

(19)



Евразийское
патентное
ведомство

(11) 037850

(13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2021.05.27

(51) Int. Cl. A61K 48/00 (2006.01)
A61P 31/12 (2006.01)
A61P 31/00 (2006.01)

(21) Номер заявки
201600226

(22) Дата подачи заявки
2014.08.29

(54) СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ РНК-НАПРАВЛЕННОГО ЛЕЧЕНИЯ ВИЧ-ИНФЕКЦИИ

(31) 61/871,626; 62/018,441; 62/026,103

(56) EBINA H. et al. "Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus", Scientific Reports, 26 August 2013, Vol. 3, Article No. 2510, internal pages 1-7, see abstract and pages 1-4.

(32) 2013.08.29; 2014.06.27; 2014.07.18

RAMALINGAM S. et al. "A CRISPR way to engineer the human genome", Genome Biology, 26 February 2013, Vol. 14, No. 2, Article No. 107, internal pages 1-4, see abstract and pages 3-4.

(33) US

RICHTER H. et al. "Exploiting CRISPR/Cas: Interference mechanisms and applications", International Journal of Molecular Sciences, 12 July 2013, Vol. 14, pp. 14518-14531, see abstract and pages 14520-14523.

(43) 2016.09.30

WO-A1-2013098244
US-A1-20030087817

(86) PCT/US2014/053441

(87) WO 2015/031775 2015.03.05

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

(72) Изобретатель:
Халили Камел, Ху Вэнъхуэй (US)

(74) Представитель:
Кузнецова С.А. (RU)

(57) Изобретение представляет способы и композиции для лечения инфекции, вызванной вирусом иммунодефицита. Композиции включают последовательности выделенных нуклеиновых кислот, содержащие CRISPR-ассоциированную эндонуклеазу и направляющую РНК, где направляющая РНК комплементарна целевой последовательности в вирусе иммунодефицита человека.

B1

037850

037850
B1

Перекрестные ссылки на родственные заявки

Данная заявка испрашивает приоритет на основании дат подачи предварительной заявки на патент США № 61/871626, которая была подана 29 августа 2013 года; предварительной заявки на патент США № 62/018441, которая была подана 27 июня 2014 года; и предварительной заявки на патент США № 62/026103, которая была подана 18 июля 2014 года. Для целей любой заявки на патент США, которая может испрашивать приоритет предварительной заявки на патент США № 61/871626, предварительной заявки на патент США № 62/018441 и предварительной заявки на патент США № 62/026103, содержание этих ранее поданных заявок настоящим включено при помощи ссылки в полном объеме.

Перечень последовательностей

В настоящую заявку включен перечень последовательностей, который предоставили в электронном виде в формате ASCII и, таким образом, включен при помощи ссылки в полном объеме. Указанная ASCII-копия, созданная 26 августа 2014 года, названа F5129-00031_SL.txt и ее размер составляет 74547 байт.

Заявление в отношении финансируемого из федерального бюджета исследования

Изобретение было разработано при поддержке правительства США в рамках грантов с номерами R01MH093271, R01NS087971 и P30MH092177, выданных Национальными институтами здравоохранения. Правительство США может обладать определенными правами на настоящее изобретение.

Область изобретения

Изобретение относится к композициям, которые специфически расщепляют целевые последовательности в ретровирусах, например вирусе иммунодефицита человека (HIV). Такие композиции, которые могут включать нуклеиновые кислоты, кодирующие эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположенные группами (CRISPR), и последовательность направляющей РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, можно вводить субъекту с ВИЧ-инфекцией или подверженому риску заражения ВИЧ-инфекцией.

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 постоянной интеграции ВИЧ-1 в геном организма-хозяина. Современная терапия (высокоактивная антиретровирусная терапия или HAART), предназначенная для контроля ВИЧ-1-инфекции и задержки развития СПИДа, существенно снижает репликацию вируса в клетках, которые поддерживают ВИЧ-1-инфекцию, и снижает до минимального уровня вирецию в плазме. Но HAART не способна подавлять экспрессию и репликацию вирусного генома, происходящих на низком уровне в тканях, и не способна целенаправленно воздействовать на латентно инфицированные клетки, например покоящиеся Т-клетки памяти, макрофаги головного мозга, микроглио и астроциты, клетки лимфоидной ткани кишечника, которые служат в качестве резервуара для ВИЧ-1. Персистирующая ВИЧ-1-инфекция также связана с сопутствующими патологическими процессами, в том числе заболеваниями сердечно-сосудистой системы и почек, остеопенией и неврологическими расстройствами. Все еще существует потребность в лечебных терапевтических стратегиях, которые нацелены на резервуары персистирующих вирусов.

Краткое описание изобретения

В данном документе предусмотрены композиции и способы, относящиеся к лечению и предупреждению ретровирусных инфекций. Ретровирусом может быть лентивирус, например вирус иммунодефицита человека, вирус иммунодефицита обезьян, вирус иммунодефицита кошачьих или вирус иммунодефицита крупного рогатого скота. Вирусом иммунодефицита человека может быть ВИЧ-1 или ВИЧ-2. В одном варианте осуществления композиции включают последовательность нуклеиновой кислоты, содержащую последовательность, кодирующую CRISPR-ассоциированную эндонуклеазу и одну или более направляющих РНК, где направляющая РНК комплементарна целевой последовательности в вирусе иммунодефицита человека. В некоторых вариантах осуществления нуклеиновая кислота входит в состав вектора экспрессии. В одном варианте осуществления композиции включают полипептид CRISPR-ассоциированную эндонуклеазу и одну или более направляющих РНК, где направляющая РНК комплементарна целевой последовательности в вирусе иммунодефицита человека. Также предусмотрены фармацевтические композиции, которые включают нуклеиновые кислоты, векторы экспрессии или полипептиды, раскрытые в данном документе. Также в данном документе предусмотрены способы лечения субъекта с инфекцией, вызванной вирусом иммунодефицита человека, или подверженного риску приобретения инфекции, вызванной вирусом иммунодефицита человека, где способ лечения включает введение субъекту терапевтически эффективного количества композиции, содержащей вектор, кодирующий CRISPR-ассоциированную эндонуклеазу и одну или более направляющих РНК, где направляющая РНК комплементарна целевой последовательности в вирусе иммунодефицита человека. Также предусмотрены способы инактивации ретровируса в клетке человека путем воздействия на клетку с помощью композиции, содержащей выделенную нуклеиновую кислоту, кодирующую комплекс редактирования генов, содержащий CRISPR-ассоциированную эндонуклеазу и одну или более направляющих РНК, где направляющая РНК комплементарна целевой последовательности нуклеиновой кислоты в ретровирусе. Комплекс редактирования генов вводит одну или более мутаций в провирусную ДНК. В некоторых вариантах осуществления мутации могут включать делецию, которая может охватывать всю или практически

всю последовательность провирусной ДНК. В другом аспекте также предусмотрен набор, содержащий отмеренное количество композиций, раскрытых в данном документе.

Подробности одного или более вариантов осуществления согласно настоящему изобретению изложены при помощи сопутствующих графических материалов и приведенного ниже описания. Другие признаки, объекты и преимущества настоящего изобретения будут очевидными из описания и графических материалов, а также из формулы изобретения.

Краткое описание графических материалов

На фиг. 1 показано, что Cas9/LTR-gRNA подавляет продуцирование репортерного вируса ВИЧ-1 в клетках микроглии CHME5, латентно инфицированных ВИЧ-1. (А) На презентативной диаграмме гейтирования проточной цитометрии в отношении EGFP видно резкое снижение TSA-индукции реактивации латентного репортерного вируса pNL4-3-ΔGag-d2EGFP под действием стабильно экспрессируемых Cas9 и LTR-A или -B в сравнении с пустым вектором экспрессии gRNA под управлением U6 (U6-CAG). (В) При анализе SURVEYOR с нуклеазой Cel-I ПЦР-продукта (от -453 до +43 в пределах LTR) из отобранных стабильных клонов, экспрессирующих LTR-A или -B, видны неожиданные паттерны мутаций по типу вставки/делеции (стрелки). (С, Д) При анализе ПЦР-фрагментов видна точная делеция участка в 190 п.о. между сайтами разрезания для LTR A и B (указатель стрелки и стрелка на D), при этом остается фрагмент в 306 п.о. (стрелка на С), подтвержденный результатами TA-клонирования и секвенирования. На фиг. 1D раскрыты SEQ ID NO: 1-3, соответственно, в порядке появления. (Е-Г) При субклонировании стабильных клонов LTR-A/B видна полная потеря реактивации репортера, определенная при помощи проточной цитометрии в отношении EGFP (Е), а также элиминация провирусного генома pNL4-3-ΔGag-d2EGFP, обнаруженная при помощи стандартной ПЦР-амплификации (F) и ПЦР-амплификации в режиме реального времени (G) геномной ДНК в отношении EGFP и Rev-чувствительного элемента ВИЧ-1 (RRE); β-актин является контролем очистки и загрузки ДНК. (Н) При ПЦР-генотипировании субклонов LTR-A/B (№ 8, 13) с использованием праймеров для амплификации ДНК-фрагмента, охватывающего участки U3/R/U5 LTR ВИЧ-1 (от -411 до +129), видны вставки/делеции (а, делеция; с, вставка) и "интактный" или комбинированный LTR (б).

На фиг. 2 показано, что Cas9/LTR-gRNA эффективно удаляет латентный вирус ВИЧ-1 из моноцитов U1. (А) Справа, на диаграмме показано вырезание полного генома ВИЧ-1 из хромосомы Xp11.4. Сайты интеграции ВИЧ-1 идентифицировали при помощи набора для ПЦР Genome-Walker link. Слева, при анализе длин ПЦР-ампликонов при помощи пары праймеров (P1/P2), нацеленной на последовательность, фланкирующую сайт интеграции в хромосоме X, видна элиминация полного генома ВИЧ-1 (9709 п.о.), при этом остаются два фрагмента (833 п.о. и 670 п.о.). (В) При TA-клонировании и секвенировании фрагмента LTR (833 п.о.) видны геномная последовательность организма-хозяина (строчные буквы, 226 п.о.) и частичные последовательности (634-27=607 п.о.) 5'-LTR (подчеркнута пунктирной линией) и 3'-LTR (первый подчеркнутый сегмент) с делецией в 27 п.о. около сайта, на который нацелен LTR-A (второй подчеркнутый сегмент). В нижней части две аллели со вставками/делециями идентифицировали из 15 секвенированных клональных ампликонов. Фрагмент в 670 п.о. состоит из последовательности организма хозяина (226 п.о.) и оставшейся последовательности LTR (634-190=444 п.о.) после вырезания 190 п.о. при помощи одновременного разрезания в целевых сайтах для LTR-A- и -B. При помощи подчеркнутых и выделенных последовательностей показаны целевой сайт для LTR-A gRNA и PAM. На фиг. 2В раскрыты SEQ ID NO: 4-13, соответственно, в порядке появления. (С) При функциональном анализе индуцированного LTR-A/B удаления генома ВИЧ-1 видно существенное блокирование TSA/PMA-индукции реактивации высвобождения вирионов p24. Клетки U1 трансфицировали с помощью pX260-LTR-A, -B или -A/B. После 2-недельного отбора с пуромицином клетки обрабатывали с помощью TSA (250 нМ)/PMA в течение 2 дней перед проведением ELISA Gag p24.

На фиг. 3 показано, что стабильная экспрессия Cas9 и LTR-A/B вакцинирует клетки TZM-blI против нового инфицирования вирусом ВИЧ-1. (А) Иммуноцитохимия (ICC) и вестерн-блоттинг (WB) анализы с антителом к Flag подтвердили экспрессию Flag-Cas9 в стабильных клонах TZM-blI, отбираемых с пуромицином (2 мкг/мл) в течение 2 недель. (Б) При ПЦР-генотипировании стабильных клонов Cas9/LTR-A/B (c1-c7) видна тесная корреляция вырезания LTR с репрессией активации репортера LTR-люцифераза. Кратные изменения представляют TSA/PMA-индукции уровни по сравнению с соответствующими уровнями без индукции. (С) Стабильно экспрессирующие Cas9/LTR-A/B клетки (c4) инфицировали псевдотипированным лентивирусом pNL4-3-Nef-EGFP при указанной множественности заражения (MOI) и эффективность инфицирования измеряли при помощи проточной цитометрии в отношении EGFP через 2 дня после инфицирования. (Д) На презентативных фазово-контрастных/флуоресцентных микрофотографиях показано, что стабильные клетки с LTR-A/B, но не контрольные (U6-CAG; черный), устойчивы к новому инфицированию (правая панель) репортерным вирусом ВИЧ-1 pNL4-3-ΔE-EGFP (серый).

На фиг. 4 показаны нецелевые эффекты Cas9/LTR-A/B в отношении генома человека. (А) В анализе SURVEYOR показано отсутствие мутаций по типу вставок/делеций в предсказанных/потенциальных нецелевых участках в клетках TZM-blI и U1. Целевой участок для LTR-A (А) использовали в качестве

положительного контроля, а пустой вектор U6-CAG (U6) - в качестве отрицательного контроля. (B-D) При полногеномном секвенировании стабильного субклона TZM-bl с LTR-A/B видно количество определенных вставок/делеций в контрольном образце U6-CAG и образце LTR-A/B (B), подробная информация о 10 определенных вставках/делециях около целевых сайтов для gRNA в обоих образцах (C), и распределение определенных нецелевых вставок/делеций (D). На фиг. 4C раскрыты SEQ ID NO: 14-15, соответственно, в порядке появления.

На фиг. 5 показана последовательность U3 LTR интегрированного репортера лентивирусный LTR-люцифераза светлячка, идентифицированная при помощи TA-клонирования и секвенирования ПЦР-продукта (от -411 до -10) из геномной ДНК клеток TZM-bl человека. Выделены последовательности протоспайсеров и PAM (NGG) для 4 gRNA (LTR-A - LTR-D) и предсказанные сайты связывания указанных факторов транскрипции. Точные сайты расщепления отмечены ножницами. +1 обозначает сайт инициации транскрипции. На фиг. 5 раскрыта SEQ ID NO: 16.

На фиг. 6 показано, что LTR-C и LTR-D существенно подавляют TSA-индуцированную реактивацию латентного вируса pNL4-3-ΔGag-d2EGFP в клетках микроглии CHME5. (A) На диаграмме схематически показан вектор pNL4-3-ΔGag-d2EGFP, содержащий Tat, Rev, Env, Vpu и Nef с репортерным геном d2EGFP. (B) При анализе SURVEYOR видны мутации по типу вставок/делеций в целевом геноме LTR у клеток, трансфицированных с помощью Cas9/LTR-D, но не Cas9/LTR-C. (C) На репрезентативной диаграмме гейтирования проточной цитометрии в отношении EGFP видно резкое снижение TSA-индуцированной реактивации латентных репортерных вирусов pNL4-3-ΔGag-d2EGFP под действием стабильной экспрессии Cas9/LTR-C или LTR-D по сравнению с пустым вектором экспрессии gRNA под управлением U6 (U6-CAG).

На фиг. 7 показано, что как LTR-C, так и LTR-D индуцировали мутации по типу вставок/делеций и значительно снижали конститутивную и TSA/PMA-индуцированную люциферазную активность в клетках TZM-bl со стабильно встроенным репортерным геном LTR ВИЧ-1-люцифераза светлячка. (A) При функциональном анализе люциферазного репортера видно значительное снижение реактивации LTR, обусловленное LTR-C, LTR-D или обоими. (B) При анализе SURVEYOR видна мутация по типу вставки/делеции в ДНК LTR (от -453 до +43), индуцированная LTR-C и LTR-D (верхняя стрелка). Комбинация LTR-C и LTR-D образует фрагмент в 194 п.о. (нижняя стрелка), полученный в результате делеции участка в 302 п.о. между LTR-C и LTR-D. (C, D) При помощи секвенирования по Сэнгеру 30 клонов подтверждена эффективность образования вставок/делеций на уровне 23% для LTR-C и 13% для LTR-D, а на иллюстративных хроматограммах видны вставки/делеции. На фиг. 7C раскрыты SEQ ID NO: 17-25, соответственно, в порядке появления. На фиг. 7D раскрыты SEQ ID NO: 26-30, соответственно, в порядке появления. (E) По результатам ПЦР-анализа полиморфизма длин рестрикционных фрагментов (RFLP) с использованием BsaI для разрезания 5 сайтов (96, 102, 372, 386, 482) ПЦР-продукта, охватывающего от -453 до +43 LTR, видны две основные полосы (96 п.о. и 270 п.о.) в контрольном образце U6-CAG, а также дополнительная полоса 372 п.о. (верхняя стрелка) после индуцированной LTR-C мутации по типу вставки/делеции в сайтах 96/102, полоса 290 п.о. (средняя стрелка) после индуцированных LTR-D мутаций в сайте 372 и фрагмент в 180 п.о. (нижняя стрелка) после индуцированного LTR-C/D вырезания. (F) На иллюстративных хроматограммах показана делеция фрагмента в 302 п.о. между LTR-C и LTR-D (вверху) и дополнительная делеция в 17 п.о. (внизу). Красные стрелки указывают на сайты соединения. *P<0,05 указывает на значимое снижение активации люциферазы, опосредованное LTR-C или LTR-D, по сравнению с контролем U6-CAG. На фиг. 7F раскрыты SEQ ID NO: 31-32, соответственно, в порядке появления.

На фиг. 8 показано TA-клонирование и секвенирование по Сэнгеру ПЦР-продуктов из субклонов CHME5 LTR-A/B и пустого контроля U6-CAG с использованием праймеров, охватывающих участки U3/R/U5 LTR ВИЧ-1 (от -411 до +129). (A) Возможная комбинация LTR-A и LTR-B разрезает как 5'-, так и 3'-LTR, образуя потенциальные фрагменты а-с, как обозначено. (B) По результатам сравнения фрагмента а (351 п.о.) с использованием BLAST видна делеция в 190 п.о. между сайтами разрезания для LTR-A и LTR-B. (C) По результатам сравнения фрагмента с (682 п.о.) с использованием BLAST видна вставка в 175 п.о. в сайте расщепления для LTR-A и делеция в 27 п.о. в сайте расщепления для LTR-B. На фиг. 8C раскрыты SEQ ID NO: 33-34, соответственно, в порядке появления.

На фиг. 9 показано, что Cas9/LTR-gRNA эффективно удаляет латентный вирус ВИЧ-1 из моноцитов U1. (A) При секвенировании по Сэнгеру фрагмента в 1,1 т.п.о., полученного в результате ПЦР длинных фрагментов с использованием пары праймеров (T492/T493), нацеленной на последовательность, flankирующую сайт интеграции в хромосоме 2 (строчные буквы, 467 п.о.), видна элиминация полного генома ВИЧ-1 (9709 п.о.), при этом остаются комбинированный 5'-LTR (подчеркнуто пунктирной линией) и 3'-LTR со вставкой в 6 п.о. (обозначена рамкой) точно в третьем нуклеотиде от сайта PAM (TGG), на который нацелен LTR-A (подчеркнут), и делеция в 4 п.о. (пппп). На фиг. 9A раскрыта SEQ ID NO: 35. (B) На репрезентативной фотографии геля электрофореза ДНК видно специфическое удаление генома ВИЧ-1. NS, неспецифическая полоса. (C, D) При количественном ПЦР-анализе с использованием пары праймеров, нацеленной на ген Gag (T457/T458), показана 85% эффективность удаления полного генома ВИЧ-1 в

клетках U1, экспрессирующих Cas9/LTR-A/B. Клетки U1 трансфицировали пустым вектором pX260 (U6-CAG) или векторами, кодирующими LTR-A/B. После 2-недельного отбора с пуромицином геномные ДНК клеток использовали для абсолютного количественного qPCR-анализа с использованием меченой pNL4-3-ΔE-EGFP геномной ДНК человека в качестве стандарта. **P<0,01 указывает на значимое снижение по сравнению с контролем U6-CAG.

На фиг. 10 показано, что Cas9/LTR-gRNA эффективно удаляетprovirus ВИЧ-1 в латентно инфицированных Т-клетках J-Lat. (А) При функциональном анализе посредством проточной цитометрии в отношении EGFP выявлено приблизительно 50% снижение PMA- и TNF α -индуцированной реактивации репортерных вирусов с EGFP. (В) При анализе SURVEYOR видны мутации по типу вставок/длекций (стрелка) в целевом геноме LTR у трансфицированных Cas9/LTR-A/B клеток. Клетки J-Lat трансфицировали пустым вектором pX260 или LTR-A и -B. После 2-недельного отбора с пуромицином клетки обрабатывали PMA или TNF α в течение 24 ч. Геномные ДНК подвергали ПЦР с использованием праймеров, охватывающих участки U3/R/U5 LTR ВИЧ-1 (от -411 до + 129) и проводили анализ SURVEYOR. **P<0,01 указывает на значимое снижение по сравнению с контролем U6-CAG. (С) При анализе ПЦР-фрагментов с использованием праймеров, охватывающих LTR ВИЧ-1 (от -374 до + 43), видна точная делекция участка в 190 п.о. между сайтами разрезания для LTR A и B, при этом остается фрагмент в 227 п.о. (стрелка). Ген домашнего хозяйства, β -актин, служит в качестве контроля очистки и загрузки ДНК.

На фиг. 11 показано, что эффективность редактирования генома зависит от присутствия Cas9 и gRNA. (А, В) При ПЦР-генотипировании выявлено отсутствие кассеты экспрессии LTR-A или LTR-B под управлением U6 (А) и отсутствие/снижение ДНК Cas9 под управлением CMV (В) в отобранных с пуромицином субклонах TZM-bI без каких-либо признаков геномного редактирования. Геномные ДНК из указанных субклонов подвергали традиционному ПЦР-анализу (А) или ПЦР-анализу в режиме реального времени (В) с использованием пары праймеров, охватывающей промотор U6 (T351) и LTR-A (T354) или -B (T356) и нацеленной на Cas9 (T477/T491). (С, D) Экспрессия белка Cas9 отсутствует в в неэффективных субклонах TZM-bI. Меченный Flag белок слияния Cas9 определяли при помощи вестерн-блоттинга (WB) и иммуноцитохимии (ICC) с помощью моноклонального антитела к Flag. Линию клеток HEK293T, стабильно экспрессирующую Flag-Cas9, использовали в качестве положительного контроля для WB (С). GAPDH служил в качестве контроля загрузки белка. Клон c6 содержал ДНК Cas9, но экспрессия белка Cas9 отсутствовала, что наводит на мысль о наличии потенциального механизма эпигенетической репрессии после отбора с пуромицином. В клоне c5 и c3 могут быть представлены усеченные Flag-Cas9 (tCas9). Ядро окрашивали при помощи Hoechst 33258 (D).

На фиг. 12 показано, что стабильная экспрессия Cas9/LTR-A/B gRNA в клетках TZM-bI вакцинирует их против псевдотипированных или нативных вирусов ВИЧ-1. (А) При проведении проточной цитометрии видно значительное снижение эффективности инфицирования нативным репортерным вирусом pNL4-3-ΔE-EGFP в экспрессирующих Cas9/LTR-A/B субклонах TZM-bI. (В, С) При ПЦР-анализе в режиме реального времени выявлены подавление или элиминация вирусных РНК (В) и ДНК (С) при помощи Cas9/LTR-A/B gRNA. (Д) При люминесцентном анализе в отношении люциферазы светлячка видно сильное ингибирование стимулированной вирусной инфекцией промоторной активности LTR под действием Cas9/LTR-A/B gRNA. Стабильно экспрессирующие Cas9/LTR-A/B gRNA клетки TZM-bI инфицировали в течение 2 ч указанными нативными вирусами ВИЧ-1 и дважды промывали PBS. На 2 день после инфицирования клетки собирали, фиксировали и анализировали при помощи проточной цитометрии в отношении экспрессии EGFP (А) или лизировали для извлечения общей РНК и подвергали RT-qPCR (Б), проводили очистку геномной ДНК для qPCR (С) и проводили измерение люминесценции (Д). *P<0,05 и **P<0,01 указывают на значимые снижения по сравнению с контролем U6-CAG.

На фиг. 13 показаны предсказанные LTR gRNA и количество нецелевых сайтов для них (100% совпадение). Смысловые и антисмысловые последовательности 5'-LTR (SEQ ID NO: 79-111 и 112-141, соответственно) (634 п.о.) из лентивирусного вектора pHR'-CMV-LacZ (AF105229) использовали для поиска целевых сайтов Cas9/gRNA, содержащих направляющую последовательность (протоспейсер) в 20 п.о. и последовательность мотива, прилегающего к протоспейсеру (NGG), с использованием инструмента Jack Lin's CRISPR/Cas9 gRNA finder (<http://spot.colorado.edu/~slin/cas9.html>). Каждую gRNA и NGG (AGG, TGG, GGG, CGG) подвергали сравнению с использованием BLAST в отношении соответствующих геномных последовательностей и последовательностей транскриптов человека, при этом показаны 1000 выровненных последовательностей. После нажатия Control + F копируют/вставляют целевую последовательность (от 1-23 до 9-23 нуклеотидов) и обнаруживают определенное количество геномных мишней со 100% совпадением. Количество нецелевых сайтов для каждого поиска делили на 3 из-за повторяющейся геномной библиотеки. Указанное количество отражает сумму по 4 поискам (NGG). Верхнее число (например, для последовательности (смысловой) gRNA: 20, 19, 18, 17, 16, 15, 14, 13, 12) указывает на наиболее удаленные от NGG целевые последовательности для gRNA. Последовательность и количества нецелевых сайтов для выбранных LTR-A/B и LTR-C/D выделены красным и зеленым соответственно.

На фиг. 14 показаны олигонуклеотиды для сайтов, на которые нацеливаются gRNA, и праймеры (SEQ ID NO: 36-78, соответственно, в порядке появления), используемые для ПЦР и секвенирования.

На фиг. 15 показаны расположения предсказанных сайтов, на которые нацеливаются gRNA LTR-A и LTR-B и раскрыты последовательности "запраш. посл." как SEQ ID NO: 142-252 и последовательности "эталон. посл." как SEQ ID NO: 253-363, все, соответственно, в порядке появления.

На фиг. 16 показано, что как LTR-C, так и LTR-D снижает конститутивную и TSA/PMA-индуцированную люциферазную активность в клетках TZMBI, со стабильно встроенным репортерным геном LTR ВИЧ-1-люцифераза светлячка, и комбинация индуцировала точное вырезание генома. Шесть мишеней для gRNA конструировали для промоторного участка LTR ВИЧ (фиг. 16А). На фиг. 16А раскрыта SEQ ID NO: 16. Клетки TZMBI совместно трансфицировали с помощью Cas9-EGFP и кассеты экспрессии химерной gRNA (ПЦР-продукты) при помощи Lipofectamine 2000. Через 3 дня положительные в отношении EGFP клетки отсортировывали при помощи FACS и для анализа на люциферазу собирали по 2000 клеток на группу (фиг. 16В). На фиг. 16В раскрыта SEQ ID NO: 31. Популяцию отсортированных клеток культивировали в течение 2 дней и обрабатывали TSA/PMA в течение 1 дня перед анализом на люциферазу (фиг. 16С). Отдельные клетки сортировали в 96-луночный планшет и культивировали до образования конфлюэнтного слоя для проведения анализа на люциферазу в отсутствие (фиг. 16Д) или присутствии (фиг. 16Е) TSA/PMA в течение 1 дня. ПЦР-продукт из популяции отсортированных клеток анализировали при помощи анализа Surveyor с нуклеазой Cel-I (фиг. 16F) и полиморфизма длин рестрикционных фрагментов с BsajI (фиг. 16G), которые показали мутацию (фиг. 16F) или неразрезанную (фиг. 16G) полосу (красная стрелка). Фрагмент в 200 п.о. (фиг. 16F, 16G, черная стрелка), полученный в результате делеции участка в 321 п.о. между LTR-C и LTR-D, как и было предсказано (фиг. 16А, красный указатель стрелки), подтверждали при помощи TA-клонирования и секвенирования, показавших точное вырезание генома (фиг. 16Н). При секвенировании по Сэнгеру ПЦР-продуктов от отдельных LTR-C и -D идентифицировали % мутаций по типу вставок/делеций и % эффективности введения мутаций по типу вставок/делеций соответственно (фиг. 16). * $p<0,05$ указывает на статистически значимое снижение, при использовании t-критерия Стьюдента, по сравнению с соответствующим контролем U6-CAG. Протоспейсер (E), протоспейсер (C), протоспейсер (A), протоспейсер (B), протоспейсер (D) и протоспейсер (F) соответствуют SEQ ID NO: 365, 367, 369, 371, 373 и 375, соответственно, в порядке появления.

