

(19)



Евразийское
патентное
ведомство

(11) 037359

(13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2021.03.17

(21) Номер заявки
201890012

(22) Дата подачи заявки
2016.06.01

(51) Int. Cl. A61P 31/18 (2006.01)
A61K 48/00 (2006.01)
C12N 5/10 (2006.01)
C12N 15/63 (2006.01)
C12N 9/22 (2006.01)

(54) СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ РНК-НАПРАВЛЯЕМОГО ЛЕЧЕНИЯ ИНФЕКЦИИ, ВЫЗЫВАЕМОЙ ВИЧ

(31) 62/169,384; 62/169,633; 62/308,320

(32) 2015.06.01; 2015.06.02; 2016.03.15

(33) US

(43) 2018.05.31

(86) PCT/US2016/035141

(87) WO 2016/196539 2016.12.08

(71)(73) Заявитель и патентовладелец:
ТЭМПЛ ЮНИВЕРСИТИ - ОФ ЗЕ
КОММОНВЭЛС СИСТЕМ ОФ ХАЙЕ
ЭДЬЮКЕЙШН (US)

(72) Изобретатель:
Халили Камел, Ху Вэньхуэй, Чжан
Юнган (US)

(74) Представитель:
Назина Е.Е., Медведев В.Н. (RU)

(56) WO-A1-2015031775

SAAYMAN et al. The therapeutic application of CRISPR/Cas9 technologies for HIV. Expert Opin Biol Ther ePub 12 April 2015, Vol 15 no 6, Pages 819-830. Especially entire article (review)

(57) Композиции для специфичного расщепления целевых последовательностей в ретровирусах, содержащие нуклеиновые кислоты, которые кодируют эндонуклеазу ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и последовательность направляющей РНК, комплементарную одной или нескольким целевым последовательностям нуклеиновой кислоты в геноме ретровируса.

B1

037359

037359

B1

Изобретение было выполнено при поддержке правительства США в рамках грантов, выданных Национальными институтами здравоохранения (NIH) Камелю Халили (Kamel Khalili) (P30MH092177) и Венхуи Ху (Wenhui Hu) (R01NS087971). Правительство США может обладать определенными правами на настоящее изобретение.

Область техники, к которой относится изобретение

Настоящее изобретение относится к композициям и способам для специфичного расщепления целевых последовательностей в ретровирусах, например в вирусе иммунодефицита человека (ВИЧ). Композиции, которые могут содержать нуклеиновые кислоты, кодирующие эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и последовательность направляющей РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, можно вводить субъекту с инфекцией, вызываемой ВИЧ, или с риском заражения ею.

Предпосылки изобретения

В течение более трех десятилетий после открытия ВИЧ-1 СПИД остается главной проблемой здравоохранения, поражая более 35,3 миллиона людей во всем мире. СПИД остается неизлечимым вследствие перманентной интеграции ВИЧ-1 в геном организма-хозяина. Существующая терапия (высокоактивная антиретровирусная терапия или HAART), предназначенная для контроля инфекции, вызываемой ВИЧ-1, и задержки развития СПИДа, существенно снижает репликацию вируса в клетках, в которых поддерживается инфекция, вызываемая ВИЧ-1, и снижает до минимального уровня вирусную нагрузку в крови. Но HAART не способна подавлять экспрессию и репликацию генома вируса, происходящих на низком уровне в тканях, и не способна целенаправленно воздействовать на латентно инфицированные клетки, например, покоящиеся Т-клетки памяти, макрофаги головного мозга, микроглию и астроциты, клетки лимфоидной ткани кишечника, которые служат в качестве резервуара для ВИЧ-1. Персистирующая инфекция, вызываемая ВИЧ-1, также связана с сопутствующими патологическими процессами, в том числе заболеваниями сердечно-сосудистой системы и почек, остеопенией и неврологическими расстройствами. По-прежнему существует потребность в лечебных терапевтических стратегиях, целенаправленно воздействующих на резервуары персистирующих вирусов.

Существующая терапия, предназначенная для контроля инфекции, вызываемой ВИЧ-1, и предупреждения прогрессирования СПИДа, значительно снизила репликацию вируса в клетках, подверженных инфекции, вызываемой ВИЧ-1, но она не устраняет репликацию вируса на низком уровне в латентно инфицированных клетках, которые содержат интегрированные копии ДНК провируса ВИЧ-1. Существует актуальная необходимость в разработке радикальных терапевтических стратегий, которые целенаправленно воздействуют на резервуары персистирующих вирусов, в том числе стратегии удаления провирусной ДНК из генома клетки-хозяина.

В последние годы было разработано несколько новых систем удаления эндогенных генов, в том числе хоуминг-эндонуклеазы (HE), нуклеазы с цинковыми пальцами (ZFN), эффекторные нуклеазы, подобные активаторам транскрипции (TALEN), и белки CRISPR-ассоциированной системы 9 (Cas9).

В способе с применением CRISPR (коротких палиндромных повторов, регулярно расположенных группами) собирают комплексы для редактирования гена, в том числе нуклеазу Cas9 и направляющую РНК (gRNA), комплементарную целевой последовательности вирусной ДНК. gRNA направляет нуклеазу Cas9 для захвата и расщепления нитей вирусной ДНК, содержащих целевую последовательность. Комплекс Cas9/gRNA редактирования генов вводит одну или более мутаций в вирусную ДНК.

Сообщали о возможности генетического разрушения интегрированного провируса ВИЧ-1 с помощью HE с нацеливанием на консервативные последовательности вирусных белков. ZFN, нацеливающиеся на ген корцептора CCR5 хозяина ВИЧ-1, вошли во 2 фазу клинических исследований по лечению ВИЧ/СПИДа. Было экспериментально показано, что TALEN эффективно расщепляет CCR5 в требуемом сайте. Комплексы редактирования Cas9/gRNA также использовали для разрушения корцепторов для проникновения ВИЧ-1 (CCR5, CXCR4) и провирусных структурных белков (Manjunath et al, *Viruses*, 14;5(11):2748-2766 (2013); Stone et al, *Curr. Opin. HIV AIDS*, 8(3):217-223 (2013); Wang et al, *PLoS One*, 26;9(12):e115987 (2014)). Однако CCR5 не является уникальным рецептором для инфекции, вызываемой ВИЧ-1, и также имеет много других клеточных функций.

Краткое описание изобретения

В настоящем изобретении предусмотрены композиции и способы, относящиеся к лечению и предупреждению ретровирусных инфекций, в частности, вируса иммунодефицита человека ВИЧ. Композиции и способы оказывают воздействие на провирус ВИЧ, который был интегрирован в геном клеток-хозяев.

В частности, в настоящем изобретении предусмотрены композиции, содержащие последовательность нуклеиновой кислоты, кодирующей CRISPR-ассоциированную эндонуклеазу, и одну или более последовательностей выделенной нуклеиновой кислот, кодирующей gRNA, где каждая gRNA является комплементарной целевой последовательности в геноме ретровируса. В предпочтительном варианте осуществления две gRNA включены в состав композиции, при этом каждая gRNA направляет нуклеазу Cas к особому целевому сайту в интегрированной ретровирусной ДНК - ДНК ВИЧ. Фрагмент ДНК между сайтами разреза удаляется, что в результате приводит к вырезанию части или всего генома ВИЧ. Наиболее эффективные комбинации gRNA включают пары, в которых одна gRNA нацеливается на сайт в

LTR-участке, а другая нацеливается на сайт в структурном гене, таком как gag или pol; и пары, в которых обе gRNA нацеливаются на сайты в LTR.

В настоящем изобретении также предусмотрен способ инактивации ретровируса в клетке млекопитающего путем воздействия на клетку с помощью композиции, включающей одну или более выделенных нуклеиновых кислот, кодирующих комплекс редактирования гена. Комплекс редактирования гена включает в себя CRISPR-ассоциированную эндонуклеазу и одну или более gRNA, где каждая gRNA является комплементарной целевой последовательности в ретровирусе.

В настоящем изобретении дополнительно предусматривается фармацевтическая композиция для инактивации интегрированной провирусной ДНК ретровируса в субъекте-млекопитающем. Композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу Cas, и по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую по меньшей мере одну gRNA, комплементарную целевой последовательности в провирусной ДНК ретровируса, такой как ДНК ВИЧ. Пары gRNA, нацеливающиеся на различные сайты в геноме ретровируса, являются предпочтительными. Последовательности выделенной нуклеиновой кислоты включены по меньшей мере в один вектор экспрессии.

В настоящем изобретении еще дополнительно предусмотрен способ лечения субъекта-млекопитающего, инфицированного ретровирусом, например ВИЧ.

Способ предусматривает стадии определения того, что субъект-млекопитающее инфицирован ВИЧ, введения эффективного количества указанной ранее фармацевтической композиции, и лечения субъекта-млекопитающего в отношении инфекции, вызываемой ВИЧ.

В настоящем изобретении также предусмотрен способ лечения, предназначенный для снижения риска инфицирования ретровирусом (например, ВИЧ) субъекта-млекопитающего, подверженного риску возникновения такой инфекции. Способ предусматривает стадии определения того, что субъект-млекопитающее подвержен риску возникновения инфекции, вызываемой ВИЧ, введения эффективного количества указанной ранее фармацевтической композиции, и снижения риска возникновения инфекции, вызываемой ВИЧ, у субъекта-млекопитающего.

В настоящем изобретении дополнительно предусмотрен набор для лечения и профилактики инфекции, вызываемой ВИЧ. Набор содержит отмеренное количество композиции, содержащей по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, и по меньшей мере одну последовательность нуклеиновой кислоты, кодирующую одну или более gRNA, где gRNA являются комплементарными целевым сайтам в ВИЧ. Альтернативно набор может содержать один или более векторов, кодирующих нуклеиновые кислоты. Набор также может содержать упаковочный материал, листок-вкладыш в упаковке с инструкциями по применению, стерильную среду, шприц и/или стерильный контейнер.

Краткое описание графических материалов

На фиг. 1А показана схема Cas9 *Streptococcus pyogenes* (SpCas9), одиночная направляющая РНК (sgRNA) и мотив, смежный с протоспейсером (PAM), при этом показан сайт разрезания в третьем нуклеотиде обеих нитей ДНК от PAM NGG или NAG, и последовательность sgRNA. sgRNA состоит из РНК CRISPR (crRNA), содержащей спейсер из 20 п. о. (затравочную или целевую последовательность), и последовательность типа "стебель-петля" из 12 п. о. (GAAA), и транскрибирующую cRNA (tracRNA).

На фиг. 1В показано схематическое изображение spCas9-экспрессирующего лентивирусного вектора (вверху) и sgRNA-экспрессирующего лентивирусного вектора (внизу). Лентивирусный репортерный вектор (вверху) приобретали в Biosettia, при этом он экспрессирует 3xFlag для иммунологического анализа, репортерный красный флуоресцентный белок (RFP) для упрощенного титрования и анализа FACS и пептид T2A для саморасщепления, предназначенного для предупреждения потенциального влияния RFP на функцию Cas9. sgRNA-экспрессирующий лентивирусный вектор модифицировали на основе вектора Addgene (#50946), имеющего сайт клонирования BbsI, селективный маркер для отбора антибиотиком пурамицином и репортерный голубой флуоресцентный белок (BFP), предназначенный для упрощенного титрования и анализа FACS.

На фиг. 1С показана диаграмма генома ВИЧ-1, демонстрирующая некоторые gRNA, нацеливающиеся на участки LTR, Gag (A-D) и Pol (A, B) ВИЧ-1.

На фиг. 1D показана подробная схема sgRNA, нацеливающих на 400 п. о. в пределах U3-участка LTR ВИЧ-1. Затравочные последовательности с зеленым PAM (подчеркнутым жирным) в смысловой нити и красным PAM (подчеркнутым жирным) в антисмысловой нити были помечены, как указано. Большинство из них также можно подвергать спариванию с нуклеазой Cas9 и РНК-направляемой нуклеазой FokI (RFN), которые могут снижать до 1500 раз включительно потенциальные нецелевые эффекты. Выбор участка из 400 п. о. был основан на его отсутствии во всех используемых в настоящее время лентивирусных векторах, которые используют для доставки генов и доставки sgRNA. Такой выбор будет предупреждать саморасщепление лентивирусных векторов Cas9/gRNA.

На фиг. 1Е показана схема вектора с репортерным геном EcoHIV, содержащим усиленную люцифразу светлячка (eLuc), полученную из вектора на основе ВИЧ-1_{NL4-3} человека. Ген eLuc был встроен между Env и Nef с саморасщепляющимся пептидом 2A перед Nef, в то время как gp120 ВИЧ-1 был замещен

gp80 от экотропного вируса мышинного лейкоза.

На фиг. 2А и 2D показан скрининг в отношении одиночных sgRNA с помощью анализа репортерного гена люциферазы EcoHIV. Клетки HEK293T подвергали котрансфекции лентивирусным вектором экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной sgRNA. Через 2 дня люциферазную активность в лизатах клеток измеряли с помощью системы для анализа люциферазы ONE-GLO™ Luciferase Assay System. Данные представляют собой среднее значение \pm SE, полученное в результате 4 независимых трансфекций.

На фиг. 2В показано, что спаренные sgRNA Gag-D с любой из sgRNA LTR снижали люциферазную активность на 64-96%. Клетки HEK293T подвергали котрансфекции лентивирусным вектором экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной sgRNA. Через 2 дня люциферазную активность в лизатах клеток измеряли с помощью системы для анализа люциферазы ONE-Glo™ Luciferase Assay System. Данные представляют собой среднее значение \pm SE, полученное в результате 4 независимых трансфекций.

На фиг. 2С показано, что спаренные sgRNA LTR-3 с любой из Gag-sgRNA или Pol-sgRNA снижали люциферазную активность на 73-93%. Клетки HEK293T подвергали котрансфекции лентивирусным вектором экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной sgRNA. Через 2 дня люциферазную активность в лизатах клеток измеряли с помощью системы для анализа люциферазы ONE-Glo™ Luciferase Assay System. Данные представляют собой среднее значение \pm SE, полученное в результате 4 независимых трансфекций.

На фиг. 3А показано, что генотипирование с помощью ПЦР подтвердило удаление ДНК ВИЧ-1 между сайтами для нацеливания 5'-LTR и сайтом разрезания Gag-D. Верхняя панель: положение праймеров для ПЦР. Нижняя панель: sgRNA GagD, спаренная с различными sgRNA, нацеливающимися на LTR. Клетки HEK293T подвергали котрансфекции векторами экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной gRNA. Через 2 дня клетки лизировали с использованием 50 mM NaOH при 95°C в течение 10 минут и нейтрализовали с использованием 1 M Tris-HCl. Неочищенные экстракты непосредственно использовали для ПЦР с применением смеси Terra PCR Direct Polymerase (Clontech) и праймеров для ПЦР T361/T458, охватывающих 5'-LTR и 5'-часть Gag (1364 п. о.), что приводило к образованию фрагмента размером 1,35 т. о. в контрольном образце, трансфицированном пустым вектором экспрессии sgRNA.

На фиг. 3В показано, что генотипирование с помощью ПЦР подтвердило удаление ДНК ВИЧ-1 между сайтами для нацеливания 3'-LTR и сайтом разрезания Gag-D. Верхняя панель: положение праймеров для ПЦР. Нижняя панель: sgRNA GagD, спаренная с различными sgRNA, нацеливающимися на LTR. Клетки HEK293T подвергали котрансфекции векторами экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной gRNA. Через 2 дня клетки лизировали с использованием 50 mM NaOH при 95°C в течение 10 мин и нейтрализовали с использованием 1 M Tris-HCl. Неочищенные экстракты использовали непосредственно для ПЦР с применением смеси Terra PCR Direct Polymerase (Clontech) и праймеров для ПЦР T758 (нуклеотиды 796-817)/T645 (нацеливающих на последовательности вектора после 3'-LTR), охватывающих 3'-LTR и весь геном ВИЧ-1 за исключением части последовательности 5' Gag, что приводит к образованию предполагаемого фрагмента размером 9,5 т. о. не выявляемого при условиях стандартной ПЦР в контрольном образце, трансфицированном пустым вектором экспрессии sgRNA.

На фиг. 3С показано, что генотипирование с помощью ПЦР подтвердило удаление ДНК ВИЧ-1 между сайтами для нацеливания LTR-3 и целевыми сайтами Gag или Pol. Верхняя панель: положение праймеров для ПЦР. Нижняя панель: GagD, спаренный с LTR-1, 2, 3, или sgRNA LTR-3, спаренная с различными sgRNA, нацеливающимися на Gag или Pol. Клетки HEK293T подвергали котрансфекции векторами экспрессии с репортерным геном EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной gRNA. Через 2 дня клетки лизировали с использованием 50 mM NaOH при 95°C в течение 10 минут и нейтрализовали с использованием 1 M Tris-HCl. Неочищенные экстракты использовали непосредственно для ПЦР с применением смеси Terra PCR Direct Polymerase (Clontech) и праймеров для ПЦР, охватывающих последовательности между 5'-LTR и 5'-геномной последовательностью (левая панель) или 3'-геномной последовательностью и 3'-LTR (правая панель).

На фиг. 3D показана схема расположения праймеров для ПЦР и сайтов для нацеливания gRNA на Gag/Pol, используемых при проверке эффективности gRNA при генотипировании с помощью ПЦР.

На фиг. 3Е, 3F представлены изображения блоттинга, на которых показаны результаты спаривания GagD с различными gRNA LTR.

На фиг. 3Г, 3Н представлены изображения блоттинга, на которых показаны результаты спаривания LTR-R с различными gRNA, нацеливающимися на Gag и Pol. Была выявлена делеция 5'-LTR-Gag или Gag-3'-LTR. Плотность полосы дикого типа (WT), делеции и вставки количественно определяли с помощью программы NIH Image J. Число под гелем указывает изменение полосы WT (%) по отношению к пустому контролю gRNA после нормализации продуктом ПЦР Cas9 с применением Cas9-специфичных праймеров T477/T478, а также кратность изменений в полосах вставки и делеции по сравнению с поло-

сой WT. Значительные изменения, индуцированные соответствующей gRNA, выделены прямоугольниками. Прямоугольники указывают на образцы, выбранные для ТА-клонирования и секвенирования по Сэнгеру.

На фиг. 3I, 3J показана дополнительная проверка эффективных gRNA с помощью генотипирования с использованием прямой ПЦР с применением дополнительных праймеров, охватывающих структурный участок и 3'-LTR. На фиг. 3I показан прямой праймер (T758) в направлении выше 5' от пар gRNA GagD с обратным праймером (T645) в векторе в направлении ниже от 3'-конца LTR. На фиг. 3J показан прямой праймер (T758) в направлении выше 5' от пар gRNA GagD с обратным праймером (T422) в R-участке 3'-конца LTR. Стрелками указаны неспецифические полосы (ns).

На фиг. 4 показано, что типичное ТА-клонирование и секвенирование по Сэнгеру подтвердило делецию размером 296 п. о. между LTR-1 и LTR-3 и дополнительную вставку размером 180 п. о. между двумя сайтами разрезания. Подготовку образца и прямую ПЦР выполняли, как описано на фиг. 3A. Фрагменты ПЦР после разрезания экстрагировали для ТА-клонирования и секвенирования по Сэнгеру. Красными стрелками указаны предполагаемые сайты разрезания в третьем нуклеотиде от PAM. Красным подчеркнута последовательность PAM.

На фиг. 5A показано схематическое представление генома ВИЧ-1, в том числе стратегия и предполагаемый результат успешного нацеливания на последовательности LTR вируса (фланкирующего интегрированного провируса) с помощью комплексов Cas9/gRNA.

На фиг. 5B показана полная структура LTR.

На фиг. 5C показана последовательность целевых сайтов и их положение в LTR.

На фиг. 6A показана диаграмма репортерной линии клеток Jurkat 2D10, в том числе изображена интегрированная репортерная последовательность ВИЧ-1.

На фиг. 6B показаны микрофотографии флуоресценции, обозначающие индуцированную PMA/TSA повторную активацию латентных провирусных последовательностей.

На фиг. 6C показаны результаты проточной цитометрии в виде гистограмм, представляющие индуцированную PMA/TSA повторную активацию латентных провирусных последовательностей.

На фиг. 7A показаны результаты скринирования одноклеточного клона.

На фиг. 7B показано подтверждение наилучших клонов в соответствии с анализом проточной цитометрии клонов, полученным в скринировании одноклеточных клонов.

На фиг. 7C показаны результаты иллюстративного вестерн-блоттинга для FLAG-Cas9 (верхняя панель) и электрофорез в агарозном геле при проведении RT-PCR (нижняя панель) в случае экспрессии gRNA.

На фиг. 8A показано положение праймеров, используемых в анализе ПЦР удаления провирусного генома ВИЧ-1 из ДНК хозяина. Праймеры были специфичными по отношению к мотиву последовательности провирусного гена Env (RRE), геномным последовательностям, фланкирующим интегрированный провирус с репортерным геном (хромосома 16, ген MSRB1), LTR и контрольному гену b-актина.

На фиг. 8B показано изображение продуктов ПЦР в агарозном геле, при этом стрелками указаны на сайты полос, отсутствующих из-за удаления последовательностей HIV-1.

На фиг. 8C показаны данные ПЦР длинных фрагментов в условиях, оптимизированных для более коротких продуктов, при этом облегчается выявление провирусных лариатных последовательностей в сайте интеграции.

На фиг. 8D показаны результаты секвенирования провирусного лариата.

На фиг. 9A показан флуоресцентный анализ одноклеточных клонов Jurkat 2D10, которые были инфицированы репортерным вирусом ВИЧ-1 NL-4-3-EGFP-P2A-Nef, при этом прогрессирование инфекции отслеживали в течение 18 дней в виде флуоресцентного анализа.

На фиг. 9B показан вестерн-блоттинг, демонстрирующий экспрессию Cas9-FLAG в исследуемых клонах.

На фиг. 9C показано изображение агарозного геля продуктов ПЦР с обратной транскрипцией в случае экспрессии gRNA в некоторых клонах.

На фиг. 10A показаны микрофотографии флуоресценции клеток Jurkat 2D10, трансдуцированных лентивирусами, экспрессирующими RFP-Cas9 и/или gRNA LTR A/B'.

На фиг. 10B показан анализ с помощью проточной цитометрии в случае экспрессии RFP-Cas9 и/или gRNA LTR A/B' и вирусной повторной активации после индукции PMA/TSA.

На фиг. 11A показана последовательность LTR A и LTR B'.

На фиг. 11B показан анализ Surveyor нецелевых мутаций по типу вставки/делеции мутаций после экспрессии Cas9 и LTR A/B'.

На фиг. 11C показаны результаты анализа секвенирования по Сэнгеру нецелевых мутаций по типу вставки/делеции.

На фиг. 11D показано положение сайта интеграции репортерного гена в ВИЧ-1 во втором экзоне гена MSRB1 в хромосоме 16, а также соседние гены.

На фиг. 11E показаны результаты сравнения экспрессии с помощью qRT-PCR соседних генов в контрольных и экспрессирующих Cas9/LTR AB' клетках.

На фиг. 12A-12C показана диаграмма ТА-клонирования и секвенирования по Сэнгеру типичных образцов, подтверждающих делецию предполагаемых фрагментов между соответствующими целевыми сайтами gRNA (фиг. 12A -12C) и различные дополнительные вставки (фиг. 12B, 12C). Последовательно-сти PAM выделены и ножницами указан третий нуклеотид от PAM. Стрелкой указан сайт связывания после расщепления и лигирования.

На фиг. 13A-13E показано получение стабильной линии клеток HEK293T с EcoHIV и люциферазой светлячка. Вирус EсоHIV-eLuc собирали через 48 ч после котрансфекции клеток HEK293T в одной лунке 6-луночного планшета с 3 мкг плазмиды pEcoHIV-eLuc и 1 мкг плазмиды VSV-G. Равный объем вирусного супернатанта (250 мкл) добавляли к клеткам HEK293T (2×10⁵/лунка) в 12-луночном планшете. Через 48 ч люциферазная активность eLuc, измеренная с помощью анализа люциферазы ONE-Glo (Promega), была в 10⁵ раз больше, чем в контрольных клетках HEK293T без обработки вирусом. Затем одиночные клетки культивировали в четырех 96-луночных планшетах при предельном разведении. Через 2-3 недели выжившие колонии клеток выделяли и исследовали в отношении репортерной активности eLuc (фиг. 13A, 13B) с помощью анализа люциферазы ONE-Glo и проверки трансгена EсоHIV-eLuc с помощью гентипирования с применением прямой ПЦР (фиг. 13C - 13E) с использованием праймера T361 (5'-gatctgttgatctaccasaca-3') и T458 (5'-ccsactgtgttagcatgtatt-3'). Половина из 13 одноклеточных клонов в первой стадии эксперимента (фиг. 13A, обозначено как 1016) и половина из 20 клонов во второй стадии (фиг. 13B, 1021) характеризовались различной степенью активности конститутивной eLuc. Лишь два из них (1016-11 и 1021-19) были значимо чувствительными к стимуляции TNFα (10 нг/мл) и PMA (10 нг/мл). Все eLuc-экспрессирующие клоны содержали трансген (фиг. 13C-13E) за исключением клона 1021-6 (зеленый квадрат), который мог утратить последовательность Gag в ходе интеграции, но все еще транскрибировать eLuc посредством альтернативного сплайсинга. Два клона (1016-6 и 1016-9) содержали трансген, однако не проявляли eLuc-активность даже после обработки средствами, возвращающими латентность, такими как TNFα /PMA и другие (красный квадрат На фиг. 13A и 13D). Таким образом, эти клоны с eLuc-активностью отбирали и сохраняли для последующих исследований. Некоторые из этих клонов нельзя было пересевать, очевидно, в связи с непрерывным образованием токсических вирусных белков, а некоторые клоны были устойчивы к трансфекции, даже при использовании набора для трансфекции с липофектаминоном.

На фиг. 14A-14D показаны результаты тестирования эффективности gRNA с помощью стабильной линии клеток HEK293T с EсоHIV и люциферазой светлячка. Фиг. 14A, 14B: клон, стабильно экспрессирующий EсоHIV-eLuc, дополнительно трансфицировали лентивирусом на основе pCW-Cas9 при 10 MOI и проводили отбор с использованием пурамицина (1 мкг/мл) в течение 2 недель. Клетки трансфицировали указанными gRNA-экспрессирующими векторами. Через 2 дня проводили анализ люциферазы ONE-Glo. Фиг. 14C, 14D: клетки, стабильно экспрессирующие EсоHIV-eLuc, котрансфицировали указанными gRNA-экспрессирующими векторами и вектором pLV-EF1α-spCas9-T2A-RFP. Люциферазную активность измеряли с помощью анализа люциферазы ONE-Glo через 48 ч. Данные представляют собой среднее значение ± SEM по 4 независимым трансфекциям с процентными изменениями eLuc-активности по сравнению с группой пустых векторов Zero gRNA.

Подробное описание

Настоящее изобретение основано, помимо всего прочего, на обнаружении того, что можно элиминировать интегрированный геном вируса иммунодефицита человека (ВИЧ) из клеток, инфицированных ВИЧ, путем применения РНК-направляемой системы коротких палиндромных повторов, регулярно расположенных группами (CRISPR)-нуклеазы Cas 9 (Cas9/gRNA) в одиночной и мультиплексной конфигурациях.

Определения

Если не определено иное, все технические и научные термины, используемые в данном документе, имеют такое же значение, которое обычно понимает специалист в области техники, к которой относится настоящее изобретение. Хотя любые способы и материалы, аналогичные или эквивалентные описанным в данном документе, можно применять на практике для тестирования настоящего изобретения, предпочтительные материалы и способы описаны в данном документе. В описании и формуле настоящего изобретения используется следующая терминология. Также следует понимать, что терминология, используемая в данном документе, служит только для цели описания конкретных вариантов осуществления и не предполагается как ограничивающая.

Предполагается, что все гены, названия генов и продукты генов, раскрытые в данном документе, соответствуют гомологам из любого вида, для которого композиции и способы, раскрытые в данном документе, являются применимыми. Следует понимать, что при раскрытии гена или продукта гена из конкретного вида предполагается, что настоящее раскрытие является только иллюстративным и не должно интерпретироваться как ограничивающее, если это четко не указано в контексте, в котором оно представлено. Таким образом, например, предполагается, что гены или продукты генов, раскрытые в данном документе, охватывают гомологичные и/или ортологичные гены и продукты генов из других видов.

Формы единственного числа используются в данном документе для обозначения одного или более

чем одного (т. е. по меньшей мере одного) грамматического объекта предмета. В качестве примера, "элемент" означает один элемент или более чем один элемент. Таким образом, упоминание "клетки", например, предусматривает множество клеток одного и того же типа. Дополнительно, в той степени, в которой термины "включающий", "включает", "имеющий", "имеет", "с" или их варианты используются в подробном описании и/или формуле изобретения, предполагается, что такие термины являются включающими аналогично термину "содержащий".

Используемые в данном документе выражения "содержащий", "содержит" или "содержал" и их варианты, применяемые в отношении определяемых или описываемых элементов объекта, композиции, устройства, способа, процесса, системы и т. д., предназначены быть включающими или открытыми, допуская наличие дополнительных элементов, что указывает тем самым на то, что определяемые или описываемые объект, композиция, устройство, способ, процесс, система и т.д. включают эти указанные элементы или, при необходимости, их эквиваленты, и что другие элементы могут быть включены и по-прежнему находиться в пределах объема/определения определяемых объекта, композиции, устройства, способа, процесса, системы и т.д.

Выражение "приблизительно", используемое в данном документе в отношении измеряемой величины, такой как количество, временная продолжительность и т.п., предназначен для охвата отклонений в $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, $\pm 1\%$ или $\pm 0,1\%$ от указанного значения, поскольку такие отклонения являются подходящими для осуществления раскрытых способов. В качестве альтернативы, особенно в отношении биологических систем или процессов, данный термин может означать нахождение в рамках порядка величины в пределах до 5-кратного, а также в пределах до 2-кратного значения величины. В случаях, когда в настоящей заявке и формуле изобретения описаны конкретные величины, то, если не указано иное, следует полагать, что термин "приблизительно" означает нахождение в пределах приемлемого интервала погрешности для конкретной величины.

