре Attune NxT (Thermo Fisher Scientific). Все данные проточной цитометрии анализировали с помощью Flowjo.

[0282] РНК-секвенирование

[0283] Клетки HEK293T, которые сохраняли 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 с планшетов с PBS, центрифугировали при $500 \times g$ в течение 5 мин и повторно промывали PBS. Общую РНК экстрагировали с использованием Direct-zol RNA MiniPrep (Zymo R2051). Приготовление библиотеки осуществляли с использованием набора для подготовки библиотеки TruSeq Stranded mRNA (Illumina RS-111-2101), начиная с 1000 нг общей РНК. Окончательные библиотеки оценивали с использованием биоанализатора 2100 (Agilent), количественно определяли с использованием набора для анализа HS-анализа Qubit dsDNA (Thermo Fisher Scientific) и секвенировали как показания с одним концом из 50 пар оснований на HiSeq 4000 (Illumina). Для обработки показаний секвенирования линкерные последовательности (SEQ ID NO: 119) удаляли с использованием FASTX-клипера (FASTX-Toolkit). Затем показания сопоставляли с геномом человека (GRCh37) с использованием выравнивателя STAR (выравнивание сплайсированных транскриптов по эталону, версия 2.5) с аннотацией транскриптома Gencode Gene V24lift37. Количественную оценку показателей проводили с помощью featureCounts (Liao et al., 2014). Все последующие анализы выполняли с помощью 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 нормализации транскриптов генов на миллион (TPM) для экспериментальных образцов со средним значением TPM контрольных (нецелевых) образцов. Анализ дифференциальной экспрессии проводили с использованием DESeq2 (Love et al., 2014).

[0284] Количественная ПЦР

[0285] Для измерений количественной ПЦР (qPCR) общую РНК сначала экстрагировали из клеток с помощью набора RNeasy Micro (Qiagen). Обратную транскрипцию 1 мкг общей РНК выполняли с использованием набора для обратной транскриптазы SuperscriptTM III (Thermo Fisher Scientific), дополненного ингибитором рекомбинантной рибонуклеазы RNaseOutTM (Thermo Fisher Scientific). Обратную транскрипцию праймировали с использованием oligo(dT)₂₀. Реакции количественной ПЦР готовили с помощью KAPA SYBR FAST qPCR Master Mix (2X) и проводили на приборе LightCycler 480 (Roche). Последовательности праймеров для экспериментов количественной ПЦР приведены в таблице 2.

[0286] Бисульфитное ПЦР-секвенирование

[0287] Для анализа метилирования CLTA CGI с помощью FACS выделяли $\sim 2 \times 10^6$ подвергнутых сайленсингу с помощью CRISPRoff клеток и TET-реактивированных клеток. Геномную ДНК выделяли из клеток в соответствии с инструкциями производителя с использованием мини-набора PureLink Genomic DNA (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 использованием

EpiMark Hot Start Taq (NEB) методом вложенной ПЦР (Liu et al., 2016). Ампликоны очищали в геле с использованием набора для гель-восстановления ДНК (Zymo) и повторно амплифицировали методом ПЦР с использованием EpiMark Hot Start Taq. Ампликоны клонировали в вектор pCR2.1 TOPO в соответствии с инструкциями производителя с использованием набора для клонирования TOPO TA (Invitrogen). Продукты клонирования трансформировали в клетки Stellar *E. coli* (Takara) и высевали на сине-белые карбенициллиновые чашки. По 20 колоний отбирали для каждого условия и секвенировали с помощью секвенирования по Сенгеру. Последовательности праймеров для бисульфитной ПЦР приведены в таблице 2. Последовательности праймеров для амплификации фрагмента GAPDH-Snrpn получали от Liu et al.

[0288] Редактирование генома Cas9 и обработки 5-aza-dC

[0289] Lentivirusные частицы, экспрессирующие Cas9 из *S. pyogenes*, трансдуцировали в клетки HEK293T, которые содержат подвергнутые сайленсингу *Snrpn-GFP* с помощью CRISPRoff или *GFP*-меченые *CLTA* и *H2B*. Cas9-экспрессирующие клетки, меченые флуоресценцией BFP в лентивирусном векторе, подвергали FACS-сортировке. Для инактивированного DNMT1 лентивирусные частицы, экспрессирующие sgRNA, которая нацелена на *DNMT1*, инфицировали в клеточные линии. Реактивацию подвергнутых сайленсингу генов оценивали по активации GFP, измеряемой с помощью проточной цитометрии. Последнюю временную точку брали через 9 дней после инфицирования sgRNA, поскольку жизнеспособность клеток после этой временной точки резко снижалась.

[0290] Для обработки 5-aza-dC в каждую лунку 24-луночного планшета высевали 1×10^5 подвергнутых сайленсингу с помощью CRISPRoff CLTA-GFP HEK293T клеток. Через 24 часа среду аспирировали и заменяли средой с добавлением водного 5-Aza-2'-дезоксцитидина (5-aza-dC) до конечного объема 500 мл на лунку. На следующий день среду, содержащую 5-aza-dC, аспирировали, клетки отделяли и анализировали в отношении жизнеспособности и активации GFP на проточном цитометре Attune NxT (Thermo Fisher Scientific). Затем клетки пассировали на свежей среде каждые 2-3 дня и анализировали на цитометре Attune.

[0291] Несмотря на то, что 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 вариации, изменения и замены будут появляться для специалистов в данной области техники без отклонения от настоящего изобретения. Необходимо понимать, что различные альтернативы вариантам осуществления настоящего изобретения, описанным в данном документе, могут быть использованы при практическом осуществлении настоящего изобретения.

[0292] Названия разделов, используемые в данном документе, представлены лишь для организационных целей, и не предполагают ограничивать описываемый предмет изобретения. Все документы или части документов, цитируемые в данной заявке,

включающие без ограничения патенты, заявки на патенты, статьи, книги и монографии, явным образом включены в данный документ посредством ссылки во всей своей полноте для любой цели.

[0293] Литературные источники

[0294] Adamson, et al. (2016). A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* 167, 1867-1882.e21. Alanis-Lobato, et al (2020). Frequent loss-of-heterozygosity in CRISPR-Cas9-edited early human embryos. *BioRxiv* 2020.06.05.135913. Amabile, et al (2016). Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* 167, 219-232.e14. Anzalone, et al (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824-844. Blomen, et al (2015). Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092-1096. Bothmer, et al. (2020). Detection and Modulation of DNA Translocations During Multi-Gene Genome Editing in T Cells. *CRISPR J.* 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. Cheng, et al (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163-1171. Choudhury, et al (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget* 7, 46545-46556. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010-1022. Dede, et al (2020). Multiplex enCas12a screens show functional buffering by paralogs is systematically absent from genome-wide CRISPR/Cas9 knockout screens. *BioRxiv* 2020.05.18.102764. Doench, J.G. (2018). Am I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet.* 19, 67-80. El-Brolosy, M.A., and Stainier, D.Y.R. (2017). Genetic compensation: A phenomenon in search of mechanisms. *PLoS Genet.* 13. ENCODE Project Consortium, Moore, J.E., et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* 583, 699-710. Ferrari, S., et al (2011). Retinitis Pigmentosa: Genes and Disease Mechanisms. *Curr. Genomics* 12, 238-249. Fulco, C.P., et al (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 354, 769-773. Gilbert, et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442-451. Gilbert, L.A., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647-661. Gong, G., et al (2004). Genetic dissection of myocilin glaucoma. *Hum. Mol. Genet.* 13 Spec No 1, R91-102. Halmai, et al. (2020). Artificial escape from XCI by DNA methylation editing of the CDKL5 gene. *Nucleic Acids Res.* 48, 2372-2387. Hanna, R.E., and Doench, J.G. (2020). Design and analysis of CRISPR-Cas experiments. *Nat. Biotechnol.* 38, 813-823. Hart, T., et al (2014). Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Mol. Syst. Biol.* 10, 733. He, Y., et al. (2020). Spatiotemporal DNA methylome dynamics of the developing mouse fetus. *Nature* 583, 752-759. Hilton, et al (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510-517. 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., et al. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *ELife* 5, e19760. Ihry, R.J., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939-946. Jia, D., et al (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449, 248-251. Josipović, G., et al. (2019). Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system. *Nucleic Acids Res.* 47, 9637-9657. Jost, M., et al. (2020). Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nat. Biotechnol.* 38, 355-364. Kearns, et al (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Dev. Camb. Engl.* 141, 219-223. Knott, G.J., and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science* 361, 866-869. Konermann, et al (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500, 472-476. Konermann, et al. (2015a). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583-588. Konermann, et al. (2015b). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583-588. Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765-771. La Spada, A.R., and Taylor, J.P. (2010). Repeat expansion disease: Progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* 11, 247-258. Leonetti, et al (2016a). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci. U. S. A.* 113, E3501-3508. Leonetti, et al (2016b). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci.* 113, E3501-E3508. Li, X.-L., 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. Liang, D., et al. (2020). Frequent Gene Conversion in Human Embryos Induced by Double Strand Breaks. *BioRxiv* 2020.06.19.162214. Liao, et al (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930. Liu, et al (2016). Editing DNA Methylation in the Mammalian Genome. *Cell* 167, 233-247.e17. Liu, et al. (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell* 172, 979-992.e6. 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, et al (2013a). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* 10, 977-979. Maeder, et al. (2013b). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat. Biotechnol.* 31, 1137-1142. Mali, P., et al (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833-838. Meyers, et al. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* 49, 1779-1784. Michlits, et al. (2020). Multilayered VBC score predicts sgRNAs that efficiently generate loss-of-function alleles. *Nat. Methods* 17, 708-716. Mlambo, et al (2018). Designer epigenome modifiers enable robust and sustained gene

silencing in clinically relevant human cells. *Nucleic Acids Res.* 46, 4456-4468. Morita, et al. (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nat. Biotechnol.* 34, 1060-1065. Morita, et al (2020). Synergistic Upregulation of Target Genes by TET1 and VP64 in the dCas9-SunTag Platform. *Int. J. Mol. Sci.* 21. O'Geen, et al (2017). dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.* 45, 9901-9916. O'Geen, H., et al (2019). Ezh2-dCas9 and KRAB-dCas9 enable engineering of epigenetic memory in a context-dependent manner. *Epigenetics Chromatin* 12, 26. Perez-Pinera, P., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973-976. Replogle, et al. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* 38, 954-961. Roth, T.L., et al. (2018). Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* 559, 405-409. Schellenberger, 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. Schumann, et al. (2015). Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci.* 112, 10437-10442. Shalem, et al (2015). High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* 16, 299-311. Shifrut, et al (2018). Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* 175, 1958-1971.e15. Stelzer, et al (2015). Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution. *Cell* 163, 218-229. Tak, et al (2017). Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods* 14, 1163-1166. Tarjan, et al (2019). Epigenome editing strategies for the functional annotation of CTCF insulators. *Nat. Commun.* 10, 4258. Tian, et al. (2019). CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron* 104, 239-255.e12. Veitia, R.A., Caburet, S., and Birchler, J.A. (2018). Mechanisms of Mendelian dominance. *Clin. Genet.* 93, 419-428. Wang, et al (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096-1101. Xu, X., and Qi, L.S. (2019). A CRISPR-dCas Toolbox for Genetic Engineering and Synthetic Biology. *J. Mol. Biol.* 431, 34-47. Zhang, et al. (2018). Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature* 554, 387-391. Zuccaro, et al. (2020). Reading frame restoration at the EYS locus, and allele-specific chromosome removal after Cas9 cleavage in human embryos. *BioRxiv* 2020.06.17.149237.

[0295] Перечень последовательностей без соблюдения формальных требований

[0296] В последовательностях, приведенных в данном документе, специалисту в данной области техники будет понятно, что метионин (M) может присутствовать на N-конце белка для инициации трансляции. Таким образом, последовательности, описанные в данном документе, могут необязательно дополнительно содержать метионин на N-конце.