На фиг. 17 показано, что Cas9/LTR-gRNA ингибирует конститутивную и индуцируемую выработку вируса ВИЧ-1, измеренную при помощи проточной цитометрии в отношении EGFP в латентно инфицированной ВИЧ-1 линии клеток микроглии CHME5. Лентивирусный вектор pHR¹, содержащий Tat, Rev, Env, Vpu и Nef с репортерным геном d2EGFP, трансдуцировали в линию фетальных клеток микроглии человека CHME5, и показана делеция в 400 п.о. в участке U3 3'-LTR (фиг. 17А). После транзиентной трансфекции Cas9/gRNA LTR-A, B, C, D ВИЧ-1 человека отдельно или в комбинации снижали интенсивность, но не процентную долю EGFP вследствие подавления промоторной активности LTR (фиг. 17В, 17С). После отбора с антибиотиком в течение 1-2 недель процентная доля клеток с EGFP также снижалась (фиг. 17Д, 17Е). ПЦР-продукт от стабильных отобранных клонов анализировали при помощи анализа Surveyor с нуклеазой Cel-I (фиг. 17F), показавшего мутацию по типу вставки/делеции отчетливо в случае LTR-A и LTR-B, но слабо в случае комбинации LTR-A/B (красная стрелка). Фрагмент в 331 п.о. (фиг. 17F, 17G, черная стрелка), полученный в результате делеции участка в 190 п.о. между LTR-A и LTR-B, как и было предсказано (фиг. 17Н, красный указатель стрелки), подтверждали при помощи TA-клонирования и секвенирования, показавших точное вырезание из генома (фиг. 17Н). На фиг. 17Н раскрыты SEQ ID NO 1-3, соответственно, в порядке появления.

На фиг. 18 показан LTR репрезентативной последовательности ВИЧ-1 (SEQ ID NO: 376). Участок U3 тянется от нуклеотида 1 до нуклеотида 432 (SEQ ID NO: 377), участок R тянется от нуклеотида 432 до нуклеотида 559 (SEQ ID NO: 378), а участок U5 тянется от 560 до нуклеотида 644 (SEQ ID NO: 379).

На фиг. 19 показан LTR репрезентативной последовательности SIV (SEQ ID NO: 380). Участок U3 тянется от нуклеотида 1 до нуклеотида 517 (SEQ ID NO: 381), участок R тянется от нуклеотида 518 до нуклеотида 693 (SEQ ID NO: 382), а участок U5 тянется от 694 до нуклеотида 818 (SEQ ID NO: 383).

Подробное описание

Изобретение основано, помимо всего прочего, на обнаружении того, что можно элиминировать интегрированный геном ВИЧ-1 из инфицированных ВИЧ-1 клеток путем использования РНК-направленной системы коротких палиндромных повторов, регулярно расположенных группами (CRISPR)-нуклеазы Cas 9 (Cas9/gRNA) в одиночной и мультиплексной конфигурациях. В пределах U3-участка LTR ВИЧ-1 идентифицировали высокоспецифичные мишени, которые эффективно редактируются при помощи Cas9/gRNA, что инактивирует экспрессию и репликацию вирусных генов в латентно инфицированных клетках микроглии, промоноцитах и Т-клетках. Cas9/gRNA не вызывает ни генотоксичность, ни нецелевое редактирование в отношении клеток-хозяев и полностью вырезает фрагмент в 9709 п.о. интегрированной провирусной ДНК, который тянется от ее 5'- до 3'-LTR. Кроме того, присутствие мультиплексных gRNA в экспрессирующих Cas9 клетках предупреждает ВИЧ-1-инфекцию. Результаты авторов настоящего изобретения позволяют предположить, что Cas9/gRNA можно конструировать для обеспечения специфического, эффективного профилактического и терапевтического подхода против СПИДа.

Соответственно, настоящее изобретение представляет композиции, содержащие нукleinовую ки-

слоту, кодирующую CRISPR-ассоциированную эндонуклеазу и направляющую РНК, которая комплементарна целевой последовательности в ретровирусе, например ВИЧ, а также фармацевтические составы, содержащие нукleinовую кислоту, кодирующую CRISPR-ассоциированную эндонуклеазу и направляющую РНК, которая комплементарна целевой последовательности в ВИЧ. Также представлены композиции, содержащие полипептид CRISPR-ассоциированную эндонуклеазу и направляющую РНК, которая комплементарна целевой последовательности в ВИЧ, а также фармацевтические составы, содержащие полипептид CRISPR-ассоциированную эндонуклеазу и направляющую РНК, которая комплементарна целевой последовательности в ВИЧ.

Также представлены способы введения композиций для лечения ретровирусной инфекции, например ВИЧ-инфекции, способы элиминации репликации вируса и способы предупреждения ВИЧ-инфекции. Терапевтические способы, описанные в данном документе, можно осуществлять совместно с другими видами антиретровирусной терапии (например, НААРТ).

Клиническое течение ВИЧ-инфекции может варьировать в зависимости от ряда факторов, в том числе генетического фона, возраста, общего состояния здоровья, питания субъекта, получаемого субъектом лечения и подтипа ВИЧ. Как правило, у большинства индивидуумов развиваются гриппоподобные симптомы в течение нескольких недель или месяцев после инфицирования. Симптомы могут включать лихорадку, головную боль, боли в мышцах, сыпь, озноб, боль в горле, язвы в полости рта или генитальные язвы, воспаленные лимфатические узлы, боль в суставах, ночную потливость и диарею. Степень выраженности симптомов может варьировать от умеренной до тяжелой в зависимости от индивидуума. Во время острой фазы вирусные частицы ВИЧ привлекаются клетками, экспрессирующими подходящие рецепторные молекулы CD4, и проникают в них. Как только вирус проник в клетку-хозяина, кодируемая ВИЧ обратная транскриптаза образует копию провиральной ДНК с РНК ВИЧ, и провиральная ДНК интегрируется в геномную ДНК клетки-хозяина. Именно она является провирусом ВИЧ, который реплицируется клеткой-хозяином, что приводит к высвобождению новых вирионов ВИЧ, которые могут затем инфицировать другие клетки. Способы и композиции согласно настоящему изобретению в целом и различным образом применимы для вырезания интегрированной провиральной ДНК ВИЧ, хотя настоящее изобретение вовсе не ограничивается этим, а композиции можно вводить субъекту на любой стадии инфекции или неинфицированным субъектам, которые подвержены риску ВИЧ-инфекции.

Первичная ВИЧ-инфекция проходит в течение от нескольких недель до нескольких месяцев, и за неё, как правило, следует длинный клинический "латентный" период, который может длиться до 10 лет. Латентный период также называют бессимптомной ВИЧ-инфекцией или хронической ВИЧ-инфекцией. Количество CD4-лимфоцитов у субъекта восстанавливается, но не до уровней, присутствующих перед инфицированием, и у большинства субъектов происходит сероконверсия, т.е. в их крови имеются обнаруживаемые уровни антител к ВИЧ, в течение 2-4 недель после инфицирования. Во время этого латентного периода в мононуклеарах периферической крови может отсутствовать обнаруживаемая репликация вируса, и в периферической крови может присутствовать небольшое количество поддающегося культивированию вируса или он отсутствует. Во время латентного периода, также называемого стадией клинической латентности, инфицированные ВИЧ люди могут не испытывать связанных с ВИЧ симптомов или только слабовыраженные симптомы. Но вирус ВИЧ продолжает репродуцироваться на очень низких уровнях. У субъектов, которые проходили лечение с помощью видов антиретровирусной терапии, этот латентный период может длиться в течение нескольких десятилетий или более. Однако субъекты на этой стадии все еще способны передавать ВРИ другим, даже если они получают антиретровирусную терапию, хотя антиретровирусная терапия снижает риск передачи. Как отмечено выше, антиретровирусная терапия не подавляет низкие уровни экспрессии вирусного генома, 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.

Клинические признаки и симптомы СПИДа (синдрома приобретенного иммунодефицита) проявляются по мере снижения количества CD4-лимфоцитов, что приводит к необратимому поражению иммунной системы. У многих пациентов также присутствуют СПИД-ассоциированные осложнения, в том числе, например, оппортунистические инфекции, такие как туберкулез, сальмонеллез, цитомегалия, кандиноз, криптококковый менингит, токсоплазмоз и криптоспоридиоз; а также определенные виды рака, в том числе, например, саркома Капоши и виды лимфомы; а также синдром истощения, неврологические осложнения и ВИЧ-ассоциированная нефропатия.

Композиции.

Композиции согласно настоящему изобретению включают нукleinовые кислоты, кодирующие CRISPR-ассоциированную эндонуклеазу, например Cas9, и направляющую РНК, которая комплементарна целевой последовательности в ретровирусе, например ВИЧ. У бактерий локусы CRISPR/Cas кодируют РНК-направленные системы адаптивного иммунитета против мобильных генетических элементов (вирусов, мобильных генетических элементов и конъюгативных плазмид). Были идентифицированы три типа (I-III) систем CRISPR. Кластеры CRISPR содержат спайсеры, последовательности, комплементарные ранее упомянутым мобильным элементам. Кластеры CRISPR транскрибируются и процессируются в зрелую РНК CRISPR (коротких палиндромных повторов, регулярно расположенных группами) (crRNA).

CRISPR-ассоциированная эндонуклеаза, Cas9, принадлежит к системам CRISPR/Cas II типа и обладает сильной эндонуклеазной активностью для разрезания целевой ДНК. Cas9 направляется зреющей crRNA, которая содержит приблизительно 20 пар оснований (п.о.) уникальной целевой последовательности (называемый спайсером) и транс-активирующую малую РНК (tracrRNA), которая служит в качестве направляющей для процессинга pre-crRNA при помощи рибонуклеазы III. Дуплекс crRNA:tracrRNA направляет Cas9 для целенаправленного воздействия на ДНК посредством комплементарного спаривания оснований между спайсером в crRNA и комплементарной последовательностью (называемой протоспайсером) в целевой ДНК. Cas9 распознает тринуклеотидный (NGG) мотив, прилегающий к протоспайсеру (PAM), для уточнения сайта разрезания (3-й нуклеотид от PAM). crRNA и tracrRNA могут экспрессироваться отдельно или могут быть сконструированы в искусственную слитую малую направляющую РНК (sgRNA) посредством синтетической структуры "стебель-петля" (AGAAAУ) для имитации естественного дуплекса crRNA/tracrRNA. Такие sgRNA, подобно shRNA, можно синтезировать, или транскрибировать *in vitro* для прямой трансфекции РНК, или экспрессировать с РНК-вектора экспрессии с промотором U6 или H1, хотя эффективности расщепления искусственной sgRNA ниже, чем эффективности систем с crRNA и tracrRNA, экспрессированными отдельно.

Композиции согласно настоящему изобретению могут включать нукleinовую кислоту, кодирующую CRISPR-ассоциированную эндонуклеазу. В некоторых вариантах осуществления CRISPR-ассоциированная эндонуклеаза может быть нуклеазой Cas9. Нуклеаза Cas9 может иметь нуклеотидную последовательность, идентичную последовательности *Streptococcus pyogenes* дикого типа. В некоторых вариантах осуществления CRISPR-ассоциированная эндонуклеаза может быть последовательностью из другого вида, например другого вида *Streptococcus*, такого как *thermophilus*; *Pseudomonas aeruginosa*, *Escherichia coli*, или других секвенированных геномов бактерий и архей, или других прокариотных микробов. Альтернативно, последовательность Cas9 *Streptococcus pyogenes* дикого типа может быть модифицирована. Последовательность нукleinовой 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 GI:669193761 или KM099233.1 GI:669193765.

Альтернативно, последовательность нуклеазы Cas9, например, может быть последовательностью, входящей в состав коммерчески доступного вектора, такого 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 Genbank KM099231.1 GI:669193757; KM099232.1 GI:669193761 или KM099233.1 GI:669193765 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 включать, например, аминокислотную последовательность, которая отличается от Cas9 дикого типа вследствие того, что содержит одну или более мутаций (например, мутацию с добавлением, делеций или заменой или комбинацию таких мутаций). Одна или более мутаций с заменой может быть заменой (например, консервативной аминокислотной заменой). 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, как правило, 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; и фенилаланин и тирозин. Аминокислотные остатки в аминокислотной последовательности Cas9 могут представлять собой не встречающиеся в природе аминокислотные остатки. Встречающиеся в природе аминокислотные остатки включают те, которые естественным образом закодированы при помощи генетического кода, а также нестандартные аминокислоты (например, аминокислоты с D-конфигурацией вместо L-конфигурации). Пептиды согласно настоящему изобретению также могут включать аминокислотные остатки, которые являются модифицированными вариантами стандартных остатков (например, пирролизин можно использовать вместо лизина, а сelenоцистеин можно использовать вместо цистеина). Не встречающиеся в природе аминокислотными остатками являются те, которые не были обнаружены в природе, но которые соответствуют основной формуле аминокислоты и могут быть включены в пептид. Таковые включают D-аллоизолейцин(2R,3S)-2-амино-3-метилпентановую кислоту и L-цикlopентилглицин-(S)-2-амино-2-цикlopентилкускусную кислоту. Что касается других примеров, можно получить информацию из пособий или всемирной сети Интернет (сайт в настоящее время поддерживается Калифорнийским технологическим институтом и на нем показаны структуры неприродных аминокислот, которые были успешно включены в функциональные белки).

Последовательность нуклеазы Cas9 может быть подвергнутой мутированию последовательностью. Например, нуклеаза Cas9 может быть подвергнута мутированию в консервативных доменах HNH и

RuvC, которые вовлечены в специфическое в отношении нити расщепление. Например, мутация аспартат-на-аланин (D10A) в каталитическом домене RuvC позволяет никакому мутанту Cas9 (Cas9n) надрезать, а не расщеплять ДНК с получением однонитевых разрывов, и последующая преимущественная ремаркация посредством HDR может потенциально снижать частоту нежелательных мутаций по типу вставки/делеции, образующихся в результате нецелевых двухнитевых разрывов.

В некоторых вариантах осуществления композиции согласно настоящему изобретению могут включать полипептид CRISPR-ассоциированную эндонуклеазу, закодированный с помощью любой из последовательностей нукleinовых кислот, описанных выше. Термины "пептид", "полипептид" и "белок" используются в данном документе взаимозаменяющими, хотя, как правило, ими называют пептидные последовательности различных размеров. Авторы настоящего изобретения могут называть аминокислотные композиции согласно настоящему изобретению "полипептидами", чтобы сказать, что они представляют собой линейные полимеры из аминокислотных остатков, и для содействия различию их с белками полной длины. Полипептид согласно настоящему изобретению может "составлять" или "включать" фрагмент CRISPR-ассоциированной эндонуклеазы, и настоящее изобретение охватывает полипептиды, которые составляют или включают биологически активные варианты CRISPR-ассоциированной эндонуклеазы. Будет понятно, что полипептиды, таким образом, могут включать только фрагмент CRISPR-ассоциированной эндонуклеазы (или ее биологически активный вариант), но также могут включать дополнительные остатки. У биологически активных вариантов будет сохраняться достаточная активность для расщепления целевой ДНК.

Связи между аминокислотными остатками могут представлять собой обычные пептидные связи или другую ковалентную связь (такую как сложноэфирная или эфирная связь), и полипептиды можно модифицировать путем амидирования, фосфорилирования или гликозилирования. Модификация может затрагивать остаток полипептида и/или одну или более боковых цепей. Химические модификации могут быть встречающимися в природе модификациями, произведенными *in vivo* после трансляции мРНК, кодирующей полипептид (например, гликозилирование в бактериальном хозяине), или искусственными модификациями, произведенными *in vitro*. Биологически активный вариант CRISPR-ассоциированной эндонуклеазы может включать одну или более структурных модификаций, полученных в результате любой комбинации встречающихся в природе (т.е. произведенных естественным путем *in vivo*) и искусственных модификаций (т.е. встречающихся в природе или не встречающихся в природе модификаций, произведенных *in vitro*). Примеры модификаций включают без ограничения амидирование (например, замещение свободной карбоксильной группы на С-конце аминогруппой); биотинилирование (например, ацилирование лизина или других реакционно-способных аминокислотных остатков молекулой биотина); гликозилирование (например, добавление гликозильной группы к остатку любого из аспарагинов, гидроксилизина, серина или треонина с получением гликопротеина или гликопептида); ацетилирование (например, добавление ацетильной группы, как правило, на N-конце полипептида); алкилирование (например, добавление алкильной группы); изопрененирование (например, добавление изопренOIDной группы); липоилирование (например, присоединение липоатного фрагмента) и фосфорилирование (например, добавление фосфатной группы к серину, тирозину, треонину или гистидину).

Один или более аминокислотных остатков в биологически активном варианте могут быть не встречающимися в природе аминокислотным остатком. Встречающиеся в природе аминокислотные остатки включают те, которые естественным образом закодированы при помощи генетического кода, а также нестандартные аминокислоты (например, аминокислоты с D-конфигурацией вместо L-конфигурации). Пептиды согласно настоящему изобретению также могут включать 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 те, которые не были обнаружены в природе, но которые соответствуют основной формуле аминокислоты и могут быть включены в пептид. Таковые включают D-аллоизолейцин(2R,3S)-2-амино-3-метилпентановую кислоту и L-цикlopентилглицин-(S)-2-амино-2-цикlopентилуксусную кислоту. Что касается других примеров, можно получить информацию из пособий или всемирной сети Интернет (сайт в настоящее время поддерживается Калифорнийским технологическим институтом, и на нем показаны структуры неприродных аминокислот, которые были успешно включены в функциональные белки).

Альтернативно или дополнительно, один или более аминокислотных остатков в биологически активном варианте может представлять собой встречающийся в природе остаток, который отличается от встречающегося в природе остатка, выявленного в соответствующем положении в последовательности дикого типа. Иначе говоря, биологически активные варианты могут включать одну или более аминокислотных замен. Авторы настоящего изобретения могут называть замену, добавление или делецию аминокислотных остатков мутацией в последовательности дикого типа. Как отмечалось, при замене встречающийся в природе аминокислотный остаток может замещаться не встречающимися в природе остатком или просто отличным встречающимися в природе остатком. Кроме того, замена может представлять собой консервативную или неконсервативную замену. Консервативные аминокислотные замены, как правило, включают замены в пределах следующих групп: глицин и аланин; валин, изолейцин и лейцин; аспараги-

новая кислота и глутаминовая кислота; аспарагин, глутамин, серин и треонин; лизин, гистидин и аргинин; и фенилаланин и тирозин.

Полипептиды, которые являются биологически активными вариантами CRISPR-ассоциированной эндонуклеазы, могут характеризоваться с учетом степени, в которой их последовательности схожи с соответствующим полипептидом дикого типа или идентичны ему. Например, последовательность биологически активного варианта может быть по меньшей мере или приблизительно на 80% идентичной по соответствующим остаткам с полипептидом дикого типа. Например, биологически активный вариант CRISPR-ассоциированной эндонуклеазы может иметь аминокислотную последовательность, характеризующуюся по меньшей мере или приблизительно 80% идентичностью последовательности (например, по меньшей мере или приблизительно 85, 90, 95, 97, 98 или 99% идентичностью последовательности) с CRISPR-ассоциированной эндонуклеазой или с ее гомологом или ортологом.

У биологически активного варианта полипептида CRISPR-ассоциированного эндонуклеазы будет сохраняться достаточная биологическая активность, чтобы он был применим в способах согласно настоящему изобретению. У биологически активных вариантов будет сохраняться достаточная активность для функционирования в целенаправленном расщеплении ДНК. Биологическую активность можно оценивать способами, известными обычному специалисту в данной области, и они включают без ограничения анализы расщепления *in vitro* или функциональные анализы.

Полипептиды можно получать рядом способов, в том числе, например, методиками рекомбинантной ДНК или химическим синтезом. После получения полипептидов их можно выделять и очищать до любой необходимой степени при помощи способов, хорошо известных в данной области техники. Например, можно применять лиофилизацию, например, после HPLC с обращенной фазой (предпочтитель но) или с нормальной фазой, или эксклюзационной хроматографии, или распределительной хроматографии на полисахаридном геле, таком как Sephadex G-25. Состав конечного полипептида можно подтверждать аминокислотным анализом после разрушения пептида при помощи стандартных способов, аминокислотным секвенированием или при помощи методик FAB-MS. Соли, в том числе кислые соли, сложные эфиры, амиды и N-ацилпроизводные аминогрупп полипептида можно получать с помощью способов, известных из уровня техники, и такие пептиды применимы с точки зрения изобретения.

Композиции согласно настоящему изобретению включают последовательность, кодирующую направляющую РНК (gRNA), содержащую последовательность, которая комплементарна целевой последовательности в ретровирусе. Ретровирусом может быть лентивирус, например, вирус иммунодефицита человека, вирус иммунодефицита обезьян, вирус иммунодефицита кошачьих или вирус иммунодефицита крупного рогатого скота. Вирусом иммунодефицита человека может быть ВИЧ-1 или ВИЧ-2. Целевая последовательность может включать последовательность из любого типа ВИЧ, например ВИЧ-1 и ВИЧ-2, и его любой циркулирующей рекомбинантной формы. Генетическая изменчивость ВИЧ отражена во множестве групп и подтипов, которые были описаны. Совокупность последовательностей ВИЧ собрана в базах данных ВИЧ Los Alamos и компендиумах (т.е. веб-сайт с базой данных последовательностей находится по адресу <http://www.hiv.lanl.gov/>). Способы и композиции согласно настоящему изобретению можно применять по отношению к ВИЧ из любого из этих различных групп, подтипов и циркулирующих рекомбинантных форм. Они включают, например, основную группу ВИЧ-1 (часто называемую группа М) и небольшие группы, группы N, O и P, а также без ограничения любой из следующих подтипов A, B, C, D, F, G, H, J и K, или группу (например, без ограничения любую из следующих групп N, O и P) ВИЧ. Способы и композиции также можно применять по отношению к ВИЧ-2 и любой из клад A, B, C, F или G (также называемых "подтипами" или "группами"), а также любой циркулирующей рекомбинантной форме ВИЧ-2.

Направляющая РНК может быть последовательностью, комплементарной кодирующей или некодирующей последовательности. Например, направляющая РНК может быть последовательностью ВИЧ, такой как последовательность длинного концевого повтора (LTR), кодирующая белок последовательность или регуляторная последовательность. В некоторых вариантах осуществления направляющая РНК содержит последовательность, которая комплементарна участку длинного концевого повтора (LTR) ВИЧ. LTR ВИЧ-1 составляет примерно 640 п.о. в длину. Иллюстративным LTR ВИЧ-1 является последовательность под SEQ ID NO: 376. Иллюстративным LTR SIV является последовательность под SEQ ID NO: 380. Длинные концевые повторы (LTR) ВИЧ-1 делятся на участки U3, R и U5. Иллюстративными участками U3, R и (J5 LTR ВИЧ-1 являются SEQ ID NO: 377, 378 и 379 соответственно. Иллюстративными участками U3, R и U5 LTR SIV являются SEQ ID NO: 381, 382 и 383 соответственно. Конфигурация участков U1, R, U5 для иллюстративных последовательностей ВИЧ-1 и SIV показана на фиг. 18 и 19 соответственно. LTR содержат все необходимые сигналы для экспрессии генов, и они вовлечены в интеграцию 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 несколько субучастков, например TAR или действующий в транс-положении чувствительный элемент, который вовлечен в активацию транскрипции; полиг А, который вовлечен в димеризацию и упаковку генома; PBS или сайт связывания праймера; Psi или сигнал упаковки; DIS или сайт инициации образования димеров.

Применимые направляющие последовательности комплементарны участкам U3, R или U5 LTR. Иллюстративные последовательности направляющей РНК, которые целенаправленно воздействуют на участок U3 ВИЧ-1, показаны на фиг. 13. Последовательность направляющей РНК может содержать, например, последовательность

LTR A: ATCAGATATCCACTGACCTTGG (SEQ ID NO: 96),

LTR B: CAGCAGTTCTTGAAGTACTCCGG (SEQ ID NO: 121),

LTR C: GATTGGCAGAACTACACACCAGG (SEQ ID NO: 87) или

LTR D: GCGTGGCCTGGGCGGGACTGGGG (SEQ ID NO: 110).

Места расположения LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87) и LTR D (SEQ ID NO: 110) в участке U3 (SEQ ID NO: 16) показаны на фиг. 5. Дополнительные иллюстративные последовательности направляющей РНК, которые целенаправленно воздействуют на участок U3, перечислены в таблице, показанной на фиг. 13, и могут иметь последовательность с любым SEQ ID NO: 79-111 и SEQ ID NO: 111-141. В некоторых вариантах осуществления направляющая последовательность может содержать последовательность, характеризующуюся 95% идентичностью с любой последовательностью под SEQ ID NO: 79-111 и SEQ ID NO: 111-141. Таким образом, последовательность направляющей РНК может содержать, например, последовательность, характеризующуюся 95% идентичностью с последовательностью

LTR A: ATCAGATATCCACTGACCTTGG (SEQ ID NO: 96),

LTR B: CAGCAGTTCTTGAAGTACTCCGG (SEQ ID NO: 121),

LTR C: GATTGGCAGAACTACACACCAGG (SEQ ID NO: 87) или

LTR D: GCGTGGCCTGGGCGGGACTGGGG (SEQ ID NO: 110).

Авторы настоящего изобретения также могут называть последовательность направляющей РНК протоспейсером, например протоспейсер(A), протоспейсер(B), протоспейсер(C) и протоспейсер(D).

Последовательность направляющей РНК может быть последовательностью, обнаруженной в пределах эталонной последовательности или консенсусной последовательности участка U3, R или U5 ВИЧ-1. Настоящее изобретение, однако, этим не ограничивается, и можно выбирать последовательности направляющей РНК для целенаправленного воздействия на любую вариантную или мутантную последовательность ВИЧ. В некоторых вариантах осуществления направляющая РНК может включать вариантную последовательность или последовательность квазивида. В некоторых вариантах осуществления направляющая РНК может быть последовательностью, соответствующей последовательности в геноме вируса, который несет субъект, проходящий лечение. Таким образом, например, можно получать последовательность конкретного участка U3, R или U5 в вирусе ВИЧ, который несет субъект, и можно применять направляющие РНК, комплементарные конкретным последовательностям пациента.

В некоторых вариантах осуществления направляющая РНК может быть последовательностью, комплементарной кодирующей белок последовательности, например последовательности, кодирующей один или более вирусных структурных белков (например, gag, pol, env и tat). Таким образом, последовательность может быть комплементарной последовательности в gag-полипротеине, например MA (матричном белке, p17); CA (капсидном белке, p24); SP1 (спейсерном пептиде 1, p2); NC (белке нуклеокапсида, p7); SP2 (спейсерном пептиде 2, p1) и P6-белке; pol, например, обратной транскриптазе (RT) и РНКазе Н, интегразе (IN) и протеазе ВИЧ (PR); env, например, gp160, или продукте расщепления gp160, например gp120 или SU, и gp41 или TM; или tat, например Tat с одним 72 аминокислотным экзоном или Tat с двумя 86-101 аминокислотными экзонами. В некоторых вариантах осуществления направляющая РНК может быть последовательностью, комплементарной последовательности, кодирующей аксессорный белок, в том числе, например, vif, n willf (отрицательный фактор) vpu (вирусного белка U) и tev.