"Кодирующий" относится к внутренне присущему свойству специфических последовательностей нуклеотидов в полинуклеотиде, таком как ген, кДНК или мРНК, служить в качестве матриц для синтеза в ходе биологических процессов других полимеров и макромолекул, имеющих определенную последовательность нуклеотидов (т.е. рРНК, тРНК и мРНК) либо определенную последовательность аминокислот и биологические свойства, обусловленные ими. Таким образом, ген кодирует белок, если в результате транскрипции и трансляции мРНК, соответствующей этому гену, в клетке или другой биологической системе продуцируется белок. Как кодирующая нить, нуклеотидная последовательность которой идентична последовательности мРНК и обычно представлена в перечнях последовательностей, так и не кодирующая нить, применяемая в качестве матрицы для транскрипции гена или кДНК, могут называться кодирующей белок или другой продукт этого гена или кДНК.

"Эффективное количество" или "терапевтически эффективное количество" соединения представляет собой количество соединения, которое является достаточным для предоставления полезного эффекта субъекту, которому вводят соединение. "Эффективное количество" средства для доставки представляет собой количество, достаточное для эффективного связывания или доставки соединения.

Термин "удаление" вируса, например ВИЧ, используемый в данном документе, означает, что вирус неспособен к репликации, геном удален, фрагментирован, разрушен, генетически инактивирован или имеется любое другое физическое, биологическое, химическое или структурное проявление, которое препятствует переносу вируса или инфицированию им любой другой клетки или субъекта, приводя к очищению от вируса *in vivo*. В некоторых случаях фрагменты генома вируса могут поддаваться выявлению, однако вирус неспособен к репликации или инфицированию и т.д.

"Вектор экспрессии" относится к вектору, содержащему рекомбинантный полинуклеотид, содержащий последовательности, контролирующей экспрессию, функционально связанные с нуклеотидной последовательностью, подлежащей экспрессии. Вектор экспрессии содержит действующие в *cis* положении элементы, достаточные для экспрессии; другие элементы для экспрессии могут быть предоставлены клеткой-хозяином или в системе экспрессии *in vitro*. Векторы экспрессии включают в себя все таковые векторы, известные из уровня техники, такие как космиды, плазмиды (например, "голые" или содержащиеся в липосомах) и вирусы (например, лентивирусы, ретровирусы, аденовирусы и аденоассоциированные вирусы), которые содержат в своем составе рекомбинантный полинуклеотид.

"Гомологичный" относится к сходству последовательности или идентичности последовательности между двумя полипептидами или между двумя молекулами нуклеиновой кислоты. Если положение в обеих из двух сравниваемых последовательностей занято одним и тем же основанием или аминокислотной мономерной субъединицей, например, если положение в каждой из двух молекул ДНК занято аденином, тогда молекулы являются гомологичными в этом положении. Процент гомологии между двумя последовательностями является функцией от числа совпадающих или гомологичных положений, общих для двух последовательностей, разделенной на количество положений по отношению к X 100. Например, если 6 из 10 положений в двух последовательностях совпадают или гомологичны, то две последовательности гомологичны на 60%. Например, ДНК-последовательности ATTGCC и TATGGC обладают гомологией, составляющей 50%. Как правило, сравнение выполняют в том случае, если две последовательности выравнивают с целью получения максимальной гомологии.

"Выделенный" означает измененный относительно естественного состояния или изъятый из него. Например, нуклеиновая кислота или пептид, в естественных условиях присутствующие в живом животном, не являются "выделенными", однако те же самые нуклеиновая кислота или пептид, частично или полностью отделенные от сопутствующих материалов в их естественном состоянии, являются "выделенными". Выделенные нуклеиновая кислота или белок могут существовать в практически очищенной форме или могут существовать в ненативном окружении, таком как, например, в клетке-хозяине.

В контексте настоящего изобретения используют следующие аббревиатуры для часто встречающихся оснований нуклеиновой кислоты. "А" означает аденозин, "С" означает цитозин, "G" означает гуанозин, "Т" означает тимидин и "U" означает уридин.

Если не указано иное, "нуклеотидная последовательность, кодирующая аминокислотную последовательность" включает в себя все нуклеотидные последовательности, которые являются вырожденными вариантами друг друга и которые кодируют одну и ту же аминокислотную последовательность. Фраза "нуклеотидная последовательность, кодирующая белок или РНК" может также включать в себя интроны в той степени, в какой некоторый вариант нуклеотидной последовательности, кодирующей белок, может содержать интрон(интроны).

Выражения "пациент", "субъект", "индивидуум" и им подобные используются взаимозаменяемо, и они относятся к любому животному или их клеткам, вне зависимости *in vitro* или *in situ*, восприимчивым к способам, описанным в данном документе. В определенных неограничивающих вариантах осуществления пациент, субъект или индивидуум представляет собой человека.

"Парентеральное" введение композиции включает, например, методики подкожной (s.c), внутривенной (i.v.), внутримышечной (i.m.) или интратеральной инъекции или инфузии.

Используемые в данном документе термины "полинуклеотид", "последовательность нуклеиновой кислоты" или "ген" используются взаимозаменяемо во всем настоящем описании и включают комплементарную ДНК (кДНК), линейные или кольцевые олигомеры или полимеры из природных и/или модифицированных мономеров или связей, в том числе дезоксирибонуклеозидов, рибонуклеозидов, их замещенных и альфа-аномерных форм, пептидных нуклеиновых кислот (PNA), "закрытых" нуклеиновых кислот (LNA), фосфоротиоатов, метилфосфонатов и им подобных. Термин "полинуклеотид", используемый в данном документе, определяется как цепь из нуклеотидов. Кроме того, нуклеиновые кислоты представляют собой полимеры из нуклеотидов. Таким образом, нуклеиновые кислоты и полинуклеотиды используются в данном документе взаимозаменяемо. Специалист в данной области имеет общее представление о том, что нуклеиновые кислоты представляют собой полинуклеотиды, которые можно гидролизовать на мономерные "нуклеотиды". Мономерные нуклеотиды могут быть гидролизованы в нуклеозиды. Полинуклеотиды включают в себя без ограничения все последовательности нуклеиновых кислот, получаемые с помощью любых способов, доступных в данной области техники, в том числе без ограничения рекомбинантных способов, т. е. клонирования последовательностей нуклеиновых кислот из библиотеки рекомбинантных молекул или генома клетки с помощью стандартной технологии клонирования, PCR™ и т.п., а также синтетических способов. Последовательности нуклеиновых кислот могут быть "химерными", то есть состоять из разных участков. В контексте настоящего изобретения "химерные" соединения представляют собой олигонуклеотиды, которые содержат два или более химических участка, например, участок(участки) ДНК, участок(участки) РНК, участок(участки) PNA и т.д. Каждый химический участок образован по меньшей мере одним мономерным звеном, т. е. нуклеотидом. Эти последовательности обычно содержат по меньшей мере один участок, в котором последовательность модифицирована в целях проявления одного или более желаемых свойств.

"Аналоги" по отношению к нуклеозидам включают синтетические нуклеозиды, имеющие модифицированные фрагменты оснований и/или модифицированные фрагменты Сахаров, например в целом описанные в Scheit, Nucleotide Analogs, John Wiley, New York, 1980; Freier & Altmann, Nucl. Acid. Res., 1997, 25(22), 4429-4443, Toulme, J.J., Nature Biotechnology 19:17-18 (2001); Manoharan M., Biochemica et Biophysica Acta 1489:117-139(1999); Freier S. M., Nucleic Acid Research, 25:4429-4443 (1997), Uhlman, E., Drug Discovery & Development, 3: 203-213 (2000), Herdewin P., Antisense & Nucleic Acid Drug Dev., 10:297-310 (2000); 2'-О, 3'-С-связанные [3.2.0] бициклоарабинонуклеозиды (см., например, N.K Christensen, et al., J. Am. Chem. Soc, 120: 5458-5463 (1998). Такие аналоги включают синтетические нуклеозиды, сконструированные с целью повышения связывающих свойств, например, стабильности дуплексов или триплексов, специфичности или им подобных.

Термин "вариант", используемый применительно к полинуклеотидной последовательности, может охватывать полинуклеотидную последовательность, родственную гену дикого типа. Данное определение также может включать в себя, например, "аллельные", "сплайс-", "видовые" или "полиморфные" варианты. Сплайс-вариант может характеризоваться значительной идентичностью с эталонной молекулой, но, как правило, будет иметь большее или меньшее количество полинуклеотидов вследствие альтернативного сплайсинга экзонов в ходе процессинга мРНК. Соответствующий полипептид может иметь дополнительные функциональные домены или не иметь доменов. Видовые варианты представляют собой полинуклеотидные последовательности, различающиеся между видами. Особенно полезными в настоящем изобретении являются варианты продуктов генов дикого типа. Варианты могут образовываться в резуль-

тате по меньшей мере одной мутации в последовательности нуклеиновой кислоты и могут обуславливать образование измененных мРНК или полипептидов, структура или функция которых может быть или может не быть изменена. Любой указанный природный или рекомбинантный ген может не иметь, иметь одну или много аллельных форм. Широко распространенные мутационные изменения, которые приводят к образованию вариантов, обычно относят к естественным делециям, добавлениям или заменам нуклеотидов. Каждый из этих типов изменений может иметь место в отдельности или в комбинации с другими один или более раз в указанной последовательности.

Если не указано иное, "нуклеотидная последовательность, кодирующая аминокислотную последовательность" включает в себя все нуклеотидные последовательности, которые являются вырожденными вариантами друг друга и которые кодируют одну и ту же аминокислотную последовательность. Фраза "нуклеотидная последовательность, кодирующая белок или РНК" может также включать в себя интроны в той степени, в какой некоторый вариант нуклеотидной последовательности, кодирующей белок, может содержать интрон(интроны).

Используемые в данном документе термины "пептид", "полипептид" и "белок" используются взаимозаменяемо и означают соединение, состоящее из аминокислотных остатков, ковалентно связанных пептидными связями. Белок или пептид должен содержать по меньшей мере две аминокислоты, и отсутствует ограничение на максимальное количество аминокислот, которые могут составлять последовательность белка или пептида. Таким образом, например, термины олигопептид, белок и фермент включены в определение полипептида или пептида независимо от того, получены ли они с помощью рекомбинантных методик, химического или ферментативного синтеза, или встречаются в природе. Полипептиды включают в себя любой пептид или белок, содержащий две или более аминокислот, соединенных друг с другом пептидными связями. Как используется в данном документе, данный термин относится как к коротким цепям, которые также зачастую называются в данной области техники, например, пептидами, олигопептидами и олигомерами, так и к более длинным цепям, которые, как правило, называются в данной области техники белками, которых существует много типов. "Полипептиды" включают в себя, например, биологически активные фрагменты, практически гомологичные полипептиды, олигопептиды, гомодимеры, гетеродимеры, варианты полипептидов, модифицированные полипептиды, производные, аналоги, белки слияния. Этот термин также включает в себя полипептиды, которые были модифицированы или дериватизированы, например, среди прочего, путем гликозилирования, ацетилирования, фосфорилирования и др. Полипептиды включают в себя природные пептиды, рекомбинантные пептиды, синтетические пептиды или их комбинацию.

Используемый в данном документе "вариант" полипептидов относится к аминокислотной последовательности, которая изменена с помощью одного или более аминокислотных остатков. Вариант может иметь "консервативные" изменения, где замененная аминокислота характеризуется аналогичными структурными и химическими свойствами (например, замена лейцина изолейцином). Реже вариант может иметь "неконсервативные" изменения (например, замена глицина трипрофаном). Аналогичные незначительные изменения могут также включать делеции или вставки аминокислот, или и первое, и второе. Руководство по определению того, какие аминокислотные остатки могут быть заменены, вставлены или удалены без нарушения биологической активности, можно найти с помощью компьютерных программ, хорошо известных в данной области, например компьютерной программы LASERGENE (DNASTAR).

"Терапевтический" препарат представляет собой препарат, вводимый субъекту, у которого проявляются признаки патологии, с целью уменьшения или устранения этих признаков.

Фраза "терапевтически эффективное количество", используемая в данном документе, относится к количеству, которое является достаточным для предупреждения или лечения (задержки или предупреждения возникновения, предупреждения прогрессирования, ингибирования, ослабления или нормализации) заболевания или состояния, в том числе облегчения симптомов таких заболеваний.

"Лечение" представляет собой мероприятие, осуществляемое с целью предупреждения развития или изменения патологии или симптомов нарушения. Соответственно, "лечение" относится к терапевтическому лечению и профилактическим или превентивным мероприятиям. "Лечение" также может быть определено как паллиативное лечение. Нуждающиеся в лечении включают тех, у которых уже имеется нарушение, а также тех, у которых нарушение можно предупредить. Соответственно, "лечение" нарушения или состояния или "лечить" нарушение или состояние предусматривает: (1) предупреждение или задержку появления клинических симптомов состояния, нарушения или расстройства, развивающихся у человека или другого млекопитающего, которые могут быть поражены или предрасположены к состоянию, нарушению или расстройству, но пока не испытывают или не проявляют клинических или субклинических симптомов нарушения, состояния или расстройства; (2) ингибирование состояния, нарушения или расстройства, т.е., приостановку, ослабление или задержку развития заболевания или его рецидива (в случае поддерживающего лечения) или по меньшей мере его одного клинического или субклинического симптома; или (3) облегчение заболевания, т.е. способствование регрессии состояния, нарушения или расстройства или по меньшей мере одного из его клинических или субклинических симптомов. Преимущество для индивидуума, подлежащего лечению, является статистически значимым или по меньшей мере различаемым пациентом или лечащим врачом.

"Вектор" представляет собой композицию, которая содержит выделенную нуклеиновую кислоту и которую можно применять для доставки выделенной нуклеиновой кислоты внутрь клетки. Многочисленные векторы известны в данной области, в том числе без ограничений линейные полинуклеотиды, полинуклеотиды, связанные с ионными или амфифильными соединениями, плазмиды и вирусы. Таким образом, термин "вектор" включает автономно реплицирующиеся плазмиду или вирус. Данный термин также следует истолковывать как включающий в себя неплазмидные и невирусные соединения, которые облегчают перенос нуклеиновой кислоты в клетки, такие как, например, полилизиновые соединения, липосомы и т. п. Примеры вирусных векторов включают в себя без ограничения аденовирусные векторы, векторы на основе аденоассоциированного вируса, ретровирусные векторы и т. п.

Диапазоны: во всем настоящем раскрытии различные аспекты настоящего изобретения могут быть представлены в формате диапазонов. Следует понимать, что описание в формате диапазонов служит только для удобства и краткости и не должно истолковываться как жесткое ограничение объема настоящего изобретения. Соответственно, следует считать, что в описании диапазона конкретно раскрыты все возможные поддиапазоны, а также отдельные числовые величины в пределах этого диапазона. Например, следует считать, что в описании такого диапазона, как 1-6, конкретно раскрыты такие поддиапазоны, как 1-3, 1-4, 1-5, 2-4, 2-6, 3-6 и т. д., а также отдельные числа в пределах этого диапазона, например, 1, 2, 2,7, 3, 4, 5, 5,3 и 6. Это применимо независимо от ширины диапазона.

В случаях, когда какая-либо аминокислотная последовательность конкретно указана под номером доступа в Swiss-Prot или GENBANK, данная последовательность включена в данный документ посредством ссылки. Информация, связанная с номером доступа, как, например, идентификация сигнального пептида, внеклеточного домена, трансмембранного домена, промоторной последовательности и места начала трансляции, также включена в данный документ посредством ссылки во всей своей полноте.

Композиции на основе CRISPR-нуклеазы

Было показано, что применение технологии на основе Cas9 в удалении резервуара ВИЧ-1, в частности нацеливание на LTR, является перспективной стратегией для лечения и возможного излечения СПИДа. Hu et al., PNAS 2014, 111:114616, раскрыли, что стабильная трансфекция культур клеток человека плазмидами, экспрессирующими Cas9/gRNA, нацеленными на сайты в LTR ВИЧ-1, приводила к успешной эрадикации части и/или всего генома ВИЧ-1 без нарушения функции клетки-хозяина. Целевые сайты были обозначены LTR-A, LTR-B, LTR-C и LTR-D. Нацеливание на два различных сайта в LTR было особенно эффективным при получении делеций, достаточно протяженных для того, чтобы привести к вырезанию всей или практически всей последовательности провирусной ДНК.

Предшествующее существование Cas9/gRNA в клетках также предупреждало новую инфекцию, вызываемую ВИЧ-1.

ВИЧ и другие ретровирусы являются высокомультиплексными, поэтому существует необходимость в более широком спектре реагентов и способов на основе Cas9/gRNA для нацеливания на интегрированный геном ВИЧ. Особенно применимыми были бы реагенты на основе Cas9/gRNA, которые эффективно нацеливаются на структурные гены ВИЧ, такие как gag и pol.

Соответственно, варианты осуществления настоящего изобретения направлены на композиции и способы лечения высокомультиплексных и/или латентных вирусов и их удаления из клетки-хозяина *in vitro* или *in vivo*. Способы по настоящему изобретению можно использовать для выведения вирусного или другого чужеродного материала из организма хозяина без нарушения целостности генетического материала хозяина. Можно использовать нуклеазу для нацеливания на вирусную нуклеиновую кислоту, при этом нарушается репликация или транскрипция вирусов или даже происходит вырезание генетического материала вирусов из генома хозяина. Нуклеазу можно специфически нацеливать для удаления лишь вирусной нуклеиновой кислоты без воздействия на материал хозяина, когда вирусная нуклеиновая кислота существует в виде частицы в клетке или когда она интегрирована в геном хозяина. Нацеливание на вирусную нуклеиновую кислоту можно выполнять с помощью специфичного в отношении последовательности фрагмента, такого как направляющая РНК, которая нацеливается на вирусный геномный материал с целью разрушения с помощью нуклеазы и не нацеливается на геном клетки-хозяина. В некоторых вариантах осуществления используется нуклеаза CRISPR/Cas и направляющая РНК (gRNA), которые совместно нацеливаются и избирательно редактируют или разрушают вирусный геномный материал. CRISPR (короткие палиндромные повторы, регулярно расположенные группами) представляет собой встречающийся в природе элемент бактериальной иммунной системы, который защищает бактерий от инфекции фагами. Направляющая РНК определяет место расположения комплекса CRISPR/Cas в вирусной целевой последовательности. Связывание комплекса определяет место расположения эндонуклеазы Cas в целевой последовательности вирусного генома, что приводит к разрывам в вирусном геноме. Можно использовать другие системы нуклеаз, в том числе, например, нуклеазы с цинковыми пальцами, эффекторные нуклеазы, подобные активаторам транскрипции (TALEN), мегануклеазы или любые другие системы, которые можно использовать для разрушения или нарушения вирусной нуклеиновой кислоты без нарушения нормальной функции генетического материала хозяина.

Композиции, включенные в данный документ, можно использовать для нацеливания на вирусную нуклеиновую кислоту в любой форме или на любой стадии жизненного цикла вируса. Целевая вирусная

нуклеиновая кислота может присутствовать в клетке-хозяине в виде независимых частиц. В предпочтительном варианте осуществления вирусная инфекция является латентной и вирусная нуклеиновая кислота интегрирована в геном хозяина. На любую подходящую вирусную нуклеиновую кислоту можно осуществлять нацеливание в целях расщепления и переваривания.

Системы CRISPR/Cas

Система CRISPR-Cas включает в себя комплекс редактирования гена, содержащий CRISPR-ассоциированную нуклеазу, например Cas9, и направляющую РНК, комплементарную целевой последовательности, расположенной на нити ДНК, такой как целевая последовательности в провирусной ДНК, интегрированной в геном млекопитающего. Иллюстративный комплекс редактирования гена показан на фиг. 1А. Комплекс редактирования гена может расщеплять ДНК в целевой последовательности. Расщепление может в свою очередь вызывать введение различных мутаций в провирусную ДНК, что приводит к инактивации стадии провируса ВИЧ. Механизм, с помощью которого такие мутации инактивируют провирус, может варьироваться. Например, мутация может нарушать репликацию провируса и экспрессию гена вируса. Мутации могут быть расположены в регуляторных последовательностях или последовательностях структурных генов и приводить к нарушению продуцированию ВИЧ. Мутация может предусматривать делецию. Размер делеции может варьировать от одной пары нуклеотидных оснований до приблизительно 10000 пар оснований. В некоторых вариантах осуществления делеция может включать в себя всю или практически всю интегрированную последовательность нуклеиновой кислоты ретровируса. В некоторых вариантах осуществления делеция может включать в себя полную последовательность интегрированной нуклеиновой кислоты ретровируса. Мутация может предусматривать вставку, то есть добавление одной или более пар нуклеотидных оснований в последовательность провируса. Размер вставленной последовательности также может варьировать, например от приблизительно одной пары оснований до приблизительно 300 пар нуклеотидных оснований. Мутация может представлять собой точечную мутацию, т.е. замещение одного нуклеотида другим нуклеотидом. Применимыми точечными мутациями являются те, которые характеризуются функциональными последствиями, например, мутации, которые приводят к превращению кодона аминокислоты в стоп-кодон или которые приводят к получению нефункционального белка.

Как правило, белки CRISPR/Cas содержат по меньшей мере один домен распознавания РНК и/или связывания РНК. Домены распознавания РНК и/или связывания РНК взаимодействуют с направляющими РНК. Белки CRISPR/Cas также могут содержать домены нуклеаз (т.е. домены ДНКаз или РНКаз), домены связывания ДНК, домены хеликаз, домены РНКаз, домены белок-белкового взаимодействия, домены димеризации, а также другие домены.

В некоторых вариантах осуществления CRISPR/Cas-подобный белок может представлять собой белок CRISPR/Cas дикого типа, модифицированный белок CRISPR/Cas или фрагмент белка CRISPR/Cas дикого типа или модифицированного белка CRISPR/Cas. CRISPR/Cas-подобный белок можно модифицировать с целью повышения аффинности связывания с нуклеиновой кислотой и/или специфичности, изменения ферментативной активности и/или изменения другого свойства белка. Например, домены нуклеазы (т.е., ДНКазы, РНКазы) CRISPR/Cas-подобного белка можно модифицировать, удалять или инактивировать. Альтернативно CRISPR/Cas-подобный белок можно подвергать усечению с целью удаления доменов, которые не являются необходимыми для функционирования слитого белка. CRISPR/Cas-подобный белок также можно подвергать усечению или модифицировать с целью оптимизации активности эффекторного домена слитого белка.

В вариантах осуществления система CRISPR/Cas может представлять собой систему I типа, II типа или III типа. Неограничивающие примеры подходящих белков CRISPR/Cas включают Cas3, Cas4, Cas5, Cas5e (или CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (или CasA), Cse2 (или CasB), Cse3 (или CasE), Cse4 (или CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csz1, Csx15, Csf1, Csf2, Csf3, Csf4 и Cu1966.

В одном варианте осуществления РНК-направляемую эндонуклеазу получают из системы CRISPR/Cas II типа. В других вариантах осуществления РНК-направляемую эндонуклеазу получают из белка Cas9. Белок Cas9 может происходить из

Streptococcus pyogenes, *Streptococcus thermophilus*, *Streptococcus* sp.,
Nocardiopsis dassonvillei, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*,
Streptomyces viridochromogenes, *Streptosporangium roseum*,
Alicyclobacillus acidocaldarius, *Bacillus pseudomycoides*, *Bacillus selenitireducens*,
Exiguobacterium sibiricum, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*,
Microscilla marina, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*,
Polaromonas sp., *Crocospaera watsonii*, *Cyanotheca* sp., *Microcystis aeruginosa*,
Synechococcus sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicellulosiruptor*

becscii, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia magna*, *Natronaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrotoga mobilis*, *Thermosipho africanus* или *Acaryochloris marina*.

В некоторых вариантах осуществления CRISPR/Cas-подобный белок можно получить из белка Cas9 дикого типа или его фрагмента. В других вариантах осуществления CRISPR/Cas-подобный белок можно получить из модифицированного белка Cas9. Например, аминокислотную последовательность белка Cas9 можно модифицировать с целью изменения одного или нескольких свойств (например, активности, аффинности, стабильности нуклеазы и т.д.) белка. Альтернативно домены белка Cas9, не вовлеченные в РНК-направляемое расщепление, можно удалить из белка таким образом, что модифицированный белок Cas9 будет меньшего размера, чем белок Cas9 дикого типа.

Иллюстративной и предпочтительной CRISPR-ассоциированной нуклеазой является нуклеаза Cas9. Нуклеаза Cas9 может иметь нуклеотидную последовательность, идентичную последовательности *Streptococcus pyogenes* дикого типа. В некоторых вариантах осуществления CRISPR-ассоциированная эндонуклеаза может представлять собой последовательность из другого вида, например другого вида *Streptococcus*, такого как *thermophilus*; *Pseudomonas aeruginosa*, *Escherichia coli* или других секвенированных геномов бактерий и архей, или других прокариотических микроорганизмов. Альтернативно последовательность Cas9 *Streptococcus pyogenes* дикого типа может быть модифицирована. Последовательность нуклеиновой кислоты может быть кодон-оптимизирована для эффективной экспрессии в клетках млекопитающих, т.е. "гуманизирована". Гуманизированная последовательность нуклеазы Cas9, например, может быть последовательностью нуклеазы Cas9, закодированной любым из векторов экспрессии, перечисленных под номерами доступа Genbank KM099231.1 GI:669193757; KM099232.1 GL669193761 или KM099233.1 GI: 669193765. В качестве альтернативы последовательность нуклеазы Cas9, например, может представлять собой последовательность, содержащуюся в коммерчески доступном векторе, таком как PX330 или PX260 от Addgene (Кембридж, Массачусетс). В некоторых вариантах осуществления эндонуклеаза Cas9 может иметь аминокислотную последовательность, которая является вариантом или фрагментом любой из последовательностей эндонуклеазы Cas9 под номерами доступа в Genbank KM099231.1 GL669193757; KM099232.1 GI:669193761 или KM099233.1 GL669193765 или аминокислотной последовательности Cas9 PX330 или PX260 (Addgene, Кембридж, Массачусетс). Нуклеотидную последовательность Cas9 можно модифицировать так, чтобы она кодировала биологически активные варианты Cas9, и эти варианты могут иметь или могут содержать, например, аминокислотную последовательность, которая отличается от Cas9 дикого типа вследствие того, что содержит одну или более мутаций (например, мутацию по типу добавления, делеции или замены или комбинацию таких мутаций). Одна или более мутаций по типу замены могут представлять собой замену (например, консервативную аминокислотную замену). Например, биологически активный вариант полипептида Cas9 может иметь аминокислотную последовательность, характеризующуюся по меньшей мере или приблизительно 50% идентичностью последовательности (например, по меньшей мере или приблизительно 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% идентичностью последовательности) с полипептидом Cas9 дикого типа. Консервативные аминокислотные замены обычно включают замены в пределах следующих групп: глицин и аланин; валин, изолейцин и лейцин; аспарагиновая кислота и глутаминовая кислота; аспарагин, глутамин, серин и треонин; лизин, гистидин и аргинин; а также фенилаланин и тирозин. Аминокислотные остатки в аминокислотной последовательности Cas9 могут представлять собой не встречающиеся в природе аминокислотные остатки. Встречающиеся в природе аминокислотные остатки включают в себя те, которые в естественных условиях кодируются с помощью генетического кода, а также нестандартные аминокислоты (например, аминокислоты, имеющие D-конфигурацию вместо L-конфигурации). Пептиды согласно настоящему изобретению также могут содержать аминокислотные остатки, которые являются модифицированными вариантами стандартных остатков (например, пирролизин можно использовать вместо лизина, а селеноцистеин можно использовать вместо цистеина). Не встречающимися в природе аминокислотными остатками являются такие, которые не были обнаружены в природе, но которые соответствуют основной формуле аминокислоты и могут быть включены в состав пептида. Такие включают D-аллоизолейцин(2R,3S)-2-амино-3-метилпентановую кислоту и L-циклопентилглицин-(S)-2-амино-2-циклопентилуксусную кислоту. Что касается других примеров, можно получить информацию из учебных пособий или всемирной сети Интернет (сайт в настоящее время поддерживается Калифорнийским технологическим институтом, и на нем представлены структуры неприродных аминокислот, которые были успешно включены в состав функциональных белков).

Последовательность нуклеазы Cas9 может быть подвергнутой мутации последовательностью. Например, нуклеаза Cas9 может быть подвергнута мутированию в консервативных доменах HNH и RuvC, которые вовлечены в специфическое в отношении нити расщепление. Например, мутация аспаратат-на-аланин (D10A) в каталитическом домене RuvC позволяет никазному мутанту Cas9 (Cas9n) надрезать, а не расщеплять ДНК с получением одностебельных разрывов, и последующая преимущественная репарация посредством HDR может потенциально снижать частоту нежелательных мутаций по типу вставки/делеции, образующихся в результате нецелевых двухстебельных разрывов.

Настоящее изобретение включает несколько преимуществ по сравнению с системой Cas9/gRNA, описанной в Hu, et al, PNAS 2014, 111:114616. В экспериментах, раскрытых в разделе "Примеры", дополнительные высокоспецифичные целевые последовательности были выявлены в LTR ВИЧ-1 и в структурных генах ВИЧ-1. Эти целевые последовательности (также обозначаемые как целевые "сайты") эффективно редактировали с помощью Cas9/gRNA, вызывая инактивацию экспрессии вирусных генов и репликацию в латентно-инфицированных клетках млекопитающих. Определенные из этих дополнительных конструкций Cas9/gRNA и их комбинации, как было обнаружено, приводят к вырезанию части или всей интегрированной провирусной ДНК ВИЧ из генома клетки-хозяина. Пары конструкций с одним участником, направленным по отношению к целевому сайту LTR, и другим участником, направленным по отношению к целевому сайту структурного гена, были особенно эффективными в обеспечении вырезания или удаления генома ВИЧ. Это является первой демонстрацией того, что комбинированное воздействие на сайт LTR и структурный ген может приводить к вырезанию промежуточного фрагмента интегрированной ДНК ВИЧ. Таким образом, в настоящем изобретении значительно расширяется спектр композиций на основе Cas9/gRNA, которые доступны для нацеливания на интегрированную ДНК ВИЧ в клетках-хозяевах.