[0297] SEQ ID NO: 1=TET1 (UniProt: Q8NFU7)

MSRSRHARPSRLVRKEDVNKKKNSQLRKTTKGANKNVASVKTLSPGKLGKLI
QERDVKKKTEPKPPVVRSLSTRAGAARMNLDRTVLFQNPESLTCNGFTMALRSTSL
RRLSQPPLVVAKSKKVPLSKGLEKQHDCDYKILPALGVKHSENDSVPMQDTQVLPDIET

LIGVQNPSLLKGSQETTQFWSQRVEDSKINIPTHSGPAAEILPGPLEGTRCGEGLFSEET
 LNDTSGSPKMFAQDTVCAFPQRATPKVTSQGNPSIQLEELGSRVESLKLSDSYLDPIKSE
 HDCYPTSSLNKVIPDLNLRNCLALGGSTSPTSVIKFLLAGSKQATLGAKPDHQEAFEATA
 NQQEVSDTTSFLGQAFGAIPHQWELPGADPVHGEALGETPDLPEIPGAIPVQGEVFGTIL
 DQQETLGMGSVVPDLPVFLPVPPNPIATFNAPSKWPEPQSTVSYGLAVQGAIQILPLGS
 GHTPQSSSNSEKNSLPPVMAISNVENEKQVHISFLPANTQGFPLAPERGLFHASLGIAQLS
 QAGPSKSDRGSSQVSVTSTVHVVNNTTVVTMPVPMVSTSSSSYTLLPTLEKKKRKRCGV
 CEPCCQKTNCGECTYCKNRKNSHQICKKRKCEELKKKPSVVVPLEVIKENKRPQREKKP
 KVLKADFDNKPVNGPKSESMDYSRCGHGEEQKLELNPHTVENVTKNEDSMTGIEVEK
 WTQNKKSQLTDHVKGDFSANVPEAEKSKNSEVDKKRTKSPKLFVQTVRNGIKHVHCLP
 AETNVSFKKFNIEEFGKTLENNSYKFLKDTANHKNAMSSVATDMSCDHLKGRSNVLF
 QQPGFNCSSIPHSSHSIINHHAISHNEGDQPKTPENIPSKEPKDGSVPQPSLLSLMKDRRLT
 LEQVVAIEALTQLSEAPSENSSPSKSEKDEESEQRTASLLNSCKAILYTVRKDLQDPNLQ
 GEPKLNHCPSLEKQSSCNTVVFNGQTTLSNSHINSATNQASTKSHEYSKVTNSLSLFI
 KSNSSKIDTNKSIAQGIITLDNCSNDLHQLPPRNNEVEYCNQLLDSSKKLDSDDLSCQDA
 THTQIEEDVATQLTQLASIINIKINYPEDKKVESTPTSLVTCNVQQKYNQEKGTIQKPPS
 SVHNNHGSSLTKQKNPTQKKTKSTPSRDRRKKKPTVVSQENDRQKWEKLSYMYGTIC
 DIWIASKFQNFQFCPHDFPTVFGKISSSTKIWKPLAQTRSIMQPKTVFPPLTQIKLQRYPE
 SAEKVKVEPLDSLFLHLKTESNGKAFTDKAYNSQVQLTVNANQKAHPLTQPSSPPNQ
 CANVMAGDDQIRFQQVVKEQLMHQRLPTLPGISHETPLPESALTLRNVNVVCSGGITVV
 STKSEEEVCSSTFGTSEFSTVDSAQKNFNDYAMNFFTNPTKNLVSITKDELPTCCLDRV
 IQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGNARIEIVVYTGKEGKSSHGCPIAKW
 VLRRSSDEEKVLCLVRQRTGHHCTAVMVVLMVWDGIPLMADRLYTELTENLKSYN
 GHPTDRRCTLNENRTCTCQIDPETCGASFSFGCSWSMYFNGCKFGRSPSPRRFRIDPSSP
 LHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARECRLGSKEGRPFSGVTAC
 LDFCAHPHRDIHNMNNGSTVVCTLTREDNRS LGVIPQDEQLHVLPLYKLSDTDEFGSKE
 GMEAKIKSGAIEVLAPRRKKRTCFTQPVPRSGKKRAAMMTEVLAHKIRAVEKKPIPRIK
 RKNNSTTTNNSKPSLPTLGSNTETVQPEVKSETEPHFILKSSDNTKTYSLMPSAPHPVKE
 ASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGRLSGANAAAADGPGISQ
 LGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLTSPQDLASSPMEEDEQHS
 EADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDEHIFLDANIGGVAIAPAHGSVLIECARR
 ELHATTPVEHPNRNHPTRLSLVFYQHKNLNKPQHGFELNKIKFEAKEAKNKKMKASEQ
 KDQAANEGPEQSSEVNE LNQIPSHKALTLTHDNVTVSPYALTHVAGPYNHWV

[0298] SEQ ID NO: 2=TET2 (UniProt Q6N021)

YGIPCMKGSQNSRVSPDFTQESRGYSKCLQNGGIKRTVSEPSLSGLLQIKKLKQD
 QKANGERRNFGVSQERNPGESSQPNVSDLSDKKESVSSVAQENAVKDFTSFSTHNCSGP
 ENPELQILNEQEGKSANYHDKNIVLLKNKAVLMPNGATVSASSVEHETHGELLEKTL SQY
 YPDCVSIQVQKTTSHINAINSQATNELSCEITHPSHTSGQINSAQTSNSELPPKPAVVSEA
 CDADDADNASKLAAMLN TCSFQKPEQLQQQKSVFEICPSAENNIQGTTKLASGEEFCS
 GSSSNLQAPGGSSERYLKQNE MNGAYFKQSSVFTKDSFSATTT PPPPSQLLLSPPPPLPQV

PQLPSEGKSTLNGGVLEEHHHYPNQSNTTLLREVKIEGKPEAPPSQSPNPSTHVCSPSPML
SERPQNNCVNRNDIQTAGTMTVPLCSEKTRPMSEHLKHNPPIFGSSGELQDNCQQLMRN
KEQEILKGRDKEQTRDLVPPTQHLYLKP GWIELKAPRFHQAESHKRN EASLPSILQYQPN
LSNQMTSKQYTGNSNMPGGLPRQAYTQKTTQLEHKSQMYQVEMNQGQSQGTVDQHL
QFQKPSHQVHF SKTDHLPKAHVQSLCGTRFHFQQRADSQTEKLMSPVLKQHLNQQASE
TEPFSNSHLLQHKPHKQAAQTQPSQSSHL PQNQQQQK LQIKNKEEILQTFPHQSNND
QQREGSFFGQTKVEECFHGENQYSKSSEFETHNVQMGLEEVQINRRNSPYSQTMKSSA
CKIQVSCSNTHLVSENKEQTTHPEL FAGNK TQNLHMQYFPNNVIPKQDLLHRCFQEQ
EQKSQQASV LQGYKNRNQDMSGQQA AQLAQQRYLIHNHANVFPVPDQGG SHTQTPPQ
KDTQKHAALRW HLLQKQEQQQTQQPQTESCHSQMHRPIKVEPGCKPHACMHTAPPEN
KTWKKVTKQENPPASCDNVQQKSIETMEQHLKQFHAKSLFDHKAL TLKSQKQVKVEM
SGPVTVLTRQTTAAELDSHTPALEQQTTSSEKTPTKRTAASVLN NFIESPSKLLDTPIKNL
LDTPVKTYDFPSCRCVEQIIEKDEGPFYTHLGAGPNVAAIREIMEERFGQK GKAIRIERV
IYTGKEGKSSQGCPIAKWVRRSSSEEKLLCLVRERAGHTCEAAVIVILILVWEGIPLSLA
DKLYSEL TETLRKYGTLNRRCALNEERTCACQGLDPETCGASFSGCSWSMYNGCK
FARSKIPRKFKLLGDDPK EEEKLESHLQNLSTLMAPTYKKLAPDAYNNQIEYEHRAPEC
RLGLKEGRPFSGVTACLDFCAHAHRDLHNMQNGSTLVCTLTREDNREFGGKPEDEQLH
VLPLYKVS DVDEFGSVEAQEEKRSGAIQVLSFRRKVRMLAEPVKTCRQRKLEAKKA
AAEKLSSLENSSNKNEKEKSAPSRTKQ TENASQAKQLAELLRLSGPVMQSSQPQPLQK
QPPQPQQQRPQQQPHHPQTESVNSYSASGSTNPYMRRPNPVSYPNSSHTSDIYGSTS
PMNFYSTSSQAAGSYLNSSNPMNPYPGLLNQNTQYPSYQCNGNLSVDNCSPYLGSYSPO
SQPMDLYRYPSQDPLSKLSLPIHTLYQPRFGNSQSFTSKYLYGNQNMQGDGFSSCTIR
PNVHHVGLPPYP THEMDGHFMGAT SRLPPNLSNPNDYKNGEHHSPSHIHNYS AAP
GMFNSSLHALHLQNKENDMLSHTANGLSKMLPALNHDRTACVQGGLHKLSDANGQEK
QPLALVQGVASGAEDNDEVWSDSEQSFLDPDIGGVA VAPTHGSILIECAKRELHATTPL
KNPNRNHPTRISLVFYQH KSMNEPKHGLALWEAKMAEKAREKEEECEKYGPDYVPQK
SHGKKVKREPAEPHETSEPTYLRFIKSLAERTMSVTTDSTVTTSPYAFTRVTGPYNRYI

[0299] SEQ ID NO:3=TET3 (Uniprot O43151)

MSQFQVPLAVQPDLPGLYDFPQRQVMVGSFPGSGLSMAGSESQLRGGGDGRKK
RKRCGTCEPCRRENC GACTSCTNRRTHQICKLRKCEVLK KKVGLLKEVEIKAGEGAGP
WGQGA AVKTGSELSPVDGPVPGQMDSGPVYHGDSRQLSASGVPVNGAREPAGPSLLGT
GGPWRVDQKPDWEAAPGPAHTARLEDAHDLVAFSAVAEAVSSYGALSTRLYETFNRE
MSREAGNNSRGP RP GPEGCSAGSELDLTLQ TALALARHGMKPPNCNCDGPECPDYLEW
LEGKIKSVVMEGGEERPRLPGPLPPGEAGLPAPSTRPLLSSEVPQISPQEGLPLSQSALSIA
KEKNISLQTAIAIEALTQLSSALPQPSHSTPQASCPLPEALSPAPFRSPQSYLRAPSWPVV
PPEEHSSFAPDSSAFPPATPRTEFPEAWGTDTPPATPRSSWPMRPSDPMAELEQLL GSA
SDYIQSVFKRPEALPTKPKVKVEAPSSSPAPAPSPVLQREAPTPSSEP DTHQKAQTALQQ
HLHHKRS LFLEQVHDT SFPAPSEPSAPGWWPPPSSVPRLPDRPPKEKKKKLPTPAGGPV
GTEKAAPGIKPSVRKPIQIKKSRPREAQPLFPPVRQIVLEGLRSPASQEVQAHPAPLPASQ
GSAVPLPPEPSLALFAPSPSRDSL PPTQEMRSPSPMTALQPGSTGPLPPADDKLEELIRQF

EAEFGDSFGLPGPPSVPIQDPENQQTCLPAPESPFATRSPKQIKIESSGAVTVLSTTCFHSE
 EGGQEATPTKAENPLTPTLSGFLESPLKYLDTPTKSLLDTPAKRAQAEFPTCDCVEQIVE
 KDEGPYYTHLGSPTVASIRELMEERYGEKKGKAIKIEKVIYTGKEGKSSRGCPIAKWVIR
 RHTLEEKLLCLVRHRAGHHCQNAVIVILILAWEGIPRSLGDTLYQELDTLRKYGNPTSR
 RCGLNDDRTCACQGKDPNTCGASFSFGCSWSMYFNGCKYARSKTPRKFRLAGDNPKEE
 EVLRKSFQDLATEVAPLYKRLAPQAYQNQVTNEEIAIDCRLGLKEGRPFAGVTACMDFC
 AHAHKDQHNLVNGCTVVCTLTKEDNRCVKGIPEDQLHVLPLYKMANDEFSGSEENQ
 NAKVGSQAIQVLTAFPREVRRLPEPAKSCRQRQLEARKAAAEKKKIQKEKLSTPEKIKQ
 EALELAGITSDPGLSLKGGLSQQGLKPSLKVEPQNHFSFKYSGNAVVESYSVLGNCRPS
 DPYSMNSVYSYHSYYAQPSLTSVNGFHSKYALPSFSYYGFPSSNPVFPSSQLGPGAWGH
 SGSSGSFEKKPDLHALHNSLSPAYGGAEF AELPSQAVPTDAHHTPHHQQPAYPGPKEY
 LLPKAPLLHSVSRDPSFAQSSNCYNRSIKQEPVDPLTQAEPVPRDAGKMGKTPLSEVSQ
 NNGPSHLWGQYSGGSPMSPKRTNGVGGSWGVSFSSGESPAIVPDKLSSFGASCLAPSHFT
 DGQWGLFPGEGQQAASHGGRLRGKPWSPCKFGNSTSALAGPSLTEKPWALGAGDFNS
 ALKGSFGFQDKLWNP MKGEEGRIPAAGASQLDRAWQSFGLPLGSSEKLF GALKSEEKL
 WDPFSLEEGPAEPPSKGAVKEEKGGGGAE EEEEEELWSDSEHNFLDENIGGVAVAPAHG
 SILIECARRELHATTPLKKNRCHPTRISLVFYQHKNLNQPNHGLALWEAKMKQLAERA
 RARQEEAARLGLGQQAELYGKKRKGWGGTVVAEPQQKEKKGVPTRQALAVPTDSA
 VTVSSYAYTKVTGPYSRWI