В некоторых вариантах осуществления последовательность может быть последовательностью, комплементарной структурному или регуляторному элементу, например LTR, который описан выше; TAR (целевой последовательности для вирусной трансактивации), сайту связывания для Tat-белка и для клеточных белков, состоящему из приблизительно 45 первых нуклеотидов вирусных mRNA в ВИЧ-1 (или первых 100 нуклеотидов в ВИЧ-2), который образует структуру шпильки по типу "стебель-петля"; RRE (Rev-чувствительному элементу), РНК-элементу, закодированному в участке env ВИЧ-1, состоящему из приблизительно 200 нуклеотидов (положения 7710-8061 от сайта инициации транскрипции в ВИЧ-1, охватывающие границу gp120 и gp41); PE (Psi-элементу), группе из 4 структур по типу "стебель-петля", предшествующих и перекрывающих старт-кодон Gag; SLIP, TTTTTT "скользящему сайту", за которым следует структура по типу "стебель-петля"; CRS (действующие в цис-положении репрессорные последовательности); INS (ингибирующие/дестабилизирующие РНК-последовательности), выявленным, например, в нуклеотидах 414-631 в участке gag ВИЧ-1.

Последовательность направляющей РНК может быть смысловой или антисмысловой последовательностью. Последовательность направляющей РНК обычно включает мотив, прилегающий к протоспейсеру (PAM). Последовательность PAM может варьировать в зависимости от требований специфичности используемой CRISPR-эндонуклеазы. В системе CRISPR-Cas, полученной из *S.puogenes*, целевая ДНК, как правило, непосредственно предшествует мотиву, прилегающему к протоспейсеру (PAM), 5'-NGG. Таким образом, для Cas9 *S.puogenes* последовательностью PAM может быть AGG, TGG, CGG или GGG. Другие ортологи Cas9 могут характеризоваться отличающейся специфичностью в отношении PAM. Например, для Cas9 из *S.thermophilus* требуется 5'-NNAGAA для CRISPR 1 и 5'-NGGNG для CRISPR 3, а для *Neiseria meningitidis* требуется 5'-NNNNNGATT. Специфическая последовательность направляющей РНК может варьировать, но независимо от последовательности, применимыми последовательностями направляющей РНК будут те, которые сводят к минимуму нецелевые эффекты, при этом достигают высокой эффективности и полного разрушения интегрированного в геном провируса ВИЧ-1. Длина последовательности направляющей РНК может варьировать от приблизительно 20 до приблизительно 60 или более нуклеотидов, например 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 более 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 биоинформационический скриптинг 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-1 (потенциально консервативные в геноме хозяина); отбор направленных на LTR-А- и -В gRNA в 30 п.о., а также системы pre-crRNA, отражающей первонаучальный бактериальный иммунный механизм, для увеличения специфичности/эффективности в противовес системе на основе gRNA в 20 п.о., химерной crRNA-tracrRNA и WGS, секвенирование по Сэнгеру и анализ SURVEYOR для идентификации и исключения потенциальных нецелевых эффектов.

Последовательность направляющей РНК может быть сконфигурирована в качестве отдельной последовательности или в качестве комбинации одной или более различных последовательностей, например мультиплексной конфигурации. Мультиплексные конфигурации могут включать комбинации двух, трех, четырех, пяти, шести, семи, восьми, девяти, десяти или более различных направляющих РНК, например любую комбинацию последовательностей в U3, R или U5. В некоторых вариантах осуществления можно применять комбинации LTR A, LTR B, LTR C и LTR D. В некоторых вариантах осуществления можно применять комбинации любых из последовательностей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87) и LTR D (SEQ ID NO: 110). В некоторых вариантах осуществления можно применять любую комбинацию из последовательностей с последовательностью под SEQ ID NO: 79-111 и SEQ ID NO: 111-141. В тех случаях, когда композиции вводят в виде вектора экспрессии, направляющие РНК могут быть закодированы в отдельном векторе. Альтернативно, несколько векторов могут быть сконструированы так, чтобы каждый включал две или более различных направляющих РНК. Применимые конфигурации будут приводить к вырезанию вирусных последовательностей между сайтами расщепления, что будет приводить к разрушению генома ВИЧ или прекращению экспрессии белка ВИЧ. Таким образом, применение двух или более различных направляющих РНК способствует вырезанию вирусных последовательностей между сайтами расщепления, распознаваемыми CRISPR-эндонуклеазой. Вырезаемый участок может варьировать по размеру от одного нуклеотида до нескольких тысяч нуклеотидов. Иллюстративные вырезаемые участки описаны в примерах.

В тех случаях, когда композиции вводят в виде нукleinовой кислоты, или они входят в состав вектора экспрессии, CRISPR-эндонуклеаза может быть закодирована той же нукleinовой кислотой или тем же вектором, что и последовательности направляющей РНК. Альтернативно или дополнительно, CRISPR-эндонуклеаза может быть закодирована в нукleinовой кислоте, физически отделенной от последовательностей направляющей РНК, или в отдельном векторе.

В некоторых вариантах осуществления молекулы РНК, например crRNA, tracrRNA, gRNA, сконструированы так, чтобы включать 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-окси-уксусную кислоту; метиловый сложный эфир уридин-5-оксикусусной кислоты; 5-карбоксиметилуридин;

5-метоксикарбонилметилуридин; 5-метоксикарбонилметил-2'-О-метилуридин; 5-метоксикарбонилметил-2'-тиоуридин; 5-карбамоилметилуридин; 5-карбамоилметил-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'-метиладенозин; 2-метиладенозин; N⁶N-метиладенозин; N⁶N⁶-диметиладенозин; N⁶2'-О-триметиладенозин; 2-метилтио-N⁶N-изопентениладенозин; N⁶-(цик-гидроксизопентенил)аденозин; 2-метилтио-N⁶-(цик-гидроксизопентенил)аденозин; N⁶-глицинилкарбамоиладенозин; N⁶-треонилкарбамоиладенозин; N⁶-гидроксипропилкарбамоиладенозин; 2-метилтио-N⁶-метил-N⁶-треонилкарбамоиладенозин; 2'-О-рибозиладенозин (фосфат); инозин; 2'-О-метилинозин; 1-метилинозин; 1,2'-О-диметилинозин; 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; недомодифицированный гидроксивибутозин; вибутозин; гидроксивибутозин; пероксивибутозин; квеузин; эпоксиквеузин; галактозил-квеузин; маннозил-квеузин; 7-циано-7-деазагуанозин; арахеозин [также называемый 7-формамидо-7-деазагуанозин] и 7-аминометил-7-деазагуанозин.

Авторы настоящего изобретения могут использовать термины "нуклеиновая кислота" и "полинуклеотид" взаимозаменяя для обозначения как РНК, так и ДНК, в том числе кДНК, геномной ДНК, синтетической ДНК и ДНК (или РНК), содержащей аналоги нуклеиновой кислоты, любая из которых может кодировать полипептид согласно настоящему изобретению, и все из которых охвачены настоящим изобретением. Полинуклеотиды могут характеризоваться практически любой трехмерной структурой. Нуклеиновая кислота может быть двунитевой или однонитевой (т.е. смысловой нитью или антисмысловой нитью). Неограничивающие примеры полинуклеотидов включают гены, фрагменты генов, экзоны, интроны, матричную РНК (мРНК) и ее части, транспортную РНК, рибосомальную РНК, миРНК, микро-РНК, рибозимы, кДНК, рекомбинантные полинуклеотиды, разветвленные полинуклеотиды, плазмида, векторы, выделенную ДНК с любой последовательностью, выделенную РНК с любой последовательностью, зонды нуклеиновых кислот и праймеры, а также аналоги нуклеиновых кислот. Применительно к настоящему изобретению нуклеиновые кислоты могут кодировать фрагмент встречающейся в природе Cas9 или ее биологически активный вариант и направляющую РНК, где направляющая РНК комплементарна последовательности в ВИЧ.

"Выделенная" нуклеиновая кислота может быть, например, встречающейся в природе молекулой ДНК или ее фрагментом, при условии что по меньшей мере одна из последовательностей нуклеиновых кислот, в норме непосредственно flankирующих эту молекулу ДНК во встречающемся в природе геноме, удалена или отсутствует. Таким образом, выделенная нуклеиновая кислота включает без ограничения молекулу ДНК, которая существует как отдельная молекула, независимо от других последовательностей (например, химически синтезированная нуклеиновая кислота, или кДНК, или фрагмент геномной ДНК, полученный при помощи полимеразной цепной реакции (ПЦР) или обработки рестрикционной эндонуклеазой). Выделенная нуклеиновая кислота также означает молекулу ДНК, которая встроена в вектор, автономно реплицирующуюся плазмиду, вирус или в геномную ДНК прокариота или эукариота. Кроме того, выделенная нуклеиновая кислота может включать сконструированную нуклеиновую кислоту, такую как молекула ДНК, которая является частью гибридной или слитой нуклеиновой кислоты.

Нуклеиновая кислота, существующая среди многих (например, десятков или сотен-миллионов) других нуклеиновых кислот, например в библиотеках кДНК, или геномных библиотеках, или пластинах геля, содержащих рестриктазные фрагменты геномной ДНК, не является выделенной нуклеиновой кислотой.

Молекулы выделенной нуклеиновой кислоты могут быть получены при помощи стандартных методик. Например, методики полимеразной цепной реакции (ПЦР) можно применять для получения выделенной нуклеиновой кислоты, содержащей нуклеотидную последовательность, описанную в данном документе, в том числе нуклеотидные последовательности, кодирующие полипептид, описанный в данном документе. ПЦР можно применять для амплификации специфических последовательностей из ДНК, а также из РНК, в том числе последовательностей из общей геномной ДНК или общей клеточной РНК. Разные способы ПЦР описаны, например, в PCR Primer: A Laboratory Manual, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Как правило, информацию о последовательности на концах представляющего интерес участка или за их пределами используют для конструирования олигонуклеотидных праймеров, которые идентичны или схожи последовательностью с противоположными нитями матрицы, подлежащей амплификации. Также доступны различные ПЦР-стратегии, при помощи 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. Напри-

мер, одну или более пар длинных олигонуклеотидов (например, >50-100 нуклеотидов) можно синтезировать так, чтобы они содержали необходимую последовательность, при этом каждая пара содержит короткий сегмент комплементарности (например, приблизительно 15 нуклеотидов), чтобы образовывался дуплекс при гибридизации олигонуклеотидной пары. ДНК-полимеразу используют для удлинения олигонуклеотидов, что приводит к одной двунитевой молекуле нукleinовой кислоты на олигонуклеотидную пару, которую затем можно лигировать в вектор. Выделенные нукleinовые кислоты согласно настоящему изобретению также можно получать при помощи мутагенеза, например, встречающейся в природе части ДНК, кодирующей Cas9 (в соответствии, например, с вышеупомянутой формулой).

Две нукleinовые кислоты или полипептиды, которые они кодируют, могут быть описаны как характеризующиеся определенной степенью идентичности друг с другом. Например, белок Cas9 и его биологически активный вариант могут быть описаны как проявляющие определенную степень идентичности. Результаты выравнивания могут быть получены путем обнаружения коротких последовательностей Cas9 на сайте "Protein Information Research" (PIR) (<http://pir.georgetown.edu>) с последующим анализом при помощи алгоритма Basic Local Alignment Search Tool (BLAST) "короткие практически идентичные последовательности" на веб-сайте NCBI (<http://www.ncbi.nlm.nih.gov/blast>).

Используемое в данном документе выражение "процентная идентичность последовательностей" означает степень идентичности между любой данной запрашиваемой последовательностью и рассматриваемой последовательностью.

Например, встречающаяся в природе Cas9 может быть запрашиваемой последовательностью, а фрагмент белка Cas9 может быть рассматриваемой последовательностью. Подобным образом, фрагмент белка Cas9 может быть запрашиваемой последовательностью, а его биологически активный вариант может быть рассматриваемой последовательностью.

Для определения идентичности последовательностей запрашиваемую последовательность нукleinовой кислоты или аминокислотную запрашиваемую последовательность можно выравнивать с одной или более рассматриваемыми последовательностями нукleinовой кислоты или рассматриваемыми аминокислотными последовательностями, соответственно, с использованием компьютерной программы ClustalW (версия 1.83, параметры по умолчанию), которая позволяет проводить выравнивания последовательностей нукleinовой кислоты или белковых последовательностей по их полной длине (глобальное выравнивание). См. Chenna et al., Nucleic Acids Res. 31:3497-3500, 2003.

ClustalW вычисляет наилучшее совпадение между запрашиваемой и одной или более рассматриваемыми последовательностями и выравнивает их таким образом, что можно определить идентичности, схожести и отличия. Гэпы из одного или более остатков могут быть вставлены в запрашиваемую последовательность, рассматриваемую последовательность или обе для получения максимальной длины выравниваний последовательностей. Для быстрого попарного выравнивания последовательностей нукleinовых кислот применяют следующие параметры по умолчанию: размер слова: 2; размер окна: 4; метод оценки: процентная доля; количество лучших диагоналей: 4 и штраф за внесение гэпа: 5. Для множественных выравниваний последовательностей нукleinовых кислот применяют следующие параметры: штраф за открытие гэпа: 10,0; штраф за продление гэпа: 5,0 и вес транзиций: да. Для быстрого попарного выравнивания белковых последовательностей применяют следующие параметры: размер слова: 1; размер окна: 5; метод оценки: процентная доля; количество лучших диагоналей: 5; штраф за внесение гэпа: 3. Для множественного выравнивания белковых последовательностей применяют следующие параметры: матрица сравнения: blosum; штраф за открытие гэпа: 10,0; штраф за продление гэпа: 0,05; гэпы вместо гидрофильных остатков: вкл.; гидрофильные остатки: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg и Lys; штрафы за специфичные к остаткам гэпы: вкл. Результатом является выравнивание последовательностей, которое отражает взаимоотношения между последовательностями. ClustalW можно проводить, например, на сайте средств запуска Бейлорского колледжа медицины (searchlauncherbcm.tmc.edu/multi-align/multi-align.html) и на сайте Европейского института биоинформатики во всемирной сети Интернет (ebi.ac.uk/clustalw).

Для определения процентной идентичности между запрашиваемой последовательностью и рассматриваемой последовательностью ClustalW делит количество идентичных совпадений из наилучшего выравнивания на количество сравниваемых остатков (гэп- положения исключаются), и результат умножается на 100. Результатом является процентная идентичность рассматриваемой последовательности относительно запрашиваемой последовательности. Следует отметить, что значение процентной идентичности может быть округлено до ближайшей десятой доли. Например, 78,11, 78,12, 78,13 и 78,14 округляют в меньшую сторону до 78,1, тогда как 78,15, 78,16, 78,17, 78,18 и 78,19 округляют в большую сторону до 78,2.

Нукleinовые кислоты и полипептиды, описанные в данном документе, можно называть "экзогенными". Термин "экзогенный" означает, что нукleinовая кислота или полипептид являются частью рекомбинантной конструкции нукleinовой кислоты или закодированы в нем, или не находятся в своем естественном окружении. Например, экзогенная нукleinовая кислота может быть последовательностью из одного вида, введенной в другой вид, т.е. гетерологичной нукleinовой кислотой. Как правило, такую экзогенную нукleinовую кислоту вводят в другой вид при помощи рекомбинантной конструкции нук-

леиновой кислоты. Экзогенная нуклеиновая кислота также может быть последовательностью, которая является нативной для организма и которая была вновь внесена в клетки этого организма. Экзогенная нуклеиновая кислота, которая включает нативную последовательность, часто может отличаться от встречающейся в природе последовательности присутствием неприродных последовательностей, связанных с экзогенной нуклеиновой кислотой, например ненативных регуляторных последовательностей, flankирующих нативную последовательность в рекомбинантной конструкции нукleinовой кислоты. Кроме того, стабильно трансформированные экзогенные нуклеиновые кислоты, как правило, интегрированы в положениях, отличных от положения, в котором обнаруживается нативная последовательность.

В данном документе также предусмотрены рекомбинантные конструкции, и их можно использовать для трансформации клеток для того, чтобы они экспрессировали Cas9 и/или направляющие РНК, комплементарные целевой последовательности в ВИЧ. Рекомбинантная конструкция нукleinовой кислоты содержит нукleinовую кислоту, кодирующую Cas9 и/или направляющую РНК, комплементарную целевой последовательности в ВИЧ, как описано в данном документе, функционально связанную с регуляторным участком, подходящим для экспрессии Cas9 и/или направляющей РНК, комплементарной целевой последовательности в ВИЧ, в клетке. Следует понимать, что целый ряд нукleinовых кислот может кодировать полипептид с конкретной аминокислотной последовательностью. Вырожденность генетического кода хорошо известна из данной области техники. Для многих аминокислот существует несколько нуклеотидных триплетов, которые служат в качестве кодона для аминокислоты. Например, кодоны в кодирующей последовательности для Cas9 можно модифицировать так, чтобы получить оптимальную экспрессию в конкретном организме, с использованием таблиц замен кодонов, подходящих для данного организма.

Также предусмотрены векторы, содержащие нукleinовые кислоты, такие как описываемые в данном документе. "Вектор" представляет собой репликон, такой как плазмида, фаг или космида, в который можно встроить другой сегмент ДНК для осуществления, таким образом, репликации встроенного сегмента. В целом, вектор способен к репликации, если он ассоциирован с соответствующими регуляторными элементами. Подходящие векторные оставы включают, например, традиционно применяемые в данной области техники, такие как плазмиды, вирусы, искусственные хромосомы, ВАС, YAC или PAC. Термин "вектор" включает векторы клонирования и векторы экспрессии, а также вирусные векторы и интегрирующие векторы. "Вектор экспрессии" представляет собой вектор, который включает регуляторный участок. Широкий спектр комбинаций организм-хозяин/вектор экспрессии можно применять для экспрессии последовательностей нукleinовых 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 (Карлсbad, Калифорния).

Векторы, предусмотренные в данном документе, также могут включать, например, точки начала репликации, участки прикрепления к ядерному матриксу (SAR) и/или маркеры. Маркерный ген может придавать клетке-хозяину селектируемый фенотип. Например, маркер может придавать устойчивость к биоциду, такую как устойчивость к антибиотику (например, канамицину, G418, блеомицину или гигромицину). Как отмечено выше, вектор экспрессии может включать последовательность-метку, разработанную для облегчения манипуляции или обнаружения (например, очистки или локализации) экспрессированного полипептида. Последовательности-метки, такие как последовательности зеленого флуоресцентного белка (GFP), глутатион-S-трансферазы (GST), полигистидина, с-тус, гемагглютинина или метки FlagTM (Kodak, Нью-Хейвен, Коннектикут), как правило, экспрессируются в качестве слияния с кодируемым полипептидом. Такие метки можно встраивать в любое место в пределах полипептида, в том числе либо на карбоксильном конце, либо на аминоконце.

Дополнительные векторы экспрессии также могут включать, например, сегменты хромосомных, отличных от хромосомных и искусственных последовательностей ДНК. Подходящие векторы включают производные SY40 и известные бактериальные плазмиды, например плазмиды E.coli col E1, pCR1, pBR322, pMal-C2, pET, pGEX, pMB9 и их производные, плазмиды, такие как RP4; фаговые ДНК, например многочисленные производные фага 1, например NM989, и другие фаговые ДНК, например M13 и однонитевую ДНК нитевидных фагов; плазмиды дрожжей, такие как плазмиды 2μ или ее производные, векторы, подходящие для эукариотических клеток, такие как векторы, подходящие для клеток насекомых или млекопитающих; векторы, полученные из комбинаций плазмид и фаговых ДНК, такие как плазмиды, которые были модифицированы, чтобы использовать фаговую ДНК или другие последовательности контроля экспрессии.

Также можно использовать системы экспрессии дрожжей. Например, вектора pYES2 не для слияний (сайты клонирования XbaI, SphI, ShoI, NotI, GstXI, EcoRI, BstXI, BamH1, SacI, KpnI и HindIII; Invitrogen) или вектор для слияний pYESHisA, B, C (сайты клонирования XbaI, SphI, ShoI, NotI, BstXI, EcoRI, BamH1, SacI, KpnI и HindIII, N-концевой пептид применялся для очистки при помощи смолы ProBond и отщеплялся при помощи энтерокиназы; Invitrogen), если говорить лишь об этих двух, можно применять в

соответствии с настоящим изобретением. Двухгибридную систему экспрессии дрожжей также можно получать в соответствии с настоящим изобретением.

Вектор также может включать регуляторный участок. Термин "регуляторный участок" означает нуклеотидные последовательности, 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, индуцильные элементы, последовательности связывания белка, 5'- и 3'-нетранслируемые участки (UTR), сайты инициации транскрипции, последовательности терминации, последовательности полиаденилирования, сигналы ядерной локализации и интроны.

Как используется в данном документе, термин "функционально связанный" относится к размещению регуляторного участка и последовательности, которую необходимо транскрибировать в нуклеиновую кислоту, чтобы повлиять на транскрипцию или трансляцию такой последовательности. Например, для приведения кодирующей последовательность под контроль промотора сайт инициации трансляционной рамки считывания полипептида, как правило, размещают на один-пятьдесят нуклеотидов ниже промотора. Однако промотор можно размещать не менее чем на приблизительно 5000 нуклеотидов выше сайта инициации трансляции или на приблизительно 2000 нуклеотидов выше сайта инициации транскрипции. Промотор, как правило, содержит, по меньшей мере, коровый (основной) промотор. Промотор также может включать по меньшей мере один регуляторный элемент, такой как энхансерная последовательность, вышерасположенный элемент или вышерасположенный участок активации (UAR). Выбор промоторов, которые следует включать, зависит от нескольких факторов, в том числе без ограничения эффективности, селективности, индуцильности, необходимого уровня экспрессии и предпочтительной в отношении клетки или ткани экспрессии. Для специалиста в данной области обычным процессом является модулирование экспрессии кодирующей последовательности путем соответствующего выбора и размещения промоторов и других регуляторных участков относительно кодирующей последовательности.

Векторы включают, например, вирусные векторы (такие как аденоовириусы ("Ad"), аденоассоциированные вирусы (AAV), и вирус везикулярного стоматита (VSV), и ретровириусы), липосомы и другие содержащие липиды комплексы, а также другие макромолекулярные комплексы, способные опосредовать доставку полинуклеотида в клетку-хозяина. Векторы также могут включать другие компоненты или функциональные элементы, которые дополнительномодулируют доставку генов и/или экспрессию генов, или которые иным образом обеспечивают благоприятные свойства для целевых клеток. Как описано и показано более подробно ниже, такие другие компоненты включают, например, компоненты, которые влияют на связывание с клетками или нацеливание на них (в том числе компоненты, которые опосредуют связывание с определенным типом клеток или тканеспецифическое связывание); компоненты, которые влияют на поглощение нуклеиновой кислоты вектора клеткой; компоненты, которые влияют на локализацию полинуклеотида в клетке после поглощения (такие как средства, опосредующие ядерную локализацию) и компоненты, которые влияют на экспрессию полинуклеотида. Такие компоненты также могут включать маркеры, такие как детектируемые и/или селектируемые маркеры, которые можно использовать для обнаружения или отбора клеток, которые поглотили и экспрессируют нуклеиновую кислоту, доставленную вектором. Такие компоненты могут обеспечиваться как естественный признак вектора (как, например, применение определенного вирусного вектора, который имеет компоненты или функциональные элементы, опосредующие связывание и поглощение), или векторы можно модифицировать так, чтобы обеспечить такие функциональные элементы. Другие векторы включают таковые, описанные в Chen et al.; BioTechniques, 34: 167-171 (2003). Широкое разнообразие таких векторов известно из уровня техники и они, в целом, являются доступными.

"Рекомбинантный вирусный вектор" означает вирусный вектор, содержащий один или более гетерологичный генный продукт или последовательность. Поскольку у многих вирусных векторов наблюдаются ограничения по размеру, связанные с упаковкой, гетерологичные генные продукты или последовательности, как правило, встраивают путем замещения одной или более частей вирусного генома. Такие вирусы могут становиться дефектными по репликации, при этом требуется, чтобы удаленный функциональный элемент(ы) обеспечивался в транс-положении во время репликации и инкапсулирования вируса (с использованием, например, вируса-помощника или пакующей клеточной линии, несущей генные продукты, необходимые для репликации и/или инкапсулирования). Также были описаны модифицированные вирусные векторы, в которых полинуклеотид, подлежащий доставке, переносится на внешней поверхности вирусной частицы (см., например, Curiel D.T. et al., PNAS 88: 8850-8854, 1991).

Подходящие системы доставки нуклеиновых кислот включают рекомбинантный вирусный вектор, как правило, последовательность, по меньшей мере, из любого из аденоовириуса, аденоовириус-ассоциированного вируса (AAV), зависимого от помощника аденоовириуса, ретровириуса или комплекса липосомы и гемагглютинирующего вируса японского энцефалита (HVJ). В таких случаях вирусный вектор содержит сильный промотор для экспрессии в эукариотических клетках, функционально связанный с полинуклеотидом, например промотор цитомегаловируса (CMV). Рекомбинантный вирусный вектор может вклю-

чать один или более полинуклеотидов, описанных в данном документе, предпочтительно приблизительно один полинуклеотид. В некоторых вариантах осуществления вирусный вектор, применяемый в способах согласно настоящему изобретению, характеризуется pfu (бляшкообразующие единицы) от приблизительно 10^8 до приблизительно 5×10^{10} pfu. В вариантах осуществления, в которых полинуклеотид должен вводиться с помощью отличного от вирусного вектора, зачастую будет приемлемым применение от приблизительно 0,1 нг до приблизительно 4000 мкг, например от приблизительно 1 нг до приблизительно 100 мкг.

Дополнительные векторы включают вирусные векторы, белки слияния и химические конъюгаты. Ретровирусные векторы включают вирусы лейкоза мышей Молони и вирусы на основе ВИЧ. Один вирусный вектор на основе ВИЧ содержит по меньшей мере два вектора, где гены gag и pol взяты из генома ВИЧ, а ген env взят из другого вируса. ДНК-вирусные векторы включают векторы на основе поксвирусов, такие как векторы на основе ортопоксвирусов или вируса Avipox, герпесвирусные векторы, такие как вектор на основе вируса простого герпеса I (HSV) [Geller A.I. et al., J. Neurochem, 64: 487 (1995); Lim F. et al., в DNA Cloning: Mammalian Systems, D. Glover, Ed, (Oxford Univ. Press, Oxford England) (1995); Geller A.I. et al., Proc Natl. Acad. Sci: U.S.A.:90 7603 (1993); Geller A.I. et al., Proc. Natl. Acad. Sci. USA: 87:1149 (1990)], Adenovirus Vectors [LeGal LaSalle et al., Science, 259:988 (1993); Davidson et al., Nat. Genet. 3: 219 (1993); Yang et al., J. Virol. 69: 2004 (1995)], и векторы на основе аденоассоциированных вирусов [Kaplitt M.G. et al., Nat. Genet. 8:148 (1994)].

Векторы на основе поксвирусов вводят ген в цитоплазму клеток. Векторы на основе вируса Avipox приводят только к краткосрочной экспрессии нуклеиновой кислоты. Аденовирусные векторы, векторы на основе аденоассоциированных вирусов и векторы на основе вируса простого герпеса (HSV) могут быть признаком для некоторых вариантов осуществления согласно настоящему изобретению. Аденовирусные векторы приводят к более краткосрочной экспрессии (например, менее приблизительно месяца), чем векторы на основе аденоассоциированного вируса, которые в некоторых вариантах осуществления могут проявлять значительно более длительную экспрессию. Конкретный выбираемый вектор будет зависеть от целевой клетки и состояния, подлежащего лечению.