Соответственно, в настоящем изобретении описываются композиции для применения в инактивации провирусной ДНК, интегрированной в клетку-хозяина, содержащие последовательность выделенной нуклеиновой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, и одну или более последовательностей выделенной нуклеиновой кислоты, кодирующих одну или более gRNA, комплементарных целевой последовательности в ВИЧ или другом ретровирусе.

gRNA включает в себя зрелую crRNA, которая содержит приблизительно 20 пар оснований (п.о.) уникальной нацеливающей последовательности (называемой спейсером), и транскриптуемой малой РНК (tracrRNA), которая служит в качестве направляющей для процессинга pre-crRNA с помощью рибонуклеазы III. Дуплекс crRNA:tracrRNA направляет Cas9 для целенаправленного воздействия на ДНК посредством комплементарного спаривания оснований между спейсером в crRNA и комплементарной последовательностью (называемой протоспейсером) в ДНК-мишени. Cas9 распознает тринуклеотидный (NGG) мотив, прилегающий к протоспейсеру (PAM), для установления сайта разрезания (3-й нуклеотид от PAM). В настоящем изобретении crRNA и tracrRNA могут экспрессироваться по отдельности, или из них может быть сконструирована искусственная слитая gRNA посредством синтетической структуры типа "стебель-петля" (AGAAAU) для имитации естественного дуплекса crRNA/tracrRNA. Такая gRNA может быть синтезирована или *in vitro* транскрибирована для прямой трансфекции РНК, или экспрессирована из вектора экспрессии U6 или вектора экспрессии на основе H1-активируемой РНК.

В композициях по настоящему изобретению каждая gRNA включает в себя последовательность, которая является комплементарной по отношению к целевой последовательности ретровируса. Иллюстративным целевым вирусом является ВИЧ, однако композиции по настоящему изобретению также применимы для нацеливания на другие ретровирусы, такие как ВИЧ-2 и вирус иммунодефицита обезьян (SIV)-1.

Некоторые из иллюстративных gRNA по настоящему изобретению являются комплементарными целевым последовательностям в участках длинного концевой повтора (LTR) ВИЧ. LTR подразделяются на U3-, R- и U5-участки. Конфигурация U1-, R-, U5-участков ВИЧ-1 показана на фиг. 3А. LTR содержат все требуемые сигналы для экспрессии генов, и они участвуют в интеграции провируса в геном клетки-хозяина. Например, основной или коровый промотор, коровый энхансер и модулирующий участок обнаружены в U3, в то время как чувствительный к трансактивации элемент обнаружен в R. В ВИЧ-1 участок U5 включает в себя несколько субучастков, например TAR или действующий в транс положении чувствительный элемент, который вовлечен в активацию транскрипции; поли А, который вовлечен в димеризацию и упаковку генома; PBS или сайт связывания праймера; Psi или сигнал упаковки; DIS или сайт инициации образования димеров. Соответственно, в некоторых вариантах осуществления целевая последовательность gRNA содержит одну или более целевых последовательностей в LTR-участке провирусной ДНК ВИЧ и одну или более мишеней в структурном гене и/или неструктурном гене провирусной ДНК ВИЧ. В других вариантах осуществления целевая последовательность gRNA содержит одну или более целевых последовательностей в LTR-участке провирусной ДНК ВИЧ и одну или более мишеней в структурном гене. В другом варианте осуществления целевая последовательность gRNA содержит одну или более целевых последовательностей в LTR-участке провирусной ДНК ВИЧ и одну или более мишеней в неструктурном гене провирусной ДНК ВИЧ. В еще одном варианте осуществления целевая последовательность gRNA содержит одну или более целевых последовательностей в структурном гене прови-

руса ВИЧ и одну или более мишеней в неструктурном гене провирусной ДНК ВИЧ. В еще одном варианте осуществления целевая последовательность gRNA содержит одну или более целевых последовательностей в некодирующем гене провируса ВИЧ и одну или более мишеней в кодирующем гене провирусной ДНК ВИЧ. В еще одном варианте осуществления целевая последовательность нуклеиновой кислоты gRNA содержит одну или более целевых последовательностей нуклеиновой кислоты в первом гене и одну или более целевых последовательностей нуклеиновой кислоты во втором гене; или одну или более целевых последовательностей нуклеиновой кислоты в первом гене и одну или более целевых последовательностей нуклеиновой кислоты в третьем гене; или одну или более целевых последовательностей нуклеиновой кислоты во втором гене и одну или более целевых последовательностей нуклеиновой кислоты в третьем гене; или одну или более целевых последовательностей во втором гене и одну или более целевых последовательностей во втором гене и одну или более целевых последовательностей в третьем гене или четвертом гене; или любую их комбинацию. Можно отметить, что использование любой комбинации целевых последовательностей нуклеиновой кислоты является возможным и они ограничены лишь изобретательностью специалиста в данной области.

В результатах экспериментов, раскрытых в разделе "Примеры", было обнаружено, что определенные последовательности в U3-, R- и 115-участках LTR являются применимыми целевыми последовательностями. gRNA, комплементарные этим целевым последовательностям, указаны На фиг. 1D и 3A примера 2, На фиг. 11A примера 3 и в табл. 1 примера 4. Они включают в себя LTR1, LTR 2, LTR 3, LTR A, LTR B, LTR B', LTR C, LTR D, LTR E, LTR F, LTR G, LTR H, LTR I, LTR J, LTR K, LTR L, LTR M, LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S и LTR T. Последовательности этих gRNA показаны На фиг. 11A, 12A, 12B и 12C. Композиции по настоящему изобретению содержат эти иллюстративные gRNA, однако не ограничены ими, и могут содержать gRNA, комплементарные любому подходящему целевому сайту в LTR ВИЧ.

Некоторые из иллюстративных gRNA по настоящему изобретению нацеливаются на последовательности в кодирующем белки геноме ВИЧ. Было обнаружено, что последовательности в гене, кодирующем структурный белок gag, являются применимыми целевыми последовательностями. gRNA, комплементарные этим целевым последовательностям, включают в себя Gag A, Gag B, Gag C и Gag D. Эти целевые сайты в геноме ВИЧ-1 указаны На фиг. 3A, и их последовательности нуклеиновой кислоты показаны в табл. 1. Применимые целевые последовательности также были обнаружены в гене, кодирующем структурный белок pol. gRNA, комплементарные этим целевым последовательностям, включают в себя Pol A и Pol B. Эти целевые сайты в геноме ВИЧ-1 указаны На фиг. 3A, и их последовательности нуклеиновой кислоты показаны в табл. 1. Последовательности для gRNA, комплементарные целевым сайтам в гене белка оболочки ВИЧ-1 env, также показаны в табл. 1.

Соответственно, композиции по настоящему изобретению содержат такие иллюстративные gRNA, однако не ограничены ими, и могут содержать gRNA, комплементарные любому подходящему целевому сайту в кодирующих белок генах ВИЧ, в том числе без ограничений такие, которые кодируют структурный белок tat, и акцессорные белки vif, nef (отрицательный фактор) vpr (вирусный белок U), vpr и tev.

Последовательности направляющей РНК в соответствии с настоящим изобретением могут представлять собой смысловые или антисмысловые последовательности. Последовательность направляющей РНК обычно включает мотив, смежный с протоспейсером (PAM). Последовательность PAM может варьироваться в зависимости от требований специфичности применяемой CRISPR-эндонуклеазы. В системе CRISPR-Cas, полученной из *S. pyogenes*, целевая ДНК, как правило, непосредственно предшествует мотиву, смежному с протоспейсером (PAM), 5'-NGG. Таким образом, для Cas9 *S. pyogenes* последовательностью PAM может быть AGG, TGG, CGG или GGG. Другие ортологи Cas9 могут характеризоваться другой специфичностью в отношении PAM. Например, для Cas9 из *S. thermophilus* требуется 5'-NNAGAA для CRISPR 1 и 5'-NGGNG для CRISPR 3, а для *Neisseria meningitidis* требуется 5'-NNNNGATT. Конкретная последовательность направляющей РНК может варьироваться, но независимо от последовательности применимыми последовательностями направляющей РНК будут те, которые сводят к минимуму нецелевые эффекты, достигая при этом высокой эффективности и полного разрушения интегрированного в геном ретровируса, например, ВИЧ. Длина последовательности направляющей РНК может варьироваться от приблизительно 20 до приблизительно 60 или более нуклеотидов, например, составлять приблизительно 20, приблизительно 21, приблизительно 22, приблизительно 23, приблизительно 24, приблизительно 25, приблизительно 26, приблизительно 27, приблизительно 28, приблизительно 29, приблизительно 30, приблизительно 31, приблизительно 32, приблизительно 33, приблизительно 34, приблизительно 35, приблизительно 36, приблизительно 37, приблизительно 38, приблизительно 39, приблизительно 40, приблизительно 45, приблизительно 50, приблизительно 55, приблизительно 60 или больше нуклеотидов. Применимые способы отбора для идентификации участков с чрезвычайно низкой гомологией между чужеродным вирусным геномом и геномом клетки-хозяина, в том числе эндогенной ретровирусной ДНК, включают биоинформационный скрининг с использованием критерия отбора мишени 12 п. о. +NGG для исключения нецелевых транскриптов человека или (хотя реже) нетранслируемых геномных участков; избегание сайтов связывания факторов транскрипции в пределах промотора

LTR ВИЧ (потенциально консервативные в геноме хозяина); и WGS, секвенирование по Сэнгеру и анализ SURVEYOR, в целях выявления и исключения потенциальных нецелевых эффектов.

Последовательность направляющей РНК может быть сконфигурирована в виде отдельной последовательности или в виде комбинации одной или более различных последовательностей, например мультиплексной конфигурации. Мультиплексные конфигурации могут включать комбинации из двух, трех, четырех, пяти, шести, семи, восьми, девяти, десяти или больше различных направляющих РНК.

В экспериментах, раскрытых в примерах 2 и 3, было обнаружено, что комбинации gRNA являются особенно эффективными при экспрессии в мультиплексном режиме, т. е. одновременно в той же самой клетке. Во многих случаях комбинации приводили к вырезанию провируса ВИЧ, расположенного между целевыми сайтами. Вырезания были вызваны делециями последовательностей между расщеплениями, индуцированными Cas9 в каждом из множественных целевых сайтов. Эти комбинации представляют собой пары gRNA, при этом один участник является комплементарным целевому сайту в LTR ретровируса, а второй участник gRNA является комплементарным целевому сайту в структурном гене ретровируса. Иллюстративные эффективные комбинации включают в себя Gag D в комбинации с одним из LTR I, LTR 2, LTR 3, LTR A, LTR B, LTR C, LTR D, LTR E, LTR F, LTR G, LTR H, LTR I, LTR J, LTR K, LTR L, LTR M, LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S или LTR T. Иллюстративные эффективные комбинации также включают в себя LTR 3 в комбинации с одним из LTR-1, Gag A; Gag B; Gag C, Gag D, Pol A или Pol B.

Комбинации LTR A и LTR B' также вызывали вырезание сегментов генома ВИЧ-1, что показано в примере 3. Композиции по настоящему изобретению не ограничены этими комбинациями, а содержат любую подходящую комбинацию gRNA, комплементарную двум или более различным целевым сайтам в провирусе ВИЧ-1.

В определенных вариантах осуществления целевая последовательность нуклеиновой кислоты предусматривает одну или более последовательностей нуклеиновой кислоты в кодирующих и некодирующих последовательностях нуклеиновой кислоты генома ретровируса. Целевая последовательность нуклеиновой кислоты может быть расположена в последовательности, кодирующей структурные белки, неструктурные белки или их комбинации. Последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации. Последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, например, Tat, Rev, акцессорные белки, например, Nef, Vpr, Vpu, Vif или их комбинации.

В определенных вариантах осуществления последовательность gRNA характеризуется по меньшей мере 75% идентичностью комплементарным целевым последовательностям нуклеиновой кислоты, кодирующим Gag, предшественник Gag-Pol, Pro, обратную транскриптазу (RT), интегразу (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif или их комбинации.

В определенных вариантах осуществления последовательность gRNA является комплементарной целевым последовательностям нуклеиновой кислоты, кодирующим Gag, предшественник Gag-Pol, Pro, обратную транскриптазу (RT), интегразу (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif или их комбинации.

В других вариантах осуществления последовательности нуклеиновой кислоты gRNA характеризуются по меньшей мере 75% идентичностью последовательностям, содержащим SEQ ID NO: 1-57 или любые их комбинации. В других вариантах осуществления последовательность нуклеиновой кислоты gRNA содержит SEQ ID NO: 1-57.

В другом варианте осуществления последовательность нуклеиновой кислоты содержит последовательность, характеризующуюся по меньшей мере 75% идентичностью последовательности, содержащей SEQ ID NO: 1-57 или любые их комбинации. В других вариантах осуществления последовательность нуклеиновой кислоты содержит последовательность, изложенную в SEQ ID NO: 1-57.

В других вариантах осуществления композиция для применения в инактивации ретровирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом, содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в интегрированной ретровирусной ДНК, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ). По меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в интегрированной ретровирусной ДНК; и вторую gRNA, которая является комплементарной другой целевой последовательности в интегрированной ретровирусной ДНК, причем промежуточные последовательности между двумя gRNA удалены.

В определенных вариантах осуществления целевая последовательность нуклеиновой кислоты предусматривает одну или более последовательностей в участке длинного концевой повтора (LTR) провирусной ДНК вируса иммунодефицита человека (ВИЧ) и одну или более мишеней в структурном и/или неструктурном гене интегрированной ДНК ВИЧ; или одну или более мишеней во втором гене; или одну или более мишеней в первом гене и одну или более мишеней во втором гене; или одну или более мише-

ней в первом гене и одну или более мишеней во втором гене и одну или более мишеней в третьем гене; или одну или более мишеней во втором гене и одну или более мишеней в третьем гене или четвертом гене; или любую их комбинацию.

В другом варианте осуществления композиция для удаления ретровируса *in vitro* или *in vivo* содержит последовательность выделенной нуклеиновой кислоты эндонуклеазу, кодирующую ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в геноме ретровируса, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ). В вариантах осуществления по меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в геноме ВИЧ; и вторую gRNA, которая является комплементарной другой целевой последовательности в геноме ВИЧ, причем промежуточные последовательности между двумя gRNA удалены.

В другом варианте осуществления композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере две направляющие РНК (gRNA), при этом gRNA являются комплементарными различным целевым последовательностям в геноме ретровируса, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ). В вариантах осуществления по меньшей мере одна направляющая РНК (gRNA) включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в геноме ВИЧ; и вторую gRNA, которая является комплементарной другой целевой последовательности в геноме ВИЧ, при этом промежуточные последовательности между двумя gRNA удалены.

В определенных вариантах осуществления целевая последовательность нуклеиновой кислоты предусматривает одну или более последовательностей нуклеиновой кислоты в кодирующих и не кодирующих последовательностях нуклеиновой кислоты генома ретровируса. Целевая последовательность нуклеиновой кислоты может быть расположена в последовательности, кодирующей структурные белки, неструктурные белки или их комбинации. Последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации. Последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, например, Tat, Rev, аксессуарные белки, например, Nef, Vpr, Vpu, Vif или их комбинации.

В определенных вариантах осуществления последовательность gRNA характеризуется по меньшей мере 75% идентичностью комплементарным целевым последовательностям нуклеиновой кислоты, кодирующим Gag, предшественник Gag-Pol, Pro, обратную транскриптазу (RT), интегразу (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif или их комбинации.

В определенных вариантах осуществления последовательность gRNA является комплементарной целевым последовательностям нуклеиновой кислоты, кодирующим Gag, предшественник Gag-Pol, Pro, обратную транскриптазу (RT), интегразу (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif или их комбинации.

В других вариантах осуществления последовательности нуклеиновой кислоты gRNA характеризуются по меньшей мере 75% идентичностью последовательностям, содержащим SEQ ID NO: 1-57 или любые их комбинации. В других вариантах осуществления последовательность нуклеиновой кислоты gRNA содержит SEQ ID NO: 1-57.

Соответственно, в настоящем изобретении также предусмотрен способ инактивации провирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом, при этом способ предусматривает стадии обработки клетки-хозяина с помощью композиции, содержащей CRISPR-ассоциированную эндонуклеазу, и по меньшей мере одну gRNA, комплементарную целевому сайту в провирусной ДНК; экспрессии комплекса редактирования гена, включающего в себя CRISPR-ассоциированную эндонуклеазу и по меньшей мере одну gRNA; и инактивации провирусной ДНК. Предпочтительными являются предварительно перечисленные gRNA и эндонуклеазы Cas9. В другом предпочтительном варианте осуществления стадия обработки клетки-хозяина *in vitro* или *in vivo* предусматривает обработку по меньшей мере двумя gRNA, где каждая по меньшей мере из двух gRNA является комплементарной определенной целевой последовательности нуклеиновой кислоты в провирусной ДНК. Особенно предпочтительными являются комбинации по меньшей мере из двух gRNA, включающие композиции, где по меньшей мере одна gRNA является комплементарной целевому сайту в LTR ретровируса, и по меньшей мере одна gRNA является комплементарной целевому сайту в структурном гене ретровируса. HIV является предпочтительным ретровирусом.

В другом варианте осуществления композиция для удаления ретровируса *in vitro* или *in vivo* содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в геноме ретровируса, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ) и gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в геноме ВИЧ; и вторую gRNA, которая является комплементарной

другой целевой последовательности в геноме ВИЧ, причем промежуточные последовательности между двумя gRNA удалены. Целевые последовательности нуклеиновой кислоты предусматривают одну или более последовательностей нуклеиновой кислоты в кодирующих и некодирующих последовательностях нуклеиновой кислоты генома ВИЧ. В одном варианте осуществления целевые последовательности предусматривают одну или более последовательностей нуклеиновых кислот в геноме ВИЧ, которые содержат последовательности нуклеиновой кислоты длинного концевой повтора (LTR), последовательности нуклеиновой кислоты, кодирующие структурные белки, неструктурные белки или их комбинации. В определенных вариантах осуществления последовательности нуклеиновой кислоты, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации. В вариантах осуществления последовательности нуклеиновой кислоты, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, акцессорные белки или их комбинации. Примеры регуляторных белков включают Tat, Rev или их комбинации. Примеры акцессорных белков предусматривают Nef, Vpr, Vpu, Vif или их комбинации. В определенных вариантах осуществления последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, характеризующейся по меньшей мере 75% идентичностью с SEQ ID NO: 1-57. В определенных вариантах осуществления последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

В определенных вариантах осуществления последовательность выделенной нуклеиновой кислоты содержит последовательность нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности нуклеиновой кислоты в геноме ретровируса, например ВИЧ.

В тех случаях, когда композиции вводят в виде нуклеиновой кислоты или они содержатся в векторе экспрессии, CRISPR-эндонуклеаза может кодироваться той же нуклеиновой кислотой или тем же вектором, что и последовательности направляющей РНК. Альтернативно или дополнительно CRISPR-эндонуклеаза может кодироваться нуклеиновой кислотой, физически отделенной от последовательностей gRNA, или отдельным вектором.

Модифицированные или мутантные последовательности нуклеиновой кислоты

В некоторых вариантах осуществления любая из последовательностей нуклеиновой кислоты, включенных в данный документ, может быть модифицирована или получена из нативной последовательности нуклеиновой кислоты, например с помощью введения мутаций, делеций, замен, модификации нуклеиновых оснований, остовов и т. д. Последовательности нуклеиновой кислоты включают векторы, средства для редактирования гена, gRNA, tracrRNA и т. д. Примеры некоторых модифицированных последовательностей нуклеиновых кислот, предусмотренные для настоящего изобретения, включают такие, которые содержат модифицированные остовы, например, фосфоротиоаты, сложные фосфотриэфиры, метилфосфонаты, короткоцепочечные алкильные или циклоалкильные межсахарные связи. В некоторых вариантах осуществления модифицированные олигонуклеотиды содержат такие с фосфоротиоатными остовами и такие с гетероатомными остовами, $\text{CH}_2\text{--NH--O--CH}_2$, $\text{CH}_2\text{--N(CH}_3\text{)--O--CH}_2$ [известные как остов метилен(метилямино) или MMI], остовы $\text{CH}_2\text{--O--N(CH}_3\text{)--CH}_2$, $\text{CH}_2\text{--N(CH}_3\text{)--N(CH}_3\text{)--CH}_2$ и $\text{O--N(CH}_3\text{)--CH}_2\text{--CH}_2$, где нативный фосфодиэфирный остов представлен в виде O--P--O--CH_2 . Амидные остовы, раскрытые в De Mesmaeker et al. Acc. Chem. Res. 1995, 28:366-374), также воплощены в данном документе. В некоторых вариантах осуществления последовательности нуклеиновой кислоты имеют структуры морфолинового остова (Summerton и Weller, патент США № 5034506), остов пептидной-нуклеиновой кислоты (PNA), где остов из сложного фосфодиэфира замещен полиамидным остовом, при этом нуклеиновые основания связаны непосредственно или опосредованно с аза-азотными атомами полиамидного остова (Nielsen et al. Science 1991, 254, 1497). Последовательности нуклеиновой кислоты могут также содержать один или более замещенных сахарных фрагментов. Последовательности нуклеиновой кислоты могут также содержать сахаромиметики, такие как циклобутилы вместо пентофуранозидовой группы.

Последовательности нуклеиновой кислоты могут также включать, дополнительно или альтернативно, модификации или замены нуклеиновых оснований (часто обозначаемые в данной области просто как "основание"). Используемые в данном документе "немодифицированные" или "встречающиеся в природе" нуклеиновые основания включают в себя аденин (A), гуанин (G), тимин (T), цитозин (C) и урацил (U). Модифицированные нуклеиновые основания включают нуклеиновые основания, встречающиеся лишь в редких случаях или случайно в нуклеиновых кислотах, например, гипоксантин, 6-метиладенин, 5-Ме-пиримидины, в частности 5-метилцитозин (также обозначаемый как 5-метил-2'-дезоксцитозин и часто обозначаемый в данной области как 5-Ме-С), 5-гидроксиметилцитозин (НМС), гликозил НМС и гентобиозил НМС, а также синтетические нуклеиновые основания, например, 2-аминоаденин, 2-(метиламино)аденин, 2-(имидазолилалкил)аденин, 2-(аминоалкиламино)аденин или другие гетерозамещенные алкиладенины, 2-тиоурацил, 2-тиотимин, 5-бромурацил, 5-гидроксиметилурацил, 8-азагуанин, 7-деазагуанин, N_6 (6-аминогексил)аденин и 2,6-диаминопурин. Kornberg, A., DNA Replication, W. H. Free-

man & Co., San Francisco, 1980, pp75-77; Gebeyehu, G., et al. Nucl. Acids Res. 1987, 15:4513). Можно включать "универсальное" основание, известное из уровня техники, например инозин. Было показано, что 5-Me-C-замены повышают стабильность дуплекса нуклеиновой кислоты на 0,6-1,2°C. (Sanghvi, Y. S., in Crooke, S. T. and Lebleu, B., eds., Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278).

Другая модификация последовательностей нуклеиновой кислоты по настоящему изобретению включает химическое связывание с последовательностями нуклеиновой кислоты, одним или более фрагментами или конъюгатами, что усиливает активность клеточного захвата олигонуклеотида. Такие фрагменты включают без ограничений липидные фрагменты, такие как холестеринный фрагмент, холестеринный фрагмент (Letsinger et al., Proc. Natl. Acad. Sci USA 1989, 86, 6553), холевую кислоту (Manoharan et al. Bioorg. Med. Chem. Lett. 1994, 4, 1053), простой тиозфир, например гексил-S-триотиол (Manoharan et al. Ann. N.Y. Acad. Sci. 1992, 660, 306; Manoharan et al. Bioorg. Med. Chem. Lett. 1993, 3, 2765), тиохолестерин (Oberhauser et al., Nucl. Acids Res. 1992, 20, 533), алифатическую цепь, например додекандиоловые или ундециловые остатки (Saison-Behmoaras et al. EMBO J. 1991, 10, 111; Kabanov et al. FEBS Lett. 1990, 259, 327; Svsmarchuk et al. Biochimie 1993, 75, 49), фосфолипид, например ди-гексадецил-глицерин или триэтиламмония 1,2-ди-О-гексадецил-глицеро-3-Н-фосфонат (Manoharan et al. Tetrahedron Lett. 1995, 36, 3651; Shea et al. Nucl. Acids Res. 1990, 18, 3777), полиаминовую или полиэтилениколевую цепь (Manoharan et al. Nucleosides & Nucleotides 1995, 14, 969), или адамантан-уксусную кислоту (Manoharan et al. Tetrahedron Lett. 1995, 36, 3651).

Отсутствует необходимость в том, чтобы все положения в данной последовательности нуклеиновой кислоты были единообразно модифицированы, и по сути, более одной из вышеупомянутых модификаций может быть включено в одну последовательность нуклеиновой кислоты или даже в один нуклеозид в последовательности нуклеиновой кислоты.

В некоторых вариантах осуществления молекулы РНК, например, crRNA, tracrRNA, gRNA, сконструированы так, чтобы они содержали одно или более модифицированных нуклеотидных оснований. Например, известные модификации молекул РНК можно найти, например, в Genes VI, Chapter 9 ("Interpreting the Genetic Code"), Lewis, ed. (1997, Oxford University Press, New York) и Modification and Editing of RNA, Grosjean and Benne, eds. (1998, ASM Press, Washington DC). Модифицированные компоненты РНК включают следующие: 2'-О-метилцитидин; N⁴-метилцитидин; N⁴-2'-О-диметилцитидин; N⁴-ацетилцитидин; 5-метилцитидин; 5,2'-О-диметилцитидин; 5-гидроксиметилцитидин; 5-формилцитидин; 2'-О-метил-5-формилцитидин; 3-метилцитидин; 2-тиоцитидин; лизидин; 2'-О-метилуридин; 2-тиоуридин; 2-тио-2'-О-метилуридин; 3,2'-О-диметилуридин; 3-(3-амино-3-карбоксипропил)уридин; 4-тиоуридин; рибозилтимин; 5,2'-О-диметилуридин; 5-метил-2-тиоуридин; 5-гидроксиуридин; 5-метоксиуридин; уридин-5-оксиуксусную кислоту; сложный метиловый эфир уридин-5-оксиуксусной кислоты; 5-карбоксиметилуридин; 5-метоксикарбонилметилуридин; 5-метоксикарбонилметил-2'-О-метилуридин; 5-метоксикарбонилметил-2'-тиоуридин; 5-карбаомилметилуридин; 5-карбаомилметил-2'-О-метилуридин; 5-(карбоксигидроксиметил)уридин; метиловый сложный эфир 5-(карбоксигидроксиметил)уридина; 5-аминометил-2-тиоуридин; 5-метиламинометилуридин; 5-метиламинометил-2-тиоуридин; 5-метиламинометил-2-селеноуридин; 5-карбоксиметиламинометилуридин; 5-карбоксиметиламинометил-2'-О-метилуридин; 5-карбоксиметиламинометил-2-тиоуридин; дигидроуридин; дигидрорибозилтимин; 2'-метил-аденозин; 2-метиладенозин; N⁶-метиладенозин; N,N-диметиладенозин; N⁶,2'-О-триметиладенозин; 2-метилтио-N-изопентениладенозин; N⁶-(цис-гидроксиизопентенил)аденозин; 2-метилтио-N⁶-(цис-гидроксиизопентенил)аденозин; N-глицинилкарбаомиладенозин; N-треонилкарбаомиладенозин; N⁶-метил-N⁶-треонилкарбаомиладенозин; 2-метилтио-N-метил-N-треонилкарбаомиладенозин; N-гидроксиинорвалилкарбаомиладенозин; 2-метилтио-N⁶-гидроксиинорвалилкарбаомиладенозин; 2'-О-рибозиладенозин (фосфат); инозин; 2'-О-метилюридин; 1-метилюридин; 1,2'-О-диметилюридин; 2'-О-метилгуанозин; 1-метилгуанозин; N2-метилгуанозин; N2,N2-диметилгуанозин; N2,2'-О-диметилгуанозин; N2,N2,2'-О-триметилгуанозин; 2'-О-рибозилгуанозин (фосфат); 7-метилгуанозин; N²,7-диметилгуанозин; N²,N²,7-триметилгуанозин; виозин; метилвиозин; немодифицированный гидроксивибутозин; вибутозин; гидроксивибутозин; пероксивибутозин; квеозин; эпоксиквеозин; галактозилквеозин; маннозилквеозин; 7-циано-7-деазагуанозин; археозин [также называемый 7-формамидо-7-деазагуанозином] и 7-аминометил-7-деазагуанозин.

Молекулы выделенных нуклеиновых кислот по настоящему изобретению могут быть получены с помощью стандартных методик. Например, методики полимеразной цепной реакции (ПЦР) можно применять для получения выделенной нуклеиновой кислоты, содержащей нуклеотидную последовательность, описанную в данном документе. Различные способы проведения ПЦР описаны, например, в PCR Primer: A Laboratory Manual, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Как правило, информацию о последовательности с концов области, представляющей интерес, или за ее пределами используют для конструирования олигонуклеотидных праймеров, идентичных или подобных по последовательности противоположным нитям матрицы, подлежащей амплификации. Также доступны различные стратегии ПЦР, с помощью которых в нуклеиновую кислоту-матрицу можно вводить сайт-специфические модификации нуклеотидной последовательности.

Выделенные нуклеиновые кислоты также могут быть химически синтезированы либо в виде одной молекулы нуклеиновой кислоты (например, с использованием автоматического синтеза ДНК в 3'-5' направлении с использованием фосфорамидитного метода), либо в виде серии олигонуклеотидов. Например, одну или более пар длинных олигонуклеотидов (например, > 50-100 нуклеотидов) можно синтезировать так, чтобы они содержали желаемую последовательность, при этом каждая пара содержит короткий сегмент комплементарности (например, приблизительно 15 нуклеотидов), чтобы образовывался дуплекс при отжиге пары олигонуклеотидов. ДНК-полимеразу применяют для удлинения олигонуклеотидов с получением в результате одной двунитевой молекулы нуклеиновой кислоты на пару олигонуклеотидов, которую затем можно лигировать в вектор.

В настоящее изобретение также включена фармацевтическая композиция для инактивации интегрированной провирусной ДНК ВИЧ в субъекте-млекопитающем. Композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу Cas, и по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую по меньшей мере одну gRNA, комплементарную целевой последовательности в провирусной ДНК ВИЧ; при этом последовательности выделенной нуклеиновой кислоты включены по меньшей мере в один вектор экспрессии. В предпочтительном варианте осуществления фармацевтическая композиция содержит первую gRNA и вторую gRNA, при этом первая gRNA нацеливается на сайт в LTR ВИЧ, и вторая gRNA нацеливается на сайт в структурном гене ВИЧ, как описано ранее.