[0300] SEQ ID NO:4 (SV40 NLS)

PKKKRKV

[0301] SEQ ID NO:5 (XTEN16 (последовательность из 16 аминокислот))

SGSETPGTSESATPES

[0302] SEQ ID NO:6 (XTEN80 (последовательность из 80 аминокислот))

GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST
 EEGTSTEPSEGSAPGTSTEPSE

[0303] SEQ ID NO:7 (метка HA)

YPYDVPDYA

[0304] SEQ ID NO:8 (BFP)

SELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTMRIK VVEGGPLPFAF
 DILATSFLYGSKTFINHTQGIPDFFKQSFPEGFTWERVTTTYEDGGVLTATQDTSLQDGCLI
 YNVKIRGVNFTSNGPVMQKKT LGWEAFTETLYPADGGLEGRNDMALKLVGGSHLIANI
 KTTYRSKKPAKNLKM PGVYYVDYRLRIKEANNETYVEQHEVAVARYCDLPSKLGHK
 LN*

[0305] SEQ ID NO:9 (dCas9)

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDL
 NPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
 NGLFGNLIALSLGLTPNFKSNFDLAEDAQLSKD TYDDDLDNLLAQIGDQYADLFLAA

KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
 SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
 HLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
 WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEYFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLK
 RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
 KGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGELSEDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAIIKKYPKLES
 EFVYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLSVKELLGITIMERSSEKPNIDFLEA
 KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
 KLKGPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKPIR
 EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQL
 GGD

[0306] SEQ ID NO:10 (ddAsCfp1)

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQGFIEEDKARNNDHYKELKPIIDR
 IYKTYADQCLQLVQLDWENLSAIDSYRKEKTEETRNALIEEQATYRNAIHDFYFIGRTD
 NLTDANKRHAIEYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYE
 NRKNVFS AEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSI
 EEVFSFPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLP
 HRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLRNENVLETAEALFNELNSIDLT
 HIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEII
 SAAGKELSEAFKQKTSEILSHAHAALDQPLPTLKKQEEKEILKSQLDSSLGLYHLLDWF
 AVDESNEVDPEFSARLTGIKLEMPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASG
 WDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEKTSEGFDMYYDYFPD
 AAKMIPKCSTQLKAVTAHFQTHHTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAK
 KTGDKQGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHI
 SFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKL
 NGQAELFYRPKSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDL
 SDEARALLPNVITKEVSHEIHKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEH
 PETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAAARQAWSV
 VGTIKDLKQGYLSQVIHEIVDLMIHYQAVVLANLNFQFKSKRTGIAEKAVYQQFEKML
 IDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTG
 FVDPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAW
 DIVFEKNETQFDAKGTPFIAGKRIVPVIEHNRFTGRYRDLYPANELIALLEEKGIVFRDGS

NILPKLLENDSDHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNP
 WPMADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELRN

[0307] SEQ ID NO:11 (ddLbCfp1)

MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLEVEDEKRAEDYKGVKLL
 DRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGN
 YKSLFKKDIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFRCI
 NENLTRYISNMDIFEKVD AIFDKHEVQEIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVY
 NAIIGGFVTESEGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLSFYGEGYTS
 EEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNPAISTISKDIFGEWNVIR
 DKWNAEYDDIHLKKKAVVTEKYEDDRRSFKKIGSFSLEQLQEYADADLSVVEKLKEII
 IQKVDEIYKVYGSSEKLFDAFVLEKSLKKNDAVVAIMKDLLDSVKSFENYKAFEGEG
 KETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWD
 KDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKML
 PKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPKWSNA
 YDFNFSETEKYKDIAGFYREVVEEQGYKVSFESASKKEVDKLVVEEGKLYMFQIYNKDFSD
 KSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNP
 DNPKKTTL SYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGI
 ARGERNLLYIVVVDGKGNIVEQYSLNEIINNFNNGIRIKTDYHSLDKKEKERFEARQNW
 SIENIKELKAGYISQVVHKICELVEKYDAVIALADLNSGFKNRSRVKVEKQVYQKFEKMLI
 DKLNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFV
 NLLKTKYTSIADSKKFISDFRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNR
 IRIFRNPKNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMAL
 MSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIAR
 KVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKH

[0308] SEQ ID NO:12 (ddFnCfp1)

MYPYDVPDYASGSGMSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDE
 KRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSA
 KDTIKKQISEYIKDSEKFKNLFNQNLDIAKKGQESDLILWLKQSKDNIELFKANSDITDI
 DEALEIISFKGWTTYFKGFHENRKNVYSSNDIPTSIYRIVDDNLPKFLENKAKYESLKD
 KAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFLDEVFEIANFNLYLNQSGITKFNTII
 GGKFNVENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDD
 SDVVTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSSQ
 VFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALFEFNKH
 RDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAI
 KDLLDQTNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYIT
 QKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYLLGVMNKKNNKIFDDKA
 IKENKGEQYKIVYKLLPGANKMLPKVFFSAKSIKFNPNSEDILRIRNHSTHTKNGSPQK
 GYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYSIDEFYREVENQGYKLT
 FENISESYIDSVVNQGLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKL
 NGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPT

INFKSSGANKFNDEINLLLKEKANDVHILSIARGERHLAYYTLVDGKGNIKQDTFNIIGN
 DRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKL VIEYNAIVV
 FEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYVFKDNEFDKTGGVLRAYQLTAPFE
 TFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKG
 YFEFSFDYKNFGDKAAKKGWTIASFGSRLINFRNSDKNHNWDTREYPTKELEKLLKD
 YSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNF
 FDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNR
 NN

[0309] SEQ ID NO:13 (p65; UniProt: Q04206)

MDELFLIFPAEPAQASGPYVEIIEQPKQRGMFRYKCEGRSAGSIPGERSTDTTK
 THPT
 IKINGYTGPQTVRISLVTKDPPHRPHHEL VGKDCRDGFYEAEPCDRCIHSFQNLGIQC
 VKKRDLEQAISQRIQTNNNPQVPIEEQRGDYDLNAVRLCFQVTVRDPSPGRPLRLPPVLS
 HPIFDNRAPNTAELKICRVNRNSGSLGGDEIFLLCDKVQKEDIEVYFTGPGWEARGSF
 QADVHRQVAIVFRTPPYADPSLQAPVRVSMQLRRPSDRELSEPMEFQYLPDTPDRHRIE
 E KRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQYPFTSSLSTINY
 DEFPTMVFPSPQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQA
 VAPPAPKPTQAGEGTLSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLLN
 QGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGPGLPNGLLSGDEDFSSIAD
 MDFSALLSQISS

[0310] SEQ ID NO:14 (p65; or Addgene)

PTQAGEGTLSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLLNQGI
 PVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGPGLPNGLLSGDEDFSSIADMDF
 SALL

[0311] SEQ ID NO:15 (Rta; or Addgene)

RDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPT
 TGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKE
 EAAICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDFTLNDECLLHAM
 HISTGLSIFDTSLF

[0312] SEQ ID NO:16 (Rta; UniProt P03209)

MRPKKDGLDFLRLTPEIKKQLGSLVSDYCNVLNKEFTAGSVEITLRSYKICKAFI
 NEAKAHGREWGGLMATLNICNFWAILRNNRVRRAENAGNDACSIACPIVMRYVLDH
 LIVVTDRFFIQAPSNRVMIPATIGTAMYKLLKHSRVRAYTYSKVLGVDRAAIMASGKQV
 VEHLNRMEKEGLLSSKFKAFCKWVFTYPVLEEMFQTMVSSKTGHLTDDVKDVRALIKT
 LPRASYSSHAGQRSYVSGVLPACLLSTKSKAVETPILVSGADRMDEELMGNDGGASHTE
 ARYSESGQFHAFTDELESLPPTMPLKPGAQSADCGDSSSSSSSDSGNSDTEQSEREERA
 EAPRLRAPKSRRTSRPNRGQTPCPSNAAEPEQPWIAAVHQESDERPIFPHPSKPTFLPPVK
 RKKGLRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPT
 PTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQK
 EAAICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDFTLNDECLLHA

MHISTGLSIFDTSLF

[0313] SEQ ID NO: 17 (VP64; от Addgene)

DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDML

[0314] SEQ ID NO: 18 (полноразмерный белок тегумента VP16; VP64; UniProt P06492)

MDLLVDELFAADMNADGASPPPPRPAGGPKNTPAAPPLYATGRLSQAQLMPSPPM
PVPPAALFNRLDDDLGFSAGPALCTMLDTWNEDLFSALPTNADLYRECKFLSTLPSDVV
EWGDAYVPERTQIDIRAHGDVAFPTLPATRDGLGLYYEALSRFFHAELRAREESYRTVL
ANFCSALYRYLRASVRQLHRQAHMRGRDRDLGEMLRATIADRYRETARLARVFLH
LYLFLTREILWAAAYAEQMMRPDLFDCLCCDLESWRQLAGLFQPFMFVNGALTVRGVPI
EARRLRELNHIREHLNPLVRSAAATEEPGAPLTTPTLHGNQARASGYFMVLIRAKLDSY
SSFTTSPSEAVMREHAYSRTARTKNNYGSTIEGLDLPDDDAPEEAGLAAPRLSFLPAGHT
RRLSTAPPTDVS LGDELHLDGEDVAMAHADALDDFDLDMLGDGDSPPGPGFTPHDSAPY
GALDMADFEFEQMFTDALGIDEYGG

[0315] SEQ ID NO:19 (петля 1 стебля MS2)

AGCCAACATGAGGATCACCCATGTCTGCAGGGC

[0316] SEQ ID NO:20 (петля 2 стебля MS2)

GGCCAACATGAGGATCACCCATGTCTGCAGGGCC

[0317] SEQ ID NO:21 (белок оболочки MS2 (MCP))

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCVSRQSS
AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMELETIPIFATNSDCELIVKAMQGL
LKDGNPIPSAIAANSIY

[0318] SEQ ID NO: 86 (каталитический домен TET1 (TET1CD))

LPTCSCLDRVIQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGNIRIEIVVYTG
KEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCPATAVMVVLIMVWDGIPLPMAD
RLYTELTENLKSYNHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNGCKF
GRSPSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARECL
GSKEGRPFSGVTA CLDFCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLHVLP
LYKLSDTDEFSGSKEGMEAKIKSGAIEVLAPRRKKRKTCTQPVPRSGKKRAAMMTEVLA
HKIRAVEKKPIPRIKRNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDNTK
TYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKNDATASCGFSERSSTPHCTMPSGRLS
GANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPTGVTEPLTPHQPNHQPSFLTSPQ
DLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDSEHIFLDANIGGVAI
APAHGSVLI ECARRELHATTPVEHPNRNHPTRLSLVFYQHKNLNKPQHGFE LN KIKFEA
KEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKALTLTHDNVVTVSPYALTHV
AGPYNHWV

[0319] SEQ ID NO:97 (TET1)

MALPTCSCLDRVIQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGNIRIEIVV
YTGKEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCPATAVMVVLIMVWDGIPLP
MADRLYTELTENLKSYNHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNG

CKFGRSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE
 CRLGSKEGRPFSGVTAACLDCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLH
 VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCFTQPVPRSGKKRAAMMTEV
 LAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
 TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
 LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLT
 PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDEHIFLDANIGGV
 AIAPAHGSVLIECARRELHATTPVEHPNRNHPTRL SLVFYQHKNLNKPQHGFELNKIKFE
 AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TLTHDNVVTVSPYAL TH
 VAGPYNHWV

[0320] SEQ ID NO:98 XTEN100

GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG
 SAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSE

[0321] SEQ ID NO:99 Слитый белок JKNp146

MALPTCSCLDRVIQKDKGPYYTHLGAGPSVA AVREIMENRYGQKGN AIRIEIVV
 YTGKEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCP TAVMVVLMVWDGIPLP
 MADRLYTELTENLKSYN GHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNG
 CKFGRSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE
 CRLGSKEGRPFSGVTAACLDCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLH
 VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCFTQPVPRSGKKRAAMMTEV
 LAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
 TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
 LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLT
 PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDEHIFLDANIGGV
 AIAPAHGSVLIECARRELHATTPVEHPNRNHPTRL SLVFYQHKNLNKPQHGFELNKIKFE
 AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TLTHDNVVTVSPYAL TH
 VAGPYNHWVGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTS
 TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEMDKKYSIGLAIGTN
 SVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
 RRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYN
 QLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKS
 NFDLAEDAQLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
 YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFL
 KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE
 RMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE
 NEDILEDIVLTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
 DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP

AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE
 LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD
 DSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERG
 GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDF
 RKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIA
 KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
 KVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV
 LVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIILPKYSLF
 ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQH
 KHYLDEIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIHLFTLTNLGAPAA
 FKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDGGPSSGAPPPSGGSPA
 GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS
 EGSGPKKKRKVAGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWAN
 RPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALRE
 MADTVIPQKEEAICGQMDL SHPPRGHLDELTTTLESMTEDLNLDSPLELNEILDFTL
 NDECLLHAMHISTGLSIFDTSLFASGSGPKKKRKV

SEQ ID NO:99 включает следующие SEQ ID NO и спейсеры:

97-98-9-6-GSG-4-AGS-15-ASGSG-4; где GSG, AGS и ASGSG представляют собой пептидные линкеры.