Можно легко осуществлять выбор подходящих промоторов. Примером подходящего промотора является промотор цитомегаловируса (CMV) из 763 пар оснований. Другие подходящие промоторы, которые можно применять для экспрессии генов, включают без ограничения промоторы вируса саркомы Рауза (RSV) (Davis et al., Hum. Gene Ther. 4:151 (1993)), участок раннего промотора SV40, промотор тими-динкиназы вируса герпеса, регуляторные последовательности гена металлотионеина (MMT), промоторы прокариотических векторов экспрессии, такие как промотор β -лактамазы, tac-промотор, промоторные элементы дрожжей или других грибов, такие как промотор Gal 4, промотор ADC (алкогольдегидрогеназы), промотор PGK (фосфоглицеринкиназы), промотор щелочной фосфатазы и транскрипционные регуляторные участки животных, которые проявляют тканевую специфичность и были использованы в трансгенных животных: регуляторный участок гена эластазы I, который активен в ацинарных клетках поджелудочной железы, регуляторный участок гена инсулина, который активен в β -клетках поджелудочной железы, регуляторный участок гена иммуноглобулина, который активен в лимфоидных клетках, регуляторный участок вируса опухоли молочной железы мышей, который активен в клетках семенников, молочной железы, лимфоидных клетках и мастоцитах, регуляторный участок гена альбумина, который активен в печени, регуляторный участок гена α -фетопротеина, который активен в печени, регуляторный участок гена α -1-антитрипсина, который активен в печени, регуляторный участок гена β -глобина, который активен в миелоидных клетках, регуляторный участок гена основного белка миелина, который активен в олигодендроцитах в головном мозге, регуляторный участок гена легкой цепи-2 миозина, который активен в скелетных мышцах, и регуляторный участок гена гонадотропного рилизинг-гормона, который активен в гипоталамусе.

Определенные белки можно экспрессировать при помощи их нативных промоторов. Также могут быть включены другие элементы, которые могут повышать экспрессию, такие как энхансер или система, которая приводит к высоким уровням экспрессии, такая как tat-ген и tag-элемент. Эту кассету затем можно встраивать в вектор, например плазмидный вектор, такой как pUC19, pUC118, pBR322, или другие известные плазмидные векторы, которые включают, например, точку начала репликации E.coli. См. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory press (1989). Плазмидный вектор также может включать селектируемый маркер, такой как ген β -лактамазы для устойчивости к ампициллину, при условии, что маркерный полипептид не влияет отрицательно на метаболизм организма, подлежащего обработке. Кассета также может быть связана с фрагментом связывания нукleinовой кислоты в синтетической системе доставки, такой как система, раскрыта в WO 95/22618.

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

Рекомбинантные аденоовирические векторы, дефектные по репликации, можно получать в соответствии с известными методиками. См. Quantin et al., Proc. Natl. Acad. Sci. USA, 89:2581-2584 (1992); Stratford-Perricaudet et al., J. Clin. Invest., 90:626-630 (1992) и Rosenfeld et al., Cell, 68:143-155 (1992).

Другим способом доставки является применение векторов, продуцирующих одноклеточные ДНК, которые могут продуцировать экспрессируемый продукт внутриклеточно. См., например, Chen et al., BioTechniques, 34: 167-171 (2003), который включен в данный документ при помощи ссылки в полном объеме.

Фармацевтические композиции.

Как описано выше, композиции согласно настоящему изобретению можно получать различными способами, известными специалисту в данной области. Независимо от их источника происхождения или способа, которым они получены, композиции согласно настоящему изобретению можно составлять в соответствии с их применением. Например, нуклеиновые кислоты и векторы, описанные выше, можно составлять в композиции для применения по отношению к клеткам в тканевой культуре или для введения пациенту или субъекту. Любую из фармацевтических композиций согласно настоящему изобретению можно составлять для использования в получении лекарственного препарата, и конкретные применения указаны ниже применительно к лечению, например лечение субъекта с ВИЧ-инфекцией или подверженного риску заражения ВИЧ-инфекцией. При применении в качестве фармацевтического препарата любые из нуклеиновых кислот и векторов можно вводить в форме фармацевтических композиций. Такие композиции можно получать способом, хорошо известным в области фармацевтических препаратов, и можно вводить с помощью различных путей введения в зависимости от того, необходимо ли локальное или системное лечение, а также от области, подлежащей лечению. Введение может быть местным (в том числе офтальмологическим и нанесением на слизистые оболочки, в том числе интраназальная, вагинальная и ректальная доставка), легочным (например, при помощи ингаляции или инсуффляции порошков или аэрозолей, в том числе при помощи распылителя; интратрахеальным, интраназальным, эпидермальным и трансдермальным), глазным, пероральным или парентеральным. Способы глазной доставки могут включать местное введение (глазные капли), субконъюнктивальную, периокулярную или интравитреальную инъекцию или внесение при помощи баллонного катетера или офтальмологических вкладок, помещаемых при помощи хирургической операции в конъюнктивальный мешок. Парентеральное введение включает внутривенную, внутриартериальную, подкожную, внутрибрюшинную или внутримышечную инъекцию или вливание; или внутричерепное, например интатекальное, или внутрижелудочковое введение.

Парентеральное введение может быть в форме единичной болюсной дозы, или его можно осуществлять, например, путем продолжительной перфузии при помощи насоса. Фармацевтические композиции и составы для местного введения могут включать трансдермальные пластиры, мази, лосьоны, кремы, гели, капли, суппозитории, распыляемые растворы, жидкости, порошки и т.п. При этом 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 действующего вещества, нуклеиновые кислоты и векторы, описанные в данном документе в комбинации с одним или более фармацевтически приемлемыми носителями. Авторы настоящего изобретения используют термины "фармацевтически приемлемый" (или "фармакологически приемлемый") 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, формообразующие, связующие вещества, скользящие вещества, гели, поверхностно-активные вещества и т.п., которые можно использовать в качестве сред для фармацевтически приемлемого вещества. При получении композиций согласно настоящему изобретению действующее вещество, как правило, смешивают с формообразующим, разбавляют формообразующим или заключают в такой носитель, например, в форме капсулы, таблетки, саше, бумаги, или в другую емкость. В тех случаях, когда формообразующее служит в качестве разбавителя, оно может представлять собой твердый, полутвердый или жидкий материал (например, физиологический раствор), который служит в качестве наполнителя, носителя или среды для действующего вещества. 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 области техники, тип разбавителя может варьировать в зависимости от предполагаемого пути введения. Полученные в результате композиции могут включать дополнительные средства, такие как консерванты. В некоторых вариантах осуществления носитель может быть липидным или полимерным коллоидом или может включать его. В некоторых вариантах осуществления материала носителя может быть коллоидом, составленным в виде липосомы, гидрогеля, микрочастицы, наночастицы или мицеллы из блок-сополимера. Как отмечалось, материал носителя может образовывать капсулу, и такой материал может

быть полимерным коллоидом.

Последовательности нуклеиновых кислот согласно настоящему изобретению могут доставляться в соответствующую клетку субъекта. Этого можно достичь, например, путем применения средства для доставки в виде полимерных биоразлагаемых микрочастиц или микрокапсул, имеющих размер для оптимизации фагоцитоза фагоцитами, такими как макрофаги. Например, можно использовать микрочастицы на основе PLGA (сополимера лактида и гликолида) размером приблизительно 1-10 мкм в диаметре. Полинуклеотид заключают в эти микрочастицы, которые поглощаются макрофагами и постепенно подвергаются биологическому разложению в клетке, высвобождая, таким образом, полинуклеотид. После высвобождения ДНК экспрессируется в клетке. Второй тип микрочастиц не предназначен для непосредственного поглощения клетками, а служит, главным образом, как резервуар с медленным высвобождением нуклеиновой кислоты, которая поглощается клетками только после высвобождения из микрочастицы посредством биологического разложения. Вследствие этого, такие полимерные частицы должны быть достаточно большими, чтобы предотвращать фагоцитоз (т.е. больше 5 мкм и предпочтительно больше 20 мкм). Другим путем, чтобы добиться поглощения нуклеиновой кислоты, является использование липосом, полученных общепринятыми способами. Нуклеиновые кислоты могут быть встроены отдельно в эти средства для доставки или встроены совместно с тканеспецифическими антителами, например антителами, которые целенаправленно воздействуют на типы клеток, которые обычно являются латентно инфицированными резервуарами ВИЧ-инфекции, например макрофаги головного мозга, микроглия, астроциты и клетки лимфоидной ткани кишечника. Альтернативно, можно получить молекулярный комплекс, состоящий из плазмида или другого вектора, прикрепленных к поли-L-лизину при помощи электростатических или ковалентных сил. Поли-L-лизин связывается с лигандом, который может связываться с рецептором на целевых клетках. 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 согласно настоящему изобретению можно составлять в виде наночастицы, например наночастиц, состоящих из ядра из высокомолекулярного линейного полизтиленимина (LPEI) в комплексе с ДНК и окруженного оболочкой из модифицированного полизтиленгликолем (ПЭГилированного) низкомолекулярного LPEI.

Нуклеиновые кислоты и векторы также можно наносить на поверхность устройства (например, катетера), или они могут входить в состав насоса, пластиря или другого устройства для доставки лекарственного средства. Нуклеиновые кислоты и векторы согласно настоящему изобретению можно вводить отдельно или в смеси в присутствии фармацевтически приемлемого формообразующего или носителя (например, физиологического солевого раствора). Формообразующее или носитель выбирают исходя из способа и пути введения. Подходящие фармацевтические носители, а также фармацевтические предписания в отношении применения в фармацевтических составах описаны в Remington's Pharmaceutical Sciences (E.W. Martin), хорошо известном справочном руководстве в данной области, и в USP/NF (Фармакопея Соединенных Штатов Америки и Национальный формуляр).

В некоторых вариантах осуществления композиции можно составлять в виде геля для местного применения для блокирования передачи ВИЧ половым путем. Гель для местного применения можно наносить непосредственно на кожу или слизистые оболочки половой области мужчины или женщины перед половым актом. Альтернативно или дополнительно, гель для местного применения можно наносить на поверхность мужского или женского презерватива или диафрагмы или он может входить в их состав.

В некоторых вариантах осуществления композиции могут быть составлены в виде наночастицы, заключающей нуклеиновую кислоту, кодирующую Cas9 или вариант Cas9, и последовательность направляющей РНК, комплементарную целевому ВИЧ, или вектор, содержащий нуклеиновую кислоту, кодирующую Cas9 и последовательность направляющей РНК, комплементарную целевому ВИЧ. Альтернативно, композиции можно составлять в виде наночастицы, заключающей полипептид CRISPR-ассоциированную эндонуклеазу, например Cas9 или вариант Cas9, и последовательность направляющей РНК, комплементарную мишени.

Составы согласно настоящему изобретению могут охватывать вектор, кодирующий Cas9 и последовательность направляющей РНК, комплементарную целевому ВИЧ. Последовательность направляющей РНК может включать последовательность, комплементарную отдельному участку, например LTR A, B, C или D, или она может включать любую комбинацию последовательностей, комплементарных LTR A, B, C и D. Альтернативно, последовательность, кодирующая Cas9, и последовательность, кодирующая последовательность направляющей РНК, могут находиться в отдельных векторах.

Способы лечения.

Композиции, раскрытие которых в данном документе, в целом и различным образом применимы для лечения субъекта с ретровирусной инфекцией, например ВИЧ-инфекцией. Авторы настоящего изобретения

могут ссылаться на субъекта, пациента или индивидуума взаимозаменяю. Способы применимы для целенаправленного воздействия на любой ВИЧ, например ВИЧ-1, ВИЧ-2, и любую его циркулирующую рекомбинантную форму. Субъект эффективно подвергается лечению во всех случаях, когда наблюдается клинически благоприятный результат. Это может означать, например, полное исчезновение симптомов заболевания, снижение степени тяжести симптомов заболевания или замедление прогрессирования заболевания. Такие способы могут дополнительно включать стадии а) выявления субъекта (например, пациента и, более конкретно, пациента-человека), у которого присутствует ВИЧ-инфекция; и б) назначения субъекту композиции, содержащей нуклеиновую кислоту, кодирующую CRISPR-ассоциированную нуклеазу, например Cas9, и направляющую РНК, комплементарную целевой последовательности ВИЧ, например LTR ВИЧ. Субъекта можно выявлять при помощи общепринятых клинических тестов, например видов иммунологического анализа для определения присутствия антител к ВИЧ или полипептида p24 ВИЧ в сыворотке субъекта, или посредством амплификационных тестов нуклеиновой кислоты 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 согласно настоящему изобретению также могут включать стадию мониторинга для облегчения оптимизации дозирования и составления плана лечения, а также предсказания исхода. В некоторых способах согласно настоящему изобретению сперва можно определить, присутствует ли у пациента латентная ВИЧ-1-инфекция, а затем принять решение относительно того, следует или не следует лечить пациента одной или более композициями, описанными в данном документе. Мониторинг также можно применять для определения появления устойчивости к лекарственному средству и для быстрого различения восприимчивых пациентов и невосприимчивых пациентов. В некоторых вариантах осуществления способы могут дополнительно включать стадию определения последовательности нуклеиновой кислоты конкретного ВИЧ, который несет пациент, и затем конструирование направляющей РНК так, чтобы она была комплементарной таким конкретным последовательностям. Например, можно определить последовательность нуклеиновой кислоты участка U3, R или U5 LTR субъекта и затем сконструировать одну или более направляющих РНК так, чтобы они были в точности комплементарны последовательностям пациента.

Композиции также применимы для лечения, например в качестве профилактического лечения, субъекта, подверженного риску приобретения ретровирусной инфекции, например ВИЧ-инфекции. Такие способы могут дополнительно включать стадии: а) выявления субъекта, подверженного риску приобретения ВИЧ-инфекции; б) назначение субъекту композиции, содержащей нуклеиновую кислоту, кодирующую CRISPR-ассоциированную нуклеазу, например Cas9, и направляющую РНК, комплементарную целевой последовательности ВИЧ, например LTR ВИЧ. Субъектом, подверженным риску приобретения ВИЧ-инфекции, может быть, например, любой ведущий половую жизнь индивидуум, участвующий в незащищенном половом контакте, т.е. участвующий в половом акте без использования презерватива; ведущий половую жизнь индивидуум с другой передаваемой половым путем инфекцией; потребитель внутривенных наркотиков или необрязанный мужчина. Субъектом, подверженным риску приобретения ВИЧ-инфекции, может быть, например, индивидуум, чья профессия может привести его или ее к контакту с ВИЧ-инфицированными людьми, например работники сферы здравоохранения или работники служб экстренного реагирования. Субъектом, подверженным риску приобретения ВИЧ-инфекции, может быть, например, лицо, содержащееся под стражей в исправительном учреждении, или лицо, занятое в секс-индустрии, т.е. индивидуум, который использует половой акт для получения заработка или неденежных благ, таких как пища, наркотики или кровь.

Композиции также можно вводить беременным или кормящим женщинам с ВИЧ-инфекцией для уменьшения вероятности передачи ВИЧ от матери к ее потомству. Беременная женщина, инфицированная ВИЧ, может передавать вирус своему потомству трансплацентарно в утробе, в момент родов через родовой канал или после родов через грудное молоко. Композиции, раскрытые в данном документе, можно вводить ВИЧ-инфицированной матери в один из моментов времени пренатально, перинатально или постнатально во время периода вскармливания грудью, либо в любой комбинации из пренатального, перинатального и постнатального введения. Композиции можно вводить матери совместно с общепринятыми видами антиретровирусной терапии, которые описаны ниже. В некоторых вариантах осуществления композиции согласно настоящему изобретению также вводят ребенку непосредственно после родов и в некоторых вариантах осуществления через определенные промежутки времени впоследствии. Ребенок также может получать общепринятую антиретровирусную терапию.

Способы и композиции, раскрытые в данном документе, применимы для лечения ретровирусных инфекций. Иллюстративные ретровирусы включают вирусы иммунодефицита человека, например ВИЧ-1, ВИЧ-2; вирус иммунодефицита обезьян (SIV); вирус иммунодефицита кошачьих (FIV); вирус иммунодефицита крупного рогатого скота (BIV); вирус инфекционной анемии лошадей (EIAV) и вирус артреоз-цефалита козлов (CAEV). Способы, раскрытые в данном документе, можно применять к широкому спектру видов, например людям, отличным от людей приматам (например, обезьянам), лошадям или другому домашнему скоту, собакам, кошкам, фреткам или другим млекопитающим, содержащимся в качестве домашних животных, крысам, мышам или другим лабораторным животным.

Способы согласно настоящему изобретению могут быть изображены с точки зрения получения лекарственного препарата. Соответственно, настоящее изобретение охватывает применение средств и композиций, описанных в данном документе, при получении лекарственного препарата. Соединения, описанные в данном документе, применимы в терапевтических композициях и схемах лечения или для производства лекарственного препарата для применения в лечении заболеваний или состояний, которые описаны в данном документе.

Любую композицию, описанную в данном документе, можно вводить в любую часть тела организма-хозяина для последующей доставки к целевой клетке. Композиция может быть доставлена без ограничения в головной мозг, спинномозговую жидкость, суставы, слизистую оболочку носа, кровь, легкие, кишечный тракт, мышечные ткани, кожу или брюшную полость млекопитающих. С точки зрения путей доставки композицию можно вводить при помощи внутривенной, внутричерепной, внутрибрюшинной, внутримышечной, подкожной, внутримышечной, внутрипрямокишечной, внутривлагалищной, интракальмной, внутритеракеальной, внутрикожной или трансдермальной инъекции, при помощи перорального или назального введения, или при помощи постепенной перфузии в течение некоторого времени. В дополнительном примере аэрозольный препарат композиции можно давать организму-хозяину при помощи ингаляции.

Необходимая дозировка будет зависеть от пути введения, природы состава, природы болезни пациента, размера, веса, площади поверхности, возраста и пола пациента, других лекарственных средств, подлежащих введению, и оценки лечащего практикующего врача. Следует ожидать большой изменчивости в необходимой дозировке с учетом разнообразия клеточных мишней и отличающейся эффективности разных путей введения. Изменчивость этих уровней дозировки можно корректировать при помощи общепринятых эмпирических процедур для оптимизации, которые широко распространены в данной области техники. Введения могут быть однократными или многократными (например, 2 или 3, 4, 6, 8, 10, 20, 50, 100, 150 или более раз). Инкапсуляция соединений в подходящее средство для доставки (например, полимерные микрочастицы или имплантируемые устройства) может повысить эффективность доставки.

Продолжительность лечения при помощи любой композиции, предусмотренной в данном документе, может составлять любой промежуток времени от такого короткого, как один день, до такого длинного, как все время жизни организма-хозяина (например, много лет). Например, соединение можно вводить один раз в неделю (например, в течение от 4 недель до многих месяцев или лет); один раз в месяц (например, в течение от трех до двенадцати месяцев или в течение многих лет) или один раз в год в течение периода в 5, 10 лет или дольше. Также следует отметить, что частоту лечения можно варьировать. Например, соединения согласно настоящему изобретению можно вводить один раз (или два раза, три раза и т.д.) в день, в неделю, в месяц или в год.

Эффективное количество любой композиции, предусмотренной в данном документе, можно вводить индивидууму, нуждающемуся в лечении. Используемый в данном документе термин "эффективный" означает любое количество, которое индуцирует необходимый ответ и при этом не связано с выраженной токсичностью для пациента. Такое количество можно определить путем оценки реакции пациента после введения известного количества конкретной композиции. Кроме того, уровень токсичности, если она присутствует, можно определять путем оценки клинических симптомов пациента перед и после введения известного количества конкретной композиции. Следует отметить, что эффективное количество конкретной композиции, вводимой пациенту, можно корректировать в соответствии с необходимым исходом, а также реакцией пациента и уровнем токсичности. Значимая токсичность может варьироваться в зависимости от каждого конкретного пациента и зависит от множества факторов, в том числе без ограничения болезненного состояния, возраста пациента и переносимости побочных эффектов.

Для определения того, индуцируется ли определенная реакция, можно применять любой способ, известный специалистам в данной области. Для определения того, индуцируется ли реакция, можно применять клинические способы, при помощи которых можно оценивать степень конкретного болезненного состояния. Определенные способы, применяемые для оценки реакции, будут зависеть от природы нарушения пациента, возраста и пола пациента, других лекарственных средств, подлежащих введению, и оценки лечащего практикующего врача.

Композиции также можно вводить с другим терапевтическим средством, например антиретровирусным средством, используемым в НААРТ. Иллюстративные антиретровирусные средства включают ингибиторы обратной транскриптазы (например, нуклеозидные/нуклеотидные ингибиторы обратной транскриптазы, зидовудин, эмтрицитабин, ламивудин и тенофовир; и ненуклеозидные ингибиторы обратной транскриптазы, такие как эфавиренз, невирапин, рилпивирин); ингибиторы протеазы, например тициравир, дарунавир, индинавир; ингибиторы проникновения в клетку, например маравирок; ингибиторы слияния, например энфувиртид; или ингибиторы интегразы, например ралтегравир, долутегравир. Иллюстративные антиретровирусные средства также могут включать средства в виде комбинаций из средств различных классов, например комбинации эмтрицитабина, эфавиренза и тенофовира; комбинации эмтрицитабина, рилпивирина и тенофовира или комбинации элвитетгравира, кобицистата, эмтрицитабина и тенофовира.

Совместное введение двух или более терапевтических средств не требует того, чтобы эти средства были введены в одно и то же время или тем же путем, поскольку имеется перекрытие в периоде времени, в течение которого средства проявляют свой терапевтический эффект. Предусматривают одновременное или последовательное введение, как и введение в различные дни или недели.

Терапевтические средства можно вводить при периодической схеме лечения, например продолжительное введение низких доз терапевтического средства.

Дозу, токсичность и терапевтическую эффективность таких композиций можно определять при помощи общепринятых фармацевтических процедур на клеточных культурах или экспериментальных животных, например для определения LD₅₀ (дозы, летальной для 50% популяции) и ED₅₀ (дозы, терапевтически эффективной для 50% популяции). Соотношение доз токсического и терапевтического эффектов представляет собой терапевтический индекс, и его можно выразить как соотношение LD₅₀/ED₅₀.

Данные, полученные в результате анализов с использованием клеточных культур и исследований на животных, можно использовать для подбора диапазона дозировки для использования у человека. Дозировка таких композиций предпочтительно находится в диапазоне циркулирующих концентраций, которые предусматривают ED₅₀ с малой токсичностью или без нее. Дозировка может варьировать в пределах данного диапазона в зависимости от применяемой лекарственной формы и используемого пути введения. Для любой композиции, применяемой в способе согласно настоящему изобретению, терапевтически эффективная доза может быть рассчитана изначально по результатам анализов с использованием клеточных культур. Дозу можно подбирать на животных моделях для достижения диапазона циркулирующих в плазме концентраций, который включает IC₅₀ (т.е. концентрацию исследуемого соединения, при которой достигается полумаксимальное ингибирование симптомов), как определено на клеточной культуре. Такую информацию можно использовать для более точного определения доз, применимых у человека. Уровни в плазме можно измерять, например, при помощи высокоэффективной жидкостной хроматографии.

Как описано, терапевтически эффективное количество композиции (т.е. эффективная дозировка) означает количество, достаточное для получения терапевтически (например, клинически) необходимого результата. Композиции можно вводить от одного или более раз в день до одного или более раз в неделю; в том числе один раз в два дня. Специалисту в данной области будет понятно, что определенные факторы могут влиять на дозировку и время введения, необходимые для эффективного лечения субъекта, в том числе без ограничения степень тяжести заболевания или нарушения, предшествующие виды лечения, общее состояние здоровья и/или возраст субъекта и другие присутствующие заболевания. Более того, лечение субъекта терапевтически эффективным количеством композиций согласно настоящему изобретению может включать единичный курс лечения или серию курсов лечения.

Композиции, описанные в данном документе, подходят для применения с целым рядом систем доставки лекарственных средств, описанных выше. Дополнительно, чтобы увеличить *in vivo* время полу жизни в сыворотке вводимого соединения, композиции можно инкапсулировать, вводить в полость липосом, получать в виде коллоида, или можно применять другие традиционные методики, которые приводят к продленному времени полу жизни в сыворотке у этих композиций. Ряд способов доступен для получения липосом, как описано, например, в Szoka et al., патентах США № 4235871, 4501728 и 4837028, каждый из которых включен в данный документ при помощи ссылки. Кроме того, можно вводить лекарственное средство в системе направленной доставки лекарственных средств, например в липосоме, покрытой тканеспецифическим антителом. Липосомы будут нацеливаться на орган и избирательно поглощаться им.

Также предусмотрены способы инактивации ретровируса, например лентивируса, такого как вирус иммунодефицита человека, вирус иммунодефицита обезьян, вирус иммунодефицита кошачьих или вирус иммунодефицита крупного рогатого скота, в клетке млекопитающего. Вирусом иммунодефицита человека может быть ВИЧ-1 или ВИЧ-2. Вирусом иммунодефицита человека может быть интегрированный в хромосомы провирус. Клеткой млекопитающего может быть любой тип клеток, инфицированных ВИЧ, в том числе без ограничения CD4+-лимфоциты, макрофаги, фибробласти, моноциты, Т-лимфоциты, В-лимфоциты, естественные киллеры, дендритные клетки, такие как клетки Лангерганса и фолликулярные дендритные клетки, гематopoэтические стволовые клетки, эндотелиоциты, клетки микроглии головного мозга и эпителиальные клетки желудочно-кишечного тракта. Такие типы клеток включают те типы клеток, которые, как правило, инфицируются во время первичной инфекции, например CD4+-лимфоцит, макрофаг или клетка Лангерганса, а также те типы клеток, которые составляют резервуары латентного ВИЧ, т.е. латентно инфицированная клетка.

Способы могут включать воздействие на клетку с помощью композиции, содержащей выделенную нуклеиновую кислоту, кодирующую комплекс редактирования генов, содержащий CRISPR-ассоциированную эндонуклеазу и одну или более направляющих РНК, где направляющая РНК комплементарна целевой последовательности нуклеиновой кислоты в ретровирусе. Стадия привидения в контакт может происходить *in vivo*, т.е. композиции можно вводить непосредственно субъекту с ВИЧ-инфекцией. Однако способы не ограничены таким образом, и стадия приведения в контакт может происходить *ex vivo*. Например, у субъекта с ВИЧ-инфекцией можно взять клетку, или некоторое количество клеток, или эксплантат ткани и поместить в культуру, а затем привести в контакт с композицией, содер-

жащей CRISPR-ассоциированную эндонуклеазу и направляющую РНК, где направляющая РНК комплементарна последовательности нукleinовой кислоты в вирусе иммунодефицита человека. Как описано выше, композиция может представлять собой нукleinовую кислоту, кодирующую CRISPR-ассоциированную эндонуклеазу и направляющую РНК, где направляющая РНК комплементарна последовательности нукleinовой кислоты в вирусе иммунодефицита человека; вектор экспрессии, содержащий последовательность нукleinовой кислоты; или фармацевтическую композицию, содержащую нукleinовую кислоту, кодирующую CRISPR-ассоциированную эндонуклеазу и направляющую РНК, где направляющая РНК комплементарна последовательности нукleinовой кислоты в вирусе иммунодефицита человека; или вектор экспрессии, содержащий последовательность нукleinовой кислоты. В некоторых вариантах осуществления комплекса редактирования генов может содержать полипептид CRISPR-ассоциированный эндонуклеазы и направляющую РНК, где направляющая РНК комплементарна последовательности нукleinовой кислоты в вирусе иммунодефицита человека.