Иллюстративные векторы экспрессии для включения в фармацевтическую композицию включают плазмидные векторы и лентивирусные векторы, однако настоящее изобретение не ограничено этими векторами. Широкий спектр комбинаций организм-хозяин/вектор экспрессии можно применять для экспрессии последовательностей нуклеиновых кислот, описанных в данном документе. Подходящие векторы экспрессии включают без ограничения плазмиды и вирусные векторы, полученные, например, из бактериофага, бакуловирусов и ретровирусов. Многочисленные векторы и системы экспрессии коммерчески доступны от таких корпораций, как Novagen (Мэдисон, Висконсин), Clontech (Пало-Альто, Калифорния), Stratagene (Ла-Хойя, Калифорния) и Invitrogen/Life Technologies (Карлсбад, Калифорния). Маркерный ген может придавать клетке-хозяину селективируемый фенотип. Например, маркер может придавать устойчивость к биоцидам, такую как устойчивость к антибиотику (например, канамицину, G418, блеомицину или гиромоцину). Вектор экспрессии может содержать последовательность метки, предназначенной для облегчения манипуляций с экспрессируемым полипептидом или его выявления (например, очистки или определения местоположения). Последовательности меток, такие как последовательности зеленого флуоресцентного белка (GFP), глутатион-S-трансферазы (GST), полигистидина, с-тус, гематоглинина или метки Flag™ (Kodak, Нью-Хейвен, Коннектикут), обычно экспрессируются слитыми с кодируемым полипептидом. Такие метки можно встраивать в полипептид в любом месте, в том числе на карбоксильном конце либо на аминоконце.

Вектор также может содержать регуляторный участок. Термин "регуляторный участок" относится к нуклеотидным последовательностям, которые влияют на инициацию и скорость транскрипции или трансляции, а также на стабильность и/или подвижность продукта транскрипции или трансляции. Регуляторные участки включают без ограничения промоторные последовательности, энхансерные последовательности, чувствительные элементы, сайты распознавания белка, индуцибельные элементы, последовательности связывания белка, 5'- и 3'-нетранслируемые участки (UTR), сайты инициации транскрипции, последовательности терминации, последовательности полиаденилирования, сигналы ядерной локализации и интроны.

При необходимости полинуклеотиды согласно настоящему изобретению также можно применять со средствами для доставки микроколичеств, таким как катионные липосомы и аденовирусные векторы. Обзор процедур для получения липосом, нацеливания и доставки содержимого см. у Mannino and Gould-Fogerite, *BioTechniques*, 6:682 (1988). См. также F Feigner and Holm, *Bethesda Res. Lab. Focus*, 11(2):21 (1989) and Maurer, R.A., *Bethesda Res. Lab. Focus*, 11(2):25 (1989).

Композиции по настоящему изобретению вызывают супрессию активации провируса ВИЧ-1, или частичное или полное вырезание интегрированного ВИЧ-1 (примеры 2 и 3), при этом в настоящем изобретении предусмотрен способ лечения субъекта-млекопитающего, инфицированного ретровирусом, например ВИЧ. Способ предусматривает стадии определения того, что субъект-млекопитающее инфицирован ретровирусом, введения эффективного количества описанной ранее фармацевтической композиции, и лечения субъекта-млекопитающего в отношении инфекции, вызываемой ретровирусом.

Способ представляет собой решение проблемы интегрированного провируса, решение, которое является необходимым для лечения и предупреждения СПИДа и других заболеваний, вызываемых ретровирусами. Во время острой фазы инфекции, вызываемой ВИЧ, вирусные частицы ВИЧ проникают в клетки, экспрессирующие подходящие рецепторные молекулы CD4. Как только вирус проникает в клетку-хозяина, обратная транскриптаза, кодируемая ВИЧ, образует копию провирусной ДНК с РНК ВИЧ, и провирусная ДНК интегрируется в геномную ДНК клетки-хозяина. Именно она представляет собой провирус ВИЧ, который реплицируется в клетке-хозяине, что в результате приводит к высвобождению новых вирионов ВИЧ, которые могут затем инфицировать другие клетки.

Первичная инфекция, вызываемая ВИЧ, затухает в течение от нескольких недель до нескольких месяцев, и за ней, как правило, следует длительный клинический "латентный" период, который может продолжаться до 10 лет. Во время этого латентного периода в мононуклеарных клетках периферической крови могут отсутствовать клинические симптомы или выявляемая репликация вируса, и в периферической крови может присутствовать в небольшом количестве или отсутствовать вирус, поддающийся культивированию. Однако вирус ВИЧ продолжает репродуцироваться при очень низких уровнях. У субъектов, которых подвергали лечению с помощью видов антиретровирусной терапии, этот латентный период может длиться в течение нескольких десятилетий или дольше. Антиретровирусная терапия не супрессирует низкие уровни экспрессии генома вируса, а также эффективно целенаправленно не воздействует на латентно инфицированные клетки, такие как покоящиеся Т-клетки памяти, макрофаги головного мозга, клетки микроглии, астроциты и клетки лимфоидной ткани кишечника. Поскольку композиции по настоящему изобретению могут инактивировать провирус ВИЧ или приводить к его вырезанию, способы лечения, в которых используются композиции, представляют собой новую перспективу для борьбы с инфекцией, вызываемой ВИЧ.

Композиции по настоящему изобретению при стабильной экспрессии в потенциальных клетках-хозяевах, ослабляют или предупреждают новую инфекцию, вызываемую ретровирусами, например ВИЧ-1 (пример 3). Соответственно, в настоящем изобретении также предусмотрен способ лечения, предназначенный для снижения риска инфекции, вызываемой ретровирусами, например инфекции, вызываемой ВИЧ, у субъекта-млекопитающего, подверженного риску возникновения такой инфекции. Способ предусматривает стадии определения того, что субъект-млекопитающее подвержен риску возникновения инфекции, вызываемой ВИЧ, введения эффективного количества описанной ранее фармацевтической композиции, и снижения риска возникновения инфекции, вызываемой ВИЧ, у субъекта-млекопитающего. Предпочтительно фармацевтическая композиция содержит вектор, который обеспечивает стабильную и/или индуцируемую экспрессию по меньшей мере одного из ранее перечисленных.

Фармацевтические композиции в соответствии с настоящим изобретением можно получать с помощью разнообразных способов, известными специалисту в данной области. Например, нуклеиновые кислоты и векторы, описанные выше, можно составлять в композиции для применения по отношению к клеткам в культуре тканей или для введения пациенту или субъекту. Такие композиции можно получать способом, хорошо известным в области фармацевтики, и можно вводить с помощью разнообразных путей введения, в зависимости от того, локальное или системное лечение является желательным, а также от области, подлежащей обработке. Введение может быть местным (в том числе офтальмическим и на слизистые оболочки, в том числе путем интраназальной, вагинальной и ректальной доставки), легочным (например, путем ингаляции или инсuffляции порошков или аэрозолей, в том числе с помощью распылителя; интратрахеальным, интраназальным, эпидермальным и трансдермальным), глазным, пероральным или парентеральным. Способы доставки в глаз могут включать в себя местное введение (глазные капли), субконъюнктивальную, периокулярную или интравитреальную инъекцию или внесение с помощью баллонного катетера или офтальмических вкладок, помещаемых хирургическим путем в конъюнктивальный мешок. Парентеральное введение включает в себя внутривенную, внутриартериальную, подкожную, внутривентральную или внутримышечную инъекцию или инфузию или внутривенное, например интратекальное или интравентрикулярное, введение. Парентеральное введение можно осуществлять в форме единичной болюсной дозы или можно осуществлять, например, с помощью насоса для непрерывной перфузии. Фармацевтические композиции и составы для местного введения могут включать в себя трансдермальные пластыри, мази, лосьоны, кремы, гели, капли, суппозитории, распыляемые растворы, жидкости, порошки и т.п. Могут быть необходимыми или желательными традиционные фармацевтические носители, водные, порошкообразные или масляные основы, загустители и т.п.

Настоящее изобретение также включает фармацевтические композиции, которые содержат в качестве активного ингредиента нуклеиновые кислоты и векторы, описанные в данном документе в комбинации с одним или более фармацевтически приемлемыми носителями. Используемые в данном документе термины "фармацевтически приемлемый" (или "фармакологически приемлемый") предусмотрены для обозначения молекулярных форм и композиций, которые не вызывают побочную, аллергическую или другую нежелательную реакцию при введении животному или человеку, в зависимости от конкретного случая. Используемый в данном документе термин "фармацевтически приемлемый носитель" включает все возможные растворители, дисперсионные среды, вещества, образующие покрытия, антибактериальные, изотонические и замедляющие всасывание средства, буферы, наполнители, связующие вещества, смазывающие вещества, гели, поверхностно-активные вещества и т.п., которые можно применять в качестве сред для фармацевтически приемлемого вещества. При получении композиций по настоящему изобретению активный ингредиент обычно смешивают с наполнителем, разбавляют наполнителем или заключают в такой носитель, например, в форме капсулы, таблетки, саше, бумажной упаковки или другого контейнера. В тех случаях, когда наполнитель служит в качестве разбавителя, он может представлять собой твердый, полутвердый или жидкий материал (например, изотонический солевой раствор), который служит в качестве основы, носителя или среды для активного ингредиента. Таким образом, композиции могут быть в форме таблеток, пилюль, порошков, пастилок, саше, крахмальных облаток,

эликсиров, суспензий, эмульсий, растворов, сиропов, аэрозолей (в виде твердого вещества или в жидкой среде), лосьонов, кремов, мазей, гелей, мягких и твердых желатиновых капсул, суппозиториев, стерильных растворов для инъекций и стерильных упакованных порошков. Как известно в данной области, тип разбавителя может варьироваться в зависимости от предполагаемого пути введения. Полученные в результате композиции могут содержать дополнительные средства, такие как консерванты. В некоторых вариантах осуществления носитель может представлять собой липидный или полимерный коллоид или может содержать его. В некоторых вариантах осуществления материал носителя может представлять собой коллоид, составленный в виде липосомы, гидрогеля, микрочастицы, наночастицы или мицеллы блок-сополимера. Как отмечается, материал носителя может образовывать капсулу, и такой материал может представлять собой полимерный коллоид.

Последовательности нуклеиновых кислот по настоящему изобретению можно доставлять в соответствующую клетку субъекта. Этого можно достичь, например, путем применения средства для доставки в виде полимерных биоразлагаемых микрочастиц или микрокапсул, доведенных до требуемого размера для оптимизации фагоцитоза фагоцитами, такими как макрофаги. Например, можно применять микрокапсулы PLGA (сополимера лактида и гликолида) диаметром, составляющим примерно 1-10 мкм. Полинуклеотид заключают в эти микрокапсулы, которые поглощаются макрофагами и постепенно подвергаются биологическому разложению в клетке, высвобождая таким образом полинуклеотид. После высвобождения ДНК экспрессируется в клетке. Второй тип микрокапсулы не предназначен для непосредственного поглощения клетками, а служит главным образом в качестве резервуара с медленным высвобождением для нуклеиновой кислоты, которая поглощается клетками только после высвобождения из микрокапсулы посредством биологического разложения. Следовательно, такие полимерные частицы должны быть достаточно большими, чтобы предотвращать фагоцитоз (т. е. больше 5 мкм и предпочтительно больше 20 мкм). Другим путем достижения поглощения нуклеиновой кислоты является применение липосом, полученных стандартными способами. Нуклеиновые кислоты могут быть включены в эти средства для доставки по отдельности или включены совместно с тканеспецифическими антителами, например антителами, которые целенаправленно воздействуют на типы клеток, которые обычно являются латентно инфицированными резервуарами ВИЧ-инфекции, например, макрофаги головного мозга, клетки микроглии, астроциты и клетки лимфоидной ткани кишечника. В качестве альтернативы можно получить молекулярный комплекс, состоящий из плазмиды или другого вектора, прикрепленных к поли-L-лизину посредством сил электростатического или ковалентного взаимодействия. Поли-L-лизин связывается с лигандом, который может связываться с рецептором на клетках-мишенях. Доставка "голой ДНК" (т. е. без средства для доставки) во внутримышечный, внутрикожный или подкожный участок, является другим средством достижения экспрессии *in vivo*. В релевантных полинуклеотидах (например, векторах экспрессии) последовательность нуклеиновой кислоты, кодирующую последовательность выделенной нуклеиновой кислоты, содержащую последовательность, кодирующую CRISPR-ассоциированную эндонуклеазу и направляющую РНК, функционально связана с промотором или комбинацией энхансер-промотор. Промоторы и энхансеры описаны выше.

В некоторых вариантах осуществления композиции по настоящему изобретению можно составлять в виде наночастицы, например наночастицы, состоящих из сердцевины из высокомолекулярного линейного полиэтиленimina (LPEI) в комплексе с ДНК, окруженной оболочкой из модифицированного полиэтиленгликолем (пегилированного) низкомолекулярного LPEI.

Нуклеиновые кислоты и векторы также могут быть нанесены на поверхность устройства (например, катетера) или содержаться в насосе, пластыре или другом устройстве для доставки лекарственного средства. Нуклеиновые кислоты и векторы согласно настоящему изобретению можно вводить по отдельности или в смеси в присутствии фармацевтически приемлемого формообразующего или носителя (например, физиологического солевого раствора). Наполнитель или носитель выбирают, исходя из способа и пути введения. Подходящие фармацевтические носители, а также необходимые фармацевтические вспомогательные вещества для применения в фармацевтических составах описаны в Remington's Pharmaceutical Sciences (E. W. Martin), справочном руководстве, хорошо известном в данной области техники, и в USP/NF (Фармакопее Соединенных Штатов Америки и Национальном формуляре).

В некоторых вариантах осуществления композиции можно составлять в виде наночастицы, инкапсулирующей нуклеиновую кислоту, кодирующую Cas9 или вариант Cas9, и по меньшей мере одной последовательности gRNA, комплементарной целевому ВИЧ; или она может включать в себя вектор, кодирующий эти компоненты. Альтернативно композиции можно составлять в виде наночастицы, инкапсулирующей CRISPR-ассоциированную эндонуклеазу, полипептиды, кодируемые одной или несколькими композициями на основе нуклеиновых кислот по настоящему изобретению.

Независимо от того, вводят ли композиции в виде нуклеиновых кислот или полипептидов, их составляют таким образом, чтобы содействовать поглощению клеткой млекопитающего. Применимые векторные системы и составы описаны выше. В некоторых вариантах осуществления вектор может доставлять композиции в конкретный тип клеток. Однако настоящее изобретение не ограничено этим, и также предусматриваются другие способы доставки ДНК, такие как химическая трансфекция с использованием, например, фосфата кальция, DEAE-декстрана, липосом, липоплексом, поверхностно-активных ве-

шеств и перфторированных жидкостей, равно как и способы физической доставки, такие как электропорация, микроинъекция, баллистическое введение частиц и применение систем "генной пушки".

В других вариантах осуществления композиции содержат клетку, которая была трансформирована или трансфицирована одним или более векторами Cas/gRNA. В некоторых вариантах осуществления способы согласно настоящему изобретению можно применять *ex vivo*. То есть клетки субъекта можно отбирать из организма и обрабатывать с помощью композиций в культуре для вырезания, например последовательностей вируса ВИЧ, и обработанные клетки возвращают в организм субъекта. Клетки могут представлять собой клетки субъекта, или они могут быть подобраны по гаплотипу или представлять собой линию клеток. Клетки могут быть подвергнуты ионизирующему облучению для предотвращения репликации. В некоторых вариантах осуществления клетки подобраны по человеческому лейкоцитарному антигену (HLA), являются аутологичными, представляют собой линии клеток или имеют комбинацию этих характеристик. В других вариантах осуществления клетка может представлять собой стволовую клетку. Например, она может представлять собой эмбриональную стволовую клетку или искусственную плюрипотентную стволовую клетку (индуцированную плюрипотентную стволовую клетку (iPS-клетку)). Эмбриональные стволовые клетки (ES-клетки) и искусственные плюрипотентные стволовые клетки (индуцированные плюрипотентные стволовые клетки, iPS-клетки) были сформированы из клеток многих видов животных, в том числе человека. Эти типы плюрипотентных стволовых клеток будут наиболее применимым источником клеток для регенеративной медицины, поскольку эти клетки способны к дифференцировке практически во все органы в результате соответствующей индукции их дифференцировки, сохраняя свою способность к активному делению при поддержании их плюрипотентности. iPS-клетки, в частности, могут быть сформированы из соматических клеток собственного организма, и, таким образом, не будут, вероятно, обуславливать этические и социальные проблемы в сравнении с ES-клетками, которые получают путем разрушения эмбрионов. Дополнительно iPS-клетки, которые являются клетками собственного организма, позволяют избежать реакций отторжения, которые являются наибольшим препятствием для регенеративной медицины или трансплантационной терапии.

Выделенные нуклеиновые кислоты можно с легкостью доставлять в субъекта способами, известными в данной области, например способами, с помощью которых доставляют siRNA. В некоторых аспектах Cas может представлять собой фрагмент, в который включены активные домены молекулы Cas, тем самым размер молекулы уменьшается. Следовательно, молекулы Cas9/gRNA можно применять в клинической практике, подобно подходам, которые используются в современной генной терапии. В частности, для применения у субъектов могут быть разработаны стволовые клетки или iPS-клетки, стабильно экспрессирующие Cas9/мультиплекс gRNA, для терапии с помощью трансплантации клеток, а также вакцинации.

Трансдуцированные клетки получают для повторной инфузии в соответствии с принятыми способами. После культивирования в течение периода приблизительно 2-4 недель, количество клеток может составлять от 1×10^5 до 1×10^7 . При этом ростовые характеристики клеток варьируют от пациента к пациенту и от типа клеток к типу клеток. Приблизительно за 72 ч до повторной инфузии трансдуцированных клеток отбирали аликвоты для анализа фенотипа и определения процентной доли клеток, экспрессирующих терапевтическое средство. Что касается введения, то клетки по настоящему изобретению можно вводить при норме, определенной по LD₅₀ для типа клеток и побочным эффектам для типа клеток при различных концентрациях, применительно к массе и общему состоянию здоровья пациента. Введение можно осуществлять посредством однократной или разделенными дозами. Стволовые клетки взрослого организма также можно мобилизовать с использованием экзогенно вводимых факторов, которые стимулируют их образование и выход из тканей или пространств, которые могут включать без ограничений костный мозг или виды жировой ткани.

Способы лечения

В определенных вариантах осуществления способ удаления ретровирусного генома из клетки или субъекта предусматривает приведение клетки в контакт с фармацевтической композицией или введение субъекту фармацевтической композиции, содержащей терапевтически эффективное количество последовательности выделенной нуклеиновой кислоты, кодирующей эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности нуклеиновой кислоты в геноме ретровируса.

В других вариантах осуществления способ ингибирования репликации ретровируса в клетке или субъекте предусматривает приведение клетки в контакт с фармацевтической композицией или введение субъекту фармацевтической композиции, содержащей терапевтически эффективное количество последовательности выделенной нуклеиновой кислоты, кодирующей эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности нуклеиновой кислоты в геноме ретровируса.

В способах лечения инфекции, вызываемой ретровирусом, например ВИЧ, субъекта можно выявлять с помощью стандартных клинических тестов, например видов иммунологического анализа для вы-

явления присутствия антител к ВИЧ или полипептида p24 ВИЧ в сыворотке крови субъекта, или посредством анализов амплификации нуклеиновых кислот ВИЧ. Количество такой композиции, назначенной субъекту, которое приводит к полному исчезновению симптомов инфекции, снижению степени тяжести симптомов инфекции или замедлению прогрессирования инфекции, считают терапевтически эффективным количеством. Способы по настоящему изобретению также могут включать стадию мониторинга для содействия оптимизации дозирования и составления схемы лечения, а также прогнозирования исхода. В некоторых способах по настоящему изобретению сперва можно определить, имеется ли у пациента латентная инфекция, вызываемая ВИЧ, а затем принять решение относительно того, следует или не следует лечить пациента одной или более композициями, описанными в данном документе. В некоторых вариантах осуществления способы могут дополнительно включать этап определения последовательности нуклеиновой кислоты конкретного ВИЧ, носителем которого является пациент, и затем конструирование направляющей РНК таким образом, чтобы она была комплементарной таким конкретным последовательностям. Например, можно определить последовательность нуклеиновой кислоты U3-, R- или U5-участки LTR или участки генов pol, gag или env субъекта, а затем сконструировать или выбрать одну или более gRNA таким образом, чтобы они были в точности комплементарны последовательностям пациента. Новые gRNA, предусмотренные настоящим изобретением, значительно усиливают вероятность составления эффективного лекарственного препарата.

В способах снижения риска возникновения инфекции, вызываемой ВИЧ, субъектом, подверженным риску возникновения инфекции, вызываемой ВИЧ, может быть, например, любой живущий половой жизнью индивидум, участвующий в незащищенных половых контактах, т. е. участвующий в половых актах без использования презерватива; ведущий половую жизнь индивидум, имеющий другую инфекцию, передаваемую половым путем; потребитель внутривенных наркотиков или не подвергнутый обрезанию мужчина. Субъектом, подверженным риску возникновения инфекции, вызываемой ВИЧ, может быть, например, индивидум, чей род деятельности может обуславливать вступление его или ее в контакт с популяциями, инфицированными ВИЧ, например работники сферы здравоохранения или работники служб экстренного реагирования. Субъектом, подверженным риску заражения инфекцией, вызываемой ВИЧ, может быть, например, лицо, содержащееся под стражей в исправительном учреждении, или лицо, занятое в секс-индустрии, то есть индивидум, который использует половые акты для получения заработка или неденежных благ, таких как пища, наркотики или кров.

В настоящем изобретении также предусмотрен набор, содержащий выделенную последовательность нуклеиновой кислоты, кодирующей CRISPR-ассоциированную эндонуклеазу, например эндонуклеазу Cas9, и по меньшей мере одну выделенную последовательность нуклеиновой кислоты, кодирующей gRNA, комплементарную целевой последовательности в провирусе ВИЧ. Альтернативно по меньшей мере одна из выделенных последовательностей нуклеиновой кислоты может быть закодирована в векторе, в таком как вектор экспрессии. Возможные варианты применения набора включают лечение или профилактику инфекции, вызываемой ВИЧ. Предпочтительно набор включает инструкции для применения, шприцы, устройства для доставки, буферы, стерильные контейнеры и разбавители, или другие реагенты для требуемого лечения или профилактики. Набор может также содержать подходящий стабилизирующий, молекулу-носитель, ароматизатор и т. п., подходящие для предполагаемого применения.

Несмотря на то, что различные варианты осуществления настоящего изобретения были описаны выше, необходимо понимать, что они были представлены лишь в качестве примера, и не предполагают ограничений. Можно провести многочисленные изменения в отношении раскрытых вариантов осуществления в соответствии с раскрытием данного документа, не выходя за пределы сущности и объема настоящего изобретения. Таким образом, охват и объем настоящего изобретения не должны быть ограничены любым из вышеописанных вариантов осуществления.

Все публикации и патентные документы, цитируемые в настоящей заявке, включены посредством ссылки для всех целей в той же степени, как если бы каждая отдельная публикация или патентный документ были упомянуты по отдельности. Цитируя различные ссылки в данном документе, авторы настоящей заявки не признают, что какая-либо определенная ссылка представляет собой "предшествующий уровень в данной области" в отношении их изобретения.

Примеры

Пример 1. Материалы и способы.

Клонирование sgRNA в лентивирусный вектор. При помощи средств биоинформационного конструирования spCas9-sgRNA для получения лучших баллов высокой эффективности и высокой специфичности разработали 20 целевых сайтов для sgRNA в участке LTR-U3 ВИЧ-1 и 4 sgRNA - для Gag, 2 sgRNA - для Pol и 2 sgRNA - для Env (табл.1). Все эти затравочные последовательности sgRNA клонировали в модифицированный лентивирусный вектор экспрессии sgRNA pKLV-Wg (фиг. 1B), модифицированный из лентивирусного вектора для sgRNA (Addgene #50946). Вкратце, вектор pKLV-Wg расщепляли с помощью BbsI и обрабатывали антарктической фосфатазой, а линейаризованный вектор очищали с помощью набора для быстрого удаления нуклеотидов (Qiagen). Равное количество смысловых и антисмысловых направляющих последовательностей (100 мкМ в 1 мкл) смешивали с полинуклеотидкиназой (PNK, 1 мкл), 1x PNK буфера и 1 мМ АТФ и инкубировали при 37°C в течение 30 мин с последующим отжигом в

устройстве для ПЦР (95°C в течение 5 мин, -1°C/цикл 15 с в течение 70 циклов). Фосфорилированный олигонуклеотидный дуплекс (10 мкМ) разбавляли при 1:100 с получением рабочего раствора (100 нМ). Затем 1 мкл олигонуклеотидного дуплекса (0,1 пмоль) смешивали с 3,5 мкл расщепляемого с помощью BbsI вектора pKLV-WG (0,015 пмоль), 5 мкл 2x T7 лигазного реакционного буфера и 0,5 мкл T7 ДНК-лигазы (NEB). Смесь инкубировали в течение 15-30 мин при комнатной температуре, охлаждали на льду и затем трансформировали в Stab13-компетентные клетки. Выбирали 2-4 колонии для ПЦР с применением T351 (U6/5') и обратного праймера для sgRNA. Разращивали 2 ПЦР-положительных клон в среде LB/Amp в течение ночи при 37°C. На следующий день плазмидные ДНК miniprep отправляли в Genewiz Inc. для секвенирования с применением T428 (hU6-последовательность/5'/F). Кассету экспрессии полной sgRNA, включающую U6-промотор, sgRNA и поли-Т-терминатор, проверяли с помощью анализа секвенирования с применением программы GeneRunner.

Анализ репортерного гена люциферазы с использованием EcoHIV. Клетки HEK293T (5×10^4 /лунка) культивировали в 96-луночном планшете в среде DMEM с высоким содержанием глюкозы, содержащей 10% FBS и антибиотики (100 Ед./мл пенициллина и 100 мкг/мл стрептомицина) в увлажненной атмосфере с 5% CO₂ при 37°C. На следующий день клетки котрансфицировали с помощью стандартного осаждения фосфатом кальция с использованием репортерного вектора EcoHIV-eLuc, вектора pLV-Cas9-RFP и указанных экспрессирующих sgRNA векторов pKLV-Wg. Через 2 дня после трансфекции получали клеточный лизат с помощью системы анализа люциферазы ONE-Glo (Promega) и измеряли люминесценцию в многоканальном ридере 2104 EnVision® (PerkinElmer). Данные представляют среднее значение \pm SE, полученное в результате 4 независимых трансфекции.

Относительные изменения в одиночных или спаренных sgRNA рассчитывали по отношению к пустому контрольному вектору для sgRNA.

Генотипирование на основе ПЦР, ТА-клонирование и секвенирование по Сэнгеру. Клетки HEK293T в 96-луночном планшете котрансфицировали векторами экспрессии репортерного гена EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и указанной gRNA. Через 2 дня клетки лизировали с помощью 90 мкл 50 mM NaOH при 95°C в течение 10 мин и нейтрализовали с помощью 10 мкл 1 M Tris-HCl. Неочищенные экстракты непосредственно использовали для ПЦР с применением смеси Terra PCR Direct Polymerase (Clontech) и указанных праймеров для ПЦР. Две стадии стандартной ПЦР проводили в течение 35 циклов с отжигом/удлинением при 68°C в течение 1 мин и при 98°C в течение 15 с. Продукты разделяли в 1,5% агарозном геле. Полосы, представляющие интерес, очищали от геля и клонировали в вектор pCRII T-A (Invitrogen) и нуклеотидную последовательность отдельных клонов определяли с помощью секвенирования в Genewiz с применением универсальных праймеров T7 и/или SP6.

Пример 2. Скрининг gRNA ВИЧ-1 и функциональная характеристика удаления ВИЧ-1 in vitro.

С целью расширения спектра gRNA, эффективных для CRISPR-опосредованного редактирования генома провируса ВИЧ, кандидатные gRNA находили и скринировали в отношении эффективности подавления экспрессии ВИЧ и в отношении способности индуцировать ликвидирование или удаление генома провируса ВИЧ-1 в клетках-хозяевах.

Кандидатные gRNA, специфичные в отношении целевых сайтов в геноме HIV-1, обнаруживали при помощи биоинформационных процедур. Кандидатные gRNA отбирали в отношении максимальной вероятности обеспечения эффективного редактирования гена, при минимальном потенциале нецелевого воздействия, т. е. потенциале вызывать повреждение сайтов в геноме хозяина. Затравочные последовательности целевого сайта для кандидатных gRNA в регуляторном U3-участке LTR показаны на фиг. 1D. gRNA включают в себя LTR 1, LTR 2, LTR 3, LTR A, LTR B, LTR C, LTR D, LTR E, LTR F, LTR G, LTR H, LTR I, LTR J, LTR K, LTR L, LTR M, LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S, LTR T. Последовательности этих gRNA показаны в табл.1. Большая часть этих кандидатных gRNA может быть спарена с нуклеотидной последовательностью РНК нуклеазой FokI, которые могут снижать потенциальные нецелевые эффекты в 1500 раз. Сайты для нацеливания кандидатных gRNA в структурных генах gag и pol, указаны стрелками на фиг. 1C и 3A. Эти gRNA включают в себя gag A, gag B, gag C, gag D; pol A и pol B. Эти последовательности показаны в табл.1. Кандидатные gRNA клонировали в лентивирусные репортерные векторы, как показано на фиг. 1B, нижняя панель. spCas9 также клонировали в лентивирусные репортерные векторы, как показано на фиг. 1B, верхняя панель.

Клетки-хозяева HEK293T котрансфицировали репортерной конструкцией на основе ВИЧ, EcoHIV-eLuc, репортерной конструкцией экспрессии для spCas9 и одной или двумя репортерными конструкциями для gRNA. Контрольные клетки получали контрольные конструкции ("LTR O"). Через 2 дня люциферазную активность в лизатах клеток измеряли с помощью системы анализа люциферазы ONE-Glo™.

Большая часть кандидатных gRNA LTR, вводимых в отдельности, были эффективными в подавлении экспрессии ВИЧ-1 в клетках-хозяевах, как определяли по снижению экспрессии гена люциферазы (фиг. 2A). Эти gRNA затем котрансфицировали в различных комбинациях с сайтами для нацеливания gRNA в генах gag или pol. Комбинации, показанные на оси абсцисс фиг. 2B и 2C, включали в себя Gag D в комбинации с одной из LTR1, LTR 2, LTR 3, LTR A, LTR B, LTR C, LTR D, LTR E, LTR F, LTR G; LTR H, LTR I, LTR J, LTR K, LTR L, LTR M; LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S или LTR T и LTR 3

в комбинации с одним из Gag A; Gag B; Gag C, Gag D, Pol A или Pol B.