[0322] SEQ ID NO:100 (p65)

SQYLPDTRDRHRIEEKRKRKYETFKSIMKKSPFSGPTDPRPPPRRIA VPSRSSASVP
 KPAPQYPFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAMVSAL
 AQAPAPVPLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDLGALLGNSTDPAV
 FTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPG
 LPNGLLSGDEDFSSIADMDFSALL

[0323] SEQ ID NO:101 Слитый белок JKNp147

MALPTCSCDRVIQKDKGPYYTHLGAGPSVA AVREIMENRYGQKGN AIRIEIVV
 YTGKEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCP TAVMVVLMVWDGIPLP
 MADRLYTELTENLKSYNHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNG
 CKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE
 CRLGSKEGRPFSGVTACLDFCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLH
 VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCFTQPVPRSGKKRAAMMTEV
 LAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
 TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
 LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLTS
 PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDEHIFLDANIGGV
 AIAPAHGSVLIECARRELHATTPVEHPNRNHPTRL SLVFYQHKNL NKPQHGFELNKIKFE
 AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TLTHDNVVTVSPYAL TH
 VAGPYNHWVGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTS
 TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGSEMDKKYSIGLAIGTN

SVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
 RRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYN
 QLFEEPNASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS
 NFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
 YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPL
 KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE
 RMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE
 NEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
 DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP
 AIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKE
 LGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD
 DSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERG
 GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDF
 RKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVVYDVRKMIA
 KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
 KVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV
 LVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF
 ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQH
 KHYLDEIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTNLGAPAA
 FKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDGGPSSGAPPPSGGSPA
 GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS
 EGSGPKKKRKVAGSSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIA
 VPSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPA
 PAPAMVSALAQAPAPVPLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDLGAL
 LGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDP
 APAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDV
 FEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAV
 TPEASHLLEDPEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHPPPRGHLDELTTTL
 ESMTEDLNLDSP LTPELNEILD TFLNDECLLHAMHISTGLSIFDTS LFASGSGPKKKRKV

SEQ ID NO:101 включает следующие SEQ ID NO и спейсеры:

97-98-9-6-GSG-4-AGS-100-GSGSGS-15-ASGSG-4; где GSG, AGS, GSGSGS и ASGSG представляют собой пептидные линкеры.

[0324] SEQ ID NO:102 Слитый белок GCP21

MALPTCSCLDRVIQKDKGPYYTHLGAGPSVA AVREIMENRYGQKGN AIRIEIVV
 YTGKEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCP TAVMVVLMVWDGIPLP
 MADRLYTELTENLKSYN GHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNG
 CKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE

CRLGSKEGRPFSGVTA CLDFCAHPHRDIHNMNNGSTVVCTL TREDNRSLGVIPQDEQLH
 VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCFTQPVPRSGKKRAAMMTEV
 LAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
 TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
 LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLTS
 PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWS DSEHIFLDANIGGV
 AIAPAHGSVLIECARRELHATTPVEHPNRNHPTRL SLVFYQHKNLNKPKQHGFE LNKIKFE
 AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TLTHDNVVTVSPYAL TH
 VAGPYNHWVGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTS
 TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEMDKKYSIGLAIGTN
 SVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
 RRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYN
 QLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKS
 NFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT
 KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
 YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFL
 KDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIE
 RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEE
 NEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRK LINGIR
 DKQSGKTILDFLKS DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP
 AIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKE
 LGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD
 DSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERG
 GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDF
 RKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIA
 KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
 KVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV
 LVVAKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKDLIILPKYSLF
 ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LKGSPE DNEQKQLFVEQH
 KHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIHLFTL TNLGAPAA
 FKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDGGGGSPKKKRKVDPK
 KKRKVDPKKR KV

SEQ ID NO:102 включает следующие SEQ ID NO и спейсеры:

97-98-9-GGGGS-4-D-4-D-4; где GGGGS, D и D представляют собой пептидные линкеры.

[0325] SEQ ID NO:103 - JKNp84: dCas9-TET1

MDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHE

RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
 NPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
 NGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA
 KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQ
 SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
 HLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITP
 WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTFEDREMIEERLKTYAHLFDDKVMKQLK
 RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ
 KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGELSEDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTAIIKKYPKLES
 EFVYGDYK VYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEA
 KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
 KLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVL SAYNKHRDKPIR
 EQAENIIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQL
 GGDGGGGSPKKKRK VDPKKKRK VDPKKKRK VGS LPTCSCLDRVIQKDKGPYYTHLGA
 GPSVA AVREIMENRYGQKGN AIRIEIVVYTGKEGKSSHGCPIAKWVLRSSDEEKVLCL
 VRQRTGHH CPTAVMVVLMVWDGIPLMADRLYTELTENLKS YNGHPTDRRCTLNENR
 TCTCQGIDPETCGASFSFGCSWSMYFNGCKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLA
 TRLAPIYKQYAPVAYQNQVEYENVARECRLGSKEGRPFSGVTACLDFCAHPHRDIHNM
 NNGSTVVCTLTREDNRSLGVIPQDEQLHVLPLYKLSDTDEFSGSKEGMEAKIKSGAIEVLA
 PRRKKRTCFTQPVPRSGKKRAAMMTEVLAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSL
 PTLGSNTETVQPEVKSETEPHFILKSSDNTKTYSLMPSAPHPVKEASPGFSWSPK TASATP
 APLKN DATASC GFSERSSTPHCTMPSGRLSGANAAAADGPGISQLGEVAPLPTLSAPVM
 EPLINSEPTGVTEPLTPHQPNHQPSFLTSPQDLASSPMEEDEQHSEADEPPSDEPLSDDPL
 SPAEEKLPHIDEYWSDSEHIFLDANIGGVAIAPA HGSVLI ECARRELHATTPVEHPNRNHP
 TRLSLVFYQHKNLNK PQHGFELNKIKFEAKEAKNKKMKASEQKDAQANEGPEQSSEVN
 ELNQIPSHKALTLTHDNVVTVSPYALTHVAGPYNHVV

SEQ ID NO:103 включает следующие SEQ ID NO и спейсеры:

9-GGGGS-4-D-4-D-4-GS-86; где GGGGS, D, D и GS представляют собой пептидные линкеры.

[0326] SEQ ID NO:104=GCPp3: MCP-XTEN80-VP64

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
 AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNME LTIPIFATNSDCELVKAMQGL

LKDGNPIPSAIAANSYIYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSDALDDFDLDM LG
SDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LASGSGPKKKR KV

SEQ ID NO:104 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-17-ASGSGPKKKR KV; где GSG, AGS и ASGSGPKKKR KV
представляют собой пептидные линкеры.

[0327] SEQ ID NO:105=GCPp4: MCP-XTEN80-VP64-p65

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNME LTIPIFATNSDCELVKAMQGL
LKDGNPIPSAIAANSYIYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSDALDDFDLDM LG
SDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LINSRSSGSPKKKRKVGSQYL
PDTDDRHRIEEKRKRTYETFKSIMKKS PFSGPTDPRPPRR IAVPSRSSASVPKPA PQPYPF
TSSLSTINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSALA QAPAPVP
VLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDP AVFTDLASVD
NSEFQQLLNQGIPV APHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSG
DEFSSIADMDFSALLASGSGPKKKR KV

SEQ ID NO:105 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-17-INSRSSGS-4-G-100-ASGSG-4; где GSG, AGS, INSRSSGS, G и
ASGSG представляют собой пептидные линкеры.

[0328] SEQ ID NO:106=GCPp5: MCP-XTEN80-VP64-p65p-Rta

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNME LTIPIFATNSDCELVKAMQGL
LKDGNPIPSAIAANSYIYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSDALDDFDLDM LG
SDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LINSRSSGSPKKKRKVGSQYL
PDTDDRHRIEEKRKRTYETFKSIMKKS PFSGPTDPRPPRR IAVPSRSSASVPKPA PQPYPF
TSSLSTINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSALA QAPAPVP
VLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDP AVFTDLASVD
NSEFQQLLNQGIPV APHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSG
DEFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPFHPP
GSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLED PDEETSQ
AVKALREMA DTVIPQKEEA AICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LTPEL
NEILD TFLNDECLLHAMHISTGLSIFDTSLFASGSGPKKKR KV

SEQ ID NO:106 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-17-INSRSSGS-4-G-100-GSGSGS-15-ASGSG-4; где GSG, AGS,
INSRSSGS, G, GSGSGS и ASGSG представляют собой пептидные линкеры.

[0329] SEQ ID NO:107=GCPp6: MCP-XTEN80-p65

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNME LTIPIFATNSDCELVKAMQGL

LKDGNIPIPSAIAANSYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
 SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSSQYLPDTDDRHRHRI
 EEKRKRITYETFKSIMKKS PFSGPTDPRPPRRIAVPSRSSASVPKPAPQYPFTSSLSTINY
 DEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPVVPVLAPGPPQA
 VAPPAPKPTQAGEGTLSEALLQLQFDDEDL GALLGNSTDP AVFTDLASVDNSEFQQLLN
 QGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPL GAPGLPNGLLSGDEDFSSIAD
 MDFSALLASGSGP KKKRKV

SEQ ID NO:107 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-100-ASGSG-4; где GSG, AGS и ASGSG представляют собой пептидные линкеры.

[0330] SEQ ID NO:108=GCPp7: MCP-XTEN80-Rta

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
 AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLN MELTIPIFATNSDCELVKAMQGL
 LKDGNIPIPSAIAANSYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
 SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSRDSREGMFLPKPE
 AGSAISDVFE GREVCQPKRIRPFHPPGSPWANRPLASLAPTPTGPVHEPVGSLTPAPVPQ
 PLDPAPA VTP EASHLLEDPDEETSQAVKALREMA DTVIPQKEEA AICGQMDLSHPPPRG
 HLDELTTTLESMTEDLNLDSP LPELNEILD TFLNDECLLHAMHISTGLSIFDTSLFASGSG
 P KKKRKV

SEQ ID NO:108 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-15-ASGSG-4; где GSG, AGS и ASGSG представляют собой пептидные линкеры.

[0331] SEQ ID NO:109=GCPp8: MCP-XTEN80-p65-Rta

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS
 AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLN MELTIPIFATNSDCELVKAMQGL
 LKDGNIPIPSAIAANSYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
 SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS GPKKKRKVAGSSQYLPDTDDRHRHRI
 EEKRKRITYETFKSIMKKS PFSGPTDPRPPRRIAVPSRSSASVPKPAPQYPFTSSLSTINY
 DEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPVVPVLAPGPPQA
 VAPPAPKPTQAGEGTLSEALLQLQFDDEDL GALLGNSTDP AVFTDLASVDNSEFQQLLN
 QGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPL GAPGLPNGLLSGDEDFSSIAD
 MDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPFHPPGSPWANRP
 LPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPA VTP EASHLLEDPDEETSQAVKALREMA
 DTVIPQKEEA AICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LPELNEILD TFLN
 DECLLHAMHISTGLSIFDTSLFASGSGP KKKRKV

SEQ ID NO:109 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGS-100-GSGSGS-15-ASGSG-4; где GSG, AGS, GSGSGS и ASGSG представляют собой пептидные линкеры.

[0332] SEQ ID NO:110=GCPp9: MCP-XTEN80-NLS

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSS

AQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQGL
 LKDGNIPIPSAIAANSYGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG
 SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSPPKPKKRKVAGSASGSGPKPKKRKV

SEQ ID NO:110 включает следующие SEQ ID NO и спейсеры:

21-6-GSG-4-AGSASGSG-4; где GSG и AGSASGSG представляют собой пептидные
 линкеры.