Независимо от того, вводят ли композиции в виде нукleinовых кислот или полипептидов, их составляют таким образом, чтобы содействовать поглощению клеткой млекопитающего. Применимые системы векторов и составы описаны выше. В некоторых вариантах осуществления вектор может доставлять композиции к специальному типу клеток. Однако настоящее изобретение не ограничено таким образом, и также предполагаются другие способы доставки ДНК, такие как химическая трансфекция, с использованием, например, фосфата кальция, DEAE-декстрина, липосом, липоплексов, поверхностно-активных веществ и перфторуглеродных жидкостей, равно как и способы физической доставки, такие как электропорация, микроАІъекция, применение баллистических частиц и систем "генной пушки".

Общепринятые способы, например виды иммунологического анализа для выявления CRISPR-ассоциированной эндонуклеазы или анализы на основе нукleinовых кислот, такие как ПЦР, для выявления gRNA, можно применять для подтверждения того, что комплекс был поглощен и экспрессируется клеткой, в которую он был введен. Сконструированные клетки, затем можно повторно вводить субъекту, от которого они были получены, как описано ниже.

Комплекс редактирования генов содержит CRISPR-ассоциированную нуклеазу, например Cas9, и направляющую РНК, комплементарную целевой последовательности ретровируса, например целевой последовательности ВИЧ. Комплекс редактирования генов может вносить разнообразные мутации в привирусную ДНК. Механизм, при помощи которого такие мутации инактивируют вирус, может варьировать, например, мутация может влиять на репликацию привируса, экспрессию вирусного гена или вырезание привируса. Мутации могут быть расположены в регуляторных последовательностях или последовательностях структурных генов и приводить к нарушенной продукции ВИЧ. Мутация может включать делецию. Размер делеции может варьировать от одной пары нуклеотидных оснований до приблизительно 10000 пар оснований. В некоторых вариантах осуществления делеция может охватывать всю или практически всю последовательность привируса. В некоторых вариантах осуществления делеция может охватывать полную последовательность привируса. Мутация может включать вставку, которая представляет собой добавление одной или более пар нуклеотидных оснований в последовательность привируса. 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 одиночного 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.

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. Т.е. клетки субъекта можно забирать из организма и обрабатывать с помощью композиций в культуре для вырезания последовательностей ВИЧ, и обработанные клетки возвращают в организм субъекта. Клетка может представлять собой клетки субъекта, или они могут быть клетками с совпадающим гаплотипом или линией клеток. Клетки могут быть подвергнуты облучению для предупреждения репликации. В некоторых вариантах осуществления клетки представляют собой аутологичные линии клеток с совпадающими человеческими лейкоцитарными антигенами (HLA) или их комбинации. В других вариантах осуществления клетки могут представлять собой стволовую клетку. Например, эмбриональную стволовую клетку или искусственную плорипotentную стволовую клетку (индивидуированную плорипotentную стволовую клетку (iPS-клетку)). Эмбриональные стволовые клетки (ES-клетки) и искусственные плорипotentные стволовые клетки (индивидуированные плорипotentные стволовые клетки, iPS-клетки) были получены из многих видов животных, в том числе человека. Эти типы плорипotentных стволовых клеток будут наиболее применимым источником клеток для регенеративной медицины, поскольку эти клетки способны к дифференциации практически во все органы путем соответствующей индукции их дифференциации, при сохранении их способности к активному делению, пока поддерживается их плорипотентность. iPS-клетки, в частности, могут быть получены из соматических клеток собственного организма и, таким образом, не будут, вероятно, вызывать проблем с точки зрения этики и социальных вопросов в сравнении с ES-клетками, кото-

рые получают путем разрушения эмбрионов. Кроме того, iPS-клетки, которые являются собственными клетками, позволяют избежать реакций отторжения, которые являются наибольшей проблемой в регенеративной медицине или трансплантационной терапии.

Кассету экспрессии gRNA можно с легкостью доставлять в субъекта способами, известными в данной области техники, например способами, при помощи 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 будут разрабатываться стволовые клетки или iPS-клетки, стабильно экспрессирующие Cas9/мультиплекс gRNA, для терапии с помощью трансплантации клеток, а также вакцинации против ВИЧ-1.

Трансдуцированные клетки получают для реинфузии в соответствии с принятыми способами. После культивирования в течение периода приблизительно 2-4 недели количество клеток может составлять от 1×10^6 до 1×10^{10} . При этом ростовые характеристики клеток варьируют от пациента к пациенту и от типа клеток к типу клеток. Приблизительно за 72 ч до реинфузии трансдуцированных клеток отбирали аликвоты для анализа фенотипа и определения процентной доли клеток, экспрессирующих терапевтическое средство. Что касается введения, клетки согласно настоящему изобретению можно вводить при норме, определенной по LD₅₀ для типа клеток и побочные эффекты 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 ограничения костный мозг или виды жировой ткани.

Готовые изделия.

Композиции, описанные в данном документе, могут быть упакованы в подходящие емкости с этикетками, например, для применения в качестве терапевтического средства для лечения субъекта с ретровирусной инфекцией, например ВИЧ-инфекцией, или субъекта, подверженного риску заражения ретровирусной инфекцией, например ВИЧ-инфекцией.

Емкости могут включать композицию, содержащую последовательность нукleinовой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, например эндонуклеазу Cas9, и направляющую РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, или вектор, кодирующий такую нукleinовую кислоту, и один или более из подходящих стабилизатора, молекулыносителя, ароматизатора и/или т.п. в зависимости от предполагаемого применения. Соответственно, упакованные продукты (например, стерильные емкости, содержащие одну или более композиций, описанных в данном документе, и упакованные для хранения, перевозки или продажи в концентрированном виде или в виде готовых к употреблению концентраций) и наборы, включающие по меньшей мере одну композицию согласно настоящему изобретению, например последовательность нукleinовой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, например эндонуклеазу Cas9, и направляющую РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, или вектор, кодирующий такую нукleinовую кислоту, и инструкции по применению, также находятся в пределах объема настоящего изобретения. Продукт может включать емкость (например, пузырек, банку, бутылку, пакет и т.п.), содержащую одну или более композиций согласно настоящему изобретению. Кроме того, готовое изделие дополнительно может включать, например, упаковочные материалы, инструкции по применению, шприцы, устройства для доставки, буферы или другие контрольные реагенты, для лечения или мониторинга состояния, которое требует профилактики или лечения.

В некоторых вариантах осуществления наборы могут включать одно или более дополнительных антиретровирусных средств, например ингибитор обратной транскриптазы, ингибитор протеазы или ингибитор проникновения в клетку. Дополнительные средства могут быть упакованы совместно в той же емкости, что и последовательность нукleinовой кислоты, кодирующая CRISPR-ассоциированную эндонуклеазу, например эндонуклеазу Cas9, и направляющую РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, или вектор, кодирующий такую нукleinовую кислоту, или они могут быть упакованы отдельно. Последовательность нукleinовой кислоты, кодирующую CRISPR-ассоциированную эндонуклеазу, например эндонуклеазу Cas9, и направляющую РНК, комплементарную целевой последовательности в вирусе иммунодефицита человека, или вектор, кодирующий такую нукleinовую кислоту, и дополнительное средство можно объединять непосредственно перед применением или вводить отдельно.

Продукт также может включать пояснения (например, напечатанную этикетку, или листок-вкладыш, или другой носитель, описывающий применение продукта (например, аудио- или видеокассету)). Пояснения могут быть объединены с емкостью (например, прикреплены к емкости), и в них может описываться способ, при помощи которого находящиеся там композиции следует вводить (например, частота и путь введения), показания для них и другие применения. Композиции могут быть готовыми для введения (например, находиться в виде единиц, соответствующих дозе) и могут включать один или

более дополнительных фармацевтически приемлемых вспомогательных средств, носителей или других разбавителей и/или дополнительное терапевтическое средство. Альтернативно, композиции могут предусматриваться в концентрированной форме с разбавителем и инструкциями для разбавления.

Примеры

Пример 1. Материалы и способы.

Получение плазмид. Векторы, содержащие кассету экспрессии Cas9 и gRNA в клетках человека, pX260 и pX330 (Addgene), использовали для создания разнообразных конструкций, LTR-A, B, C и D.

Клеточная культура и стабильные линии клеток. Линию репортерных клеток TZM-bl и линию клеток U1 получали из NIH AIDS Reagent Program, а клетки микроглии CHME5 известны в данной области техники.

Иммуногистохимия и вестерн-блоттинг. Использовали общепринятые способы иммуноцитохимического исследования клеток и оценки экспрессии белка при помощи вестерн-блоттинга.

Анализ с использованием люциферазы светлячка. Клетки лизировали через 24 ч после обработки с использованием Passive Lysis Buffer (Promega) и анализировали при помощи набора Luciferase Reporter Gene Assay (Promega) в соответствии с протоколом производителя. Люциферазную активность нормализовали относительно количества клеток, определенного при помощи параллельного MTT-анализа (Vybrant, Invitrogen).

ELISA p24. После инфицирования или реактивации уровня вирусной нагрузки ВИЧ-1 в надосадочных жидкостях количественно определяли при помощи ELISA Gag p24 (Advanced BioScience Laboratories, Inc.), следуя протоколу производителя. Для оценки жизнеспособности клеток после обработок параллельно проводили MTT-анализ в соответствии с инструкцией эксплуатации от производителя (Vybrant, Invitrogen).

Проточная цитометрия в отношении EG.FP. Клетки трипсинизировали, промывали PBS и фиксировали в 2% параформальдегиде в течение 10 мин при комнатной температуре, затем промывали дважды PBS и анализировали с помощью проточного цитометра Guava EasyCyte Mini (Guava Technologies).

Получение репортерного вируса ВИЧ-1 и инфицирование. Клетки HEK293T трансфицировали с помощью pNL4-3-ΔE-EGFP (NIH AIDS Research and Reference Reagent Program) с использованием реагента Lipofectamine 2000 (Invitrogen). Через 48 ч надосадочную жидкость собирали, фильтровали через 0,45 мкм фильтр и определяли титры на клетках HeLa с использованием EGFP в качестве маркера инфицирования. Для инфицирования вирусом стабильные клетки TZM-bl с Cas9/gRNA инкубировали 2 ч с разбавленным исходным раствором вируса и затем промывали дважды PBS. В день 2 и 4 после инфицирования клетки собирали, фиксировали и анализировали при помощи проточной цитометрии в отношении экспрессии EGFP или осуществляли очистку геномной ДНК для ПЦР и полногеномного секвенирования.

Амплификация геномной ДНК, ПЦР, TA-клонирование и секвенирование по Сэнгеру, ПЦР GenomeWalker link. Использовали общепринятые способы манипуляций с ДНК для клонирования и секвенирования. Для идентификации сайтов интеграции ВИЧ-1 использовали набор Lenti-X™ Integration Site Analysis.

Анализ Surveyor. Присутствие мутаций в ПЦР-продуктах проверяли при помощи SURVEYOR Mutation Detection Kit (Transgenomic) в соответствии с протоколом от производителя. Вкратце, гетерогенный ПЦР-продукт денатурировали в течение 10 мин при 95°C и гибридизировали при помощи постепенного охлаждения с использованием термоциклира. Затем 300 нг гибридизированной ДНК (9 мкл) подвергали расщеплению при помощи 0,25 мкл нуклеазы SURVEYOR в присутствии 0,25 мкл SURVEYOR Enhancer S и 15 мМ MgCl₂ в течение 4 ч при 42°C. Затем добавляли стоп-раствор и образцы разделяли в 2% агарозном геле вместе с равными количествами контролен в виде нерасщепленных ПЦР-продуктов.

Некоторые ПЦР-продукты использовали для анализа полиморфизма длин рестрикционных фрагментов. Равные количества ПЦР-продуктов расщепляли при помощи BsaI. Расщепленную ДНК разделяли на агарозном геле, содержащем этидия бромид (2%). Для секвенирования ПЦР-продукты клонировали при помощи TA Cloning® Kit Dual Promoter с использованием вектора pCR™II (Invitrogen). Вставку подтверждали расщеплением с помощью EcoRI и положительные клоны посыпали в Genewiz для секвенирования по Сэнгеру.

Отбор целевых сайтов LTR, полногеномное секвенирование и биоинформационический и статистический анализ. Для первоначальной идентификации потенциальных целевых сайтов в LTR использовали инструмент Jack Lin's CRISPR/Cas9 gRNA finder.

Получение плазмиды. Сегмент ДНК, экспрессирующий LTR-A или LTR-B для pre-crRNA, клонировали в вектор pX260, который содержит ген для отбора с пуромицином (Addgene, плазмиды № 42229). Сегмент ДНК, экспрессирующий LTR-C или LTR-D для химерной crRNA-tracrRNA, клонировали в вектор pX330 (Addgene, плазмиды № 42230). Оба вектора содержали кодирующую последовательность гуманизированного Cas9 под управлением промотора CAG и кассету экспрессии gRNA под управлением промотора U6 человека. Векторы расщепляли при помощи BbsI и обрабатывали антарктической фосфатазой и линеаризованный вектор очищали при помощи набора для быстрого удаления нуклеотидов (Qi-

agen). Пару нуклеотидов для каждого сайта нацеливания (фиг. 14, AlphaDNA) гибридизировали, фосфорилировали и лигировали в линеаризованный вектор. Кассету экспрессии gRNA секвенировали с праймером для секвенирования U6 (фиг. 14) в GENEWIZ. Для векторов pX330 конструировали пару универсальных ПЦР-праймеров с перекрывающимися сайтами расщепления (фиг. 14), при помощи которых вырезали кассету экспрессии gRNA (U6-gRNA-crRNA-стебель-tracrRNA) для непосредственной трансфекции или субклонирования в другие векторы.

Клеточная культура. Линию репортреных клеток TZM-bI от Dr. John C. Kappes, Dr. Xiaoyun Wu и Tranzyme Inc, линию клеток U1/Hiv-1 от Dr. Thomas Folks и клон J-Lat полной длины от Dr. Eric Verdin получали благодаря NIH AIDS Reagent Program, подразделение по изучению СПИДа, NIAID, NIH. Линию фетальных клеток микроглии CHME5/HIV получали как описано выше. Клетки TZM-bI и CHME5 культивировали в минимальной питательной среде Дульбекко с высоким содержанием глюкозы, дополненной 10% термоинактивированной фетальной бывшей сыворотки (FBS) и 1% пенициллина/стрептомицина. Клетки U1 и J-Lat культивировали в RPMI 1640, содержащей 2,0 мМ L-глутамина, 10% FBS и 1% пенициллина/стрептомицина.

Стабильные линии клеток и субклонирование. Клетки TZM-bI или CHME5/HIV высевали в 6-луночные планшеты при плотности $1,5 \times 10^5$ клеток/лунка и трансфицировали с использованием реагента Lipofectamine 2000 (Invitrogen) с помощью 1 мкг плазмиды pX260 (для LTR-A и B) или 1 мкг/0,1 мкг плазмид pX330/pX260 (для LTR-C и D). На следующий день клетки переносили в 100-мм чашки и инкубировали в среде для выращивания, содержащей 1 мкг/мл пуромицина (Sigma). Через две недели выделяли выжившие клеточные колонии при помощи цилиндров для клонирования (Comring). Клетки U1 ($1,5 \times 10^5$) подвергали электропорации с 1 мкг ДНК при помощи 10 мкЛ наконечника, импульсов 3×10 мс 1400 В на системе для трансфекции The NeonTM (Invitrogen).

Клетки отбирали с помощью 0,5 мкг/мл пуромицина в течение двух недель. Стабильные клоны субкультивировали при помощи метода предельного разведения в 96-луночных планшетах и полученные из одной клетки субклоны сохраняли для дальнейших исследований.

Иммуноцитохимия и вестерн-блоттинг. Клетки TZM-bI, стабильно экспрессирующие Cas9/gRNA, культивировали в 8-луночных предметных стеклах в течение 2 дней и фиксировали в течение 10 мин в 4% параформальдегиде/PBS. После трех промывок клетки обрабатывали 0,5% Triton X-100/PBS в течение 20 мин и блокировали в 10% сыворотке осла в течение 1 ч. Клетки инкубировали на протяжении ночи при 4°C с первичным антителом мыши к Flag M2 (1:500, Sigma). После промывания три раза клетки инкубировали в течение 1 ч со вторичными антителами осла к антителу мыши с Alexa-Fluor-594 и инкубировали с Hoechst 33258 в течение 5 мин. После трех промывок PBS клетки накрывали покровными стеклами с применением водной среды для заключения с добавками против выцветания (Biomedica) и анализировали под флуоресцентным микроскопом Leica DMI6000B.

Клетки TZM-bI, культивируемые в 6-луночном планшете, солюбилизировали в 200 мкЛ буфера для лизиса на основе Triton X-100, содержащего 20 мМ Tris-HCl (pH 7,4), 1% Triton X-100, 5 мМ этилендиаминететраусусной кислоты, 5 мМ дитиотреитола, 150 мМ NaCl, 1 мМ фенилметилсульфонилфторида, 1× смесь ингибиторов протеиназ для экстракта ядер (Cayman Chemical, Анн-Арбор, Мичиган), 1 мМ орто-ванадата натрия и 30 мМ NaF. Клеточные лизаты перемешивали в ротаторе при 4°C в течение 30 мин. Ядерный и клеточный дебрис очищали путем центрифугирования при 20000g в течение 20 мин при 4°C. Равные количества белков из лизата (20 мкг) денатурировали путем кипячения в течение 5 мин в буфере для образцов с додецилсульфатом натрия (SDS), разделяли на фракции при помощи электрофореза в SDS-полиакриламидном геле в Tris-глициновом буфере и переносили на нитроцеллюлозную мембрану (BioRad). Стандарты, предварительно окрашенные SeeBlue (Invitrogen), использовали в качестве эталона молекулярной массы. Блоты блокировали в 5% BSA/Tris-буферизованном солевом растворе (pH 7,6) и 0,1% Tween-20 (TBS-T) в течение 1 ч и затем инкубировали на протяжении ночи при 4°C с моноклональным антителом мыши к Flag M2 (1:1000, Sigma) или моноклональным антителом мыши к GAPDH (1:3000, Santa Cruz Biotechnology). После отмычки TBS-T блоты инкубировали с коньюгированным с IRDye 680LT антителом к мыши в течение 1 ч при комнатной температуре. Мембранны сканировали и анализировали при помощи инфракрасной системы формирования изображений Odyssey (LI-COR Biosciences).

Анализ с использованием люциферазы светлячка. Клетки лизировали через 24 ч после обработки с использованием Passive Lysis Buffer (Promega) и анализировали при помощи набора Luciferase Reporter Gene Assay (Promega) в соответствии с протоколом производителя. Люциферазную активность нормализовали относительно количества клеток, определенного при помощи параллельного MTT-анализа (Vybrant, Invitrogen).

ELISA p24. После инфицирования или реактивации уровня вирусной нагрузки ВИЧ-1 в надосадочных жидкостях количественно определяли при помощи ELISA Gag p24 (Advanced BioScience Laboratories, Inc.), следяя протоколу производителя. Для оценки жизнеспособности клеток после обработок параллельно проводили MTT-анализ в соответствии с протоколом производителя (Vybrant, Invitrogen).

Проточная цитометрия в отношении EGFP. Клетки трипсинизировали, промывали PBS и фиксировали в 2% параформальдегиде в течение 10 мин при комнатной температуре, затем промывали дважды PBS и анализировали с помощью проточного цитометра Guava EasyCyte Mini (Guava Technologies).

Получение репортерного вируса ВИЧ-1 и инфицирование. Клетки HEK293T трансфицировали с помощью pNL4-3-ΔE-EGFP, SF162 и JRFL (NIH AIDS Research and Reference Reagent Program) с использованием реагента Lipofectamine 2000 (Invitrogen). Для псевдотипирования pNL4-3-ΔE-EGFP совместно трансфицировали вектор VSVG. Через 48 ч надосадочную жидкость собирали, фильтровали через 0,45 мкм фильтр и определяли титры на клетках HeLa с использованием экспрессируемого EGFP в качестве маркера инфицирования. Для инфицирования вирусом стабильные клетки TZM-bl с Cas9/gRNA инкубировали 2 ч с разбавленным исходным раствором вируса и промывали дважды PBS. В дни 2 и 4 после инфицирования клетки собирали, фиксировали и анализировали при помощи проточной цитометрии в отношении экспрессии EGFP или осуществляли очистку геномной ДНК для ПЦР и полногеномного секвенирования.

Очистка геномной ДНК, ПЦР, ТА-клонирование и секвенирование по Сэнгеру. Геномную ДНК выделяли из клеток при помощи набора для очистки ДНК из клетки/ткани ArchivePure (5PRIME) в соответствии с протоколом, рекомендованным производителем. 100 нг экстрагированной ДНК подвергали ПЦР с использованием набора для высокоточной ПЦР FailSafe (Epicentre) при помощи праймеров, перечисленных на фиг. 14. Три стадии стандартной ПЦР осуществляли в течение 30 циклов с гибридизацией при 55°C и элонгацией при 72°C. Продукты разделяли в 2% агарозном геле. Полосы, представляющие интерес, очищали от геля и клонировали в вектор pCRII T-A (Invitrogen) и нуклеотидную последовательность отдельных клонов определяли при помощи секвенирования в Genewiz с использованием универсальных праймеров T7 и/или SP6.

Традиционная ПЦР и ПЦР с обратной транскрипцией (RT-PCR) в режиме реального времени. Для экстракции общей РНК клетки подвергали обработке при помощи набора RNeasy Mini kit (Qiagen) согласно инструкциям производителя. Потенциальную остаточную геномную ДНК удаляли благодаря расщеплению в колонке с использованием ДНКазы при помощи набора RNase-Free DNase Set (Qiagen). Один мкг РНК для каждого образца подвергали обратной транскрипции в кДНК с использованием случайных гексануклеотидных праймеров при помощи High Capacity cDNA Reverse Transcription Kit (Invitrogen. Гранд-Айленд, Нью-Йорк). Традиционную ПЦР осуществляли с использованием стандартного протокола.

Анализы количественной ПЦР (qPCR) проводили в LightCycler480 (Roche) с использованием набора SYBR® Green PCR Master Mix (Applied Biosystems). RT-реакции разводили до 5 нг общей РНК на микролитр реакционных смесей и 2 мкл использовали в 20-мкл ПЦР-реакции. Для qPCR-анализа провироусов ВИЧ-1 использовали 50 нг геномной ДНК. Праймеры синтезировали в AlphaDNA, и они показаны на фиг. 14. Праймеры для генов домашнего хозяйства GAPDH и RPL13A человека получали от Real-TimePrimers (Elkins Park, Пенсильвания). Каждый образец исследовали в трех повторностях. Значения порогового цикла (Ct) получали графически для целевых генов и генов домашнего хозяйства. Разность Ct-значений для гена домашнего хозяйства и целевого гена выражали как ΔCt-значения. ΔΔCt-значения получали путем вычитания ΔCt-значений контрольных образцов из таковых экспериментальных образцов. Относительные кратное или процентное изменение рассчитывали как 2-ΔΔCt. В некоторых случаях абсолютный количественный анализ выполняли при помощи геномной ДНК человека с введенной плазмидой pNL4-3-ΔE-EGFP в качестве стандарта. Количество вирусных копий ВИЧ-1 рассчитывали, исходя из калибровочной кривой после нормализации относительно гена домашнего хозяйства.

ПЦР Genome Walker link и ПЦР длинных фрагментов. Сайты интеграции ВИЧ-1 в клетках-хозяевах идентифицировали при помощи набора Lenti-X™ Integration Site Analysis (Clontech), следуя инструкциям производителя. Вкратце, геномные ДНК высокого качества экстрагировали из клеток U1 при помощи NucleoSpin Tissue (Clontech). Для конструирования библиотек сайтов интеграции вирусов каждый образец геномной ДНК расщепляли с помощью образующих тупой конец ферментов расщепления Dra I, Ssp I или HpaI отдельно на протяжении ночи при 37°C. Эффективность расщепления подтверждали при помощи электрофореза в 0,6% агарозе. Расщепленную ДНК очищали при помощи геля NucleoSpin и набора PCR Clean-Up с последующим лигированием расщепленных фрагментов геномной ДНК в адаптер GenomeWalker™ при 16°C на протяжении ночи. Реакцию лигирования останавливали путем инкубации при 70°C в течение 5 мин и разведения в 5 раз TE-буфером. Первичную ПЦР выполняли на ДНК-сегментах при помощи праймера-адаптера 1 (API) и LTR-специфичного праймера 1 (LSP1) с использованием Advantage 2 Polymerase Mix, за которой следовала вторичная (гнездовая) ПЦР при помощи праймеров AP2 и LSP2 (фиг. 14). Продукты вторичной ПЦР разделяли на агарозном геле, содержащем 1,5% этидия бромида. Главные полосы очищали от геля и клонировали в вектор pCRII T-A (Invitrogen), и нуклеотидную последовательность отдельных клонов определяли при помощи секвенирования в Genewiz с использованием универсальных праймеров T7 и SP6. Риды последовательностей анализировали при помощи BLAST-поиска в NCBI. В клетках U1 идентифицировали два сайта интеграции ВИЧ-1 в хромосомах X и 2. Пару праймеров, охватывающую каждый сайт интеграции (фиг. 14), синтезировали в AlphaDNA. ПЦР

длинных фрагментов с использованием геномной ДНК U1 выполняли при помощи набора для высокоточной ПЦР Phusion (New England Biolabs), следуя протоколу производителя. ПЦР-продукты визуализировали на 1% агарозном геле и подтверждали при помощи секвенирования по Сэнгеру.

Анализ Surveyor. Присутствие мутаций в ПЦР-продуктах проверяли при помощи набора SURVEYOR Mutation Detection (Transgenomic) в соответствии с протоколом производителя. Вкратце гетерогенные ПЦР-продукты денатурировали в течение 10 мин при 95°C и гибридизировали путем постепенного охлаждения с использованием термоциклира. Затем 300 нг гибридизированной ДНК (9 мкл) подвергали расщеплению при помощи 0,25 мкл нуклеазы SURVEYOR в присутствии 0,25 мкл SURVEYOR Enhancer S и 15 мМ MgCl₂ в течение 4 ч при 42°C. Затем добавляли стоп-раствор и образцы разделяли в 2% агарозном геле вместе с равными количествами контролей в виде нерасщепленных ПЦР-продуктов.

Некоторые ПЦР-продукты использовали для анализа полиморфизма длин рестрикционных фрагментов. Равное количество ПЦР-продуктов расщепляли при помощи BsaI. Расщепленную ДНК разделяли на агарозном геле, содержащем этидия бромид (2%). Для секвенирования ПЦР-продукты клонировали при помощи TA Cloning® Kit Dual Promoter с использованием вектора pCR™II (Invitrogen). Вставку подтверждали расщеплением с помощью EcoRI и положительные клоны посыпали в Genewiz для секвенирования по Сэнгеру.

Отбор целевых сайтов LTR и предсказание потенциальных нецелевых сайтов. Для первоначальных исследований получали промоторную последовательность LTR (от -411 до -10) из интегрированного репортера лентивирусный LTR-люцифераза при помощи ТА-клонирования и секвенирования ПЦР-продуктов из генома клеток TZM-bl человека из-за потенциальной мутации LTR во время пассирования. Данная промоторная последовательность характеризовалась 100% соответствием с 5'-LTR из лентивирусного вектора pHRL-CMV-LacZ (AF105229). Таким образом, смысловую и антисмысловую последовательности 5'-LTR pHRL' полной длины (634 п.о.) использовали для поиска в отношении целевых сайтов Cas9/gRNA, содержащих нацеливающую последовательность gRNA в 20 п.о. и последовательность PAM (NRG), с использованием инструмента Jack Lin's CRISPR/Cas9 gRNA finder (<http://spot.colorado.edu/~slin/cas9.html>). Количество потенциальных нецелевых сайтов с точным совпадением предсказывали при помощи сравнения с использованием BLAST каждой нацеливающей последовательности gRNA и NRG (AGG, TGG, GGG и CGG; AAG, TAG, GAG, CAG) против всех доступных геномных последовательностей и последовательностей транскриптов человека с использованием набора программ NCBI/blastn с отсечением по E-величине 1000 и длиной слова 7. После нажатия Control + F копировали/вставляли целевую последовательность (от 1-23 до 9-23 нуклеотидов) и обнаруживали определенное количество геномных мишней со 100% совпадением с целевой последовательностью. Количество нецелевых сайтов для каждого поиска делили на 3 по причине повторяющейся геномной библиотеки.