Большая часть одиночных gRNA (фиг. 2A) ингибировала конститутивную люциферазную активность, однако некоторые усиливали репортерную активность или даже не оказывали никакого воздействия. Данные указывают на то, что редактирование одного сайта в LTR индуцирует мутации по типу вставки/делеции, которые могут нарушать связывание активаторов или репрессоров транскрипции с целью промоторной активности LTR ВИЧ-1.

Однако спаренные gRNA Gag-D с любой из LTR-sgRNA снижали люциферазную активность на 64-96% (фиг. 2B). Эти данные указывают на то, что все разработанные gRNA эффективно индуцировали редактирование гена в целевых сайтах на плазмидных уровнях. Комбинация Gag с любыми LTR-sgRNA будет удалять коровый промотор LTR и, таким образом, индуцировать значительное снижение промоторной активности. Оставшуюся активность репортерного гена люциферазы может отражать небольшая группа экспрессирующих репортерный ген клеток EcoHIV, которые не содержат spCas9 или двух sgRNA. Необходимо обратить внимание, что для снижения эффективности требуется присутствие всех 4 плазмид в одних и тех же клетках: EcoHIV-eLuc, для spCas9, для двух sgRNA.

gRNA LTR-3, спаренная с любой из разработанных gRNA Gag или Pol, также значительно снижала активность репортерного гена люциферазы на 73-93% (фиг. 2C). Эти данные указывают на то, что все разработанные gRNA, нацеливающиеся на участок Gag или Pol, эффективно индуцировали редактирование гена в целевых сайтах на плазмидных уровнях. Комбинация LTR-3 с любыми sgRNA в структурном участке будет удалять коровый промотор LTR и структурный геном между сайтами разрезания и, таким образом, индуцировать значительное снижение промоторной активности. Такое доказательство концепции применимо к любым LTR-sgRNA, а также к любым другим sgRNA, нацеливающимся на структурные участки, такие как Gag, Pol и Env, или любой другой неструктурный участок между обеими концевыми LTR. Остаточную активность репортерного гена люциферазы может отражать небольшая группа экспрессирующих репортерный ген клеток EcoHIV, которые не содержат векторов на основе spCas9 или двух sgRNA. Необходимо обратить внимание, что для снижения эффективности требуется присутствие всех 4 плазмид в одних и тех же клетках: EcoHIV-eLuc, для spCas9, для двух sgRNA.

Затем определяли, отражало ли подавление экспрессии ВИЧ-1 делецию сегментов генома провируса ВИЧ-1. Клетки HEK293T котрансфицировали векторами экспрессии репортерного гена EcoHIV-eLuc, pLV-EF1a-spCas9-T2A-RFP и gRNA. Через 2 дня клетки лизировали с помощью 50 mM NaOH при 95°C в течение 10 минут и нейтрализовали с помощью 1 M Tris-HCl. Неочищенные экстракты непосредственно использовали для ПЦР с применением смеси Terra PCR Direct Polymerase (Clontech) и указанных праймеров для ПЦР.

При использовании первого набора праймеров (фиг. 3A) фрагменты ПЦР после разрезания разработанных сайтов для нацеливания характеризовались предполагаемым размером в 17 из 20 sgRNA LTR. Лишь LTR-F, LTR-G и LTR-K не характеризовались удалением, что свидетельствовало о том, что они неспособны расщеплять 5'-LTR или являются менее эффективными. Это согласуется с неэффективным или менее эффективным результатом в случае трансфекции одиночной sgRNA LTR-F, LTR-G и LTR-K, как показано на фиг. 2A. Среди эффективных sgRNA эффективность разрезания, выявляемая отношением разрезанного фрагмента к соответствующему неразрезанному фрагменту, была максимальной у sgRNA LTR. Интересным является то, что дополнительную полосу большего размера, чем предполагаемый фрагмент, наблюдали в большинстве пар.

При использовании второго набора праймеров (фиг. 3B) во всех образцах наблюдали две слабые полосы, указывающие на неспецифические продукты ПЦР. Однако уникальные фрагменты ПЦР после разрезания разработанных целевых сайтов характеризовались предполагаемым размером в 16 из 20 sgRNA LTR. Среди эффективных sgRNA эффективность разрезания, выявляемая по интенсивности предполагаемого фрагмента ПЦР, была максимальной у sgRNA LTR-Q, L, B, S, O, C, I.

Затем исследовали дополнительные комбинации sgRNA, нацеливающиеся на LTR и различные структурные гены. Комбинации указаны на фиг. 3C. Среди sgRNA LTR-1, 2, 3 максимальной эффективностью характеризовалась LTR-1. Gag-A, C, D и Pol-A, B характеризовались высокой эффективностью удаления. Gag-B не характеризовался расщеплением (подтверждено по повтору ПЦР, правая панель). Как и в предыдущем случае, интересным является то, что дополнительную полосу большего размера, чем предполагаемый фрагмент, наблюдали в парах в случае расщепления 5'-LTR.

Также было обнаружено, что пары gRNA, комплементарные целевым сайтам в U3-участке LTR, также приводили к образованию делеций в геноме провируса ВИЧ-1. Приготовление образца и прямую ПЦР выполняли, как описано ранее. Фрагменты ПЦР после разрезания экстрагировали для ТА-клонирования и секвенирования по Сэнгеру (фиг. 4). Типичное ТА-клонирование и секвенирование по Сэнгеру подтвердило делецию размером 296 п. о. между LTR-1 и LTR-3 и дополнительную вставку размером 180 п. о. между двумя сайтами разрезания. Секвенирование 15 клонов показало аналогичный характер точного лигирования без каких-либо мутаций по типу вставки/делеции между двумя сайтами разрезания. Дополнительная последовательность вставки размером 180 п. о. соответствует последовательности вектора, исходя из NCBI Blast.

В совокупности результаты показывают, что большая часть кандидатных gRNA является эффек-

тивной для удаления предполагаемой последовательности генома ВИЧ-1 между выбранными двумя сайтами для нацеливания и для подавления экспрессии провируса, как показано по активности репортерного гена люциферазы. В частности, комбинация вирусных структурных gRNA с одной или двумя gRNA LTR обеспечивала более высокую эффективность удаления генома. Результаты этих экспериментов расширяют спектр gRNA, которые можно использовать в системе CRISPR для обеспечения эффективного расщепления генома ВИЧ-1.

Пример 3. Новая комбинация gRNA и векторная система для удаления генома ВИЧ-1 из латентно инфицированных Т-клеток с помощью системы редактирования генов CRISPR/Cas9.

Эксперименты выполняли с целью дальнейшего расширения спектра эффективных gRNA против ВИЧ-1 и с целью повышения гибкости доставки gRNA и Cas9 в клетки-хозяева.

Впервые исследовали стратегию на основе комбинированного лечения. При лечении использовали новую комбинацию gRNA LTR B' совместно с LTR A, которую ранее раскрыли в Hu, et al, 2014. Последовательности LTR A и LTR B' показаны на фиг. 11A, а их положения в LTR ВИЧ-1 указаны на фиг. 5A-5D.

Комбинированная экспрессия Cas9, gRNA LTR A и LTR B' подавляет активацию латентного провируса ВИЧ и вызывает вырезание провирусных последовательностей. Эксперименты проводили в линии репортерных Т-клеток Jurkat2D10, которая представлена на фиг. 6A. Эта линия клеток содержит интегрированный транскрипционно латентный провирус ВИЧ-1 с eGFP вместо Nef в качестве репортерного гена активации провируса. Активация индуцируется обработкой форбол-12-мирикат-13-ацетатом (PMA) или трихостатином А (TSA). Интегрированная репортерная последовательность ВИЧ-1 показана на фиг. 6A. Графики флуоресценции индуцированных (правая панель) и неиндуцированных (левая панель) клеток 2D10 показаны на фиг. 6B. Анализы проточной цитометрии индуцированных (правая панель) и неиндуцированных (левая панель) клеток 2D10 показаны на фиг. 6C.

Репортерные клетки 2D10 (2×10^6 /состояние) электропорировали с помощью 10 мкг контрольной плазмиды pX260 или плазмид pX260 LTR-A и pX260 LTR-B', по 5 мкг каждой (Neon System, Invitrogen, импульс 3 раза по 10 мс/1350 В). Через 48 ч среду заменяли средой, содержащей пуромицин при 0,5 мкг/мл. Через одну неделю после отбора пуромицин удаляли и клетки выращивали еще в течение одной недели. Клетки также экспрессировали меченый FLAG Cas9.

Затем клетки разбавляли до концентрации 10 клеток/мл и помещали в 96-луночные планшеты при 50 мкл/луночку. Через 2 недели одноклеточные клоны скринировали в отношении наличия меченой GFP реактивации репортерного гена ВИЧ-1 (12 ч обработки PMA 25 нМ/TSA, 250 нМ) с применением проточного цитометра Guava EasyCyte Mini.

Клоны, экспрессирующие Cas9, LTR A и LTR B', сравнивали с клонами, экспрессирующими только Cas9 (фиг. 7A - 7C). Было обнаружено, что только клоны, экспрессирующие Cas9, LTR A и LTR B', не приводили к реактивации репортерного гена ВИЧ-1 после индукции PMA (фиг. 7B, нижняя панель). В отличие от этого клоны, экспрессирующие только Cas9, не характеризовались реактивацией репортерного гена ВИЧ-1, о чем свидетельствовала экспрессия EGFP (фиг. 7B, верхняя панель).

Затем определяли, было ли связано подавление реактивации латентного репортерного провируса ВИЧ-1 с успешным вырезанием провирусных последовательностей из генома хозяина. ДНК, полученную из клонов, проанализированную в предыдущем эксперименте, подвергали ПЦР с целью амплификации мотива RRE последовательности гена env провируса или геномных последовательностей, фланкирующих интегрированный репортерный провирус (ген MSRB1). Положение праймеров показано на фиг. 8A.

Анализ ПЦР показал, что клоны, экспрессирующие Cas9 и комбинацию LTR A/B', не приводили к образованию продуктов ПЦР, включающих RRE и MSRB1, указывая на вырезание ДНК, включающих эти последовательности. В отличие от этого RRE и MSRB1 амплифицировали и с легкостью выявляли в клонах, экспрессирующих только Cas9 (фиг. 8B). Широкий диапазон генотипирования ПЦР подтвердил, что экспрессия комбинации LTR A/B' приводила к вырезанию последовательности размером 652 п. о., расположенной между 5'-U3-участком и 3'-U3-участком (фиг. 8C). Вырезание дополнительно подтверждали секвенированием расщепляемого лариата из локуса интеграции в хромосоме 16 (фиг. 8D). В совокупности результаты подтверждают, что подавление реактивации латентного провируса ВИЧ-1 путем экспрессии LTR A/B' и Cas9 было вызвано вырезанием провирусных последовательностей из генома хозяина.

Также было обнаружено, что стабильная экспрессия LTR A/B' и Cas9 защищала клоны 2D10 от новых инфекций, вызываемых ВИЧ-1. Клоны характеризовали на наличие экспрессии Cas9 с помощью вестерн-блоттинга (фиг. 9B) и на наличие экспрессии LTR B' - с помощью анализа обратной ПЦР (фиг. 9C). Клоны инфицировали репортерным вирусом ВИЧ-1 NL4-3-EGFP-P2A-NEF и отслеживали прогрессирование инфекции с помощью анализа FACS. Лишь клоны, которые экспрессировали LTR A/B' и Cas9, были устойчивыми к инфекции, вызываемой ВИЧ-1. Это показано по значительно более низким уровням экспрессии EGFP по сравнению с клонами, у которых отсутствовали LTR A/B' (ctrl 7) или Cas 9 (AB 8) (фиг. 9A).

Лентивирусная доставка Cas9/gRNA обеспечивает эффективное и регулируемое по времени нацеливание на провирусные последовательности. Затем определяли, можно ли использовать лентивирусные

вирусы для экспрессии компонентов Cas9/gRNA в клетках-хозяевах. Лентивирусные векторы обеспечивают универсальные и гибкие средства экспрессии, и доступен ряд индуцируемых лекарственными препаратами лентивирусных векторов. Клетки Jurkat 2D10 трансдуцировали лентивирусами, экспрессирующими RFP-Cas9 (красная флуоресценция) и/или gRNA LTR A/B' (маркер BFP, синяя флуоресценция) при MOI 5 (фиг. 10A). Через 72 ч репортерный вирус GFP реактивировали с помощью обработки PMA/TSA и количественно определяли с помощью проточной цитометрии. Анализ точечной диаграммы показан на фиг. 10B. Активация ВИЧ выявляется в виде сдвига вверх (зеленый) в точечной диаграмме при индукции (например, фиг. 10B, левая верхняя панель по сравнению с левой нижней панелью). Можно было заметить, что только клетки, трансдуцированные Cas9 и LTR A/B', характеризуются значительной клеточной фракцией, которая не проявляется при сдвиге в сторону зеленой флуоресценции (фиг. 10B, правая нижняя панель). Результаты подтверждают, что компоненты Cas9/gRNA можно эффективно доставлять лентивирусными векторами.

Экспрессия Cas9/LTR A/B' не вызывает выявляемых нецелевых эффектов и минимальных изменений в экспрессии прилегающих генов. Эффективное вырезание провируса ВИЧ-1 с помощью редактирования CRISPR является малополезным, если оно сопровождается индуцированными мутациями в нормальных генах хозяина, содержащих последовательности, аналогичные целевым последовательностям. Шесть предполагаемых/возможных нецелевых сайтов для LTR A/B' исследовали в клонах Jurkat, в которых геном ВИЧ-1 подвергся успешному удалению. Последовательности LTR A и LTR B' показаны на фиг. 11A. Не было показано мутаций по типу вставки/делеции ни в реакциях анализа Surveyor (фиг. 11B), ни в секвенировании по Сэнгеру (фиг. 11C). Положение сайта интеграции репортерного гена в ВИЧ-1 во втором экзоне гена MSRB1 в хромосоме 16 и соседние гены показаны на фиг. 11D. Уровни экспрессии генов, прилегающих к сайту интеграции после удаления последовательности ВИЧ-1, измеряли с помощью qRT-PCR и сравнивали с уровнями экспрессии в контрольных клетках. Результаты на фиг. 11E показывают, что эффективная обработка с помощью Cas9 и LTR A/B' не оказывает значительного влияния на экспрессию генов, расположенных рядом с вырезанными последовательностями ВИЧ.

Пример 4. Функциональный скрининг направляющих РНК, нацеливающих на регуляторный и структурный геном вируса ВИЧ-1 с целью излечения СПИДа.

В настоящем исследовании идентифицировали наиболее эффективные gRNA, нацеливающиеся на LTR ВИЧ-1 и структурный участок вируса, и оптимизировали спаривание gRNA, которое может приводить к эффективному удалению генома ВИЧ-1.

Высокоспецифичные gRNA разрабатывали с помощью биоинформационных средств и их способность направлять Cas9 с целью расщепления ДНК провируса ВИЧ-1 оценивали с помощью высокопродуктивного анализа репортерного гена люциферазы ВИЧ-1 и быстрого генотипирования на основе прямой ПЦР.

Таблица 1. Олигонуклеотиды для gRNA, нацеливающих на LTR, Gag и Pol ВИЧ-1 и праймеры для ПЦР

Название мишени	Направление	Последовательности (от 5' к 3')
LTR-A	T353: прямой	aaacAGGGCCAGGGATCAGATATCCACTGACCTTgt (SEQ ID NO: 1)
	T354: обратный	taaacAAGGTCAGTGGATATCTGATCCCTGGCCCT (SEQ ID NO: 2)
LTR-B	T355: прямой	aaacAGCTCGATGTCAGCAGTTCTGAAGTACTCgt (SEQ ID NO: 3)
	T356: обратный	taaacGAGTACTTCAAGAACTGCTGACATCGAGCT (SEQ ID NO: 4)
LTR-C	T357: прямой	caccGATTGGCAGAACTACACACC (SEQ ID NO: 5)
	T358: обратный	aaacGGTGTGTAGTTCTGCCAATC (SEQ ID NO: 6)
LTR-D	T359: прямой	caccGCGTGGCCTGGGCGGGACTG (SEQ ID NO: 7)
	T360: обратный	aaacCAGTCCCGCCAGGCCACGC (SEQ ID NO: 8)
LTR-E	T361: прямой	caccGATCTGTGGATCTACCACACACA (SEQ ID NO: 9)
	T362: обратный	aaacTGTGTGTGGTAGATCCACAGATC (SEQ ID NO: 10)
LTR-F	T363: прямой	caccGCTGCTTATATGCAGCATCTGAG (SEQ ID NO: 11)
	T364: обратный	aaacCTCAGATGCTGCATATAAGCAGC (SEQ ID NO: 12)
LTR-G	T530: прямой	caccGTGTGGTAGATCCACAGATCA (SEQ ID NO: 13)
	T531: обратный	aaacTGATCTGTGGATCTACCACAC (SEQ ID NO: 14)
LTR-H	T532: прямой	caccGCAGGAAGTAGCCTTGTGTG (SEQ ID NO: 15)
	T533: обратный	aaacCACACAAGGCTACTTCCTGC (SEQ ID NO: 16)

LTR-I	T534: прямой	caccGATCAGATATCCACTGACCTT (SEQ ID NO: 17)
	T535: обратный	aaacAAGGTCAGTGGATATCTGATC (SEQ ID NO: 18)
LTR-J	T536: прямой	caccGCACACTAATACTTCTCCCTC (SEQ ID NO: 19)
	T537: обратный	aaacGAGGGAGAAGTATTAGTGTGC (SEQ ID NO: 20)
LTR-K	T538: прямой	caccGCCTCCTAGCATTTTCGTACA (SEQ ID NO: 21)
	T539: обратный	aaacTGTGACGAAATGCTAGGAGGC (SEQ ID NO: 22)
LTR-L	T540: прямой	caccGCATGGCCCGAGAGCTGCATC (SEQ ID NO: 23)
	T541: обратный	aaacGATGCAGCTCTCGGGCCATGC (SEQ ID NO: 24)
LTR-M	T542: прямой	caccGCAGCAGTCTTTGTAGTACTC (SEQ ID NO: 25)
	T543: обратный	aaacGAGTACTACAAAGACTGCTGC (SEQ ID NO: 26)
LTR-N	T544: прямой	caccGCTGACATCGAGCTTTTCTACA (SEQ ID NO: 27)
	T545: обратный	aaacTGTAGAAAGCTCGATGTCAGC (SEQ ID NO: 28)
LTR-O	T546: прямой	caccGTCTACAAGGGACTTTCCGCT (SEQ ID NO: 29)
	T547: обратный	aaacAGCGGAAAGTCCCTTGTAGAC (SEQ ID NO: 30);
LTR-P	T548: прямой	caccGCTTTCCGCTGGGGACTTTCC (SEQ ID NO: 31);
	T549: обратный	aaacGGAAAGTCCCCAGCGGAAAGC (SEQ ID NO: 32);
LTR-Q	T687: прямой	caccGCCTCCCTGGAAGTCCCCAG (SEQ ID NO: 33)
	T688: обратный	aaacCTGGGGACTTTCCAGGGAGGC (SEQ ID NO: 34)
LTR-R	T689: прямой	caccGCCTGGGCGGGACTGGGGAG (SEQ ID NO: 35)
	T690: обратный	aaacCTCCCCAGTCCCCGCCAGGC (SEQ ID NO: 36)
LTR-S	T691: прямой	caccGTCCATCCCATGCAGGCTCAC (SEQ ID NO: 37)
	T692: обратный	aaacGTGAGCCTGCATGGGATGGAC (SEQ ID NO: 38)
LTR-T	T548: прямой	caccGCGGAGAGAGAAGTATTAGAG (SEQ ID NO: 39)
	T549: обратный	aaacCTCTAATACTTCTCTCTCCGC (SEQ ID NO: 40)
Gag-A	T687: прямой	caccGGCCAGATGAGAGAACCAAG (SEQ ID NO: 41)
	T688: обратный	aaacCTTGGTTCTCTCATCTGGCC (SEQ ID NO: 42)
Gag-B	T714: прямой	caccGCCTTCCCACAAGGGAAGGCCA (SEQ ID NO: 43)
	T715: обратный	aaacTGGCCTTCCCTTGTGGGAAGGC (SEQ ID NO: 44);
Gag-C	T758: прямой	caccGCGAGAGCGTCGGTATTAAGCG (SEQ ID NO: 45)
	T759: обратный	aaacCGCTTAATACCGACGCTCTCGC (SEQ ID NO: 46)
Gag-D	T760: прямой	caccGGATAGATGTAAAAGACACCA (SEQ ID NO: 47)
	T761: обратный	aaacTGGTGTCTTTTACATCTATCC (SEQ ID NO: 48)
Pol-A	T689: прямой	caccGCAGGATATGTAAGTACAG (SEQ ID NO: 49)
	T690: обратный	aaacCTGTCAGTTACATATCCTGC (SEQ ID NO: 50)
Pol-B	T716: прямой	caccGCATGGGTACCAGCACACAA (SEQ ID NO: 51)
	T717: обратный	aaacTTGTGTGCTGGTACCCATGC (SEQ ID NO: 52)
PCR	T422	caccGCTTTATTGAGGCTTAAGCAG (SEQ ID NO: 53)
	T425	aaacGAGTCACACAACAGACGGGC (SEQ ID NO: 54)
	T645	TGGAATGCAGTGGCGCGATCTTGGC (SEQ ID NO: 55)
	T477	CACAGCATCAAGAAGAACCTGAT (SEQ ID NO: 56)
	T478	TGAAGATCTCTGCAGATAGCAG (SEQ ID NO: 57)

Результаты

Биоинформационный скрининг sgRNA с высокой эффективностью и низким нецелевым воздействием. Эффективность и специфичность целевых gRNA является важной проблемой для применения Cas9/gRNA при инфекционных заболеваниях. Были созданы несколько компьютерных программ для разработки и отбора gRNA для системы spCas9-gRNA, где использовали затравочную последовательность размером 20 п. о. и NRG PAM. Несмотря на то, что программы по разработке gRNA были созданы для прогноза нецелевых эффектов, очень незначительное число программ были способны прогнозировать эффективность расщепления. Двадцать (20) gRNA, нацеливающихся на LTR HIV-1, разрабатывали с высоким баллом эффективности расщепления и специфичности в отношении генома человека (таблица 1) с использованием следующих критериев: (1) нацеливание на участок промотора LTR-U3 размером от -18 до -418 п. о. с целью нарушения исходной транскрипции ВИЧ-1 (и подавления образования вирусов), и при этом данный участок размером 400 п. о. заранее исключен в большинстве LV, препятствуя таким образом саморасщеплению LV; (2) исключение сайтов связывания факторов транскрипции, которые могут повлиять на экспрессию генов клеток хозяина в результате высокой гомологии (фиг. 1C); (3) совпадение обоих концевых LTR для обеспечения удаления всей провирусной ДНК между LTR; (4) балл нецелевого воздействия не более 50% и (5) применимость в отношении двойной нуклеотидной последовательности spCas9 или димерных РНК-направляемых нуклеаз FokI. Обнаружили несколько gRNA, нацеливающихся на структурный участок Gag и Pol (фиг. 1C) с перспективой получения наиболее эффективной комбинации gRNA для удаления всего генома ВИЧ-1. Структурный участок Env не отбирали в связи с более низкой консервативностью этой структурной последовательности между различными штаммами. Систему доставки гена LV выбирали отдельно для экспрессии spCas9 и gRNA (фиг. 1B) по следующим причинам: (1) сам LV обеспечивает много преимуществ для высокоэффективной генной терапии в линиях клеток с латентным ВИЧ, тяжело поддающихся трансфекции, исследований на животных и перспективного клинического применения (LV, не способные к интеграции; Hu P, et al. Mol Ther Methods Clin Dev 2015,2:15025; Liu KC, et al. Curr Gene Ther 2014,14:352-364); (2) отдельный LV с spCas9 обеспечивает высокую эффектив-

ность упаковки гена *spCas9* большого размера; (3) экспрессирующий *gRNA* отдельный *LV* можно использовать для клонирования экспрессирующих мультиплексную *gRNA* кассет в один вектор в целях высокой эффективности упаковки.

Функциональный скрининг в клетках HEK293T в целях идентификации эффективных *gRNA*. Для быстрого функционального скрининга наиболее эффективных мишеней анализ репортерного гена *EcoHIV-eLuc* выполняли с помощью высокопродуктивного многоканального планшетного ридера Envision. Выбирали репортерный ген *EcoHIV-eLuc*, поскольку (1) он содержит все компоненты, необходимые для репликации ВИЧ-1 помимо *Env* ВИЧ, (2) является удобным при работе в контейнерах II уровня биобезопасности в результате делеции *Env*, и (3) биолюминесценция является более чувствительной, чем флуоресценция, а репортер *eLuc* можно использовать для выявления менее 10 отдельных клеток (Song J, et al. *J Gen Virol* 2015,96:3131-3142). Выбирали линию клеток HEK293T в связи с высокой эффективностью трансфекции с помощью экономически эффективного способа осаждения фосфатом кальция. С помощью трансфекции одиночной *gRNA* было обнаружено, что большая часть *gRNA*, нацеливающихся на промотор LTR и структурный участок, могли приводить только к предельному снижению образования репортерного гена провируса *EcoHIV*, однако некоторые повышали промоторную активность или не оказали никакого воздействия (фиг. 2А, 2В). Повышение промоторной активности согласуется с недавним отчетом о том, что индуцированный *spCas9/gRNA* DSB в промоторе генов раннего ответа нейронов стимулировал их экспрессию (Madabhushi R, et al. *Cell* 2015,161:1592-1605). Была выдвинута гипотеза, что одиночная *gRNA* индуцировала один разрез целевых сайтов, приводя к образованию мутаций по типу вставки/делеции в целевых участках и/или делеции всей провирусной ДНК между обеими концевыми LTR. Мутация в промоторе может нарушать функциональную активность активаторов и/или репрессоров транскрипции, которые приводят к повышению или снижению транскрипционной активности. Мутация в структурном участке может приводить к сдвигу открытой рамки считывания структурных белков ВИЧ-1 и, таким образом, снижать экспрессию репортерного гена *eLuc*.

В целях обеспечения более надежного и чувствительного скрининга эффективных *gRNA* в отношении функционального расщепления спаренные *gRNA* котрансфицировали: каждую *gRNA* LTR относительно одной из *gRNA*, нацеливающихся на структурный участок. При такой стратегии наблюдали более значительное снижение репортерных вирусов благодаря делеции большого фрагмента между 5'- или 3'-LTR и структурным участком. В качестве примера, показанного на фиг. 2В, все комбинации между *GagD* и любой из LTR-*gRNA* значительно снижали экспрессию *eLuc* (64-96%), что является более надежным, чем использование одиночной *gRNA*. Половина из LTR-*gRNA* (10/20) характеризовалась >90% снижением активности *eLuc*. Аналогично в качестве другого примера, показанного на фиг. 2С, LTR-R *gRNA*, спаренные с любой из *GagA-D* или *PolA-B* значительно снижали активность репортерного гена люциферазы до 7-23%. Выбор *GagD* или LTR-R для спаривания также был основан на их сайте PAM, применимом в отношении системы *Cas9 Staphylococcus aureus*, и их сайтах для нацеливания, применимых в отношении линий клеток с латентным ВИЧ-1 (Jadlowsky JK, et al. *Mol Cell Biol* 2014,34:1911-192) и трансгенных мышей Tg26 (Kopp JB, et al. *Proc Natl Acad Sci U S A* 1992,89:1577-1581), где часть последовательностей *Gag* и всю последовательность *Pol* удаляли. Эти данные доказали, что комбинация LTR-*gRNA* со структурными *gRNA* была более эффективной и более легкой стратегией для скринирования эффективных *gRNA* с помощью высокопродуктивного анализа репортерного гена *eLuc* ВИЧ-1. Все разработанные *gRNA* являются функциональными в целях снижения экспрессии репортерного гена *EcoHIV-eLuc*, что согласуется с высоким баллом прогноза эффективности с помощью биоинформационного анализа.