[0333] SEQ ID NO:111=GCPp11: dCas9-XTEN16-TET1

MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDL
 NPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
 NGLFGNLIASLGLTPNFKSNFDLAEDAQLSKD TYDDDLDNLLAQIGDQYADFLAA
 KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
 SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
 HLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
 WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEG
 MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
 TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLK
 RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQ
 VSGQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ
 KGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLES
 EFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLSVKELLGITIMERSSEKPNIDFLEA
 KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
 KLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVL SAYNKH RDKPIR
 EQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQL
 GGDGGGGSPKPKKRKVDPKPKKRKVDPKPKKRKVGSGSETPGTSESATPESSLPTCSCLDRV
 IQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGN AIRIEIVVYTGKEGKSSHGCPIAKW
 VLRRSSDEEKVLCLVRQRTGHH CPTAVMVVLIMVWDGIPLMADRLYTELTENLKSYN
 GHPTDRRCTLNENRTCTCQGIDPETCGASFSGCSWSMYFNGCKFGRSPSPRRFRIDPSSP
 LHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARECRLGSKEGRPFSGVTAC
 LDFCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLHVLPLYKLSDTDEFGSKE
 GMEAKIKSGAIEVLAPRRKKRTCF TQPVPRSGKKRAAMMTEVLAHKIRAVEKKPIPRIK
 RKNNSTTTNNSKPSLPTLGSNTETVQPEVKSETEPHFILKSSDNTKTYSLMPSAPHPVKE
 ASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGRLSGANAAAADGPGISQ
 LGEVAPLPTLSAPVMEPLINSE PSTGVTEPLTPHQPNHQPSFLTSPQDLASSPMEEDEQHS

EADPPSDEPLSDDPLSPAEEKLPHIDEYWSDSEHIFLDANIGGVAIAPAHGSVLIECARR
ELHATTPVEHPNRNHPTRL SLV FYQHKNLNKPKQHGFELNKKIFEAKEAKNKKMKASEQ
KDQAANEGPEQSSEVNELNQIPSHKAL TL THDNVVTVSPYAL THVAGPYNHWV

SEQ ID NO:111 включает следующие SEQ ID NO и спейсеры:

9-GGGGS-4-D-4-D-4-G-5-86; где GGGGS, D, D и G представляют собой пептидные
линкеры.

[0334] SEQ ID NO:112=GCPp16: TET1-XTEN16-dCas9

MALPTCSCLDRVIQKDKGPYYTHLGAGPSVA AVREIMENRYGQKGN AIRIEIVV
YTGKEGKSSHGCPIAKWVLRSSDEEKVLCLVRQRTGHHCP TAVMVV LIMVWDGIPLP
MADRLYTELTENLKS YNGHPTDRRCTLNENRTCTCQ GIDPETCGASFSFGCSWSMYFNG
CKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE
CRLGSKEGRPFSGVTACLDFCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLH
VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCF TQPVPRSGKKRAAMMTEV
LAHKIRAVEKKPIPIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLT S
PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDSEHIFLDANIGGV
AIAPAHG SVLIECARRELHATTPVEHPNRNHPTRL SLV FYQHKNLNKPKQHGFELNKKIFE
AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TL THDNVVTVSPYAL TH
VAGPYNHWVSGSETPGTSESATPESMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LI
YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDKAILSA
RLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TL
LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLA
RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLY
EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIER
LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKVMG
RHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGKELGSQILKEHPVENTQLQNEKLY
LYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV
PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITK
HVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDA
YLNAAVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF
KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFS
KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGN

ELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLAD
 ANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
 ATLIHQ SITGLYETRIDLSQLGGDGGGGSPKKKRKVDPKKKRKYDPKKKRK

SEQ ID NO:112 включает следующие SEQ ID NO и спейсеры:

97-5-9-GGGGS-4-D-4-D-4; где GGGGS, D и D представляют собой пептидные
 линкеры.

[0335] SEQ ID NO:113=GCP20: TET1-XTEN80-dCas9

MALPTCSCLDRVIQKDKGPYYTHLGAGPSVAAVREIMENRYGQKGNIRIEIVV
 YTGKEGKSSHGCPYAKWVLRSSDEEKVLCLVRQRTGHHCPYAVMVLIMVWDGIPLP
 MADRLYTELTENLKSNGHPTDRRCTLNENRTCTCQGIDPETCGASFSFGCSWSMYFNG
 CKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLATRLAPIYKQYAPVAYQNQVEYENVARE
 CRLGSKEGRPFSGVTAACDFCAHPHRDIHNMNNGSTVVCTLTREDNRSLGVIPQDEQLH
 VLPLYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTCFVQVPRSGKKRAAMMTEV
 LAHKIRAVEKKPIPIKRKNNSTTTNNSKPSSLPTLGSNTETVQPEVKSETEPHFILKSSDN
 TKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASCGFSERSSTPHCTMPSGR
 LSGANAAAADGPGISQLGEVAPLPTLSAPVMEPLINSEPSTGVTEPLTPHQPNHQPSFLTS
 PQDLASSPMEEDEQHSEADEPPSDEPLSDDPLSPAEEKLPHIDEYWSDEHIFLDANIGGV
 AIAPAHGSVLI ECARRELHATTPVEHPNRNHPTRL SLVFYQHKNLNKPQHGFELNKIKFE
 AKEAKNKKMKASEQKDQAANEGPEQSSEVNELNQIPSHKAL TLTHDNVVTVSPYAL TH
 VAGPYNHWVGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPA
 GSPTSTEEGTSTEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFK
 VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV
 DDSFFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
 IYLALAHMIKFRGHFLIEGDLNPDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILS
 ARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAQLSKD TYDD
 DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
 LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL
 NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPL
 ARGNSRF AWMTRKSEETITPWNFEEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSL
 YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI
 ECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRN
 FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKVM
 GRHKPENIVIAMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKL
 YLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDN
 VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYK VREINNYHHAHD
 AYLNAVVGTA LIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMN
 FFKTEITLANGEIRK RPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGG
 FSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL

LGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKG
 NELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILA
 DANLDKVL SAYNKHRDKPIREQAENIHLFTLTNLGAPAAFKEYFDTTIDRKRYTSTKEVL
 DATLIHQ SITGLYETRIDLSQLGGDGGGGSPKKKRKVDPKKKRKVDPKKKRKV

SEQ ID NO:113 включает следующие SEQ ID NO и спейсеры:

97-6-9-GGGGS-4-D-4-D-4; где GGGGS, D и D представляют собой пептидные
 линкеры.

[0336] SEQ ID NO:114

GACGCTCAAATTTCCGCAGTGTTTAAGAGCTAAGCTGGAAACAGCATAGCAA
 GTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTT
 TTTT

[0337] SEQ ID NO:115

GTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT
 CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

[0338] SEQ ID NO:116

GACGCTCAAATTTCCGCAGT

[0339] SEQ ID NO:117 (последовательность ДНК, кодирующая каркас MS2-sgRNA)
 5'-

GTTTAAGAGCTAaGCCAACATGAGGATCACCCATGTCTGCAGGGCaTAGCAAGTTTA
 AATAAGGCTAGTCCGTTATCAACTTGGCCAACATGAGGATCACCCATGTCTGCAGGG
 CCAAGTGGCACCGAGTCGGTGCTTTTTTTT-3'

[0340] SEQ ID NO:118 (промоторная последовательность T7)

5'-TAATACGACTCACTATAGG-3'

[0341] SEQ ID NO:119

AGATCGGAAGAGCACACGTCTGAACTC

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Слитый белок, содержащий от N-конца к С-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемую РНК.
2. Слитый белок по п. 1, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.
3. Слитый белок по п. 2, отличающийся тем, что домен деметилирования представляет собой домен TET1.
4. Слитый белок по п. 2, отличающийся тем, что домен TET1 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 1, SEQ ID NO: 86 или SEQ ID NO: 97.
5. Слитый белок по п. 1, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, dCpf1, Cas-phi, мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен НМВ-бокс, домен Wor3, домен OB-складка, домен иммуноглобулина или домен В3.
6. Слитый белок по п. 5, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9.
7. Слитый белок по п. 1, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.
8. Слитый белок по п. 7, отличающийся тем, что линкер XTEN содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 5, SEQ ID NO: 6 или SEQ ID NO: 98.
9. Слитый белок по п. 1, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.
10. Слитый белок, содержащий от N-конца до С-конца, РНК-связывающую последовательность, линкер XTEN и по меньшей мере один активатор транскрипции.
11. Слитый белок по п. 10, отличающийся тем, что активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них.
12. Слитый белок по п. 11, отличающийся тем, что p65 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 13, SEQ ID NO: 14 или SEQ ID NO: 100.
13. Слитый белок по п. 11, отличающийся тем, что Rta содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 15 или SEQ ID NO: 16.
14. Слитый белок по п. 11, отличающийся тем, что VP64 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 17 или SEQ ID NO: 18.
15. Слитый белок по п. 10, отличающийся тем, что РНК-связывающая

последовательность представляет собой РНК-связывающую последовательность MS2.

16. Слитый белок по п. 15, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

17. Слитый белок по п. 10, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

18. Слитый белок по п. 10, имеющий аминокислотную последовательность с по меньшей мере 90% идентичностью последовательности с SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109 или SEQ ID NO:110.

19. Слитый белок по п. 10, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

20. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, первый линкер XTEN, фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемую РНК, второй линкер XTEN и активатор транскрипции.

21. Слитый белок по п. 20, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

22. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемую РНК.

23. Слитый белок по п. 20, дополнительно содержащий последовательность ядерной локализации.

24. Слитый белок по п. 20, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

25. Слитый белок по п. 24, отличающийся тем, что домен деметилирования представляет собой домен TET1.

26. Слитый белок по п. 20, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9, dCas12a, dCpf1, Cas-phi, домен лейциновой молнии, домен «крылатой спирали», мотив спираль-поворот-спираль, домен спираль-петля-спираль, домен HMB-бокс, домен Wor3, домен OB-складка, домен иммуноглобулина или домен В3.

27. Слитый белок по п. 26, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы, направляемая РНК, представляет собой dCas9.

28. Слитый белок по п. 20, отличающийся тем, что каждый из первого линкера XTEN и второго линкера XTEN независимо содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

29. Слитый белок по п. 20, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, флуоресцентную белковую метку или комбинацию двух или более из них.

30. Слитый белок, содержащий аминокислотную последовательность, имеющую по

меньшей мере 90% идентичность последовательности с SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:111, SEQ ID NO:112 или SEQ ID NO:113.

31. Слитый белок по п. 30, содержащий аминокислотную последовательность, имеющую по меньшей мере 95% идентичности последовательности с SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 102, SEQ ID NO: 111, SEQ ID NO: 112, или SEQ ID NO:113.

32. Слитый белок по п. 31, содержащий SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 102, SEQ ID NO: 111, SEQ ID NO: 112 или SEQ ID NO: 113.

33. Способ активации или реактивации 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: (a) sgRNA или (b) cr:tracrRNA;

за счет чего активируется или реактивируется последовательность нуклеиновой кислоты-мишени в клетке.

34. Способ по п. 32, отличающийся тем, что последовательность нуклеиновой кислоты-мишени содержит островок CpG.

35. Способ по п. 32, отличающийся тем, что последовательность нуклеиновой кислоты-мишени содержит отличный от островка CpG островок.

36. Способ по п. 32, отличающийся тем, что второй полинуклеотид содержит sgRNA.

37. Способ по п. 32, отличающийся тем, что sgRNA содержит по меньшей мере одну петлю стебля MS2.

38. Способ по п. 37, отличающийся тем, что sgRNA содержит две петли стебля MS2.

39. Способ по п. 32, отличающийся тем, что второй полинуклеотид кодирует активатор транскрипции.

40. Способ по п. 39, отличающийся тем, что активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них.

41. Способ по п. 32, отличающийся тем, что второй полинуклеотид дополнительно кодирует РНК-связывающую последовательность MS2.

42. Способ по п. 41, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

43. Способ по п. 32, отличающийся тем, что второй полинуклеотид дополнительно кодирует линкер XTEN, метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

44. Способ по п. 32, дополнительно включающий доставку в клетку третьего полинуклеотида, кодирующего второй слитый белок, который содержит активатор транскрипции.

45. Способ по п. 44, отличающийся тем, что активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них.

46. Способ по п. 44, отличающийся тем, что второй слитый белок дополнительно

содержит РНК-связывающую последовательность MS2.

47. Способ по п. 46, отличающийся тем, что РНК-связывающая последовательность MS2 содержит аминокислотную последовательность SEQ ID NO: 21.

48. Способ по п. 44, отличающийся тем, что второй слитый белок дополнительно содержит линкер XTEN, метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

49. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

50. Слитый белок по п. 49, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более из них.