Полногеномное секвенирование и биоинформационический анализ. Контрольный субклон C1 и экспериментальный субклон AB7 клеток TZM-bl валидировали в отношении эффективности вырезания мишени и функциональную супрессию LTR-люциферазного репортера. Геномную ДНК выделяли при помощи набора NucleoSpin Tissue (Clontech). Образцы ДНК отправляли в лабораторию секвенирования NextGen при онкологическом центре Фокс Чейз Темпльского университета (NextGen sequencing facility at Temple University Fox Chase Cancer Center). Библиотеки геномных ДНК в двух повторностях получали от каждого субклона при помощи NEBNext Ultra DNA Library Prep Kit для Illumina (New England Biolab), следуя инструкции производителя. Все библиотеки подвергали секвенированию с ридами в 141 п.о. со спаренными концами в двух проточных ячейках Illumina Rapid Run на устройстве HiSeq 2500 (Illumina). Данные демультиплексированных ридов из секвенированных библиотек посыпали в AccuraScience, LLC (<http://www.accurascience.com>) для профессионального биоинформационического анализа. Вкратце, необработанные риды картировали относительно генома человека (hg19) и генома ВИЧ-1 при помощи Bowtie2. Набор программ для анализа генома (GATK, версия 2.8.1) использовали для удаления повторяющихся ридов, локального выравнивания, перекалибровки основного качества и определения вставки/делеции. Баллы достоверности 10 и 30 представляли собой пороговые значения для определений с низким качеством (LowQual) и с высокой достоверностью (PASS). Потенциальные нецелевые сайты LTR-A и LTR-B с различными несовпадениями предсказывали при помощи набора программ NCBI/blastn, как описано выше, и при помощи инструмента CRISPR Design (<http://crispr.mit.edu/>). Все потенциальные целевые сайты gRNA (фиг. 15) использовали для картирования участков в пределах ±300 п.о. около каждой вставки/делеции, выявленной при помощи GATK. Расположения перекрывающихся участков в геноме человека и геноме ВИЧ-1 сравнивали между контрольным C1 и экспериментальным AB7.

Статистический анализ. Количественные данные представляли собой среднее ± стандартное отклонение на основании 3-5 независимых экспериментов, и их оценивали при помощи t-критерия Стьюдента или ANOVA, а также критерия множественного сравнения Ньюмана-Келса. Р-значение, составляющее <0,05 или 0,01, считали статистически значимым отличием.

Пример 2. Cas9/LTR-gRNA подавляет продуцирование репортерного вируса ВИЧ-1 в клетках микроглии CHME5, латентно инфицированных ВИЧ-1.

Оценивали способность направленных на ВИЧ-1 направляющих РНК (gRNA) прекращать транс-

крипционную активность LTR и удалять провирусную ДНК из геномов латентно инфицированных миелоидных клеток, которые служат резервуарами ВИЧ-1 в головном мозге, особенно труднообрабатываемой целевой популяции. Стратегия авторов настоящего изобретения была посвящена целенаправленному воздействию на промоторный участок U3 LTR ВИЧ-1. При помощи биоинформационического скрининга и предсказания эффективности/нечелевых сайтов идентифицировали четыре мишени gRNA (протоспейсеры; LTR A-D), которые исключают консервативные сайты связывания факторов транскрипции, что сводит к минимуму вероятность изменения экспрессии гена-хозяина (фиг. 5 и 13). ДНК-фрагменты, комплементарные gRNA A-D, вставляли в вектор экспрессии гуманизированной Cas9 (A/B в pX260; C/D в pX330) и проверяли их индивидуальные и комбинированные свойства в отношении изменения активности интегрированного генома ВИЧ-1. Сперва использовали линию клеток микроглии CHME5, которая несла интегрированные копии вектора однократного инфицирования ВИЧ-1, который включал 5'- и 3'-LTR и ген, кодирующий репортер, улучшенный зеленый флуоресцентный белок (EGFP), замещающий Gag (pNL4-3-ΔGag-d2EGFP). Обработка клеток CHME5 трихостатином А (TSA), ингибитором деацетилазы гистонов, реактивировала транскрипцию большей части интегрированных провирусов и приводила к экспрессии EGFP и остального протеома ВИЧ-1. Экспрессия gRNA и Cas9 заметно снижала долю TSA-индуцированных положительных по EGFP клеток CHME5 (фиг. 1А и 6). С использованием специфичного к гетеродуплексам анализа SURVEYOR на основе нуклеазы Cel I обнаружили мутации генов по типу вставок/длекций (вставки/длекции) в случае LTR A-D (фиг. 1В и 6В). Подобным образом, экспрессия gRNA, нацеливающаяся на LTR C и D, в полученных из HeLa клетках TZM-bl, которые содержали стабильно встроенные копии LTR ВИЧ-1, управляющие репортерным геном люциферазы светлячка, подавляла вирусную промоторную активность (фиг. 7А) и вызывала вставки/длекции в участке U3 LTR (фиг. 7В-Д), что показано при помощи SURVEYOR и секвенирования по Сэнгеру. Более того, комбинированная экспрессия gRNA, нацеливающаяся на LTR C/D, в этих клетках вызывала вырезание предсказанной последовательности вирусной ДНК в 302 п.о. и появление остаточного фрагмента в 194 п.о. (фиг. 7Е-Ф).

Мультиплексная экспрессия LTR-A/B gRNA в смешанных клональных клетках CHME5 вызывала удаление фрагмента в 190 п.о. между целевыми сайтами А и В и приводила к вставкам/длекциям в различной степени (фиг. 1С-Д). Среди >20 стабильных субклонов, отобранных с помощью пуромицина, обнаружили популяции клеток с полным блокированием TSA-индуцированной реактивации провируса ВИЧ-1, как определено при помощи проточной цитометрии в отношении EGFP (фиг. 1Е). Анализ на основе ПЦР в отношении EGFP и Rev-чувствительного элемента ВИЧ-1 (RRE) в провирусном геноме подтверждал удаление генома ВИЧ-1 (фиг. 1F, G). Кроме того, при секвенировании ПЦР-продуктов обнаружили, что был удален полный вирусный геном, охватывающий 5'-3' LTR, что дало фрагмент в 351 п.о. в результате вырезания 190 п.о. между сайтами расщепления А и В (фиг. 1Г и 8) и фрагмент в 682 п.о. со вставкой в 175 п.о. и длекцией в 27 п.о. в сайтах LTR-A и -B соответственно (фиг. 8С). Остаточный геном ВИЧ-1 (фиг. 1F-Н) может отражать присутствие следовых количеств негативных по Cas9/gRNA клеток. Данные результаты указывают, что нацеливающаяся на LTR Cas9/gRNA A/B удаляет геном ВИЧ-1 и блокирует его реактивацию в латентно инфицированных клетках микроглии.

Пример 3. Cas9/LTR-gRNA эффективно удаляет латентный вирус ВИЧ-1 в моноцитах U1.

Субклон U1 клеток-промоноцитов U-937, модель латентного ВИЧ-1 для инфицированных периваскулярных макрофагов и моноцитов, являются хронически инфицированными ВИЧ-1 и проявляют конститутивную экспрессию и репликацию вирусных генов на низком уровне. При помощи картирования с использованием Genome Walker в клетках U1 обнаружили две интегрированные копии провирусной ДНК в хромосомах Xp11-4 (фиг. 2А) и 2р21 (фиг. 9А). ДНК-фрагмент в 9935 п.о., представляющий собой полную провирусную ДНК ВИЧ-1 в 9709 п.о. и фланкирующую последовательность в 226 п.о., полученную из хромосомы X (фиг. 2А), и фрагмент в 10176 п.о., содержащий геном ВИЧ-1 в 9709 п.о. и его фланкирующие 467 п.о., полученные из хромосомы 2 (фиг. 9А, В), идентифицировали при помощи ПЦР-анализа длинных фрагментов из исходных контрольных клеток U1 или клеток U1 с пустым вектором (U6-CAG). Фрагменты в 226 п.о. и 467 п.о. представляют собой предсказанный сегмент из другой копии хромосомы X и 2 соответственно, в которых отсутствует интегрированная провирусная ДНК. В клетках U1, экспрессирующих LTR-A/B gRNA и Cas9, обнаружили два дополнительных ДНК-фрагмента в 833 п.о. и 670 п.о. в хромосоме X и один дополнительный фрагмент в 1102 п.о. в хромосоме 2. Таким образом, gRNA A/B позволяли Cas9 вырезать сегмент вирусного генома, охватывающий 5'-3' LTR ВИЧ-1, в обоих хромосомах. Фрагмент в 833 п.о. включал ожидаемые 226 п.о. из генома организма-хозяина и последовательность вирусного LTR в 607 п.о. с длекцией в 27 п.о. около сайта LTR-A (фиг. 2А-В). Фрагмент в 670 п.о. охватывал последовательность организма-хозяина в 226 п.о. и последовательность вирусного LTR в 444 п.о., оставшуюся после вырезания фрагмента в 190 п.о. (фиг. 1Д), вызванного направленным gRNA-A/B расщеплением в обоих LTR (фиг. 2А). Не появлялись дополнительные фрагменты, обусловленные интеграцией кольцевого LTR, поскольку он отсутствовал в исходных клетках U1, и такая конфигурация вирусного генома с кольцевым LTR возникает сразу же после инфицирования ВИЧ-1, но является кратковременной и не переносит многократные пассирования. Такие клетки проявляли существенно сниженную вирусную нагрузку ВИЧ-1, как показано при помощи функционального анализа репликации ELISA p24 (фиг. 2С) и ПЦР-анализа в режиме реального времени (фиг. 9С, D). Обнаруживаемая,

но низкая остаточная вирусная нагрузка и реактивация могут быть результатом гетерогенности клеточной популяции и/или неполного редактирования генома. Также подтверждало разрушение генома ВИЧ-1 при помощи Cas9/LTR-A/B gRNA в латентно инфицированных Т-клетках J-Lat, несущих интегрированный HIV-R7/E-/EGFP, при помощи анализа проточной цитометрии, анализа SURVEYOR и ПЦР-генотипирования (фиг. 10), которые подтверждают результаты предыдущих сообщений об удалении провируса ВИЧ-1 в Т-клетках Jurkat при помощи Cas9/gRNA и ZFN. Взятые в совокупности результаты позволяют предположить, что мультиплексная система LTR-gRNA/Cas9 эффективно подавляет репликацию и реактивацию ВИЧ-1 в латентно инфицированных ВИЧ-1 клетках-“резервуарах” (микроглия, моноциты и Т-клетки), типичных для латентной ВИЧ-1-инфекции человека, и в клетках TZM-bl, очень чувствительных для определения транскрипции и реактивации ВИЧ-1. Одиночные или мультиплексные gRNA, нацеливающиеся на 5'- и 3'-LTR, эффективно удаляли полный геном ВИЧ-1.

Пример 4. Стабильная экспрессия Cas9 и LTR-A/B вакцинирует клетки TZM-blI против новой вирусной ВИЧ-1 инфекции.

Далее поверили, могут ли комбинированные Cas9/LTR gRNA иммунизировать клетки против ВИЧ-1-инфекции с использованием стабильных экспрессирующих Cas9/gRNA-A и -B клонов на основе TZM-bl (фиг. 3A). Два из 7 отобранных с помощью пуромицина субклонов проявляли эффективное вырезание ДНК-фрагмента, охватывающего сайт LTR-A/B в 190 п.о. (фиг. 3B). Однако оставшиеся 5 субклонов не проявляли вырезания (фиг. 3B) и мутаций по типу вставок/делеций, что подтверждено при помощи секвенирования по Сэнгеру. ПЦР-генотипирование с использованием праймеров, нацеливающихся на Cas9 и U6-LTR, показало, что ни один из этих неэффективных субклонов не удерживал интегрированные копии кассет экспрессии Cas9/LTR-A/B gRNA (фиг. 11 A, B). Вследствие этого, не обнаруживали экспрессии Cas9 полной длины (фиг. 11C, D). Долговременная экспрессия Cas9/LTR-A/B gRNA не влияла отрицательно на рост или жизнеспособность клеток, что позволяет предположить о редкой встречаемости нецелевого вмешательства в геном организма-хозяина или Cas9-индуцированной токсичности в данной модели. Репликацию de novo ВИЧ-1 оценивали путем инфицирования клеток псевдотипированным VSVG репортерным вирусом pNL4-3-ΔE-EGFP, при этом положительность в отношении EGFP, определенная при помощи проточной цитометрии, указывала на репликацию ВИЧ-1. В отличие от контрольных клеток U6-CAG клетки, стабильно экспрессирующие Cas9/gRNA LTR-A/B, не поддерживали репликацию ВИЧ-1 на 2 день после инфицирования, что указывало на то, что они были эффективно иммунизированы против новой ВИЧ-1 инфекции (фиг. 3C-D). Подобный иммунитет против ВИЧ-1 наблюдали у экспрессирующих Cas/LTR-A/B gRNA клеток, инфицированных нативным Т-тропным репортерным вирусом штамма X4 pNL4-3-ΔE-EGFP (фиг. 12A) или нативными М-тропными штаммами R5, такими как SF162 и JRFL (фиг. 12B-D).

Пример 5. Нецелевые эффекты Cas9/LTR-A/B в отношении генома человека.

Привлекательность Cas9/gRNA в качестве способа вмешательства опирается на ее высокоспецифическое расщепление с образованием целевых вставок/делеций, но мультиплексные gRNA могут потенциально служить причиной мутагенеза и хромосомных нарушений в геноме организма-хозяина, вызывать цитотоксичность, генотоксичность или онкогенез. Достаточно низкая гомология генома вируса и человека уменьшает этот риск, но геном человека содержит множество эндогенных ретровирусных геномов, которые потенциально восприимчивы к направленным на ВИЧ-1 gRNA. В связи с этим оценивали нецелевые эффекты выбранных gRNA для LTR ВИЧ-1 в отношении генома человека. Поскольку затравочная последовательность в 12-14 п.о., ближайшая к участку (NGG) мотива, прилегающего к протоспейсеру (PAM), является крайне важной для специфичности расщепления, то проводили поиск последовательности > затравки в 14 п.о. + NGG и не обнаружили нецелевых кандидатных сайтов для воздействия LTR gRNA A-D (фиг. 13). Неудивительно, что постепенное укорочение сегментов gRNA приводило к увеличению количества нецелевых сайтов расщепления, на 100% совпадающих с соответствующими целевыми последовательностями (т.е. NGG+13 п.о. приводило к 6, 0, 2 и 9 нецелевым сайтам соответственно, тогда как NGG+12 п.о. приводило к 16, 5, 16 и 29; фиг. 13). При помощи высокоточной ПЦР из геномной ДНК человека получали последовательность в 500-800 п.о., охватывающую один из предсказанных нецелевых сайтов, и анализировали потенциальные мутации при помощи SURVEYOR и секвенирования по Сэнгеру. Мутаций не обнаружили (см. репрезентативные нецелевые сайты № 1, 5 и 6 в клетках TZM-blI и U1; фиг. 4A).

Для всесторонней оценки риска нецелевых эффектов выполняли полногеномное секвенирование (WGS) с использованием стабильных клеток TZM-blI, экспрессирующих Cas9/gRNA A/B и контрольных U6-CAG (фиг. 4B-D). При помощи набора программ для анализа генома (GATK, v.2.8.1) с геномами человека (hg19) и ВИЧ-1 в качестве эталонных последовательностей идентифицировали 676105 вставок/делеций. Среди вставок/делеций 24% обнаружены в контроле U6-CAG, 26% - в субклоне LTR-A/B и 50% - в обоих (фиг. 4B). Такое существенное расхождение в определении вставок/делеций между образцами позволяет предположить возможные нецелевые эффекты, но наиболее вероятно является результатом его ограниченной достоверности, ограниченной охватом WGS (15-30X), и гетерогенности клеток. В GATK сообщалось только о достоверно идентифицированных вставках/делециях: некоторые обнаруже-

ны в контроле U6-CAG, но не в субклоне LTR-A/B, а другие в LTR-A/B, но не в U6-CAG. Вследствие ограниченного охвата WGS ожидали многочисленные результаты определения отсутствующих вставок/делеций для обоих образцов. Такая ограниченная достоверность определения вставок/делеций также подразумевает возможность ложноотрицательных результатов: отсутствующие вставки/делеции встречаются в LTR-A/B, но не в контролах U6-CAG. Гетерогенность клеток может отражать изменчивость эффективности редактирования при помощи Cas9/gRNA и эффекты пассирования. Таким образом, проверяли, каждая ли вставка/делеция была индуцирована LTR-A/B gRNA посредством анализа ± 300 п.о., фланкирующих каждую вставку/делецию, относительно целевых для LTR-A/B сайтов генома ВИЧ-1 и предсказанных/потенциальных нецелевых сайтов для gRNA генома организма-хозяина (фиг. 15). Для последовательностей, на 100% совпадающих с последовательностью, содержащей затравку (12 п.о.) и NRG, идентифицировали только 8 перекрывающихся участков из 92 потенциальных нецелевых сайтов относительно 676105 вставок/делеций: 6 вставок/делеций встречаются в обоих образцах, и 2 встречаются только в контроле U6-CAG (фиг. 4C, D). Также идентифицировали 2 вставки/делеции в LTR ВИЧ-1, которые обнаружили только в субклоне LTR-A/B, но, как и ожидалось, не в контроле U6-CAG (фиг. 4C). Результаты позволяют предположить, что LTR-A/B gRNA индуцируют указанные целевые вставки/делеции, но не индуцируют нецелевых вставок/делеций, что согласуется с предыдущими данными, полученными с использованием глубокого секвенирования ПЦР-продуктов, охватывающих предсказанный/потенциальный нецелевой сайт.

Комбинированные подходы свели к минимуму нецелевые эффекты, в то же время, с достижением высокой эффективности и полного разрушения интегрированного в геном провируса ВИЧ-1. В дополнение к чрезвычайно низкой гомологии между чужеродным вирусным геномом и геномом клетки-хозяина, включающим эндогенную ретровирусную ДНК, ключевые признаки конструктивного решения в данном исследовании включали: биоинформационический скрининг с использованием самого строгого критерия выбора мишени 12 п.о.+NGG для исключения нецелевых транскриптомов человека или (хотя реже) не-транслируемых участков генома; с исключением сайтов связывания факторов транскрипции в промоторе LTR ВИЧ-1 (потенциально консервативные в геноме организма хозяина); отбор направленных на LTR-A- и -B gRNA в 30 п.о. и также системы pre-crRNA, отражающей первоначальный бактериальный иммунный механизм, для увеличения специфиности/эффективности в противовес системе на основе gRNA в 20 п.о., химерной crRNA-tracRNA; а также WGS, секвенирование по Сэнгеру и анализ SURVEYOR для идентификации и исключения потенциальных нецелевых эффектов. Действительно, применение недавно разработанных вносящей двойные однонитевые разрывы Cas9 и направленной РНК нуклеазы FokI может дополнительно содействовать идентификации новых мишеней в пределах различных консервативных участков ВИЧ-1 с уменьшением нецелевых эффектов.

По результатам видно, что система Cas9/gRNA ВИЧ-1 обладает способностью целенаправленно воздействовать на несколько копий LTR, которые размещены на разных хромосомах, что позволяет предположить, что данная система для редактирования генома может изменять последовательность ДНК ВИЧ-1 в латентно инфицированных клетках пациента, несущих несколько провирусных ДНК. Для того чтобы дополнительно убедиться в высокой эффективности редактирования и стабильности данной технологии, можно рассматривать наиболее стабильный участок генома ВИЧ-1 в качестве мишени для удаления ВИЧ-1 в образцах от пациента, которые могут нести не только один штамм ВИЧ-1. Альтернативно, можно разработать индивидуальные способы лечения, основанные на данных, полученных в результате глубокого секвенирования вирусного генома, полученного от пациента, перед конструированием терапевтических молекул Cas9/gRNA.

По результатам также видно, что редактирование генома при помощи Cas9/gRNA можно применять для иммунизации клеток против ВИЧ-1-инфекции. Превентивная вакцинация является независимой от разнообразия штаммов ВИЧ-1, поскольку система целенаправленной воздействует на геномные последовательности независимо от того, как вирусы проникают в инфицированные клетки. Предварительное существование системы Cas9/gRNA в клетках ведет к быстрой элиминации из организма нового ВИЧ-1 до того, как он интегрируется в геном организма-хозяина. Можно исследовать различные системы доставки Cas9/LTR-gRNA для иммунизации субъектов с высоким риском заражения, например виды генной терапии (вирусный вектор и наночастица) и трансплантацию аутологичных модифицированных при помощи Cas9/gRNA стволовых клеток/клеток-предшественников костного мозга или индуцируемых плюрипотентных стволовых клеток для удаления ВИЧ-1-инфекции.

В данной заявке продемонстрирована высокая специфичность Cas9/gRNA при редактировании целевого генома ВИЧ-1. По результатам, полученным из данных субклонов, выявлена строгая зависимость редактирования генома от присутствия как Cas9, так и gRNA. Более того, всего одно несовпадение нуклеотида в сконструированной мишени для gRNA повредит эффективности редактирования. В дополнение, все из 4 gRNA LTR, сконструированных в данном документе, хорошо работали на разных линиях клеток, что указывает на то, что редактирование было более эффективным в отношении генома ВИЧ-1, чем в отношении генома клетки-хозяина, где не все сконструированные gRNA были функциональными, что может быть обусловлено отличающейся эпигенетической регуляцией, различной доступностью генома или другими причинами. Учитывая простоту и быстроту разработки Cas9/gRNA, даже если мутации

ВИЧ-1 придают устойчивость к одному виду терапии на основе Cas9/gRNA, как описано выше, варианты ВИЧ-1 можно генотипировать для обеспечения индивидуальной терапии для отдельных пациентов.

Был описан целый ряд вариантов осуществления изобретения. Тем не менее, будет понятно, что различные модификации могут быть сделаны без отступления от сущности и объема изобретения. Соответственно, другие варианты осуществления находятся в пределах объема следующей формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ удаления всей или практически всей лентивирусной последовательности, при этом способ включает приведение клетки в контакт с:

а) Cas9 эндонуклеазой, ассоциированной с короткими палиндромными повторами, регулярно расположеными группами (CRISPR), или последовательностью нукleinовой кислоты, кодирующей CRISPR-ассоциированную Cas9 эндонуклеазу;

б) первой направляющей РНК или нукleinовой кислотой, кодирующей первую направляющую РНК, причем первая направляющая РНК комплементарна первой целевой последовательности нукleinовой кислоты в пределах 5'-длинного концевого повтора (LTR) лентивирусной последовательности; и

с) второй направляющей РНК или нукleinовой кислотой, кодирующей вторую направляющую РНК, причем вторая направляющая РНК комплементарна второй целевой последовательности нукleinовой кислоты в пределах 3'-длинного концевого повтора (LTR) лентивирусной последовательности,

где клетку приводят в контакт по меньшей мере с двумя разными направляющими РНК и где способ позволяет удалить из клетки всю или практически всю лентивирусную последовательность.

2. Способ по п.1, где лентивирусная последовательность выбрана из группы, включающей вирус иммунодефицита человека, вирус иммунодефицита обезьян, вирус иммунодефицита кошачьих и вирус иммунодефицита крупного рогатого скота.

3. Способ по п.2, где вирус иммунодефицита человека (ВИЧ) представляет собой ВИЧ-1 или ВИЧ-2.

4. Способ по любому из пп.2, 3, где ВИЧ интегрирован в клетку.

5. Способ по любому из пп.1-4, где клетка представляет собой клетку, латентно инфицированную ВИЧ.

6. Способ по п.5, где латентно инфицированная клетка представляет собой CD4+ Т-клетку, макрофаг, моноцит, клетку лимфоидной ткани кишечника, клетку микроглии или астроцит.

7. Способ по любому из пп.1-6, где последовательность CRISPR-ассоциированной Cas9 эндонуклеазы оптимизирована для экспрессии в клетке человека.

8. Способ по любому из пп.1-7, где первая целевая последовательность нукleinовой кислоты находится в пределах участков U3, R или U5 5'-LTR.

9. Способ по любому из пп.1-8, где первая направляющая РНК содержит последовательность, которая по меньшей мере на 95% идентична последовательности, комплементарной последовательности, выбранной из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

10. Способ по любому из пп.1-9, где первая направляющая РНК содержит последовательность, выбранную из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

11. Способ по любому из пп.1-10, где вторая целевая последовательность нукleinовой кислоты находится в пределах участков U3, R или U5 LTR.

12. Способ по любому из пп.1-11, где вторая направляющая РНК содержит последовательность, характеризующуюся 95% идентичностью с последовательностью, выбранной из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

13. Способ по любому из пп.1-12, где вторая направляющая РНК содержит последовательность, выбранную из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

14. Способ по любому из пп.1-13, где клетку дополнительно приводят в контакт с последовательностью нукleinовой кислоты, кодирующей транскриптирующую малую РНК (tracrRNA).

15. Способ по п.14, где последовательность нукleinовой кислоты транскриптирующей малой РНК (tracrRNA) слита с последовательностью первой направляющей РНК, последовательностью второй направляющей РНК или с обеими указанными последовательностями.

16. Способ по любому из пп.1-15, дополнительно включающий последовательность нукleinовой кислоты, кодирующую сигнал ядерной локализации.

17. Способ по любому из пп.1-16, где первая направляющая РНК и вторая направляющая РНК содержатся в одном и том же векторе экспрессии.

18. Способ по п.17, где вектор экспрессии представляет собой лентивирусный вектор, аденоавирусный вектор или вектор на основе аденоассоциированного вируса.

19. Способ лечения инфекции, вызванной вирусом иммунодефицита человека (ВИЧ), при этом спо-

соб включает приведение клетки в контакт с:

- a) CRISPR-ассоциированной Cas9 эндонуклеазой или последовательностью нукleinовой кислоты, кодирующей CRISPR-ассоциированную Cas9 эндонуклеазу;
- b) первой направляющей РНК или нукleinовой кислотой, кодирующей первую направляющую РНК, причем первая направляющая РНК комплементарна первой целевой последовательности нукleinовой кислоты в пределах 5'-длинного концевого повтора (LTR) последовательности ВИЧ; и
- c) второй направляющей РНК или нукleinовой кислотой, кодирующей вторую направляющую РНК, причем вторая направляющая РНК комплементарна второй целевой последовательности нукleinовой кислоты в пределах 3'-длинного концевого повтора (LTR) последовательности вируса иммунодефицита человека, причем клетку приводят в контакт по меньшей мере с двумя разными направляющими РНК.