Идентификация эффективных *gRNA* с помощью генотипирования на основе прямой ПЦР. Для проверки того, являются ли эти кандидатные *gRNA* функциональными, чтобы расщепить соответствующие мишени, согласно разработке, выполняли анализ генотипирования на основе прямой ПЦР с использованием образцов ДНК со спаренными *gRNA* и соответствующими праймерами для ПЦР, как указано (фиг. 3D). Подход на основе прямой ПЦР не требует экстракции и очистки ДНК и, таким образом, является более удобным для скрининга генотипов. При использовании одной из *gRNA*, нацеливающихся на структурные участки, с целью спаривания с сайтами для нацеливания LTR, генотипирование на основе ПЦР, по-видимому, приводило к образованию новых фрагментов (обозначенных "делеция" для удобства описания), полученных из оставшейся (остаточной) вирусной последовательности LTR и *Gag* после удаления фрагментов между 5'-LTR и *Gag* (фиг. 3Е, праймер T361/T458). В соответствии с анализом репортерного гена *eLuc* почти все *gRNA* индуцировали различную степень уменьшения полосы дикого типа, которую легко амплифицировали при помощи стандартного условия ПЦР в 5'-LTR/*Gag*, благодаря размеру (1,3 т. о.). После расщепления различную степень обозначенной "делеции" выявляли в большинстве комбинаций (фиг. 3Е). Примечательно, что дополнительные фрагменты (обозначенные "вставка"), более крупные, чем прогнозируемая "делеция", наблюдали в большинстве комбинаций при расщеплении 5'-LTR-*Gag* (фиг. 3Е). Количественное определение интенсивности полосы дикого типа показало, что LTR-О характеризуется максимальной эффективностью, за ней следуют I, С, А, как показано в прямоугольнике (фиг. 3Е). Такой характер эффективности расщепления полосы дикого типа не полностью коррелировал с характером снижения активности репортерного гена *eLuc* (фиг. 2В), вероятно в связи с тем, что для

амплификации продукта ПЦР в смешанной группе обычно предпочтительным является продукт небольшого размера. С другой стороны, слабое уменьшение полосы дикого типа в некоторых парах могло происходить в результате различной степени незначительных мутаций по типу вставки/делеции (в пределах нескольких нуклеотидов) в пределах целевых сайтов для gRNA без какой-либо фрагментарной делеции или вставки. Для избежания потенциального влияния предпочтительной амплификации при ПЦР, генотипирование на основе ПЦР выполняли с помощью праймеров, охватывающих 3'-LTR и Gag (фиг. 3E, 3I, 3J), что предположительно приводило к образованию продукта ПЦР дикого типа размером 7 т. о., который маловероятно подвергался амплификации при используемых условиях ПЦР. Фрагментарный характер "делеции" в gRNA, выявляемый с помощью одного продукта ПЦР (фиг. 3E, 3I, 3J), соответствовал обнаруженной по изменениям относительного показателя (фиг. 3E). Пара LTR-K и GagD не характеризовалась фрагментарной полосой "делеции" или "вставки" во всех четырех сериях реакций генотипирования на основе ПЦР (фиг. 3E, 3F, 3I, 3J), коррелируя только с 7% снижением в полосе дикого типа (фиг. 3E). Пара LTR-F по сравнению с GagD характеризовалась слабой полосой "делеции" в одной серии реакций ПЦР (3I), коррелируя с 17% снижением в полосе дикого типа (фиг. 3E). Пары LTR-G, P по сравнению с GagD характеризовались примерно 50% снижением интенсивности полос дикого типа в результате образования делеции в 5'-LTR-Gag (фиг. 3E) или расщепления Gag-3'-LTR (фиг. 3F, 3I, 3J). При использовании LTR-R для спаривания с любой из GagA-D и PolA-B генотипирование на основе ПЦР с указанными соответствующими праймерами также прогнозировало образование новых фрагментов ("делеция") (фиг. 3G, 3H) и дополнительных вставок в 5'-LTR/Gag (фиг. 3G) в различной степени во всех исследуемых gRNA помимо gRNA Gag-B, которая характеризовалась очень слабой способностью редактирования (фиг. 3G). Однако все эти комбинации со слабым генотипированием делеций или его отсутствием по-прежнему характеризовались значительным снижением активности репортерного гена EcoHIV-eLuc (фиг. 2B). Это могло быть не связано ни с одной или связано только с одной из двух плазмид для gRNA, трансфицированных в те же самые клетки, где одиночная gRNA остается высокоэффективной в индукции небольшой мутации по типу вставки/делеции. В совокупности эти данные свидетельствуют о том, что генотипирование на основе прямой ПЦР обеспечивает надежное и быстрое средство проверки наличия фрагментарной "делеции" и/или "вставки". Однако для оценки эффективного удаления ВИЧ-1 с помощью различных gRNA требуется комбинация функционального снижения с помощью анализа репортерного гена вируса и фрагментарного вырезания провирусной ДНК с помощью 5'-LTR- или 3'-LTR-направленного генотипирования на основе ПЦР.

Проверка фрагментарной мутации "вставка"/"делеция" с помощью ТА-клонирования и секвенирования по Сэнгеру. Для дальнейшей проверки эффективности расщепления spCas9/gRNA и исследования характера мутации "вставка/делеция" выбирали три типичных образца генотипирования на основе ПЦР для ТА-клонирования и секвенирования по Сэнгеру. Парная экспрессия LTR-R/GagA вызывала "делецию" фрагмента размером 519 п. о. между целевыми сайтами LTR-R и GagA (фиг. 12A). Кокспрессия LTR-L/GagD (фиг. 12B) и LTR-M/GagD (фиг. 12C) приводила к "делеции" фрагмента размером 772 п. о. или 763 п. о. между каждой парой целевых сайтов соответственно. Кроме того, они приводили к образованию небольших мутаций по типу вставки/делеции различных размеров или типов. В некоторых случаях идентифицировали крупную "вставку" из дополнительных последовательностей (например, размером 159-359 п. о.) (фиг. 3E, 12B, 12C). Анализ NCBI Blast показал, что эти дополнительные последовательности происходили из экзогенных векторов, а не эндогенных генов клетки-хозяина. Эти результаты указывают на то, что большая часть этих кандидатных gRNA может эффективно опосредовать целевой разрыв интегрированного генома ВИЧ с помощью вырезания или вставки/делеции.

Обсуждение

Мультиплексные gRNA могли индуцировать делецию больших фрагментов между целевыми сайтами (Hu W, et al. Proc Natl Acad Sci U S A 2014,111:11461 -11466), что обеспечивает надежное средство оценки эффективности расщепления ДНК с помощью Cas9/gRNA. В настоящем исследовании такое доказательство концепции далее подтверждали с помощью скрининга различных мультиплексов из 26 gRNA. Было показано, что большая часть разработанных gRNA являются высокоэффективными при удалении прогнозируемой последовательности генома ВИЧ-1 между двумя выбранными сайтами для нацеливания, приводя к значительному вырезанию репортерного вируса ВИЧ-1. В частности, комбинация вирусных структурных gRNA с одной или двумя gRNA LTR обеспечивала гораздо более высокую эффективность удаления генома и более легкий подход с помощью генотипирования на основе прямой ПЦР и высокопродуктивного скрининга репортерного гена. Эффективность и специфичность gRNA, выбранных в настоящем исследовании для вырезания ДНК провируса ВИЧ-1, обеспечивают успешность доклинических исследований на животных и клинических исследований пациентов с помощью технологии Cas9/gRNA, поскольку: (1) эти gRNA могут служить в качестве готового к применению источника отбора для разработки вирусной и невирусной генной терапии; (2) для отдельного пациента с ВИЧ-1 эти gRNA можно применять в качестве образца для скринирования новых gRNA, разработанных специально для изолятов ВИЧ-1, несмотря на высокий уровень мутаций ВИЧ-1; (3) легкое клонирование gRNA, быстрой скрининг репортеров и надежное генотипирование на основе прямой ПЦР обеспечивают реализуемость для практического применения Cas9/gRNA в области персонализированной медицины.

Не все разработанные gRNA характеризуются необходимыми активностями расщепления предполагаемых целевых сайтов. К настоящему времени было разработано несколько подходов для оценки эффективности редактирования генома, индуцированного с помощью технологии Cas9/gRNA. Постоянно совершенствуемые компьютерные программы для прогноза эффективности тестировали с помощью геномов клеток хозяина в качестве разрабатываемой мишени (Doench JG, et al. *Nat Biotechnol* 2014,32:1262-1267; Gagnon JA, et al. *PloS One* 2014,9:e98186; Liu H, et al. *Bioinformatics* 2015), однако они могут быть ненадежными для применения в отношении экзогенных геномов, таких как вирусы, вызывающие инфекционные заболевания. Секвенирование по Сэнгеру целевого участка посредством клонирования с помощью ПЦР обеспечивает высокую чувствительность и специфичность для определения эффективности редактирования генома (Sander JD, et al. *Nat Biotechnol* 2011, 29:697-698), однако является трудоемким для высокопродуктивного скрининга. Анализ на основе ошибочных спариваний Surveyor (Qiu P, et al. *Biotechniques* 2004,36:702-707; Kim JM, et al. *Nat Commun* 2014,5:3157; Dahlem TJ, et al. *PLoS Genet* 2012,8:e1002861) и анализ плавления с высоким разрешением (Bassett AR, Liu JL. *J Genet Genomics* 2014,41:7-19) являются чувствительными для выявления небольших мутаций по типу вставки/делеции, однако слабая специфичность делает их подверженными к генерации ложно-положительных результатов. Анализ полиморфизма длин рестрикционных фрагментов (RFLP) требует наличия сайта для рестрикционного фермента, который является ограниченным в большинстве случаев (Kim JM, et al. *Nat Commun* 2014,5:3157). Секвенирование следующего поколения обеспечивает надежное и специфичное измерение, однако является дорогостоящим и времязатратным (Guell M, et al. *Bioinformatics* 2014,30:2968-2970). В последнее время несколько анализов на основе ПЦР обеспечивают легкий и надежный способ количественно оценить эффективность редактирования, однако для них требуется разработка надежных праймеров, разложение трассировки или капиллярный секвенатор (Brinkman EK, et al. *Nucleic Acids Res* 2014,42:e168; Carrington B, et al. *Nucleic Acids Res* 2015; Yu C, et al. *PLoS One* 2014,9:e98282). В данном документе быструю, рентабельную и надежную платформу для скрининга создавали в целях идентификации эффективных gRNA с помощью высокочувствительного высокопродуктивного биолуминесцентного анализа репортерного гена совместно с быстрым генотипированием на основе прямой ПЦР. Анализ репортерного гена основан на удалении больших фрагментов между двумя целевыми сайтами для gRNA, а также небольших мутаций по типу вставки/делеции в каждом сайте gRNA.

Фрагментарное удаление подавляет промоторную активность или репортерную экспрессию, в то время как мутации по типу вставки/делеции могут изменить регуляцию промотора или индуцировать сдвиг открытой рамки считывания вирусных белков. Все эти события будут в дальнейшем нарушать активность репортерного гена. Генотипирование на основе ПЦР основано на фрагментарном расщеплении и эффективном повторном лигировании между оставшимися концевыми ДНК. Наличие повторно лигированных фрагментов ПЦР обеспечивает подтверждающее доказательство эффективности обеих gRNA. Эффективность повторного лигирования зависит от клеточного деления, поэтому генотипирование на основе ПЦР может быть ограничено в случае неделящихся клеток. Кроме того, условие ПЦР для некоторых праймеров требует оптимизации для достижения максимальной эффективности генотипирования.

Целью настоящего исследования была проверка и идентификация эффективных gRNA с помощью надежных и чувствительных высокопродуктивных анализов. Транзистентную трансфекцию репортерного гена EcoHIV-eLuc в клетки HEK293T выбирали в качестве платформы тестирования, поскольку незначительное количество репортерной плазмиды по сравнению с компонентами spCas9/gRNA (1:20) может обеспечить целевое расщепление во всех экспрессирующих репортерный ген клетках и, таким образом, повысить до максимума эффективность выявления анализа репортерного гена люциферазы и генотипирования на основе ПЦР. В отличие от этого стабильная линия клеток EcoHIV-eLuc на основе клеток HEK293T (фиг. 13A-13E, 14A-14D), которая может быть ближе к реальной ситуации латентного состояния ВИЧ-1, характеризовалась слабой чувствительностью выявления как при анализе репортерного гена люциферазы, так и генотипировании на основе ПЦР. Это объясняется тем, что экспрессирующие EcoHIV-eLuc клетки без какой-либо плазмиды для gRNA всегда существуют после трансфекции и, таким образом, репортер eLuc постоянно экспрессируется, даже в случае, если эффективность трансфекции может составлять до 80-90%.

Дополнительные преимущества транзистентной трансфекции репортерного гена включают легкую установку, рентабельную трансфекцию и высокопродуктивное измерение люминесценции. Важно, что идентифицированные gRNA остаются эффективными в реальных латентно инфицированных ВИЧ клетках или линиях клеток и могут дополнительно применяться для исследований на животных и клинических вариантов применения. Несмотря на то, что транзистентно трансфицированный репортерный ген EcoHIV-eLuc (эписомальная ДНК) не отражает провирусную ДНК латентного ВИЧ в геноме хозяина (ядре), опосредованное spCas9/gRNA редактирование гена функционирует с аналогичной эффективностью между эписомальной и ядерной ДНК провируса ВИЧ (Hu W, et al. *Proc Natl Acad Sci U S A* 2014,111:11461-11466) и других вирусов (Ramanathan V, et al. *Sci Rep* 2015,5:10833; Yuen KS, et al. *J Gen Virol* 2015,96:626-636). Кроме того, эффективное расщепление эписомальной ДНК в добавление к интегрированной провирусной ДНК ВИЧ-1 обеспечивает новое профилактическое лечение новой инфекции,

вызываемой ВИЧ (Hu W, et al. Proc Natl Acad Sci U S A 2014,111:11461 -11466) и других вирусов, вызывающих инфекционные заболевания (Peng C, Lu M, Yang D. Virol Sin 2015,30:317-325).

Некоторые искажающие факторы могут нарушать эффективность транзientной трансфекции и экспрессию трансгенов для сравнительного анализа различных gRNA. В целях минимизации этого, предпринимали несколько мер предосторожности: 1) мастер-микс репортерного гена и плазмид на основе spCas9 готовили для обеспечения равного количества этих совместных плазмид в каждой группе gRNA; 2) люциферазный репортер Renilla (1:100) использовали для нормализации эффективности трансфекции; 3) масштабную трансфекцию выполняли в 96-луночной планшете для всех gRNA в 4-6 повторах в одно и то же время и 4) все данные выражали в качестве относительных изменений по сравнению с пустым контролем с gRNA в каждом эксперименте.

Одна gRNA, нацеливающаяся на LTR-участок, может элиминировать всю провирусную ДНК, благодаря расщеплению обоих концевых LTR, однако эффективность удаления не была заметной, как показано с помощью анализа репортерного гена EcoHIV-eLuc. Для этого также требуется ПЦР длинных фрагментов в целях проверки удаления всей провирусной ДНК ВИЧ-1, поскольку с помощью стандартной ПЦР с праймерами, охватывающими LTR, нельзя отличить 5'-LTR от 3'-LTR после удаления фрагмента между двумя целевыми сайтами LTR (Hu W, et al. Proc Natl Acad Sci USA 2014,111:11461-11466). Две gRNA, нацеливающиеся на LTR-участок, индуцировали фрагментарное расщепление в пределах каждой LTR-области, которое будет подавлять промоторную активность LTR и снижать стабильность РНК ВИЧ-1, тем самым повышая эффективность полного удаления, как показывали ранее (Hu W, et al. 2014). В настоящем исследовании проверяли новое доказательство принципа, заключающегося в том, что любая пара gRNA между LTR и структурными участками обеспечивает более эффективный подход для оценки эффективности удаления ВИЧ-1. С помощью этого способа значительное функциональное снижение образования репортерного вируса ВИЧ-1 происходит в результате трех возможных расщеплений 5'LTR+Gag, Gag+3'LTR и обоих концевых LTR и может легко отслеживаться с помощью чувствительного и высокопродуктивного биолуминесцентного анализа репортерного гена. Эти расщепления можно эффективно и надежно выявлять с помощью стандартного и быстрого генотипирования на основе ПЦР с использованием праймеров, охватывающих LTR и структурные участки. Аналогично коктейль из двух gRNA LTR совместно с одним или двумя структурными gRNA может обеспечивать оптимальное и экономичное средство для удаления генома ВИЧ-1 в доклинических и клинических условиях.

Возможность нецелевых эффектов, включающих технологию Cas9/gRNA, представляла собой большую проблему в области редактирования генома. Для повышения специфичности редактирования генома создали строгую разработку gRNA, функциональный скрининг и технологическую модификацию Cas9. Выявили очень редкие случаи нецелевых эффектов, связанных с spCas9/gRNA, в культивируемых клетках с помощью секвенирования в масштабе всего генома (WGS) (Hu W, et al. 2014; Zuckermann M et al. Nat Commun 2015,6:7391; Smith C, et al. Cell Stem Cell 2014,15:12-13; Veres A, et al. Cell Stem Cell 2014,15:27-30; Yang L, et al. Nat Commun 2014,5:5507). В недавно разработанных несмещенных методах профилирования дополнительно проверяют высокую эффективность этой системы spCas9-gRNA (Ran FA, et al. Nature 2015,520:186-191; Tsai SQ, et al. Nat Biotechnol 2015,33:187-197; Frock RL, et al. Nat Biotechnol 2015,33:179-186). Ожидается, что нецелевой эффект in vivo будет гораздо меньшим вследствие эпигенетической защиты. В настоящем исследовании экзогенную вирусную ДНК анализировали против генома хозяина на наличие максимального балла эффективности и специфичности. Во время скрининга gRNA токсичности в клетках не наблюдали. Было показано, что двойные никазы spCas9 и РНК-направляемые нуклеазы FokI снижали потенциальные нецелевые эффекты до 1500 раз включительно (Ran FA, et al. Cell 2013,154:1380-1389; Mali P, et al. Nat Biotechnol 2013,31:833-838; Wyvekens N, et al. Hum Gene Ther 2015,26:425-431; Tsai SQ, et al. Nat Biotechnol 2014,32:569-576).

В заключение необходимо отметить, что большая часть разработанных gRNA является высокоэффективной для удаления прогнозируемой последовательности генома ВИЧ-1 между определенными двумя сайтами для нацеливания и нарушает активность репортерного гена eLuc. В частности, комбинация вирусных структурных gRNA с одной или двумя gRNA LTR обеспечивала более высокую эффективность удаления генома и более легкий подход с помощью генотипирования на основе ПЦР. Скрининг с помощью анализа репортерного гена eLuc ВИЧ-1 и генотипирования на основе прямой ПЦР обеспечивает надежный, быстрый и удобный подход для скрининга эффективных gRNA ВИЧ-1. Этим можно воспользоваться в целях выполнения высокопродуктивного тестирования библиотеки gRNA для любых новых изолятов ВИЧ-1 и других вирусов, вызывающих инфекционные заболевания, во время новой эры персонализированной/точной медицины.

Изобретение было описано в иллюстративном виде, и следует понимать, что используемая терминология предназначена для описания происхождения слов, а не ограничения. Очевидно, что многие модификации и варианты настоящего изобретения возможны в свете вышеприведенных идей. Поэтому следует понимать, что в объеме прилагаемой формулы изобретения настоящее изобретение может быть осуществлено иначе, чем конкретно описано.

Характерные признаки настоящего изобретения изложены в следующих пунктах.

1. Композиция для применения в инактивации ретровирусной ДНК, интегрированной в геном клет-

ки-хозяина, латентно инфицированной ретровирусом, при этом композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в интегрированной ретровирусной ДНК, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ).

2. Композиция по п.1, где указанная по меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в интегрированной ретровирусной ДНК, и вторую gRNA, которая является комплементарной другой целевой последовательности в интегрированной ретровирусной ДНК, причем промежуточные последовательности между двумя gRNA удалены.

3. Композиция по п.2, где целевые последовательности нуклеиновой кислоты предусматривают одну или более последовательностей нуклеиновой кислоты в кодирующих и не кодирующих последовательностях нуклеиновой кислоты генома ВИЧ.

4. Композиция по п.3, где целевые последовательности предусматривают одну или более последовательностей нуклеиновой кислоты в ВИЧ, которые содержат последовательности нуклеиновой кислоты длинного концевой повтора (LTR), последовательности нуклеиновой кислоты, кодирующие структурные белки, неструктурные белки или их комбинации.

5. Композиция по п.4, где последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации.

6. Композиция по п.4, где последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, акцессорные белки или их комбинации.

7. Композиция по п.6, где регуляторные белки предусматривают Tat, Rev или их комбинации.

8. Композиция по п.6, где акцессорные белки предусматривают Nef, Vpr, Vpu, Vif или их комбинации.

9. Композиция по п.3, где целевая последовательность нуклеиновой кислоты предусматривает одну или более последовательностей в участке длинного концевой повтора (LTR) провирусной ДНК вируса иммунодефицита человека (ВИЧ) и одну или более мишеней в структурном и/или неструктурном гене интегрированной ДНК ВИЧ; или одну или более мишеней во втором гене; или одну или более мишеней в первом гене и одну или более мишеней во втором гене; или одну или более мишеней в первом гене, и одну или более мишеней во втором гене, и одну или более мишеней в третьем гене; или одну или более мишеней во втором гене и одну или более мишеней в третьем гене или четвертом гене; или любую их комбинацию.

10. Композиция по п.3, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 75% идентичностью последовательности с SEQ ID NO: 1-57.

11. Композиция по п.10, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

12. Способ инактивации интегрированной ретровирусной ДНК в клетке млекопитающего, предусматривающий стадии воздействия на клетку с помощью композиции, содержащей по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую комплекс редактирования гена, содержащий эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), комплементарную целевой последовательности нуклеиновой кислоты в интегрированной ретровирусной ДНК; экспрессии комплекса редактирования гена; и инактивации интегрированной ретровирусной ДНК.

13. Способ по п.12, где интегрированная ретровирусная ДНК представляет собой ДНК вируса иммунодефицита человека (ВИЧ), при этом по меньшей мере одна последовательность выделенной нуклеиновой кислоты кодирует первую gRNA, комплементарную целевой последовательности нуклеиновой кислоты в LTR-участке ДНК ВИЧ, и вторую gRNA, комплементарную целевой последовательности нуклеиновой кислоты в структурном гене ДНК ВИЧ.

14. Способ по п.12, где последовательность нуклеиновой кислоты gRNA содержит последовательность, характеризующуюся по меньшей мере 75% идентичностью последовательности с последовательностями нуклеиновой кислоты, содержащими SEQ ID NO: 1-57.

15. Способ по п.14, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

16. Фармацевтическая композиция для инактивации интегрированной ретровирусной ДНК у субъекта-млекопитающего, содержащая последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR); по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую по меньшей мере одну направляющую РНК (gRNA), которая является комплементарной це-

левой последовательности в ретровирусной ДНК; при этом указанные последовательности выделенной нуклеиновой кислоты включены по меньшей мере в один вектор экспрессии.

17. Фармацевтическая композиция по п.16, где интегрированная ретровирусная ДНК представляет собой ДНК вируса иммунодефицита человека (ВИЧ), и указанная по меньшей мере одна gRNA включает в себя первую gRNA, которая является комплементарной первой целевой последовательности в ДНК ВИЧ, и вторую gRNA, которая является комплементарной второй целевой последовательности в ДНК ВИЧ.

18. Фармацевтическая композиция по п.16, где указанная первая целевая последовательность расположена в LTR-участке ДНК ВИЧ, и указанная вторая целевая последовательность расположена в структурном гене в ДНК ВИЧ.

19. Фармацевтическая композиция по п.16, где последовательность нуклеиновой кислоты gRNA содержит последовательность, характеризующуюся по меньшей мере 75% идентичностью последовательности с последовательностями нуклеиновой кислоты, содержащими SEQ ID NO: 1-57.

20. Фармацевтическая композиция по п.19, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

21. Способ лечения субъекта-млекопитающего, инфицированного вирусом иммунодефицита человека (ВИЧ), предусматривающий стадии определения того, что субъект-млекопитающее инфицирован ВИЧ, введения эффективного количества фармацевтической композиции по п.16, и лечения субъекта-млекопитающего в отношении инфекции, вызываемой ВИЧ.

22. Способ лечения для снижения риска возникновения инфекции, вызываемой вирусом иммунодефицита человека (ВИЧ), у субъекта-млекопитающего, предусматривающий стадии определения того, что субъект-млекопитающее подвержен риску возникновения инфекции, вызываемой ВИЧ, введения эффективного количества фармацевтической композиции по п.16, и снижения риска возникновения инфекции, вызываемой ВИЧ, у субъекта-млекопитающего.

23. Набор для лечения или профилактики инфекции, вызываемой вирусом иммунодефицита человека (ВИЧ), содержащий отмеренное количество композиции, содержащей по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, и по меньшей мере одну последовательность нуклеиновой кислоты, кодирующую одну или более направляющих РНК (gRNA), где каждая из указанных одной или более gRNA является комплементарной целевому сайту в геноме ВИЧ; или вектор, кодирующий одну или более из указанных выделенных нуклеиновых кислот; и один или более продуктов, выбранных из группы, состоящей из упаковочного материала, листка-вкладыша в упаковке с инструкциями по применению, стерильной жидкости, шприца и стерильного контейнера.

24. Набор по п.23, где указанная одна или более gRNA включает в себя первую gRNA, которая является комплементарной первой целевой последовательности в геноме ВИЧ, и вторую gRNA, которая является комплементарной второй целевой последовательности в геноме ВИЧ.

25. Набор по п.23, где указанная первая целевая последовательность расположена в LTR-участке генома ВИЧ, и указанная вторая целевая последовательность расположена в структурном и/или неструктурном гене генома ВИЧ.

26. Вектор экспрессии для инактивации провирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом, где вектор экспрессии содержит по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в провирусной ДНК, для инактивации провирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом.

27. Вектор экспрессии по п.26, где указанная по меньшей мере одна gRNA включает в себя по меньшей мере первую направляющую gRNA, которая является комплементарной первой целевой последовательности в провирусной ДНК, и вторую gRNA, которая является комплементарной второй целевой последовательности в провирусной ДНК.

28. Вектор экспрессии по п.26, где указанная первая целевая последовательность расположена в LTR-участке генома ВИЧ, и указанная вторая целевая последовательность расположена в структурном и/или неструктурном гене генома ВИЧ.

29. Полинуклеотид для инактивации ретровирусного генома, интегрированного в геном клетки-хозяина, латентно инфицированной ретровирусом, где полинуклеотид содержит по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в интегрированном ретровирусном геноме, для инактивации ретровирусного генома, интегрированного в геном клетки-хозяина, латентно инфицированной ретровирусом.

30. Композиция для удаления ретровируса *in vitro* или *in vivo*, при этом композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с

короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в ретровирусном геноме, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ).

31. Композиция по п.30, где указанная по меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в геноме ВИЧ, и вторую gRNA, которая является комплементарной другой целевой последовательности в геноме ВИЧ, причем промежуточные последовательности между двумя gRNA удалены.

32. Композиция по п.30, где целевые последовательности нуклеиновой кислоты предусматривают одну или более последовательностей нуклеиновой кислоты в кодирующих и не кодирующих последовательностях нуклеиновой кислоты генома ВИЧ.

33. Композиция по п.32, где целевые последовательности предусматривают одну или более последовательностей нуклеиновой кислоты в геноме ВИЧ, которые содержат последовательности нуклеиновой кислоты длинного концевой повтора (LTR), последовательности нуклеиновой кислоты, кодирующие структурные белки, неструктурные белки или их комбинации.

34. Композиция по п.33, где последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации.

35. Композиция по п.33, где последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, акцессорные белки или их комбинации.

36. Композиция по п.35, где регуляторные белки предусматривают Tat, Rev или их комбинации.

37. Композиция по п.35, где акцессорные белки предусматривают Nef, Vpr, Vpu, Vif или их комбинации.

38. Композиция по п.30, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 75% идентичностью последовательности с SEQ ID NO: 1-57.

39. Композиция по п.38, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

40. Композиция для удаления ретровируса *in vitro* или *in vivo*, при этом композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере две направляющие РНК (gRNA), при этом gRNA являются комплементарными различным целевым последовательностям в ретровирусном геноме, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ).

41. Композиция по п.40, где по меньшей мере одна направляющая РНК (gRNA) включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в геноме ВИЧ, и вторую gRNA, которая является комплементарной другой целевой последовательности в геноме ВИЧ, причем промежуточные последовательности между двумя gRNA удалены.

42. Композиция по п.40, где целевые последовательности нуклеиновой кислоты предусматривают одну или более последовательностей нуклеиновой кислоты в кодирующих и не кодирующих последовательностях нуклеиновой кислоты генома ВИЧ.

43. Композиция по п.42, где целевые последовательности предусматривают одну или более последовательностей нуклеиновой кислоты в геноме ВИЧ, которые содержат последовательности нуклеиновой кислоты длинного концевой повтора (LTR), последовательности нуклеиновой кислоты, кодирующие структурные белки, неструктурные белки или их комбинации.

44. Композиция по п.43, где последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации.

45. Композиция по п.43, где последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, акцессорные белки или их комбинации.

46. Композиция по п.45, где регуляторные белки предусматривают Tat, Rev или их комбинации.

47. Композиция по п.45, где акцессорные белки предусматривают Nef, Vpr, Vpu, Vif или их комбинации.

48. Композиция по п.40, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 75% идентичностью последовательности с SEQ ID NO: 1-57.

49. Композиция по п.48, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Композиция для применения в инаktivации ретровирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом, при этом композиция содержит последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в интегрированной ретровирусной ДНК, где ретровирус представляет собой вирус иммунодефицита человека (ВИЧ), при этом указанная по меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной целевой последовательности в интегрированной ретровирусной ДНК, содержащей последовательность нуклеиновой кислоты длинного концевой повтора (LTR), и вторую gRNA, которая является комплементарной другой целевой последовательности в интегрированной ретровирусной ДНК, причем промежуточные последовательности между двумя gRNA удалены.

2. Композиция по п.1, где целевые последовательности нуклеиновой кислоты предусматривают одну или более последовательностей нуклеиновой кислоты в кодирующих и не кодирующих последовательностях нуклеиновой кислоты генома ВИЧ.

3. Композиция по п.2, где целевые последовательности предусматривают одну или более последовательностей нуклеиновой кислоты в ВИЧ, которые содержат последовательности нуклеиновой кислоты длинного концевой повтора (LTR), последовательности нуклеиновой кислоты, кодирующие структурные белки, неструктурные белки или их комбинации.

4. Композиция по п.3, где последовательности, кодирующие структурные белки, содержат последовательности нуклеиновой кислоты, кодирующие Gag, предшественник Gag-Pol, Pro (протеазу), обратную транскриптазу (RT), интегразу (In), Env или их комбинации.

5. Композиция по п.3, где последовательности, кодирующие неструктурные белки, содержат последовательности нуклеиновой кислоты, кодирующие регуляторные белки, акцессорные белки или их комбинации.

6. Композиция по п.5, где регуляторные белки предусматривают Tat, Rev или их комбинации.

7. Композиция по п.5, где акцессорные белки предусматривают Nef, Vpr, Vpu, Vif или их комбинации.

8. Композиция по п.2, где целевая последовательность нуклеиновой кислоты предусматривает одну или более последовательностей в участке длинного концевой повтора (LTR) ретровирусной ДНК вируса иммунодефицита человека (ВИЧ) и одну или более мишеней в структурном и/или неструктурном гене интегрированной ДНК ВИЧ; или одну или более мишеней во втором гене; или одну или более мишеней в первом гене и одну или более мишеней во втором гене; или одну или более мишеней в первом гене, и одну или более мишеней во втором гене, и одну или более мишеней в третьем гене; или одну или более мишеней во втором гене и одну или более мишеней в третьем гене или четвертом гене; или любую их комбинацию.

9. Композиция по п.2, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, характеризующуюся по меньшей мере 75% идентичностью последовательности с SEQ ID NO: 1-57.