51. Слитый белок по п. 49, отличающийся тем, что домен деметилирования представляет собой домен TET1.

52. Слитый белок по п. 51, отличающийся тем, что домен TET1 содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 1, SEQ ID NO: 86 или SEQ ID NO: 97.

53. Слитый белок по п. 49, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы представляет собой домен цинкового пальца.

54. Слитый белок по п. 49, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы представляет собой TALE.

55. Слитый белок по п. 49, отличающийся тем, что линкер XTEN содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

56. Слитый белок по п. 55, отличающийся тем, что линкер XTEN содержит аминокислотную последовательность, имеющую по меньшей мере 90% идентичности последовательности с SEQ ID NO: 5, SEQ ID NO: 6 или SEQ ID NO: 98.

57. Слитый белок по п. 49, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, метку флуоресцентного белка, сигнальный пептид ядерной локализации или комбинацию двух или более из них.

58. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, первый линкер XTEN, фермент ДНК-эндонуклеазу с дефицитом нуклеазы, второй линкер XTEN и активатор транскрипции.

59. Слитый белок по п. 58, отличающийся тем, что активатор транскрипции представляет собой VP64, р65, Rta или комбинацию двух или более из них.

60. Слитый белок, содержащий от N-конца к C-концу, домен деметилирования, линкер XTEN и фермент ДНК-эндонуклеазу с дефицитом нуклеазы.

61. Слитый белок по п. 58, дополнительно содержащий последовательность ядерной локализации.

62. Слитый белок по п. 58, отличающийся тем, что домен деметилирования представляет собой домен TET1, домен TET2, домен TET3 или комбинацию двух или более

из них.

63. Слитый белок по п. 62, отличающийся тем, что домен деметилирования представляет собой домен TET1.

64. Слитый белок по п. 58, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы представляет собой домен цинкового пальца.

65. Слитый белок по п. 58, отличающийся тем, что фермент ДНК-эндонуклеаза с дефицитом нуклеазы представляет собой TALE.

66. Слитый белок по п. 58, отличающийся тем, что каждый из первого линкера XTEN и второго линкера XTEN независимо содержит от приблизительно 10 аминокислотных остатков до приблизительно 864 аминокислотных остатков.

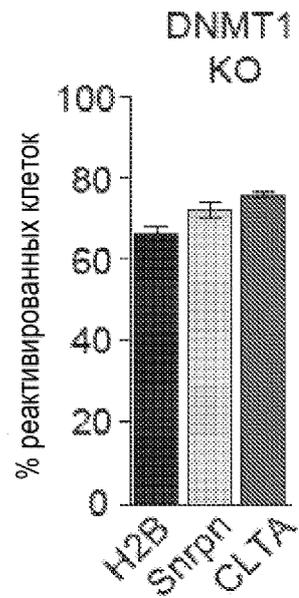
67. Слитый белок по п. 58, причем слитый белок дополнительно содержит метку эпитопа, пептид 2A, флуоресцентную белковую метку или комбинацию двух или более из них.

68. Способ активации или реактивации последовательности нуклеиновой кислоты-мишени в клетке, включающий доставку полинуклеотида, кодирующего слитый белок по п. 58, в клетку, содержащую нуклеиновую кислоту-мишень; за счет чего активируется или реактивируется последовательность нуклеиновой кислоты-мишени в клетке.

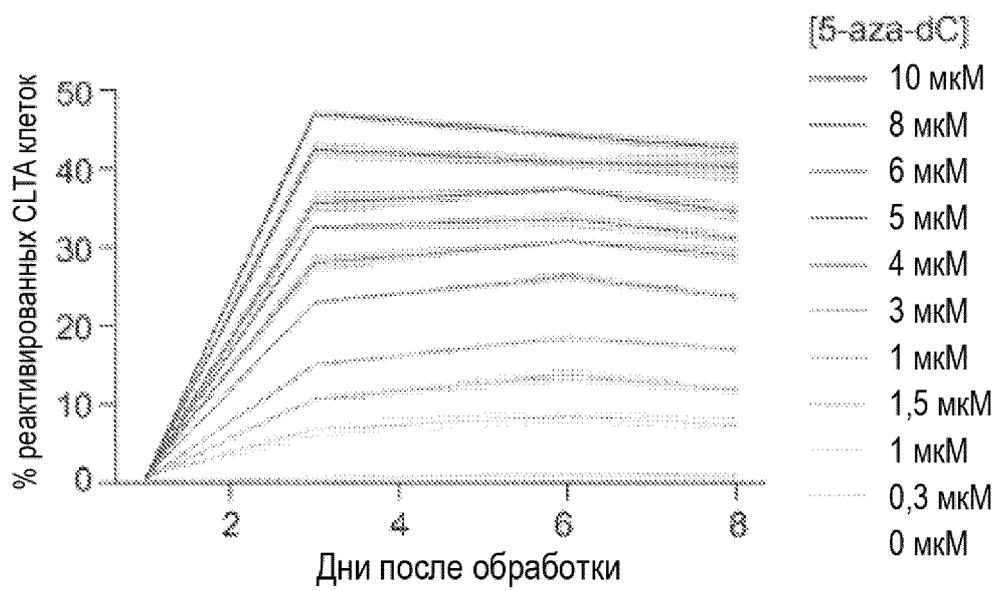
69. Способ по п. 68, отличающийся тем, что активатор транскрипции представляет собой VP64, p65, Rta или комбинацию двух или более из них.