20. Способ лечения инфекции, вызванной вирусом иммунодефицита человека, при этом способ включает приведение клетки в контакт с:

- a) CRISPR-ассоциированной Cas9 эндонуклеазой или нукleinовой кислотой, кодирующей CRISPR-ассоциированную Cas9 эндонуклеазу;
- b) первой направляющей РНК или нукleinовой кислотой, кодирующей первую направляющую РНК, причем первая направляющая РНК комплементарна первой целевой последовательности нукleinовой кислоты в пределах 5'-LTR последовательности вируса иммунодефицита человека; и
- c) второй направляющей РНК или нукleinовой кислотой, кодирующей вторую направляющую РНК, причем вторая направляющая РНК комплементарна второй целевой последовательности нукleinовой кислоты в пределах 3'-LTR последовательности вируса иммунодефицита человека, где первая направляющая РНК или вторая направляющая РНК содержит последовательность, выбранную из группы, состоящей из LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) или их комбинацию, и

где способ приводит к вырезанию генома вируса иммунодефицита человека из клетки.

21. Нукleinовая кислота, кодирующая:

- (a) CRISPR-ассоциированную Cas9 эндонуклеазу;

(b) первую направляющую РНК, при этом первая направляющая РНК комплементарна первой целевой последовательности нукleinовой кислоты в пределах 5'-LTR последовательности вируса иммунодефицита человека; и

(c) вторую направляющую РНК, при этом вторая направляющая РНК комплементарна второй целевой последовательности нукleinовой кислоты в пределах 3'-LTR последовательности вируса иммунодефицита человека, при этом нукleinовая кислота содержит по меньшей мере две разные направляющие РНК и при этом первая направляющая РНК или вторая направляющая РНК выбрана из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинации.

22. Нукleinовая кислота по п.21, где вирус иммунодефицита человека представляет собой ВИЧ-1 или ВИЧ-2.

23. Нукleinовая кислота по любому из пп.21, 22, где последовательность нукleinовой кислоты, кодирующей CRISPR-ассоциированную Cas9 эндонуклеазу, оптимизирована для экспрессии в клетке человека.

24. Нукleinовая кислота по любому из пп.21-23, дополнительно содержащая последовательность, кодирующую трансактивирующую малую РНК (tracrRNA).

25. Нукleinовая кислота по п.24, где последовательность трансактивирующей малой РНК (tracrRNA) слита с последовательностью первой направляющей РНК, последовательностью второй направляющей РНК или с обеими указанными последовательностями.

26. Нукleinовая кислота по любому из пп.21-25, дополнительно содержащая последовательность, кодирующую сигнал ядерной локализации.

27. Вектор экспрессии, содержащий последовательность нукleinовой кислоты по любому из пп.21-26.

28. Вектор экспрессии по п.27, где вектор экспрессии представляет собой лентивирусный вектор, аденоавирусный вектор или вектор на основе аденоассоциированного вируса.

29. Клетка-хозяин, содержащая вектор экспрессии по любому из пп.27, 28.

30. Фармацевтическая композиция, содержащая:

- а) нукleinовую кислоту, содержащую:

(i) первую последовательность нукleinовой кислоты, кодирующую эндонуклеазу, ассоциированную с короткими палиндромными повторами, регулярно расположеными группами (CRISPR);

(ii) вторую последовательность нукleinовой кислоты, кодирующую первую направляющую РНК, причем первая направляющая РНК комплементарна первой целевой последовательности нукleinовой кислоты в пределах 5'-LTR последовательности вируса иммунодефицита человека (ВИЧ);

(iii) третью последовательность нукleinовой кислоты, кодирующую вторую направляющую РНК, при этом вторая направляющая РНК комплементарна второй целевой последовательности нукleinовой кислоты в пределах 3'-LTR последовательности ВИЧ,

при этом нуклеиновая кислота содержит по меньшей мере две разные направляющие РНК; и
б) фармацевтически приемлемое формообразующее.

31. Фармацевтическая композиция по п.30, где фармацевтически приемлемое формообразующее включает липидный или полимерный колloid.

32. Фармацевтическая композиция по п.31, где колloid представляет собой липосому, гидрогель, микрочастицу, наночастицу или мицеллу блок-сополимера.

33. Фармацевтическая композиция по любому из пп.30-32, где фармацевтическая композиция составлена для местного применения.

34. Фармацевтическая композиция по любому из пп.31-33, в которой первая целевая последовательность нуклеиновой кислоты находится в пределах участков U3, R или U5 LTR.

35. Фармацевтическая композиция по любому из пп.31-34, в которой вторая целевая последовательность нуклеиновой кислоты находится в пределах участков U3, R или U5 LTR.

36. Фармацевтическая композиция по любому из пп.31-35, в которой первая направляющая РНК содержит последовательность, которая по меньшей мере на 95% идентична последовательности, выбранной из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

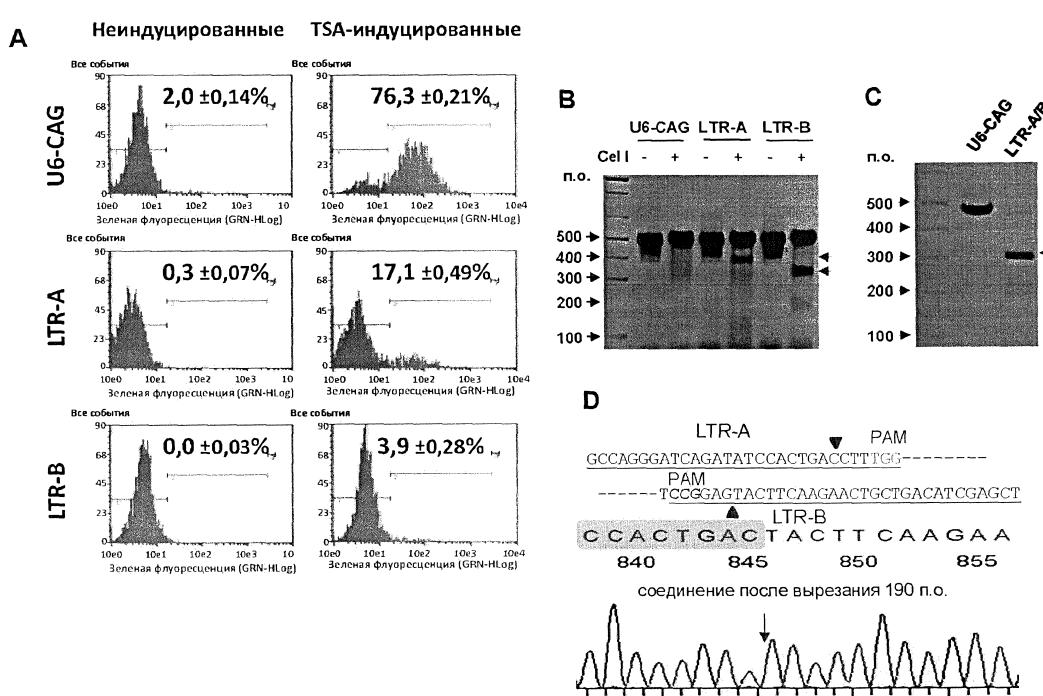
37. Фармацевтическая композиция по любому из пп.31-36, в которой вторая направляющая РНК содержит последовательность, по меньшей мере на 95% идентичную последовательности, которая комплементарна последовательности, выбранной из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

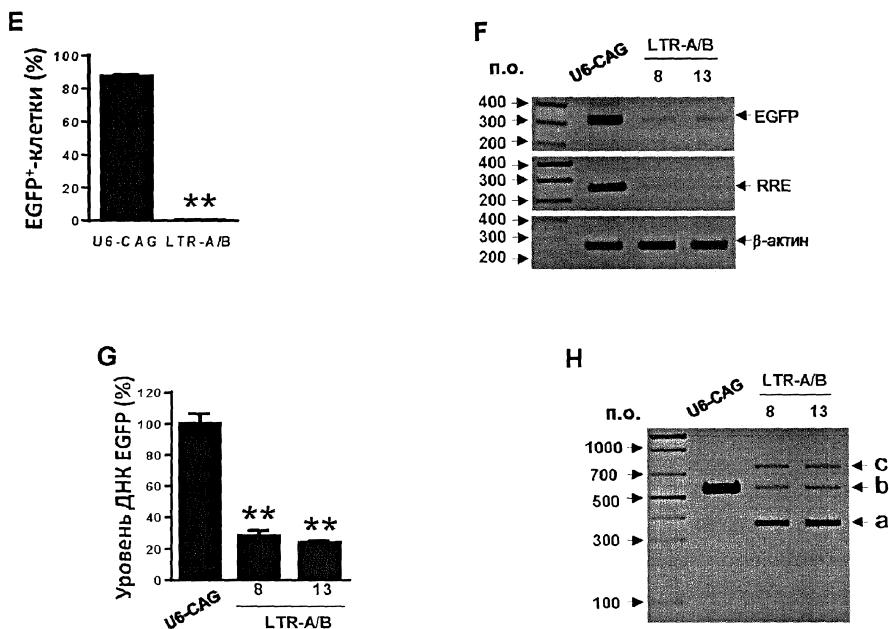
38. Фармацевтическая композиция по любому из пп.31-37, в которой первая направляющая РНК содержит последовательность, выбранную из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

39. Фармацевтическая композиция по любому из пп.31-38, в которой вторая направляющая РНК содержит последовательность, выбранную из группы, включающей LTR A (SEQ ID NO: 96), LTR B (SEQ ID NO: 121), LTR C (SEQ ID NO: 87), LTR D (SEQ ID NO: 110) и их комбинацию.

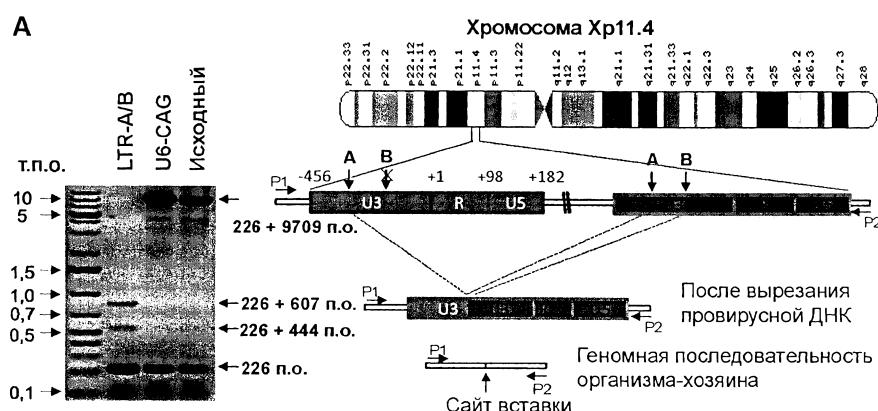
40. Фармацевтическая композиция по любому из пп.31-39, где фармацевтическая композиция составлена для парентерального введения.

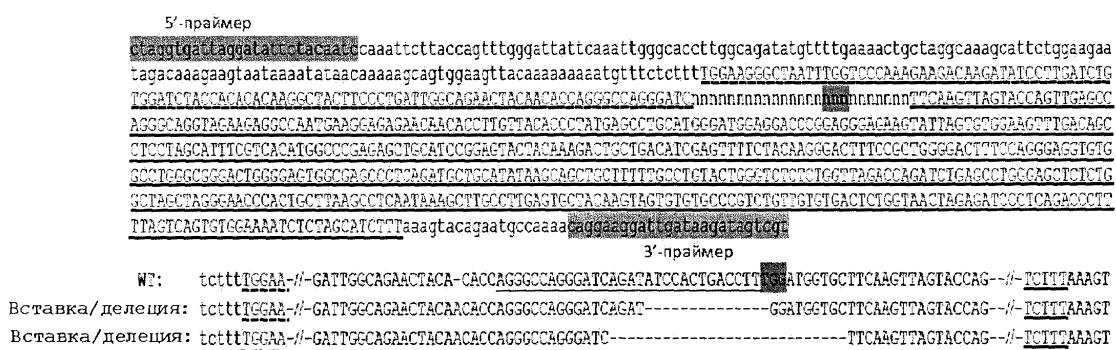
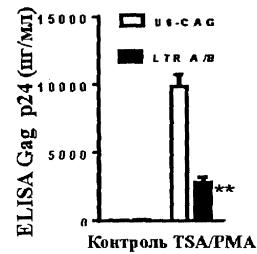
41. Способ по любому из пп.19, 20, где способ приводит к вырезанию последовательности ВИЧ из генома клетки.