10. Композиция по п.9, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

11. Способ инаktivации интегрированной ретровирусной ДНК в клетке млекопитающего, предусматривающий стадии воздействия на клетку с помощью композиции, содержащей по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую комплекс редактирования гена, содержащий эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), комплементарную целевой последовательности нуклеиновой кислоты в интегрированной ретровирусной ДНК; при этом интегрированная ретровирусная ДНК представляет собой ДНК вируса иммунодефицита человека (ВИЧ), при этом по меньшей мере одна последовательность выделенной нуклеиновой кислоты кодирует первую gRNA, комплементарную целевой последовательности нуклеиновой кислоты в LTR-участке ДНК ВИЧ, и вторую gRNA, комплементарную целевой последовательности нуклеиновой кислоты в структурном гене ДНК ВИЧ; экспрессии комплекса редактирования гена; и инаktivации интегрированной ретровирусной ДНК.

12. Способ по п.11, где последовательность нуклеиновой кислоты gRNA содержит последовательность, характеризующуюся по меньшей мере 75% идентичностью последовательности с последовательностями нуклеиновой кислоты, содержащими SEQ ID NO: 1-57.

13. Способ по п.12, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

14. Фармацевтическая композиция для инаktivации интегрированной ретровирусной ДНК у субъекта-млекопитающего, содержащая последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными

группами (CRISPR); по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую по меньшей мере одну направляющую РНК (gRNA), которая является комплементарной целевой последовательности в ретровирусной ДНК; при этом указанные последовательности выделенной нуклеиновой кислоты включены по меньшей мере в один вектор экспрессии, при этом ретровирус представляет собой вирус иммунодефицита человека (ВИЧ), и указанная по меньшей мере одна gRNA включает в себя первую gRNA, которая является комплементарной первой целевой последовательности ВИЧ, и вторую gRNA, которая является комплементарной второй целевой последовательности ВИЧ, при этом указанная первая целевая последовательность расположена в LTR-участке ДНК ВИЧ, и указанная вторая целевая последовательность расположена в структурном гене в ДНК ВИЧ.

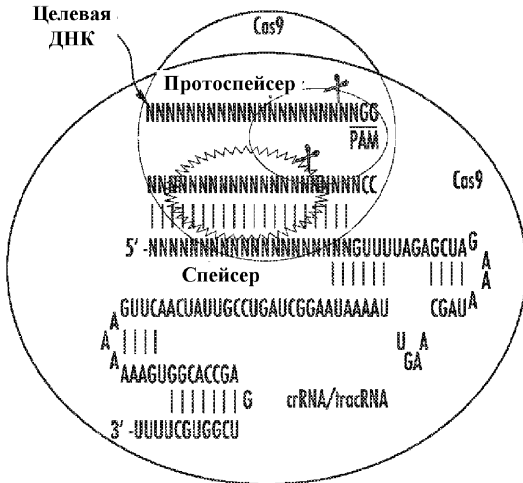
15. Фармацевтическая композиция по п.14, где последовательность нуклеиновой кислоты gRNA содержит последовательность, характеризующуюся по меньшей мере 75% идентичностью последовательности с последовательностями нуклеиновой кислоты, содержащими SEQ ID NO: 1-57.

16. Фармацевтическая композиция по п.15, где последовательность нуклеиновой кислоты gRNA содержит последовательность нуклеиновой кислоты, содержащую SEQ ID NO: 1-57.

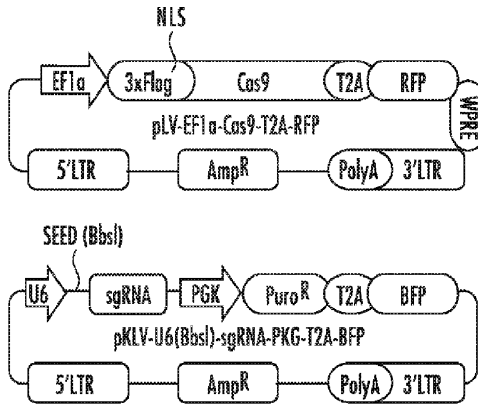
17. Применение фармацевтической композиции по п.14 для лечения субъекта-млекопитающего в отношении инфекции, вызываемой ВИЧ, или для снижения риска возникновения инфекции, вызываемой вирусом иммунодефицита человека (ВИЧ), у субъекта-млекопитающего, подверженного риску возникновения такой инфекции.

18. Вектор экспрессии для инактивации провирусной ДНК, интегрированной в геном клетки-хозяина, латентно инфицированной ретровирусом, при этом ретровирус представляет собой вирус иммунодефицита человека (ВИЧ), где вектор экспрессии содержит по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и первую направляющую РНК (gRNA), при этом первая gRNA является комплементарной первой целевой последовательности в провирусной ДНК, при этом указанная первая целевая последовательность расположена в LTR-участке генома ВИЧ; вторую gRNA, которая является комплементарной второй целевой последовательности в провирусной ДНК, при этом указанная вторая целевая последовательность расположена в структурном и/или неструктурном гене генома ВИЧ, при этом вторая целевая последовательность расположена в последовательности, кодирующей предшественник Gag-Pol, Gag или Pol, при этом промежуточные последовательности между первой целевой последовательностью и второй целевой последовательностью удалены, таким образом инактивируя провирусную ДНК, интегрированную в геном клетки-хозяина, латентно инфицированной ретровирусом.

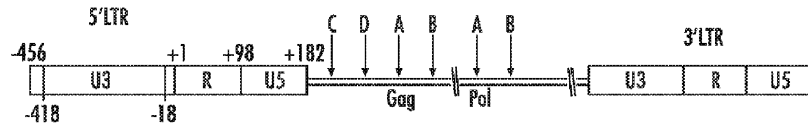
19. Полинуклеотид для инактивации ретровирусного генома, интегрированного в геном клетки-хозяина, латентно инфицированной ретровирусом, где полинуклеотид содержит по меньшей мере одну последовательность выделенной нуклеиновой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенными группами (CRISPR), и по меньшей мере одну направляющую РНК (gRNA), при этом gRNA является комплементарной целевой последовательности в интегрированном ретровирусном геноме, для инактивации ретровирусного генома, интегрированного в геном клетки-хозяина, латентно инфицированной ретровирусом, при этом ретровирус представляет собой ВИЧ, при этом указанная по меньшей мере одна gRNA включает в себя по меньшей мере первую gRNA, которая является комплементарной первой целевой последовательности генома ВИЧ, и вторую gRNA, которая является комплементарной второй целевой последовательности генома ВИЧ, при этом указанная первая целевая последовательность расположена в LTR-участке генома ВИЧ, а указанная вторая целевая последовательность расположена в структурном и/или неструктурном гене генома ВИЧ, при этом промежуточные последовательности между двумя gRNA удалены.



Фиг. 1А



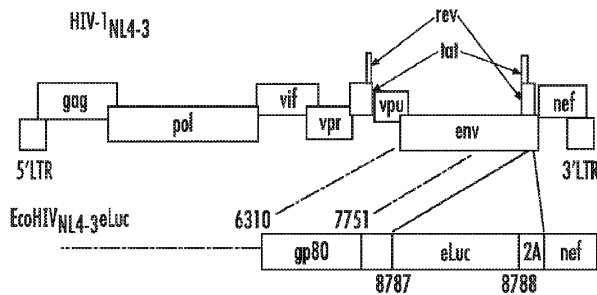
Фиг. 1В



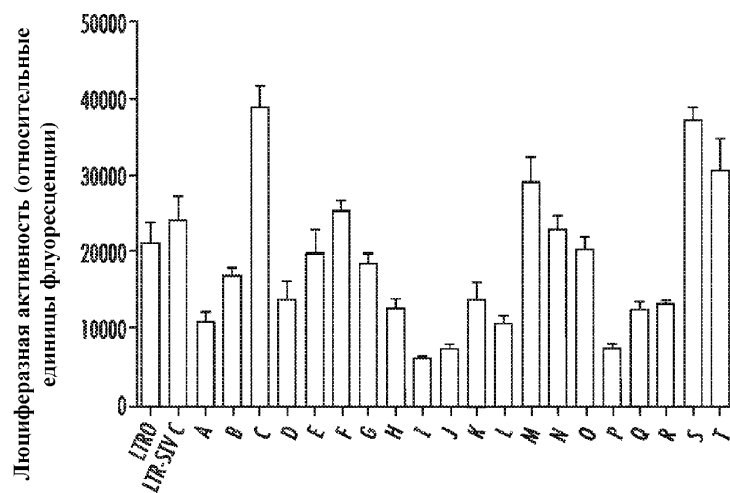
Фиг. 1С

^G
T C C T T G A T C T G T G G A T C T A C ^H C A C A C A A G G C T A C T T C C C T G A T T G G C A G A A C T A C A C A C C A G G G C C A G G G A T C A G A T A T
^E
C C A C T G A C C T T T G G A T G T G C T T C A G T T A G T A C C A G T T G A A C ^S C A G A G C A A G T A G A A G A G G C C A A T G A A G G A G A G A A C A A
^{A,I}
C A G C T T G T T A C A C C C T A T G A G C C A G C A T G G G A T G G A G G A C C ^J C G G A G G G A G A A G T A T T A G T G T G G A A G T T T G A C A G C C T C C
^{2,T} ^{B,M}
T A G C A T T T C G T C A C A T ^K G G C C G A G A G C T G C A T C C G G A G T A C T A C A A A G A C T G C T G A C A T C G A G C T T T C T A C A A G G G A C T T
^Q ^L ^F ^N
T C C G C T G G G G A C T T T C A G G G A G G T G T G G C C T G G G C G G A C T G G G G A G T G G C G A G C C C T C A G A T G C T A C A T A A G C A G C
^O ^P ^{D,3,R}

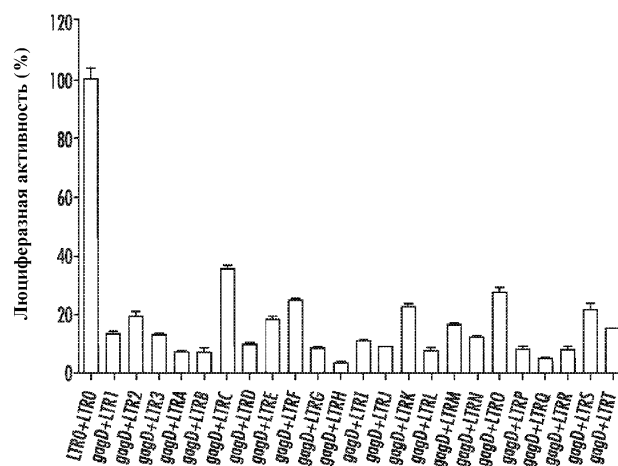
Фиг. 1D



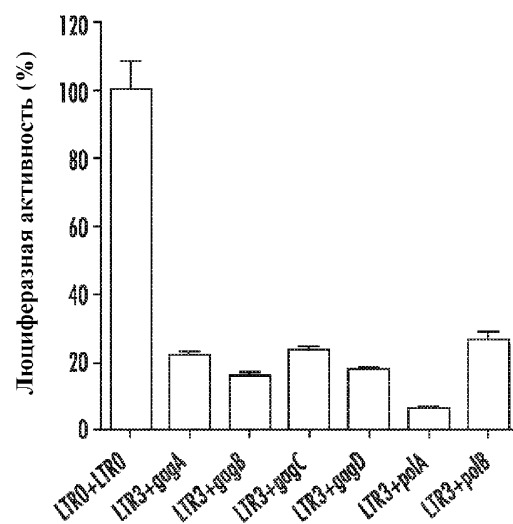
Фиг. 1Е



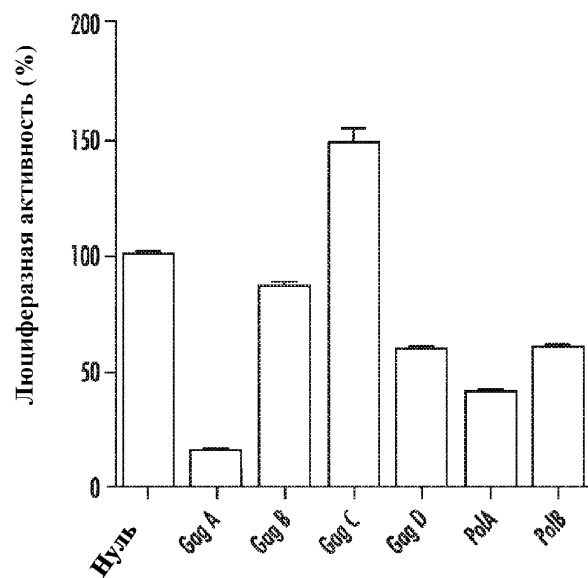
Фиг. 2А



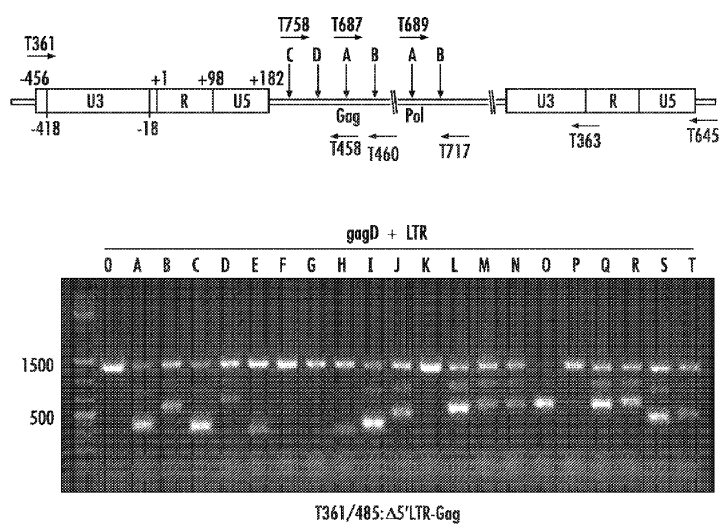
Фиг. 2В



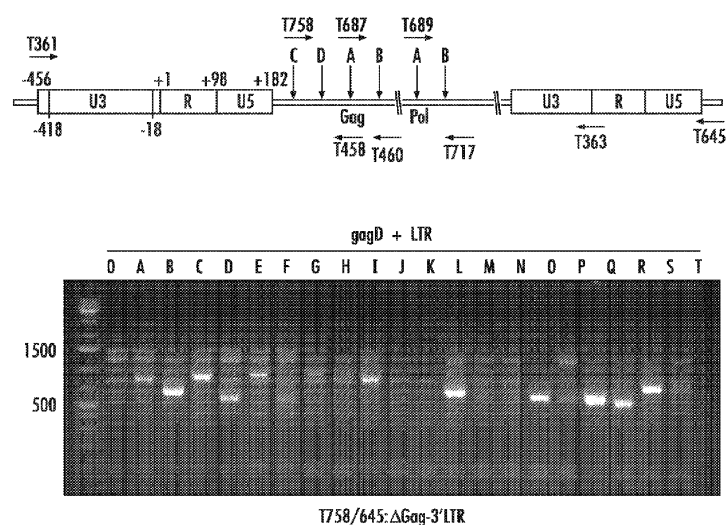
Фиг. 2С



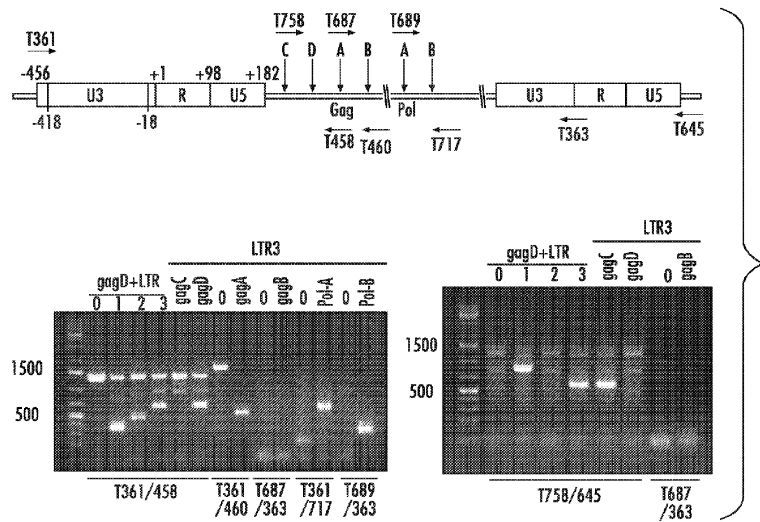
Фиг. 2D



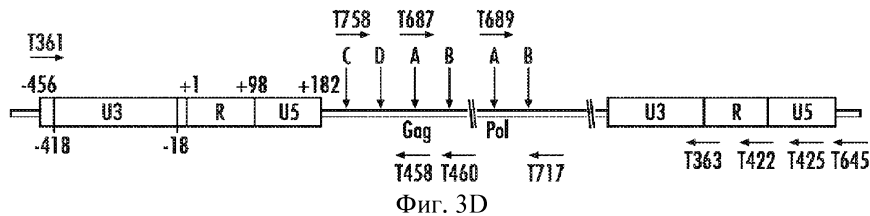
Фиг. 3A



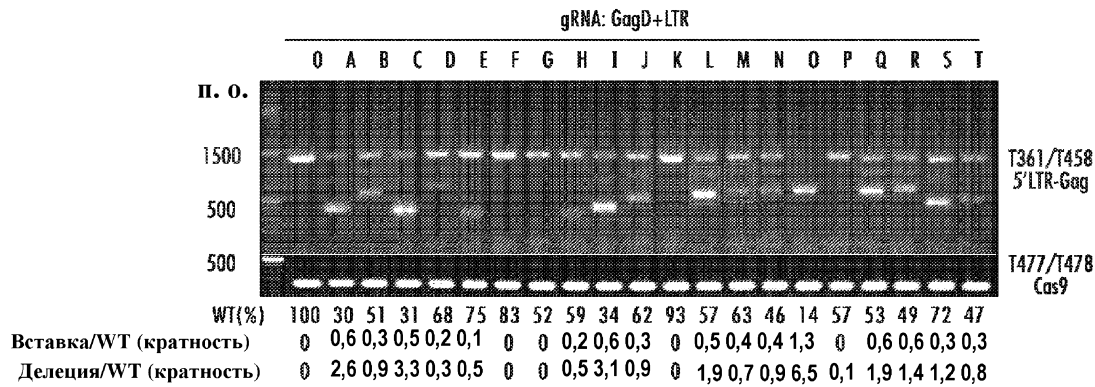
Фиг. 3B



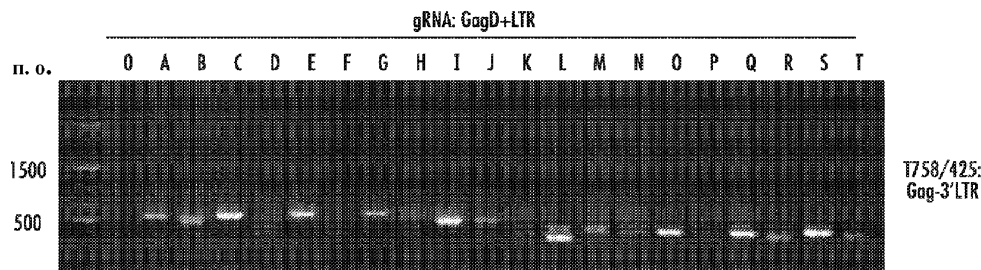
Фиг. 3С



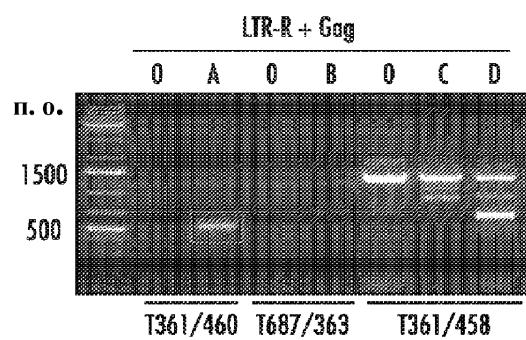
Фиг. 3D



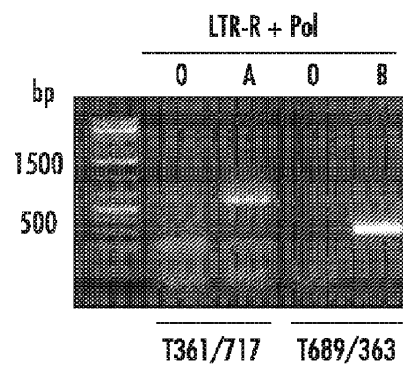
Фиг. 3Е



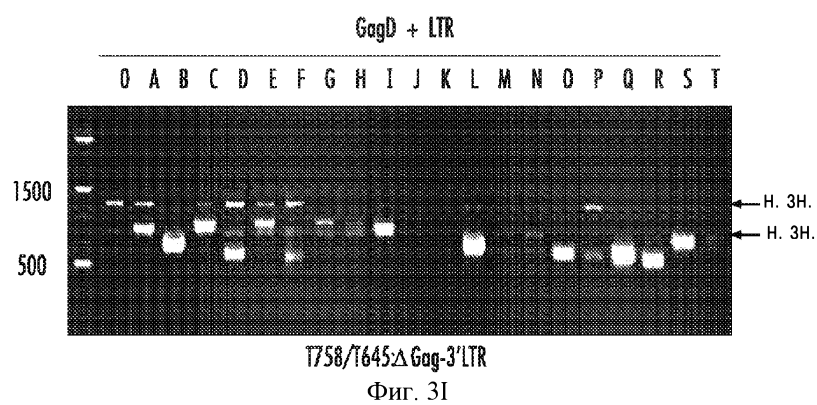
Фиг. 3F



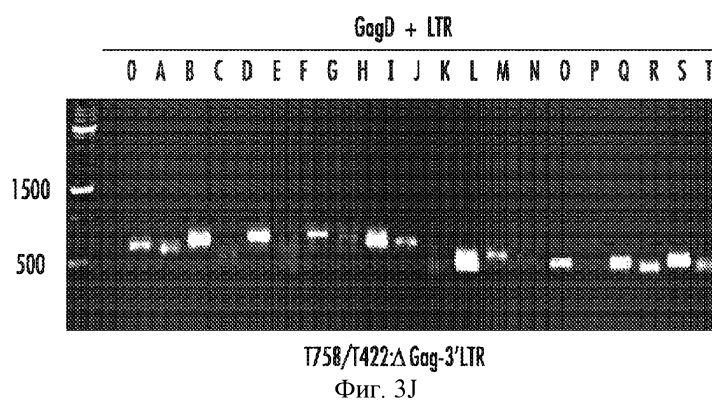
Фиг. 3G



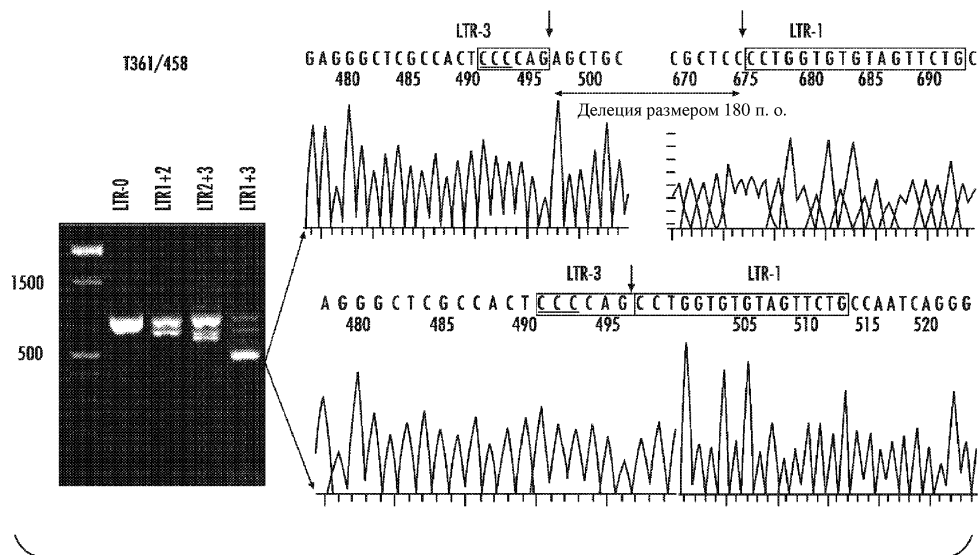
Фиг. 3H



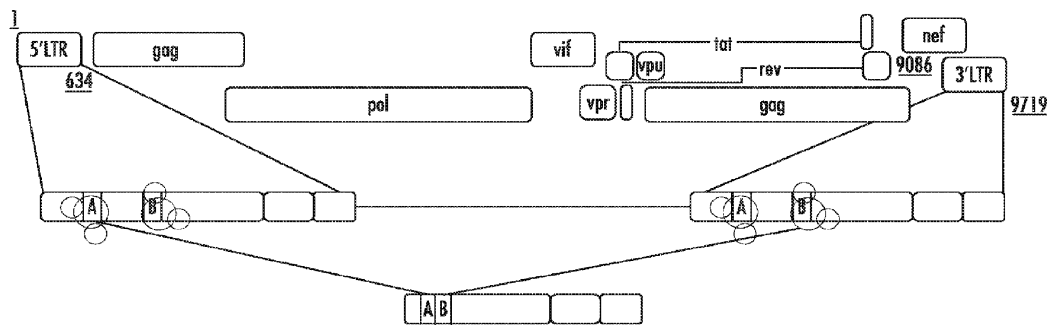
Фиг. 3I



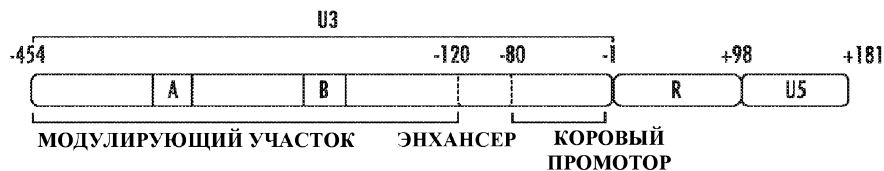
Фиг. 3J



Фиг. 4



Фиг. 5A

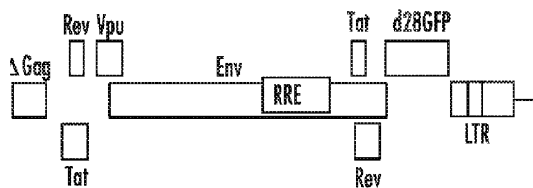


Фиг. 5B

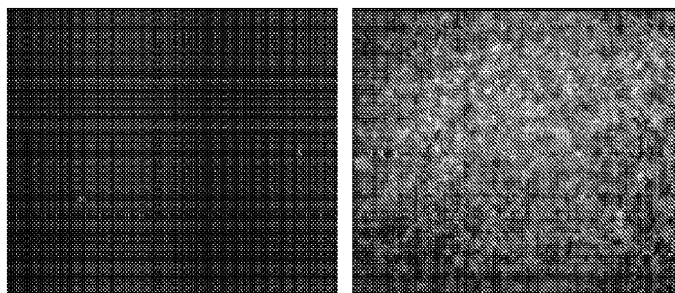
-454 TGG AAGGGCTAATTTGGTCCCAAAAAGACAAGAGATCCTTGATCTGTGGATCTACCAVAVAVAAAGGCTACTTCCCTGATTGGCAGAACTACAC
ACCAAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTTCAAGTTAGTACCAGTTGAACAGAGCAAGTAGAAGAGGCCAAATAAG
GAGAGAAGAAGAGCTTTGTTACACCTATGAGCCAGCATGGGATCCAGGACCCGGAGGGACAAGTATTAGTGTGGAAGTTTACAGCCTCCTA
GCATTTCTGCATGAGCCGAGAGCTGCATCCGAGTACTACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTCCGCTGGGGACTTTC
CAGGGAGGTGTGGCTGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTACATATAAGCAGCTGCTTTTGGCTGTACTGGTCTCTCTGG
TTAGACCAGATCTGAGCCTGGGAGCTCTTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTACCTTGAGTGTCTAAAGTAGTGT
GTGCCCGTCTGTGTGTACTCTGGTAACTAGAGATCCCTCAACCTTTTAGTCAGTGTGGAATCTCTAGCA +181

Фиг. 5C

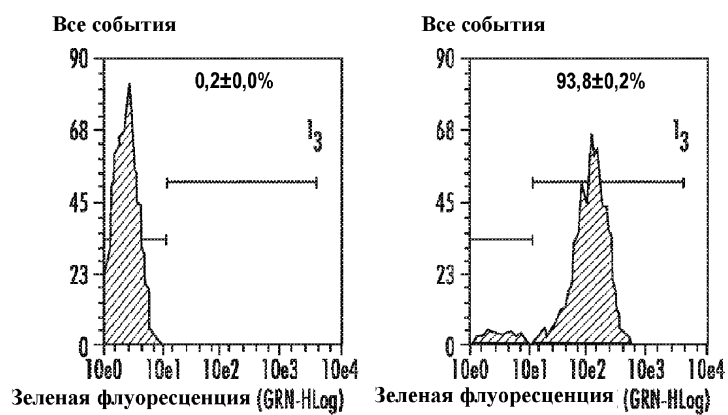
Репортерная линия клеток
JURKATS 2D10
(PEARSON 2008 J VIROLOGY)