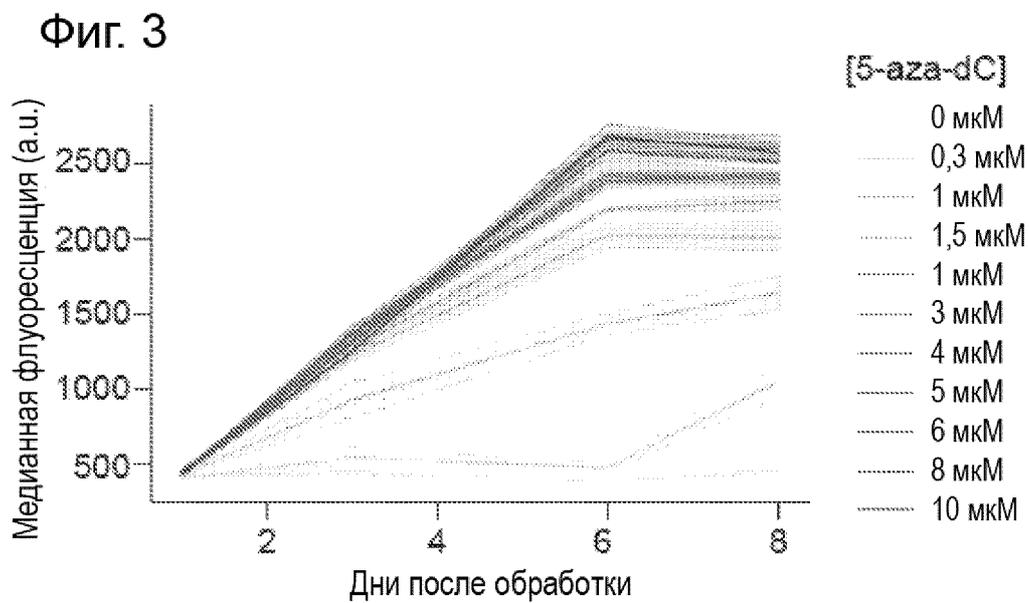
По доверенности

Фиг. 1

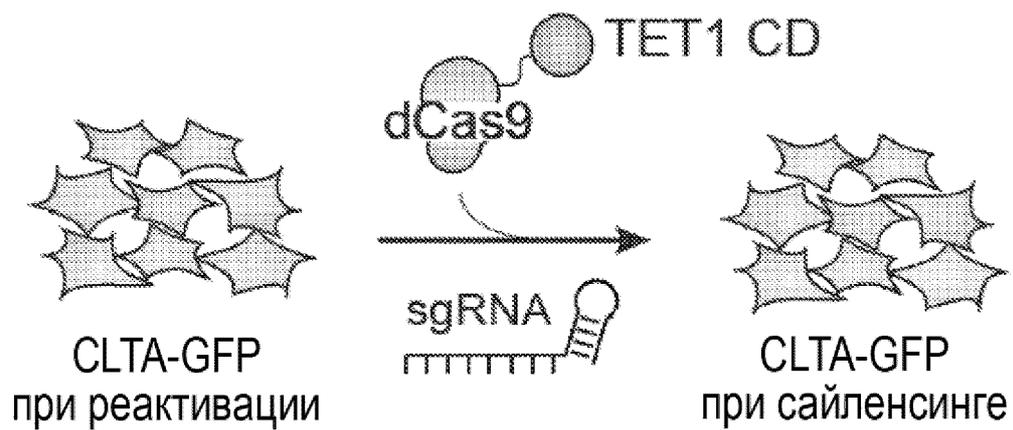


Фиг. 2

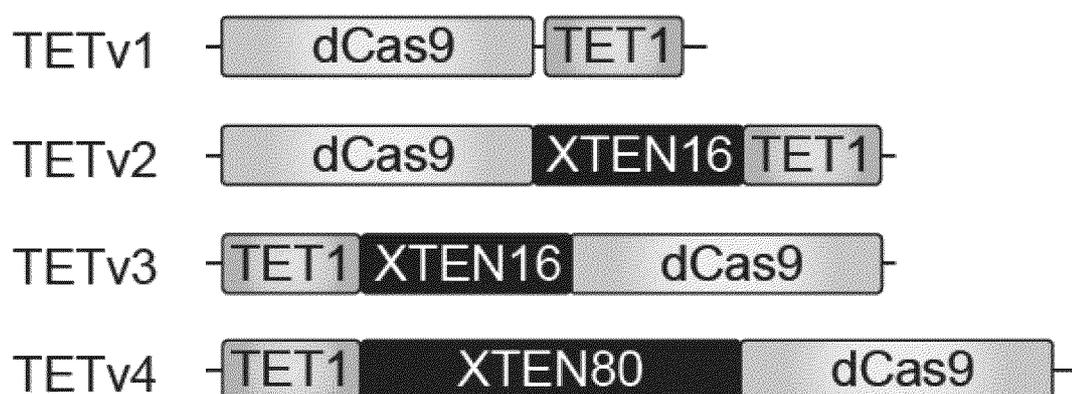




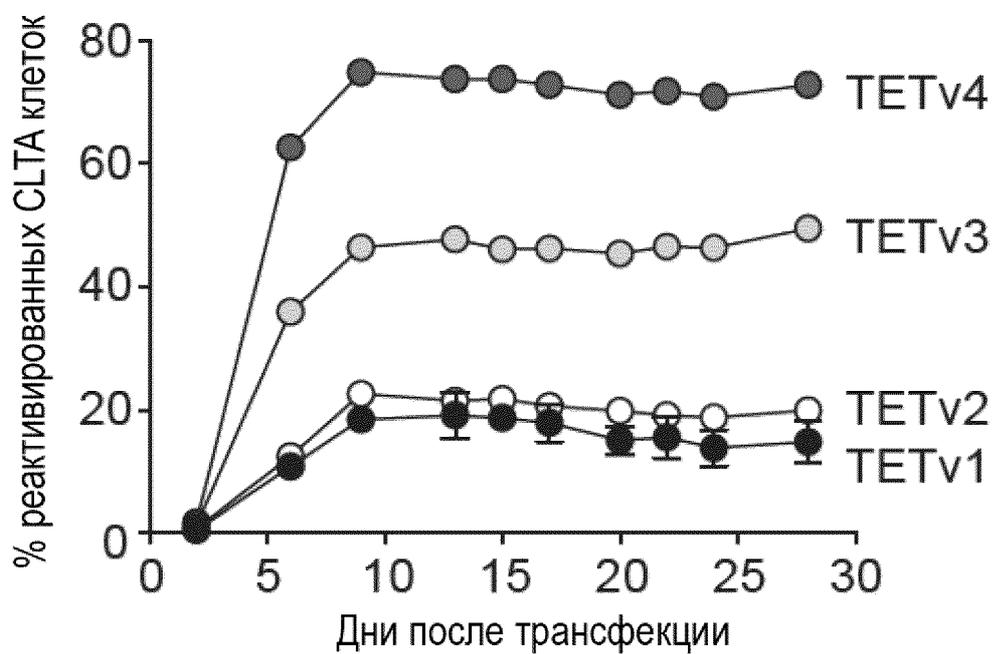
Фиг. 4



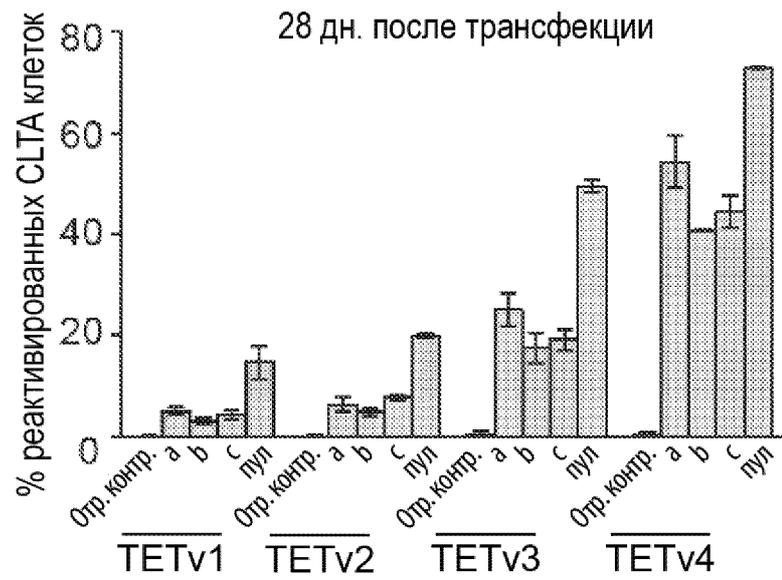
Фиг. 5



Фиг. 6

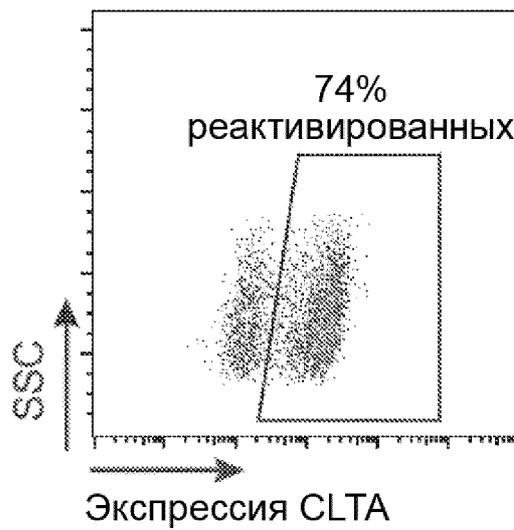


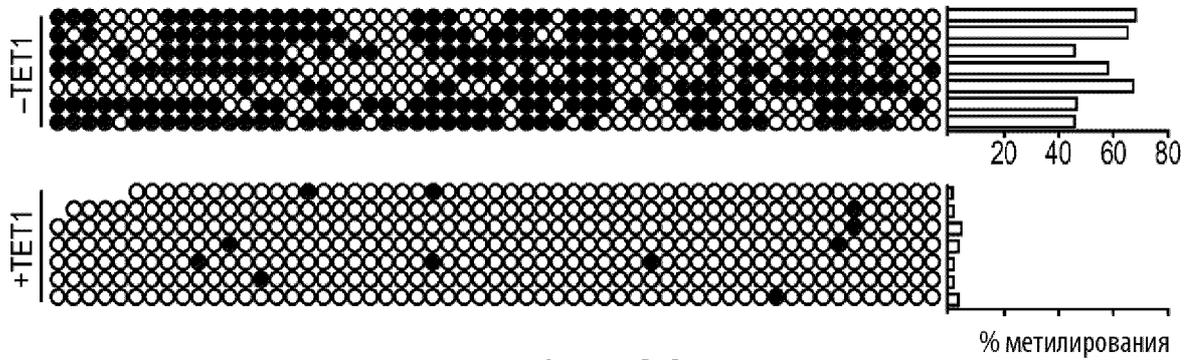
Фиг. 7



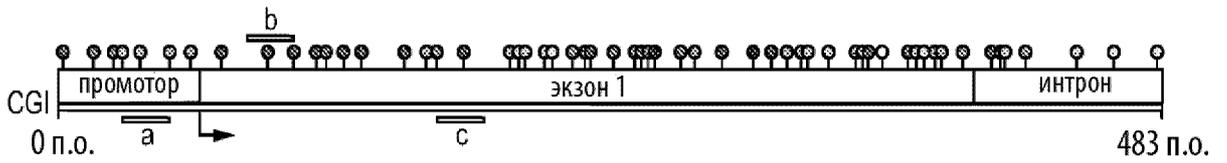
Фиг. 8

TETv4, 28 дн. после трансфекции

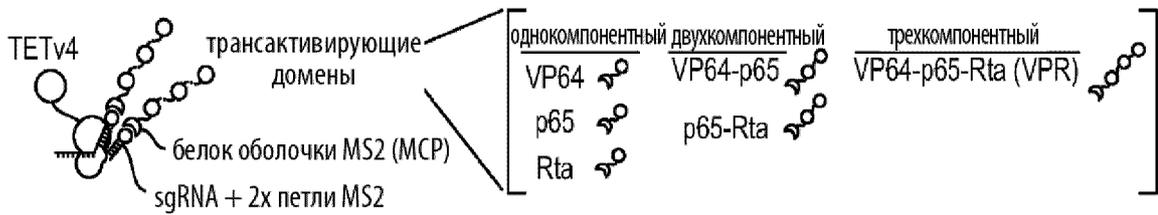




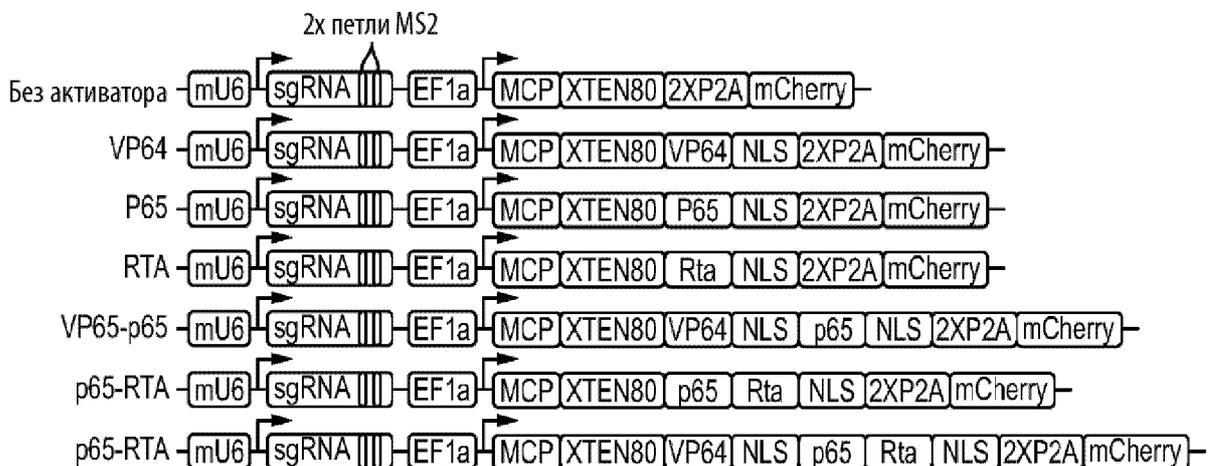
Фиг. 9А



Фиг. 9В

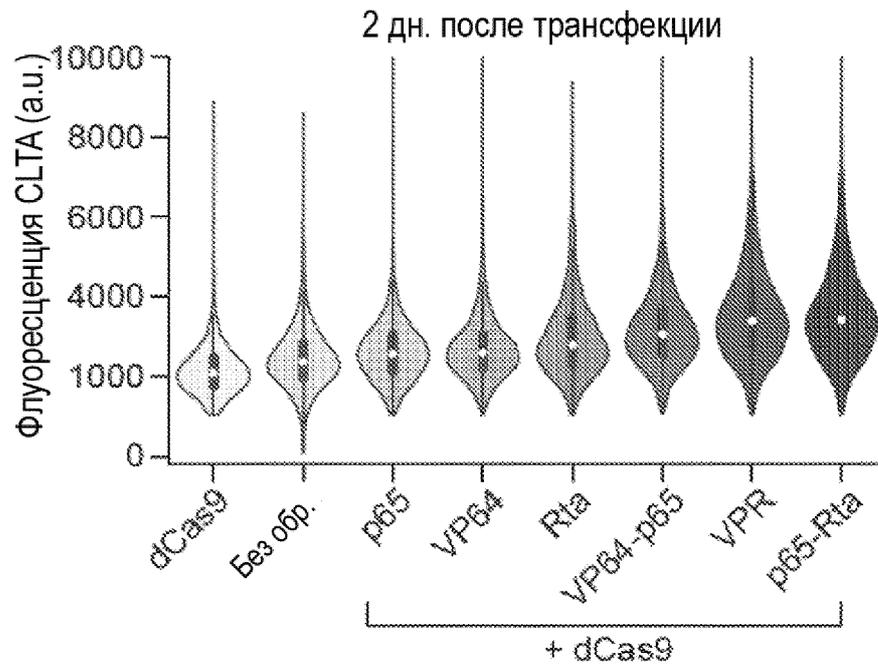


Фиг. 10

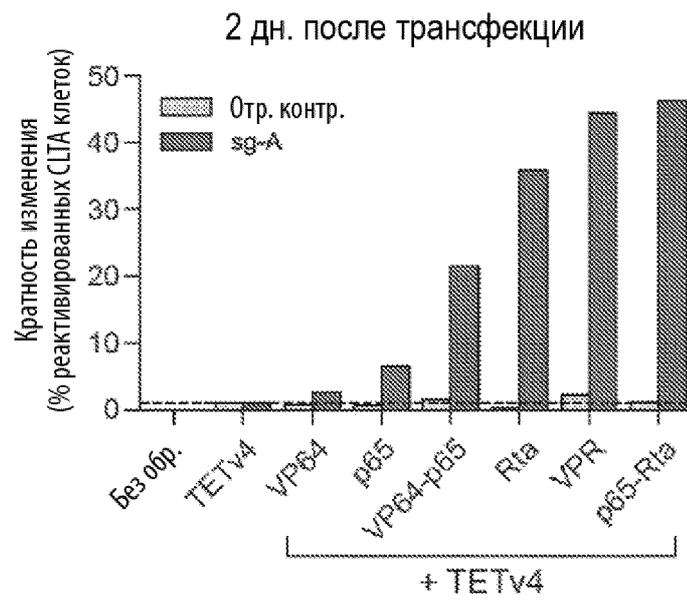


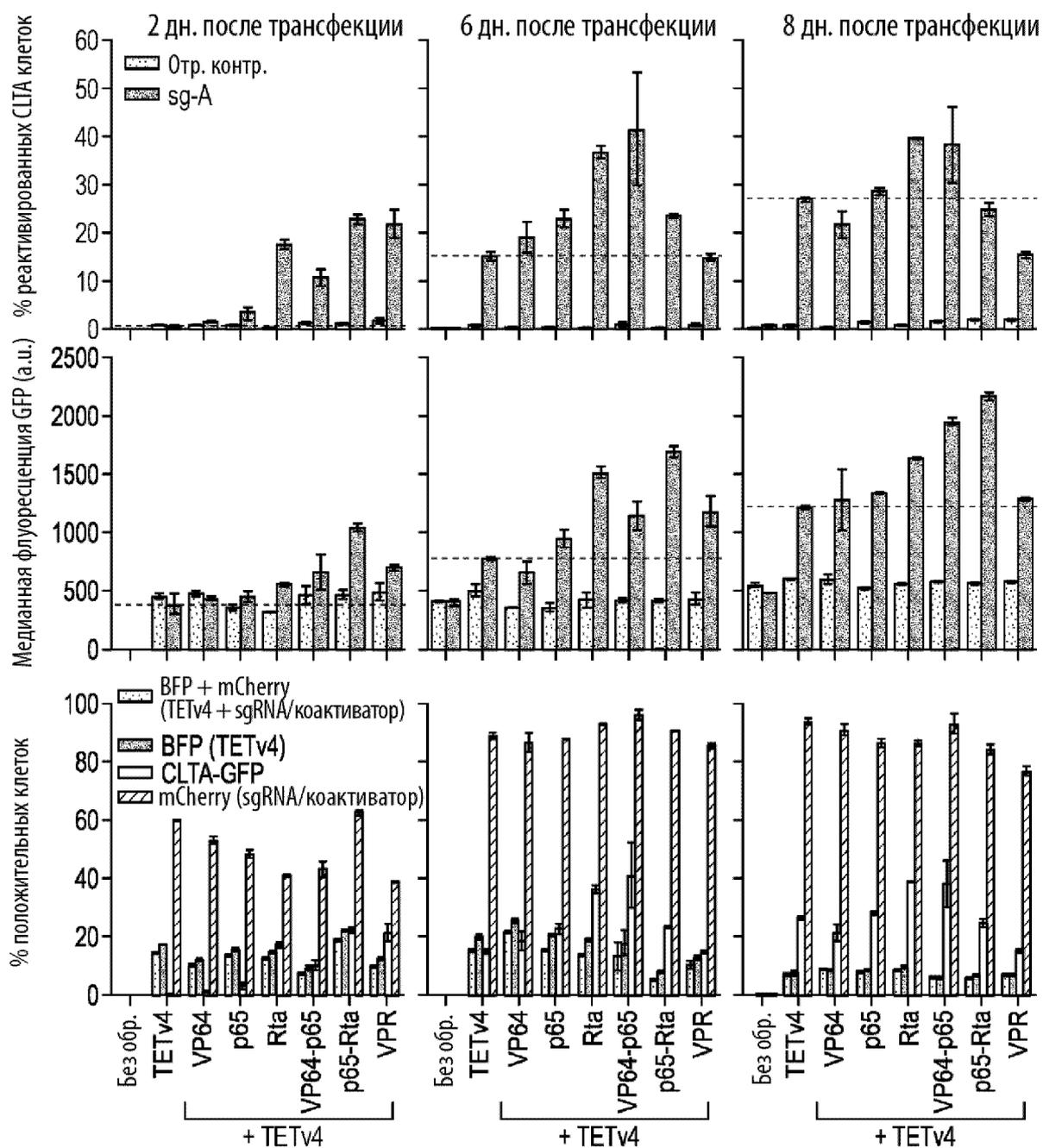
Фиг. 11

Фиг. 12



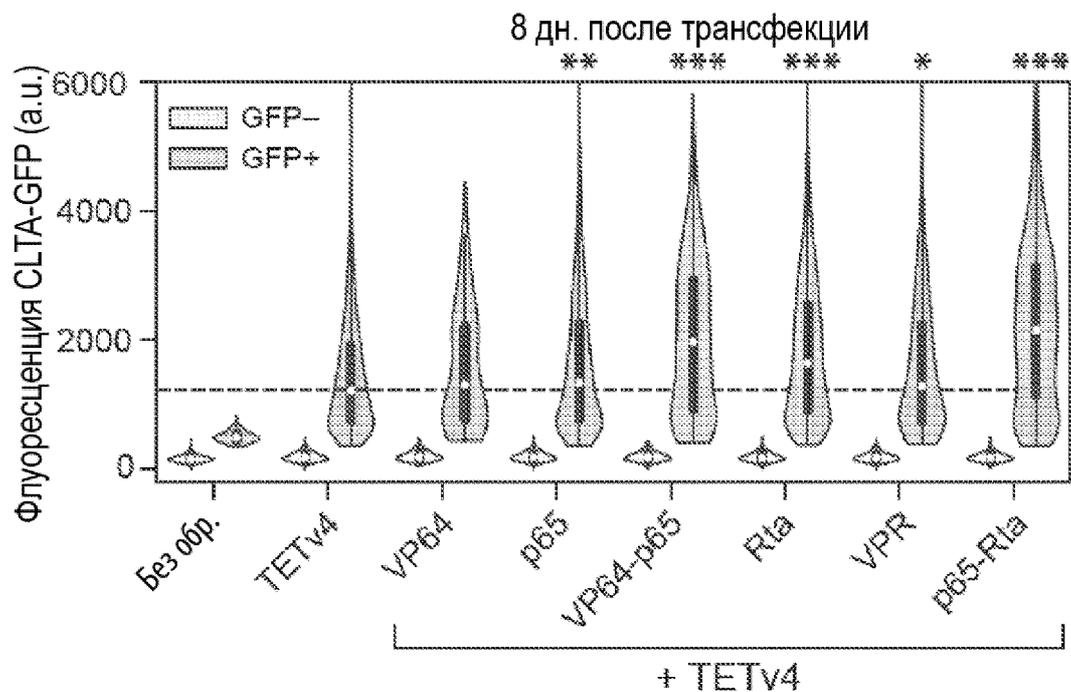
Фиг. 13



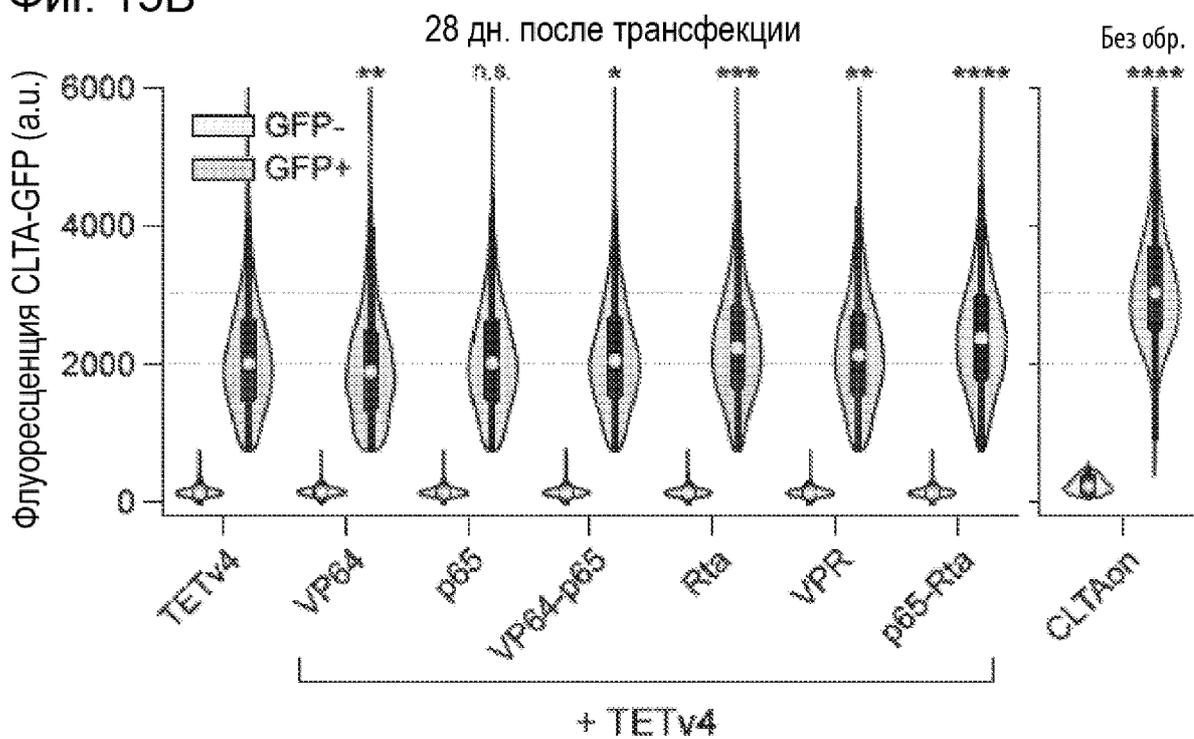


Фиг. 14

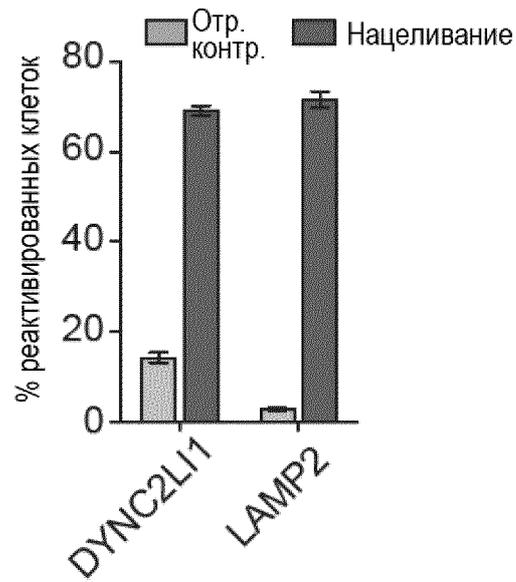
Фиг. 15А



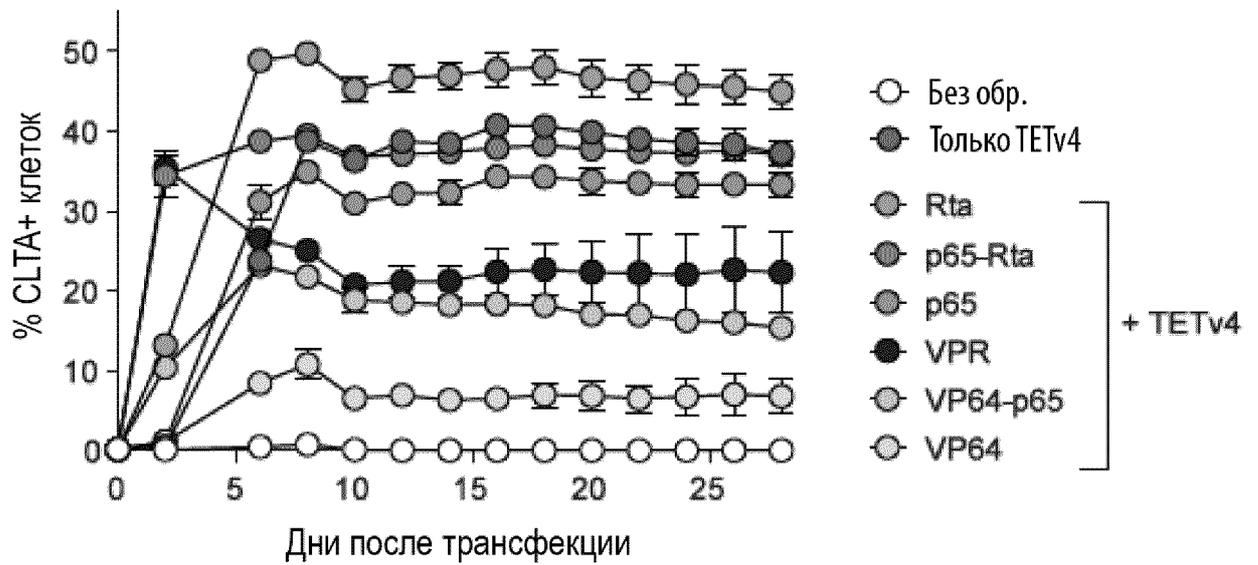
Фиг. 15В



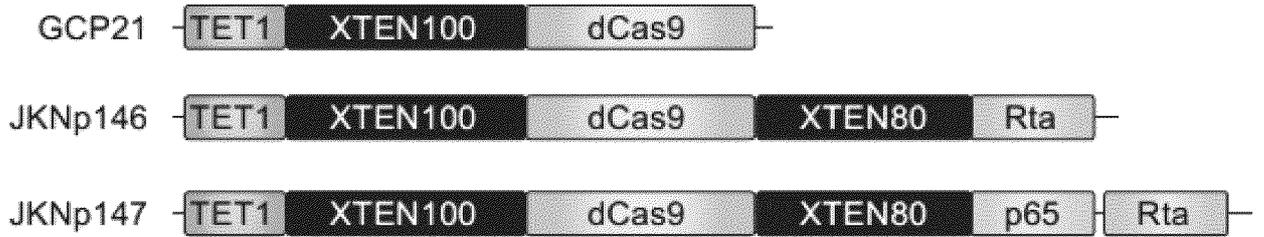
Фиг. 16



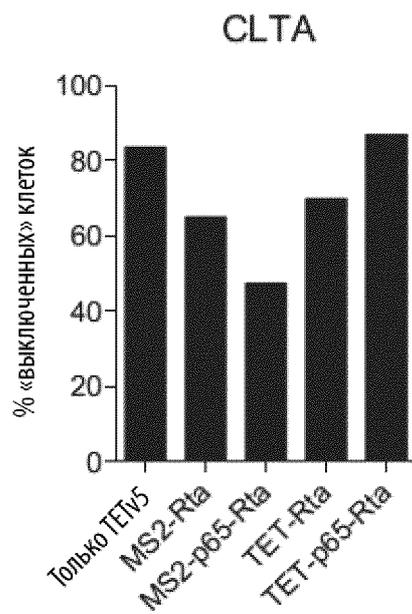
Фиг. 17



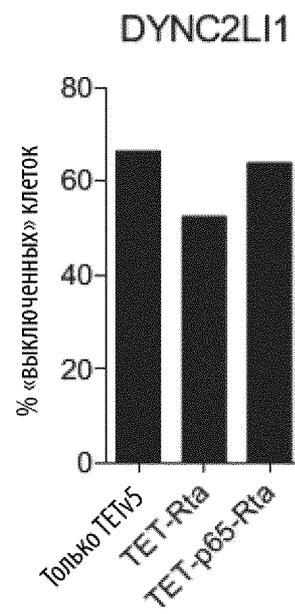
Фиг. 19А



Фиг. 19В



Фиг. 19С



Фиг. 19D