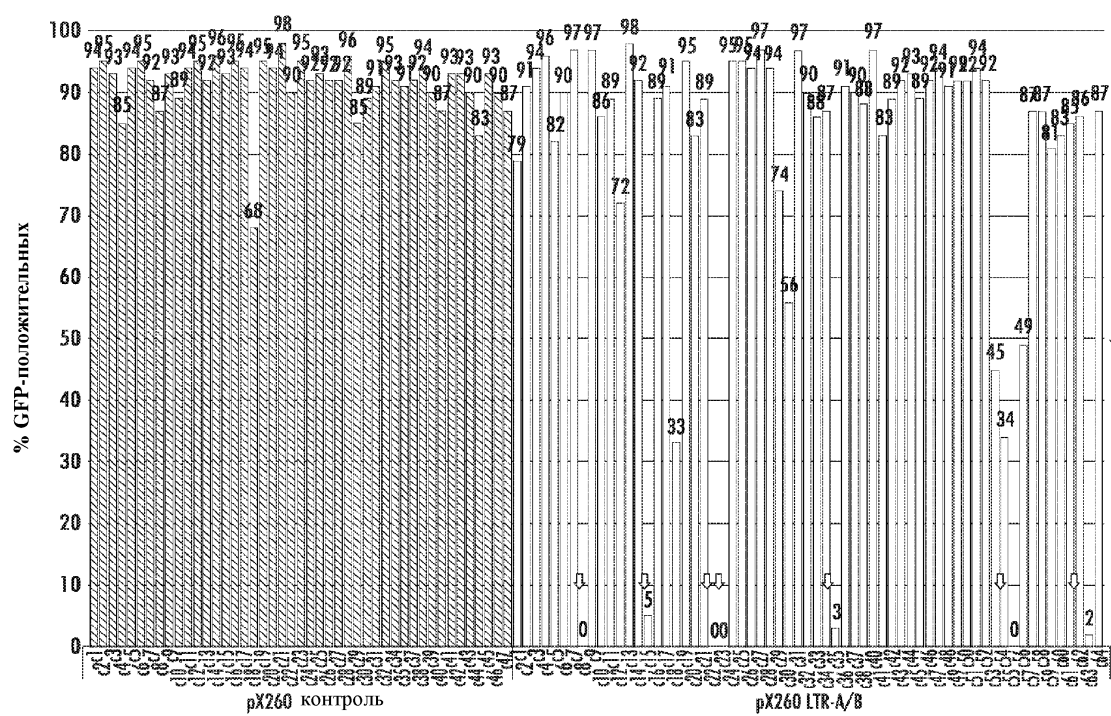
Фиг. 6A



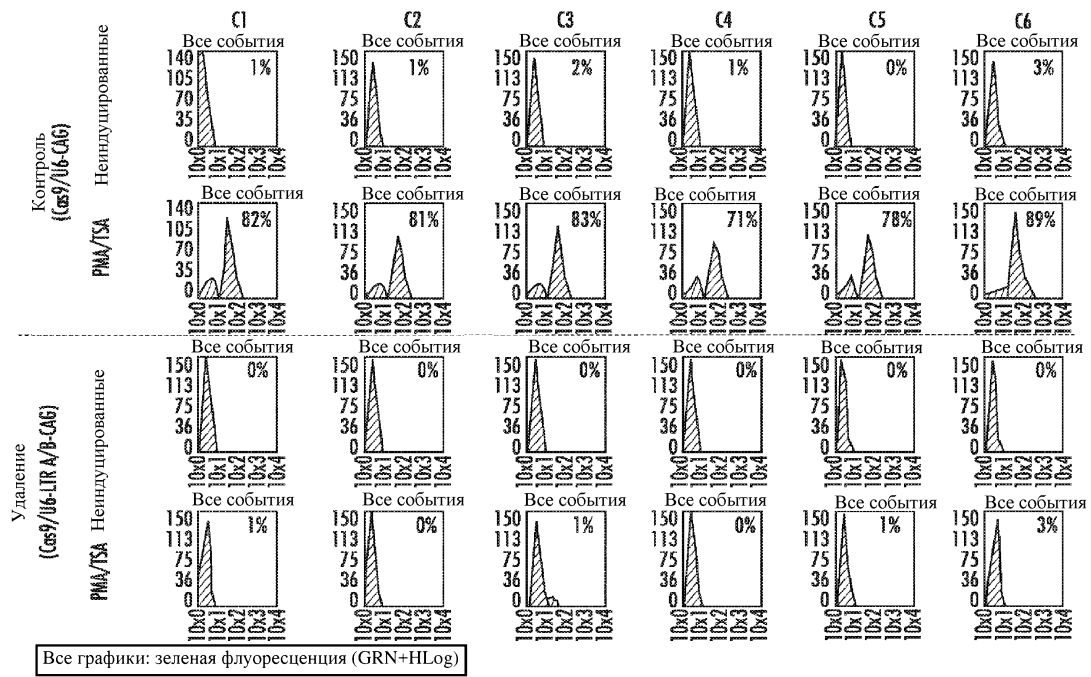
Фиг. 6B



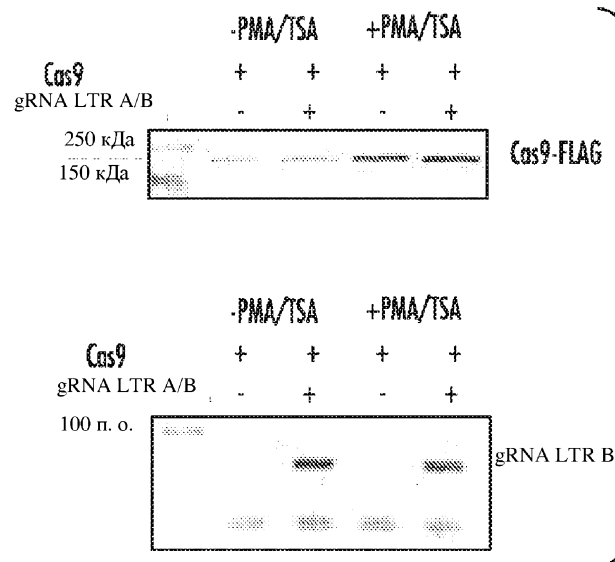
Фиг. 6C



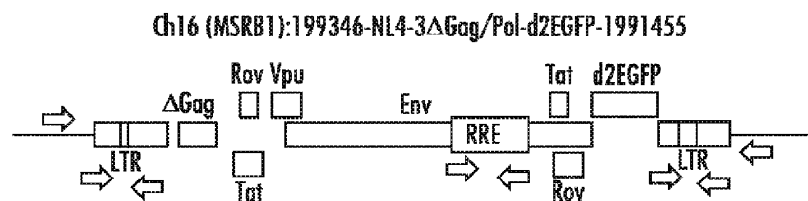
Фиг. 7A



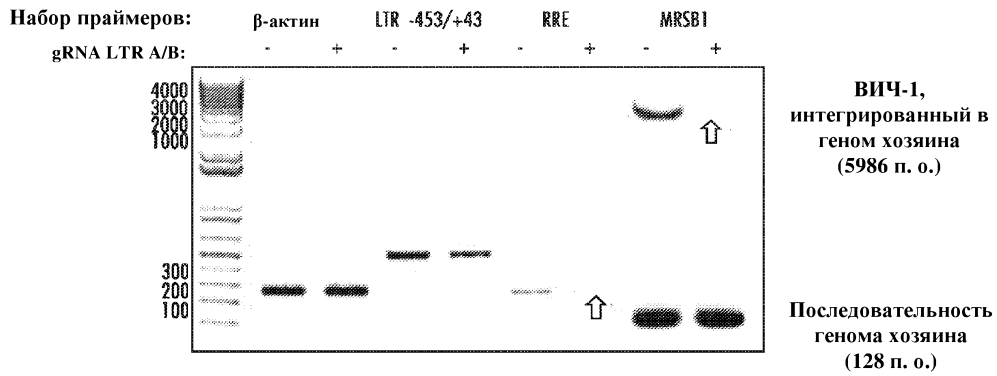
Фиг. 7B



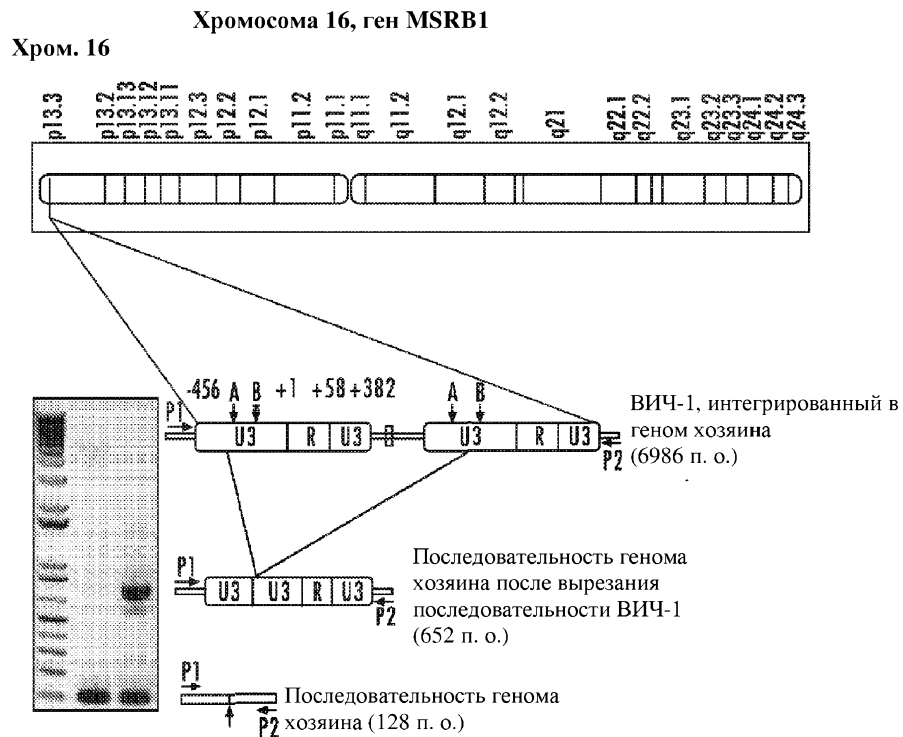
Фиг. 7C



Фиг. 8A



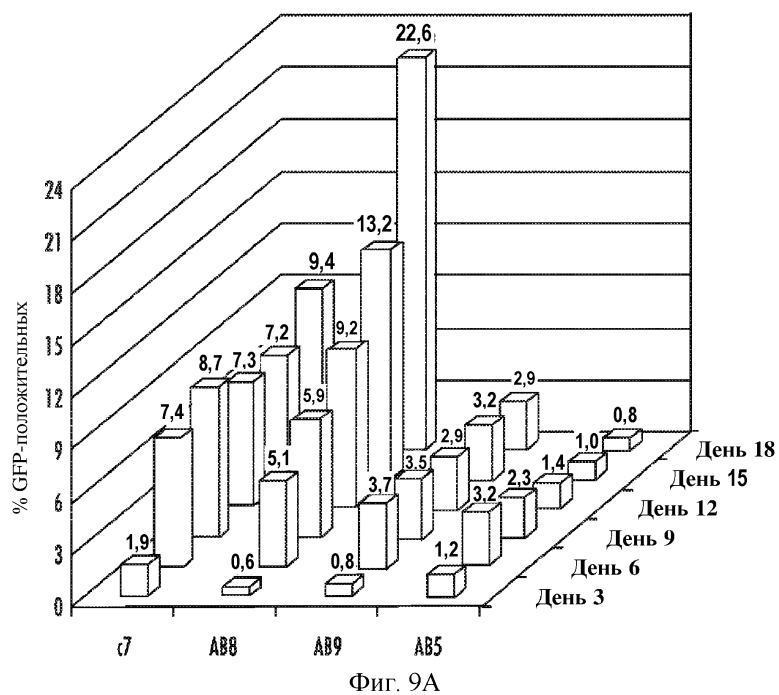
Фиг. 8В



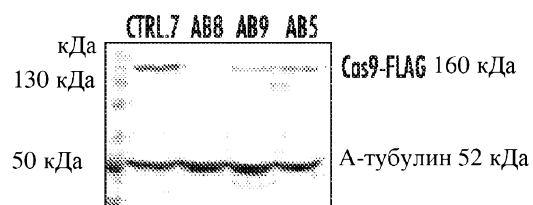
Фиг. 8С



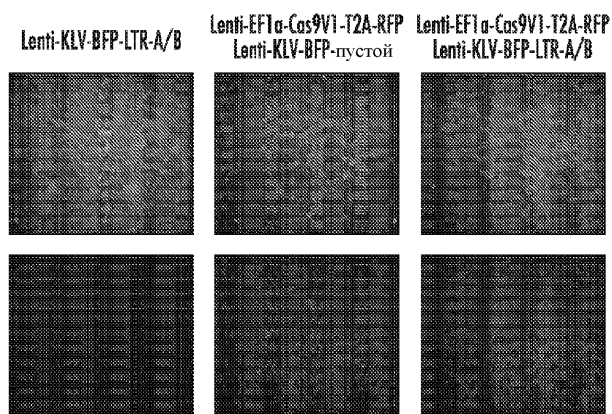
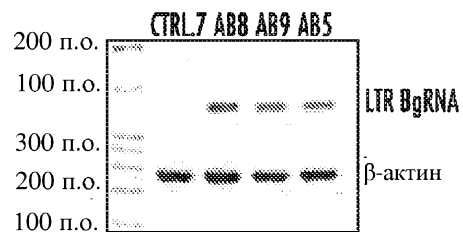
Фиг. 8D

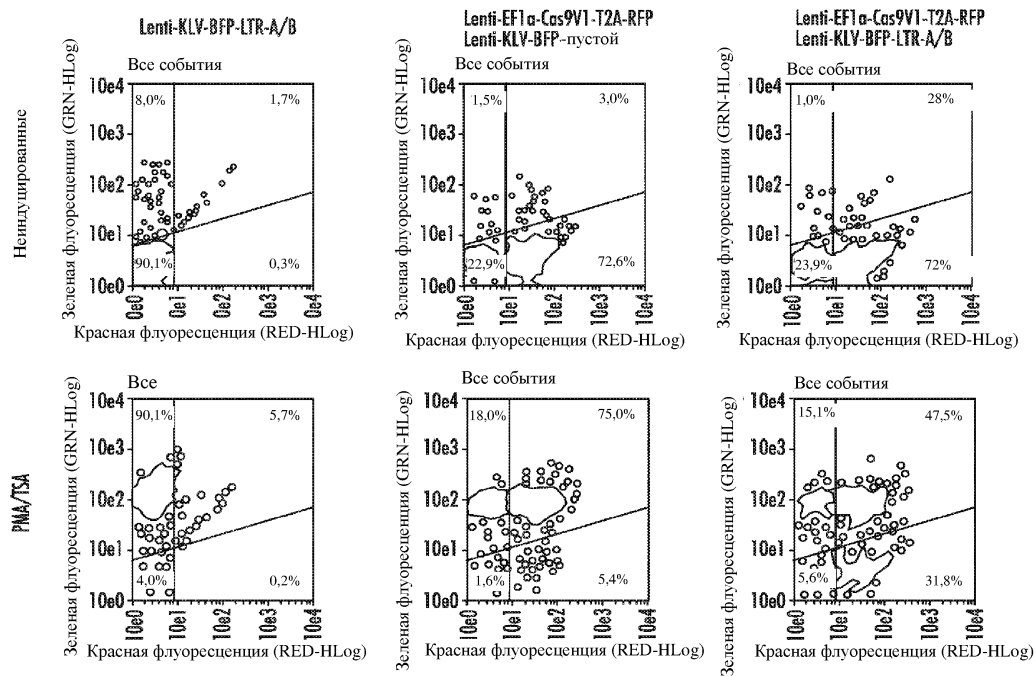


Одноклеточные клоны Jurkat 2D10



Одноклеточные клоны Jurkat 2D10

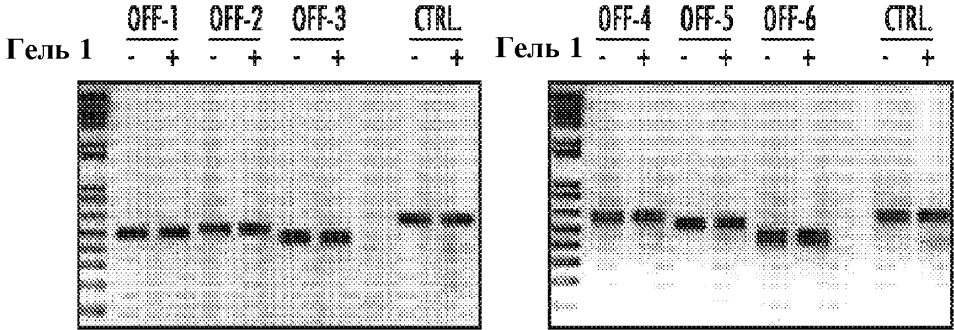




Фиг. 10В

		Длина затравочной последовательности, предшествующей PAM (п. о.)								
	Последовательность gRNA	20	19	18	17	16	15	14	13	12
LTR A	ATCAGATATCCACTGACCTTTGG	0	0	0	0	0	0	1	6	16
LTR B	CAGCAGTTCGAAGTACTCCCGG	0	0	0	0	0	0	0	0	5

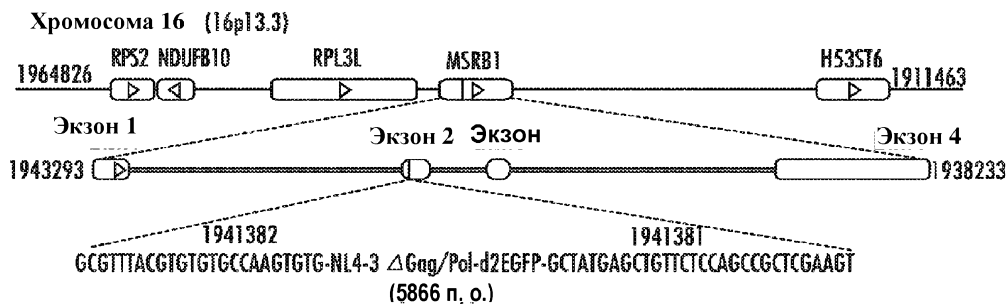
Фиг. 11А



Фиг. 11В

	Положение	Последовательность	Длина затравочной последовательности (п. о.)	Выявленные мутации по типу вставки/делеции
Нецелевой эффект-1	chr2	TCCACTGACCTTTGG	12	0
Нецелевой эффект-2	chr4	TCCACTGACCTTTGG	12	0
Нецелевой эффект-3	chr5	TCCACTGACCTTTGG	12	0
Нецелевой эффект-4	chr6	ATCCACTGACCTTTGG	13	0
Нецелевой эффект-5	chr11	TCCACTGACCTTTGG	12	0
Нецелевой эффект-6	chr16	TCCACTGACCTTTGG	12	0

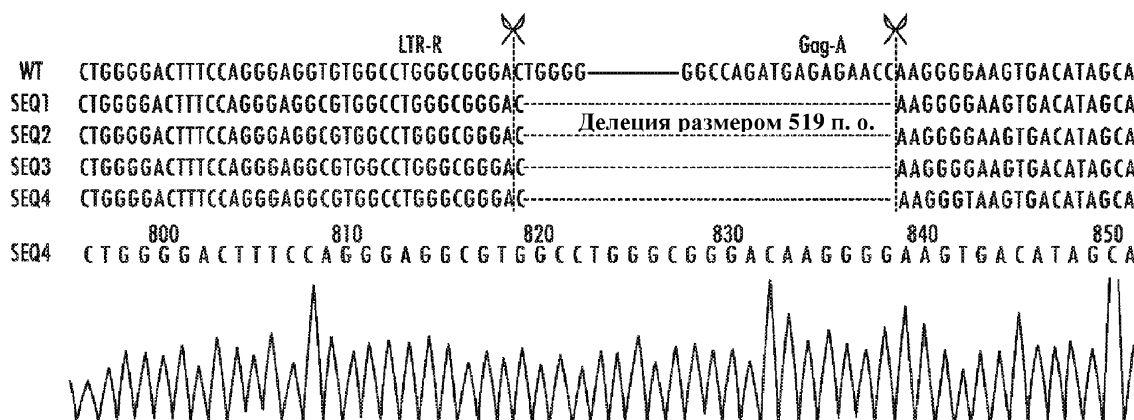
Фиг. 11С



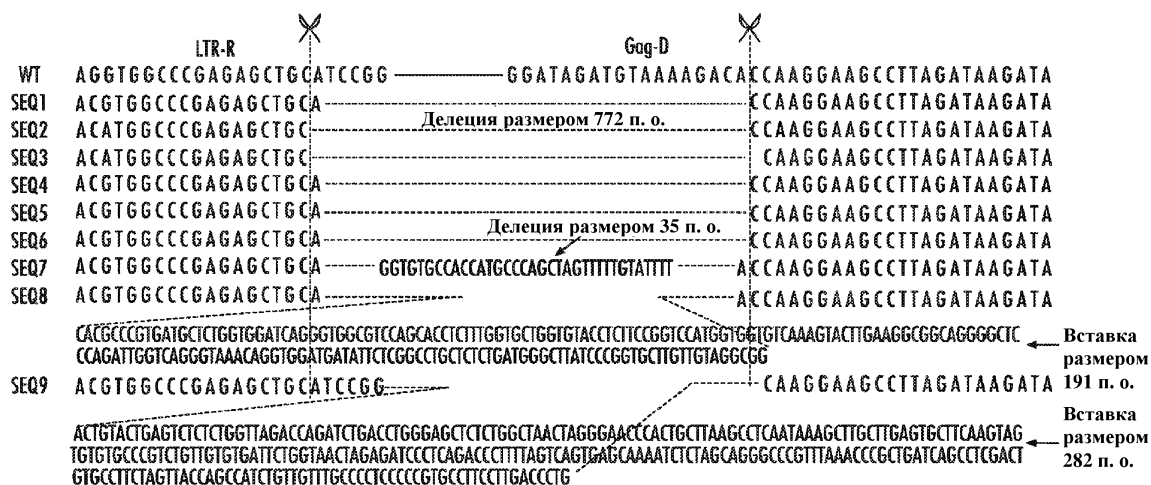
Фиг. 11D

gRT-PCR		Соотношение мишень/эталон		
Ген	Расстояние от сайта интеграции	Cas9 (КОНТРОЛЬ)	Cas9+gRNA (УДАЛЕНИЕ)	Изменение относительно контроля
RPS2	+23444	3,6E-01	4,3E-01	+0,21
NDUFB10	+18134	2,9E-02	3,4E-02	+0,17
RPL3L	+13296	3,3E-06	4,1E-0,6	+0,24
MSRB1	0	1,2E-03	1,4E-0,3	+0,21
HS35T6	-23152	4,8E-07	1,7E-07	-0,65

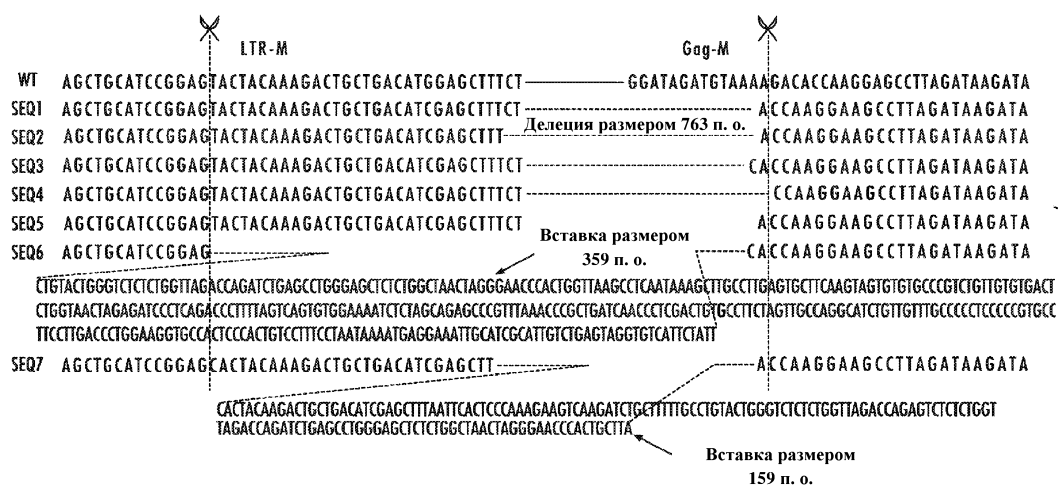
Фиг. 11Е



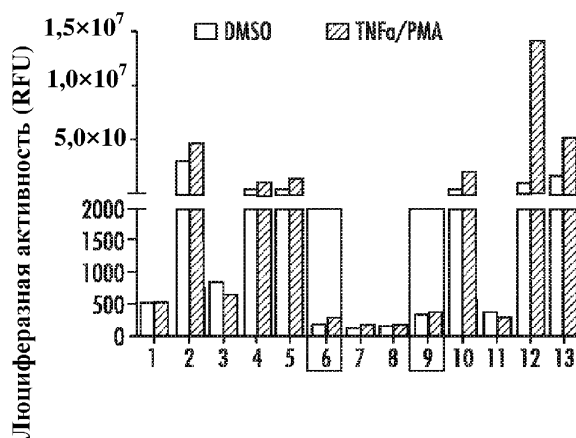
Фиг. 12А



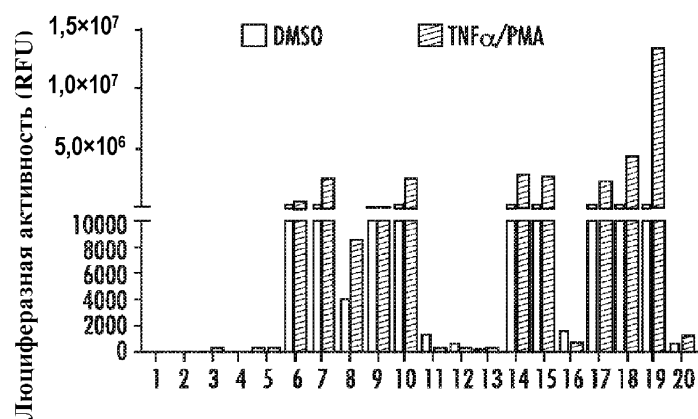
Фиг. 12В



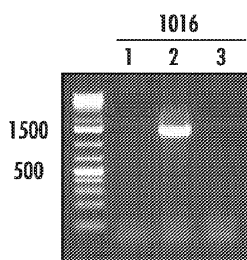
Фиг. 12С



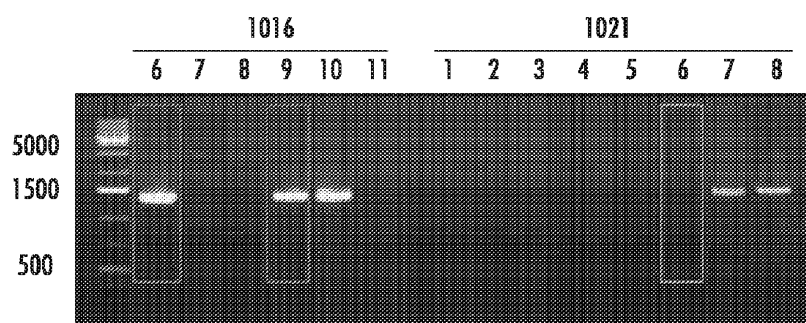
Фиг. 13А



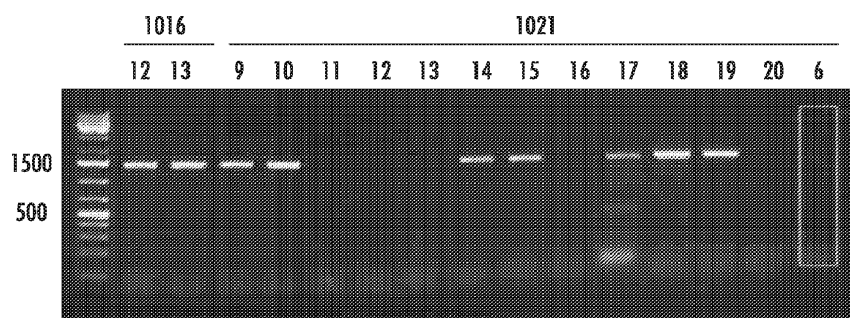
Фиг. 13В



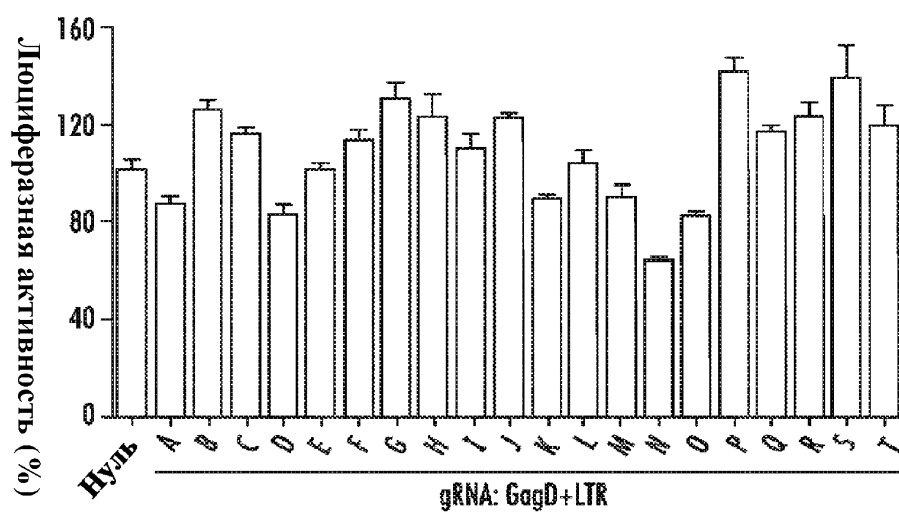
Фиг. 13С



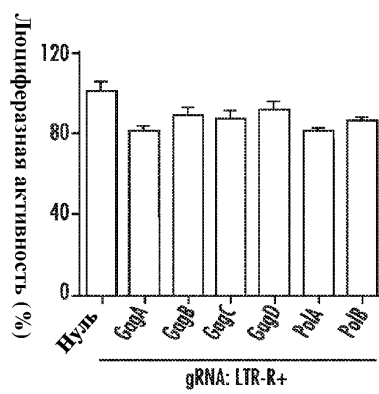
Фиг. 13D



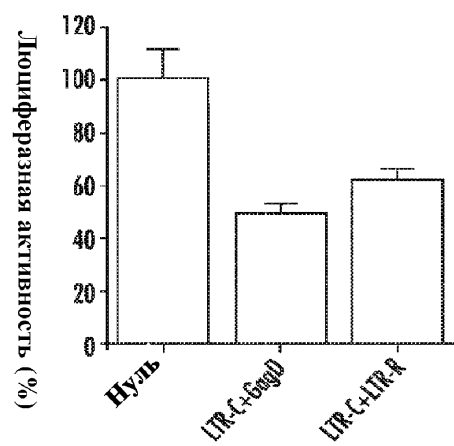
Фиг. 13E



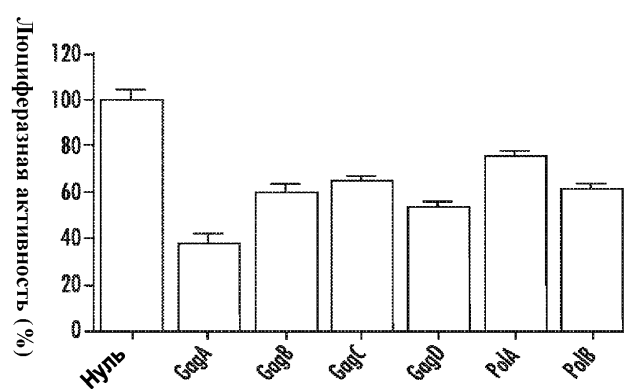
Фиг. 14A



Фиг. 14B



Фиг. 14С



Фиг. 14D



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2