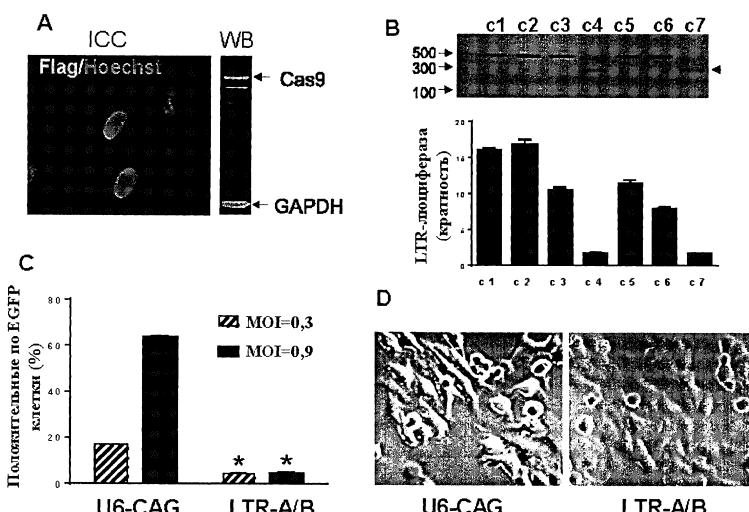


ФИГ. 1

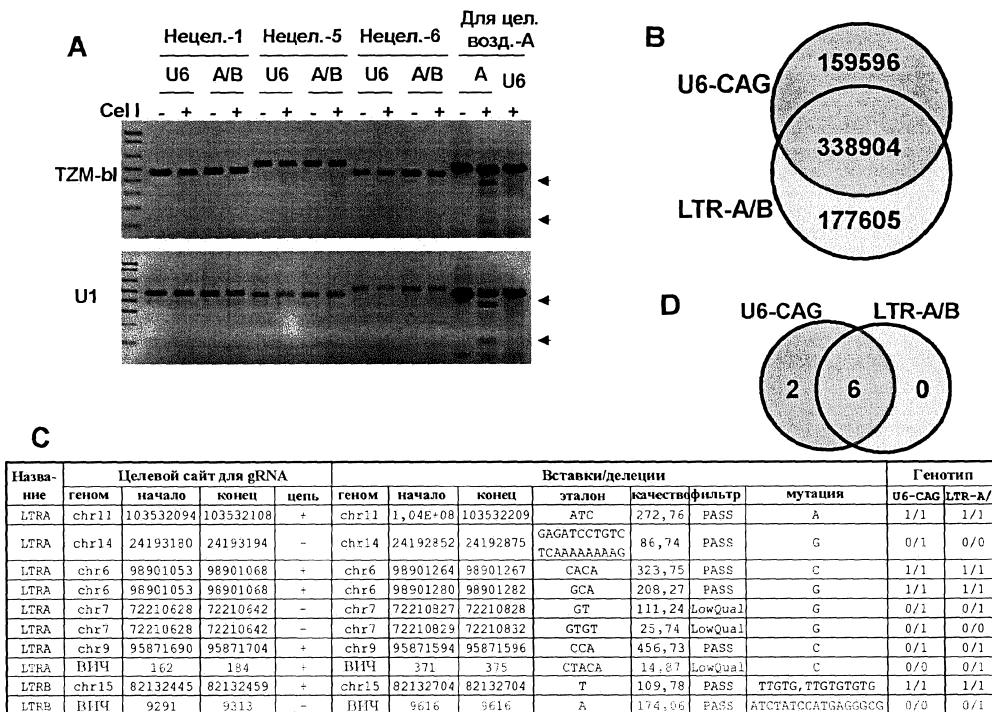


B**C**

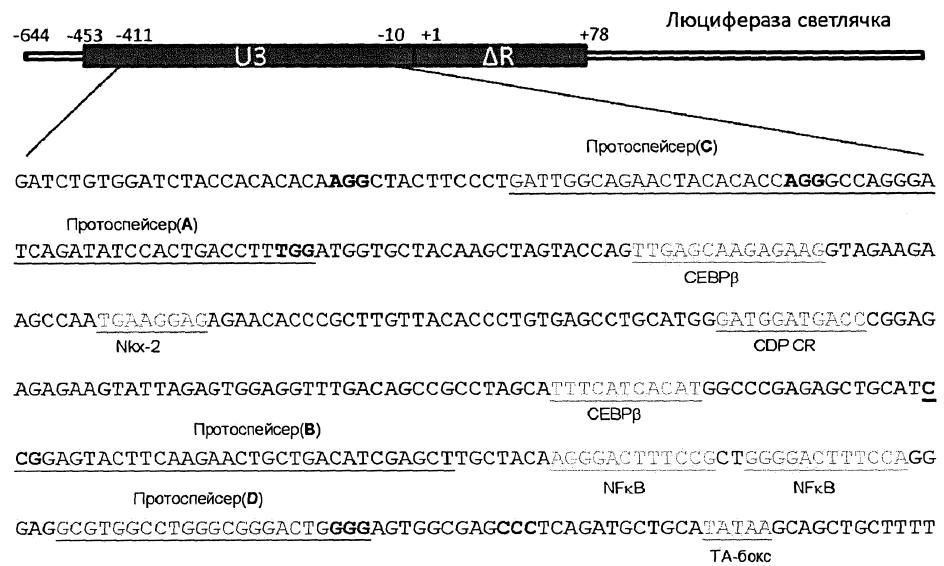
Фиг. 2



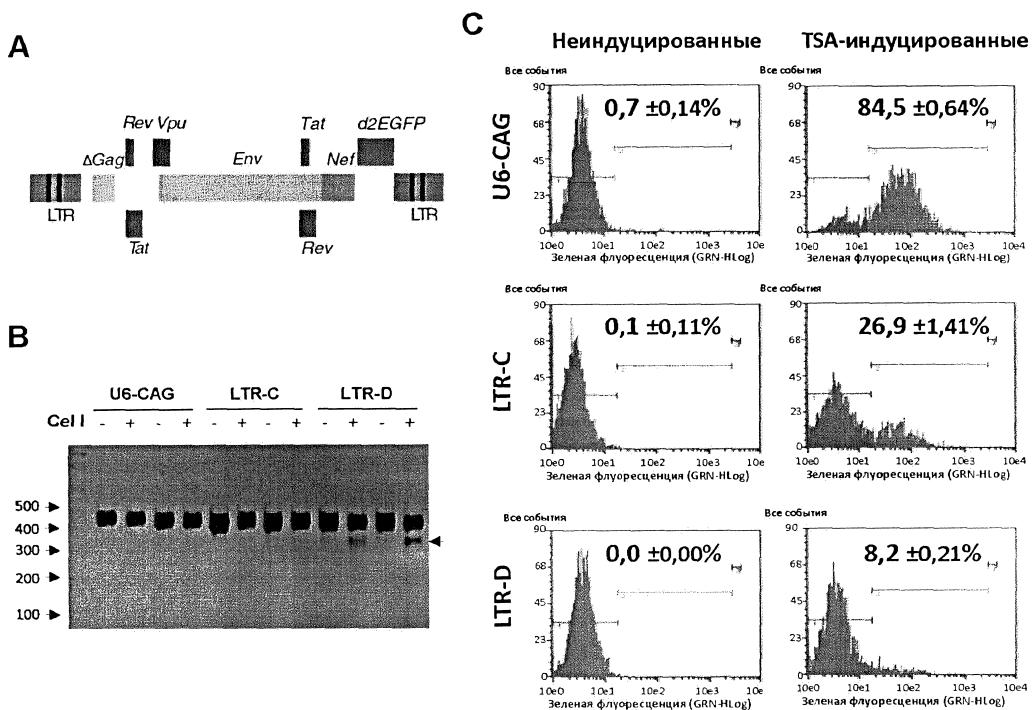
Фиг. 3



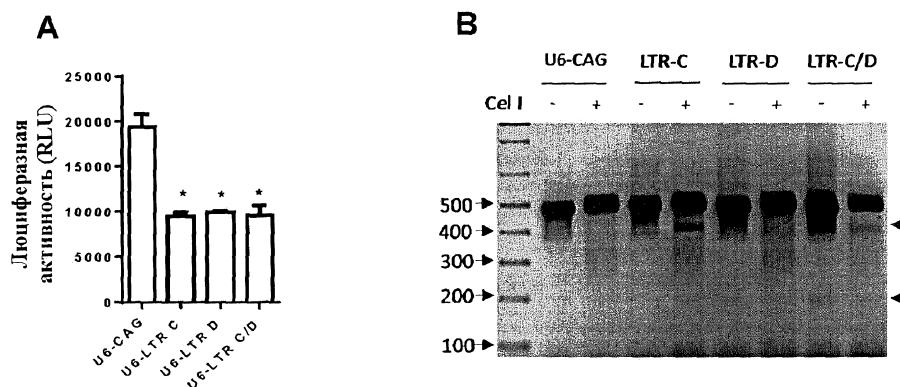
Фиг. 4

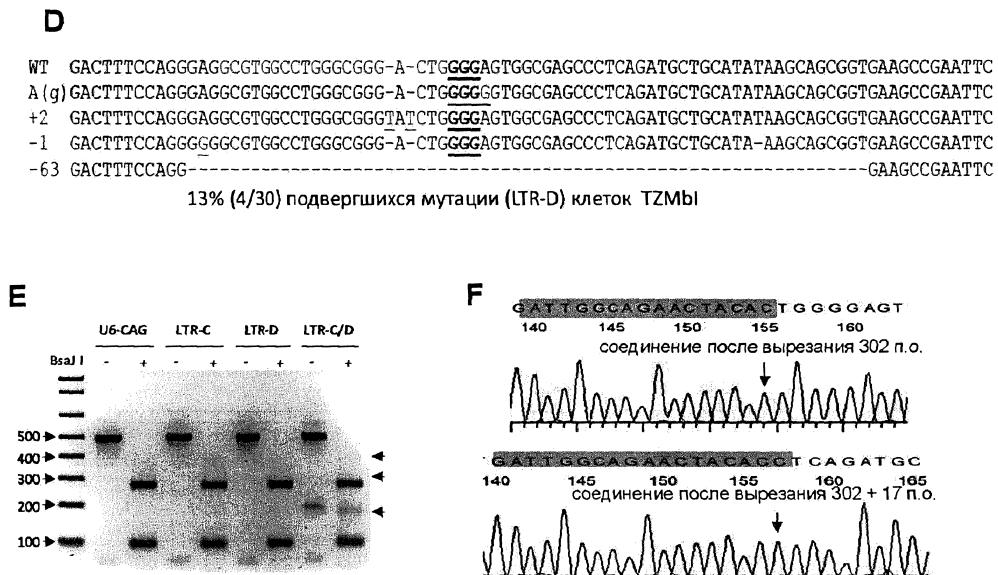


Фиг. 5

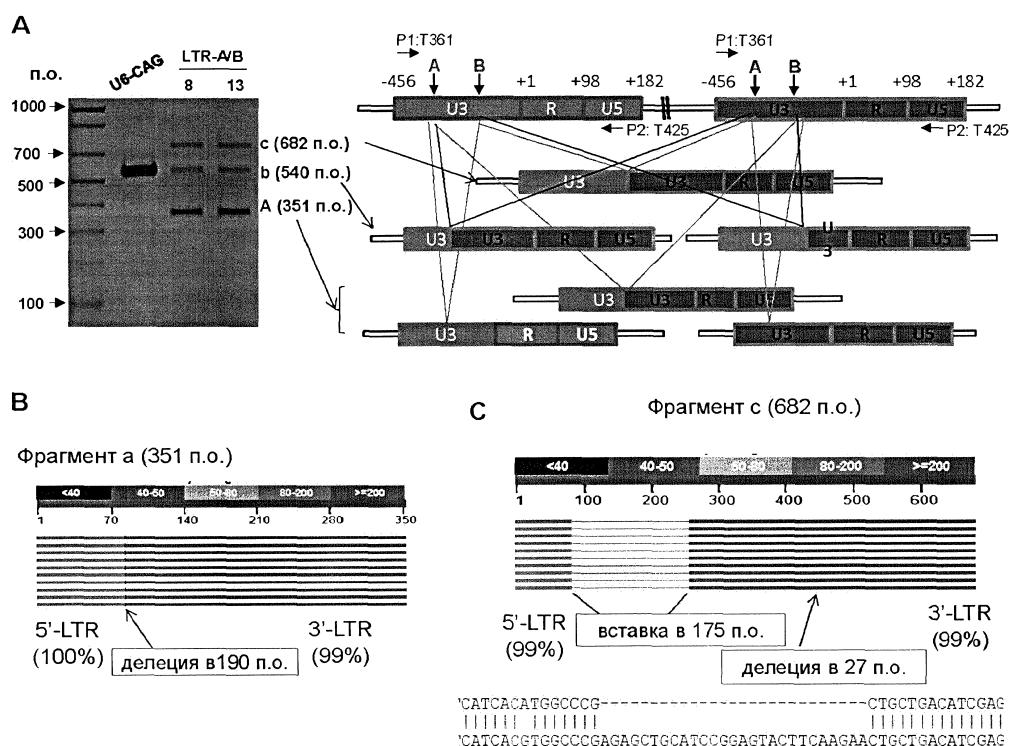


Фиг. 6

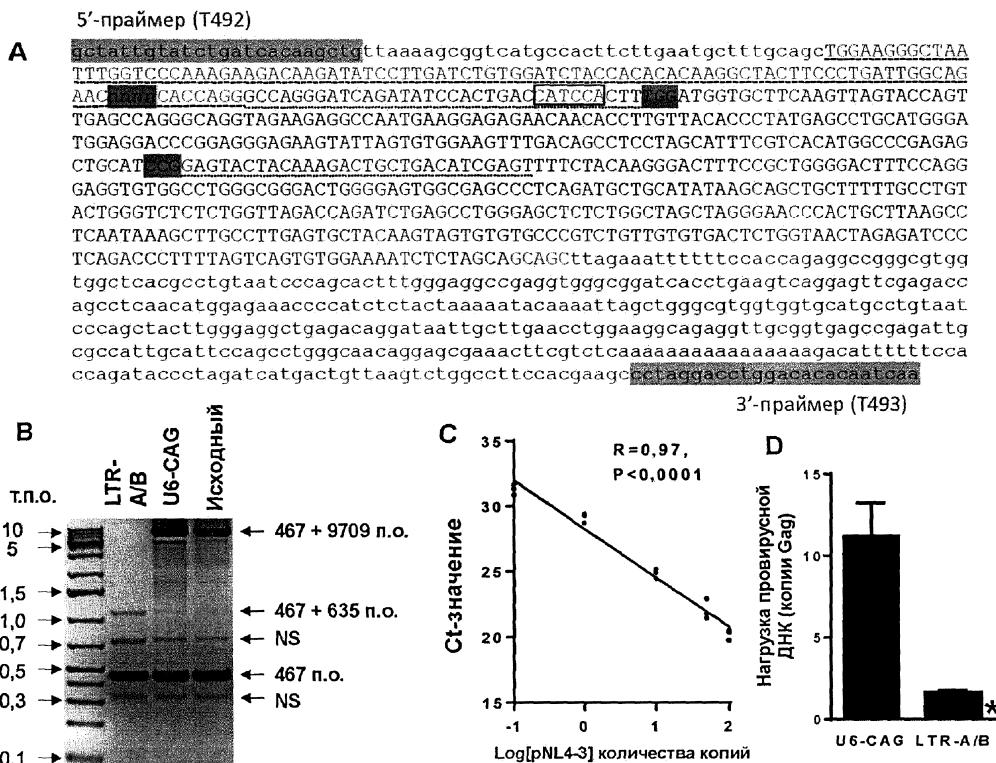




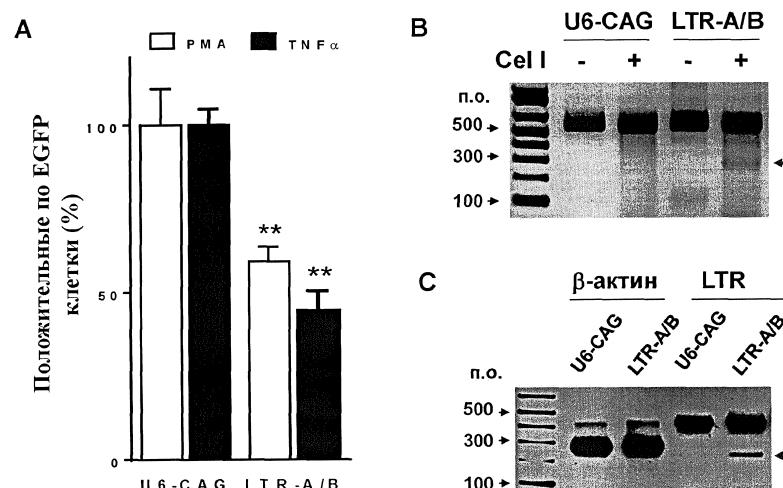
Фиг. 7



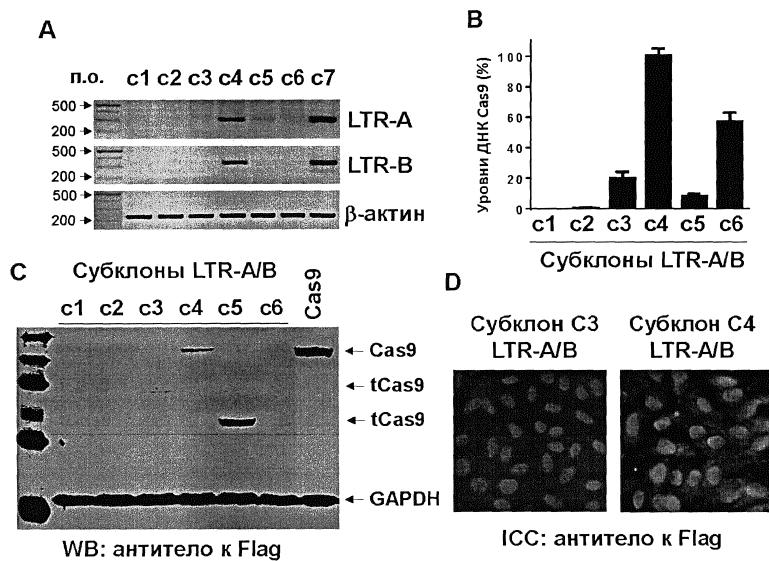
Фиг. 8



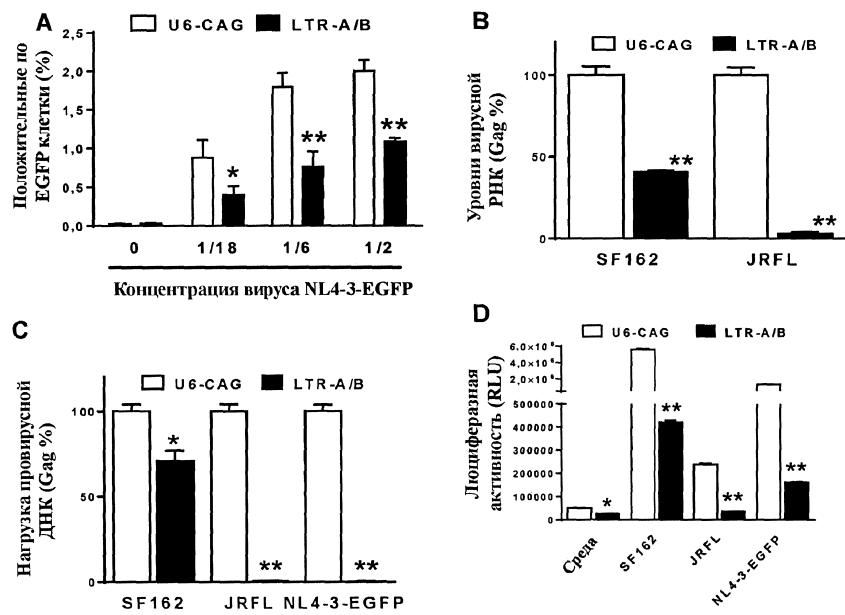
Фиг. 9



Фиг. 10



Фиг. 11



Фиг. 12

Таблица S1. Предсказанные LTR gRNA и количества нецелевых сайтов для них (100% совпадение)

	Последовательность gRNA (смысловая)	20	19	18	17	16	15	14	13	12
1	TCAGACCCCTTTAGTCAGTGTGG	0	0	0	0	0	0	0	2	25
2	TTGCTTGACTGGCTCTCTGG	0	0	0	0	0	0	0	3	10
3	CAGCTGCTTTGCTTGACTGG	0	0	0	0	0	0	2	4	11
4	CTGACATCGAGCTTGTACAAGG	0	0	0	0	0	0	0	0	17
5	CCGCCTAGCATTCATCACATGG	0	0	0	0	0	1	2	9	55
6	CGGAGAGAAGTATTAGAGTGG	0	0	0	0	0	0	0	3	18
7	ACTACCACTTGAGCAAGAGAAGG	0	0	0	0	0	0	0	1	18
8	GATATCCACTGACCTTGGATGG	0	0	0	0	0	1	3	22	211
LTR-C	9 CAGCTGCTTTGCTTGACTGG	0	0	0	0	0	0	0	2	12
10	CACAAGGCTACTTCCCTGATTGG	0	0	0	0	0	0	0	2	11
11	CTGTGGATCTACCACACAAAGG	0	0	0	0	0	2	5	12	37
12	TGGGAGCTCTGGCTAACTAGG	0	0	0	0	0	0	0	4	14
13	GGTTAGACCAAGATCTGAGCCTGG	0	0	0	0	0	0	0	9	33
14	TGCTACAAGGGACTTCCGCTGG	0	0	0	0	0	0	0	2	5
15	AGAGAGAAAGTATTAGAGTGGAGG	0	0	0	0	0	0	5	7	16
16	TTACACCCGTGAGCCTGATGG	0	0	0	0	0	0	3	14	35
17	AAGGTAAGAGAACCAATGAAGG	0	0	0	0	0	0	4	36	75
LTR-A	18 ATCTGAGCTGGGACTTCCAGG	0	0	0	0	0	0	0	6	16
19	GACAAGATATCCTGATCTGG	0	0	0	0	0	0	0	0	10
20	SCCCGCTGTTGCTGACTCTGG	0	0	0	0	0	0	3	7	35
21	ATCTGAGCTGGGAGCTCTGG	0	0	0	0	1	2	9	32	78
22	CTTC CGCTGGGACTTCCAGG	0	0	0	0	0	0	0	0	4
23	CAGAACTACACACCAAGGCCAGG	0	0	0	0	0	0	0	3	20
24	CCTGCATGGATGGATGACCCGG	0	0	0	0	2	2	2	5	21
25	CCCTGTGAGCCTGCATGGATGG	0	0	0	0	1	1	2	9	30
26	CTTC CAGGGAGGCCTGGCCTGG	0	0	0	0	0	8	15	32	75
27	GGGGACTTTCCAGGGAGGCCTGG	0	0	0	0	0	0	0	5	24
28	CCGCTGGGACTTTCCAGGGAGG	0	0	0	0	0	1	3	9	25
29	CATGGCCCCAGAGCTGCATCCGG	0	0	0	0	0	0	0	0	16
30	GCCTGGGGGGACTGGGAGTGG	0	0	0	0	1	2	7	18	250
31	AGCGTGGCCTGGGGACTGG	0	0	0	0	0	0	0	2	6
LTR-D	32 CCAGGGAGGCCCTGGGGAGG	0	0	0	0	0	0	6	9	29
33	CCAGGGAGGCCCTGGGGAGG	0	0	0	0	2	2	2	11	22
	Последовательности gRNA (антисмысловая)									
1	TGTGGTAGATCCACAGATCAAGG	0	0	0	0	0	0	0	3	13
2	GGTGTGTTAGTCTGCCAACTCAGG	0	0	0	0	0	0	0	8	11
3	GTCAGTGGATATCTGATCCCTGG	0	0	0	0	0	0	0	2	10
4	TAGCACCATCAAAGGTGAGTGG	0	0	0	0	0	0	2	3	10
5	TAGCTTGAGCACCATCAAAGG	0	0	0	0	0	1	1	6	12
6	TCTACCTCTCTGCTCAACTGG	0	0	0	0	0	0	1	1	6
7	CACTCTAAACTTCTCTCCGG	0	0	0	0	0	0	0	0	26
8	CCATGTGATGAAATGCTAGGCGG	0	0	0	0	0	1	1	5	17
9	GGGCCATGTGATGAAATGCTAGG	0	0	0	0	0	0	0	3	48
LTR-B	10 GACCACTGAGCTGGGGAGG	0	0	0	0	0	0	0	0	5
11	CTGCTTATATGAGCATCTGGAGG	0	0	0	0	0	0	0	4	39
12	CACACTACTTGAAAGCACTCAAGG	0	0	0	0	0	0	0	3	19
13	TACCAAGGTACACAAACAGACGG	0	0	0	0	0	0	0	2	12
14	ACACTGACTAAAGGTCTGGAGG	0	0	0	0	1	2	3	4	15
15	CAAGGATATCTGTTCTGGTGG	0	0	0	0	0	0	0	1	5
16	CAGGGAGGTAGCCCTGGTGG	0	0	0	0	0	1	2	4	17
17	GGGGGTGTTCTCTCCTTCATTGG	0	0	0	0	0	2	2	15	49
18	TAGTTAGCCAGAGAGCTCCAGG	0	0	0	0	0	2	4	24	93
19	CTTTATTGAGGCTTAAGCAGTGG	0	0	0	0	0	0	0	4	25
20	ACTCAAGGCAGGCTGGAGTGG	0	0	0	0	0	0	0	1	28
21	GGATATCTGATCCCTGGCCCTGG	0	0	0	0	0	1	2	8	43
22	GGCTCACAGGGTGAACAAGCGG	0	0	0	0	0	0	0	2	5
23	TCCATCCCAGCAGGCTCACAGG	0	0	0	0	0	3	3	8	20
24	AGTACTCCGGATGCAGCTCTGG	0	0	0	0	0	0	1	5	48
25	AGAGCTCCAGGCTCAGATCTGG	0	0	0	0	0	0	4	15	38
26	GATTTTCACACTGACTAAAAGG	0	0	0	0	0	0	0	8	21
27	CCGGGTCATCCATCCATGCAGG	0	0	0	0	0	0	0	4	36
28	CCTCCCTGGAAAGTCCCAGGG	0	0	0	0	0	0	3	14	37
29	GCCACTCCCCAGTCCCAGGG	0	0	0	0	0	0	1	4	14
30	CCGCCAGGCCACGCCCTGG	0	0	0	0	0	1	1	6	19

Фиг. 13

Таблица S2. Олигонуклеотиды для сайтов, на которые нацеливаются gRNA, и праймеры, используемые для ПЦР и секвенирования

Название мишени	Направление	Последовательности (от 5' к 3')
LTR-A	T353: Смыловая	aaacAGGGCCAGGGATCAGATACTCCACTGACCTT _{el}
	T354: Антисмыловая	taaacAAGGTCACTGGATATCTGATCCCTGCCCT
LTR-B	T355: Антисмыловая	aaacAGCTCGATGTCAGCAGTTCTGAAGTACTCgt
	T356: Смыловая	taaacGAGTACTCAAGAACTGCTGACATCGAGCT
LTR-C	T357: Смыловая	caccGATTGGCAGAACATCACACC
	T358: Антисмыловая	aaacGGTGTGTAGTTCTGCCAATC
LTR-D	T359: Смыловая	caccGCGTGGCTGGGCGGACTG
	T360: Антисмыловая	aaacCAGTCCCAGGCCACGC
ПЦР-праймер		
LTR -453/F	Смыловая	TGGAAGGGCTAATTCACTCCAAAC
LTR +43/R	Антисмыловая	CCGAGAGCTCCAGGCTCAGATCT
LTR -411/F	T361: Смыловая	caccGATCTGTGGATCTACCACACACA
LTR +129/R	T363: Антисмыловая	aaacGAGTCACACAACAGACGGGC
Cas-hU6/5'/XhoBm	T351: Смыловая	cgctcgaggatccAGGGCCTATTCCCCATGATTCC
Cas-CAG/3'/EcoR	T352: Антисмыловая	tggatccAGGGGGCCATTACCGTAAGTTATG
U1-хромосома X	T485: Смыловая	ACGACTATCTTATCAATCCTCCTG
	T486: Антисмыловая	CTAGGTGATTAGGATATTCTACAATC
U1-хромосома 2	T492: Смыловая	GCTATTGTATCTGATCACAGCTG
	T493: Антисмыловая	TTGATTGTGTGTCAGGTCTTAGG
d2EGFP	T494: Смыловая	GCAAGGGCGAGGAGCTGTTCAACC
	T495: Антисмыловая	TTGTAGTTGCCGTCGTCCTGAAG
Gag	T457: Смыловая	AATGGTACATCAGGCCATATCAC
	T458: Антисмыловая	CCCACTGTGTTAGCATGGTATT
Cas9	T477: Смыловая	CACAGCATCAAAGAAGAACCTGAT
	T491: Антисмыловая	TCTCCGTCTGGTGTATCTCTTC
RRE	Смыловая	CGCCAAGCTGAATAGGAGCTTTGTCC
	Антисмыловая	CTAGGATCCAGGAGCTGTGATCCTTTAGG
LTR-A	T353: Смыловая	aaacAGGGCCAGGGATCAGATACTCCACTGACCTT _{el}
	T354: Антисмыловая	taaacAAGGTCACTGGATATCTGATCCCTGCCCT
LTR-B	T355: Антисмыловая	aaacAGCTCGATGTCAGCAGTTCTGAAGTACTCgt
	T356: Смыловая	taaacGAGTACTCAAGAACTGCTGACATCGAGCT
LTR-C	T357: Смыловая	caccGATTGGCAGAACATCACACC
	T358: Антисмыловая	aaacGGTGTGTAGTTCTGCCAATC
LTR-D	T359: Смыловая	caccGCGTGGCTGGGCGGACTG
	T360: Антисмыловая	aaacCAGTCCCAGGCCACGC
ПЦР-праймер		
LTR -453/F	Смыловая	TGGAAGGGCTAATTCACTCCAAAC
LTR +43/R	Антисмыловая	CCGAGAGCTCCAGGCTCAGATCT
LTR -411/F	T361: Смыловая	caccGATCTGTGGATCTACCACACACA
LTR +129/R	T363: Антисмыловая	aaacGAGTCACACAACAGACGGGC
Cas-hU6/5'/XhoBm	T351: Смыловая	cgctcgaggatccAGGGCCTATTCCCCATGATTCC
Cas-CAG/3'/EcoR	T352: Антисмыловая	tggatccAGGGGGCCATTACCGTAAGTTATG
U1-хромосома X	T485: Смыловая	ACGACTATCTTATCAATCCTCCTG
	T486: Антисмыловая	CTAGGTGATTAGGATATTCTACAATC
U1-хромосома 2	T492: Смыловая	GCTATTGTATCTGATCACAGCTG
	T493: Антисмыловая	TTGATTGTGTGTCAGGTCTTAGG
d2EGFP	T494: Смыловая	GCAAGGGCGAGGAGCTGTTCAACC
	T495: Антисмыловая	TTGTAGTTGCCGTCGTCCTGAAG
Gag	T457: Смыловая	AATGGTACATCAGGCCATATCAC
	T458: Антисмыловая	CCCACTGTGTTAGCATGGTATT
Cas9	T477: Смыловая	CACAGCATCAAAGAAGAACCTGAT
	T491: Антисмыловая	TCTCCGTCTGGTGTATCTCTTC
RRE	Смыловая	CGCCAAGCTGAATAGGAGCTTTGTCC
	Антисмыловая	CTAGGATCCAGGAGCTGTGATCCTTTAGG

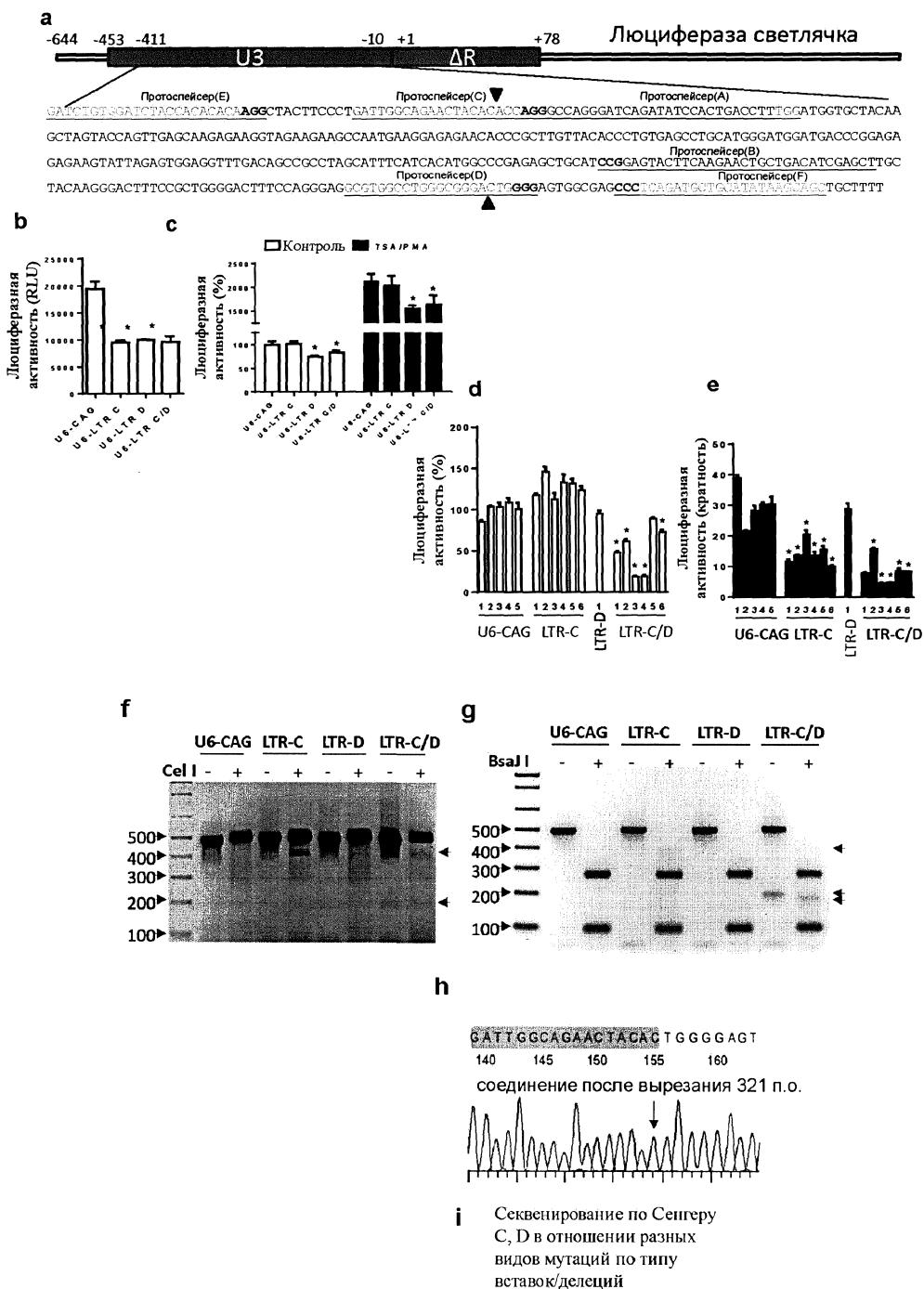
Фиг. 14

Таблица S3. Расположения предсказанных сайтов, на которые нацеливаются gRNA, LTR-A и LTR-B

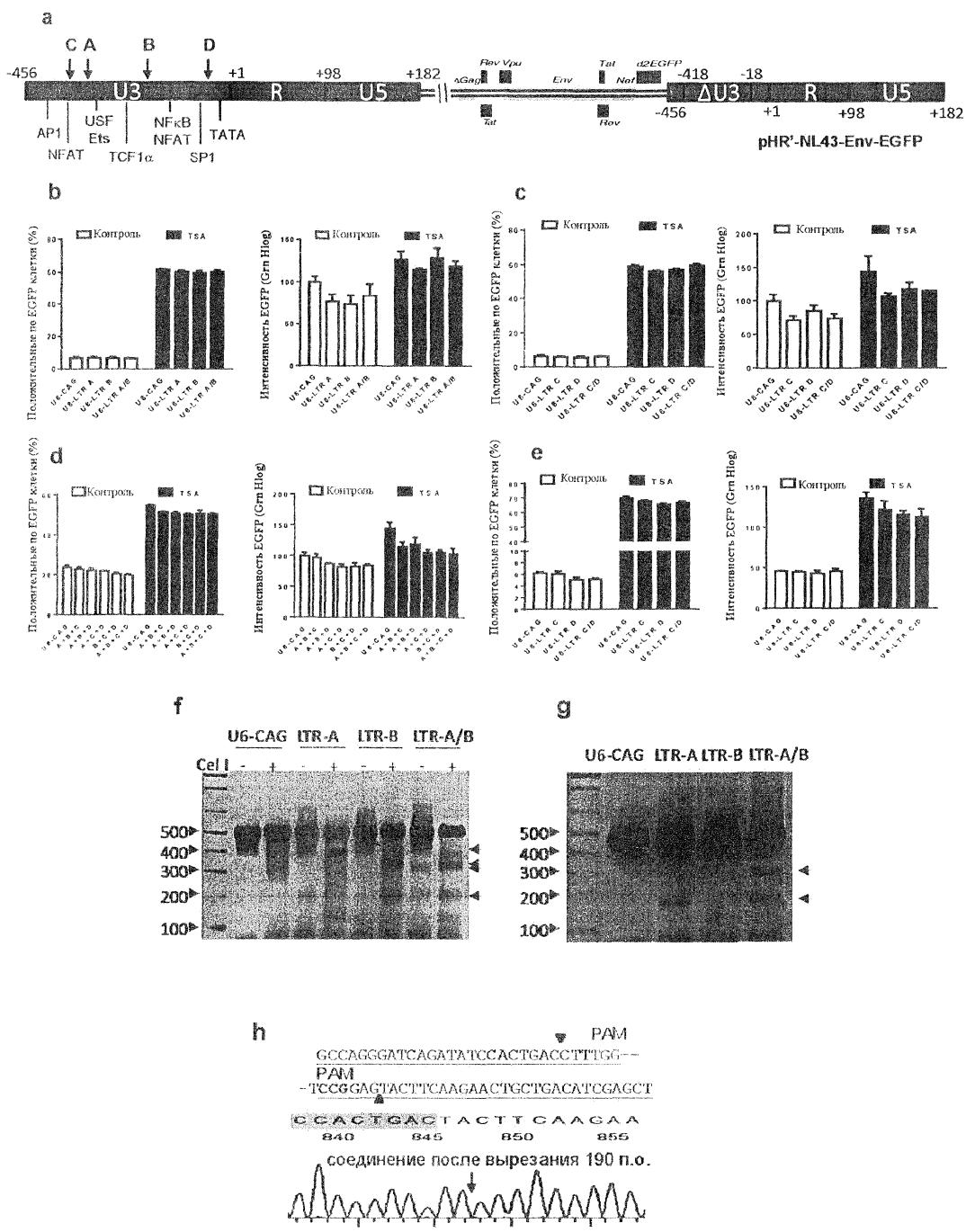
Назна- ние	Запраш. посл.	ID субъекта	Иденти- чность (%)	Е-вели- чина	начало	конец	ниль	Эталон. посл.	исориентир. (заправка 12 по)
LTRA	ATCAGATATCCACTGACCTTGG	ВИЧ	100	7,00E-04	162	184	+	ATCAGATATCCACTGACCTTGG	0
LTRA	TCAAGATATCCACTGACCTTGG	ВИЧ	100	0,003	163	184	+	TCAAGATATCCACTGACCTTGG	0
LTRA	TCAAGATATCCACTGACCTTGG	ВИЧ	100	0,003	9091	9112	+	TCAAGATATCCACTGACCTTGG	0
LTRA	CAGATATCCACTGACCTTGG	ВИЧ	100	0,005	164	184	+	CAGATATCCACTGACCTTGG	0
LTRA	CAGATATCCACTGACCTTGG	ВИЧ	100	0,009	9092	9112	+	CAGATATCCACTGACCTTGG	0
LTRA	AGATATCCACTGACCTTGG	ВИЧ	100	0,033	165	184	+	AGATATCCACTGACCTTGG	0
LTRA	AGATATCCACTGACCTTGG	ВИЧ	100	0,033	9093	9112	+	AGATATCCACTGACCTTGG	0
LTRA	GATATCCACTGACCTTGG	ВИЧ	100	0,12	166	184	+	GATATCCACTGACCTTGG	0
LTRA	GATATCCACTGACCTTGG	ВИЧ	100	0,12	9094	9112	+	GATATCCACTGACCTTGG	0
LTRA	ATATCCACTGACCTTGG	ВИЧ	100	0,42	167	184	+	ATATCCACTGACCTTGG	0
LTRA	ATATCCACTGACCTTGG	ВИЧ	100	0,42	9095	9112	+	ATATCCACTGACCTTGG	0
LTRA	TATCCACTGACCTTGG	chr5	100	1,5	21926317	21926333	+	TATCCACTGACCTTGG	0
LTRA	TATCCACTGACCTTGG	ВИЧ	100	1,5	168	184	+	TATCCACTGACCTTGG	0
LTRA	TATCCACTGACCTTGG	ВИЧ	100	1,5	9096	9112	+	TATCCACTGACCTTGG	0
LTRA	TATCCACTGACCTTAAG	chr3	100	1,5	116712577	116712593	+	TATCCACTGACCTTAAG	0
LTRA	TATCCACTGACCTTGG	chr6	100	1,5	32460607	32460623	+	TATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTAG	chr3	100	5,4	2669092	2669107	+	ATCCACTGACCTTAG	0
LTRA	ATCCACTGACCTTAG	chr3	100	5,4	158293369	158293384	+	ATCCACTGACCTTAG	0
LTRA	ATCCACTGACCTTGG	chr20	100	5,4	46918344	46918359	+	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	chr14	100	5,4	86310067	86310082	-	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	chr5	100	5,4	21926318	21926333	+	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	chr4	100	5,4	95491921	95491936	-	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	ВИЧ	100	5,4	169	184	+	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	ВИЧ	100	5,4	9097	9112	+	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTGG	chr6	100	5,4	98901053	98901068	+	ATCCACTGACCTTGG	0
LTRA	ATCCACTGACCTTAAG	chr7	100	5,4	155511293	155511308	+	ATCCACTGACCTTAAG	0
LTRA	ATCCACTGACCTTAAG	chr3	100	5,4	116712578	116712593	+	ATCCACTGACCTTAAG	0
LTRA	ATCCACTGACCTTCAG	chr5	100	5,4	152371289	152371304	+	ATCCACTGACCTTCAG	0
LTRA	ATCCACTGACCTTCAG	chr4	100	5,4	110823169	110823184	-	ATCCACTGACCTTCAG	0
LTRA	ATCCACTGACCTTGAG	chrX	100	5,4	74260260	74260275	+	ATCCACTGACCTTGAG	0
LTRA	ATCCACTGACCTTGAG	chr6	100	5,4	32460608	32460623	+	ATCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTAGG	chr12	100	20	14485012	14485026	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr7	100	20	72210628	72210642	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr6	100	20	160845640	160845654	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr3	100	20	2669093	2669107	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr3	100	20	158293370	158293384	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr2	100	20	237551230	237551244	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr20	100	20	46918345	46918359	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTGGG	chr14	100	20	86310067	86310081	-	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr12	100	20	116054688	116054702	+	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr11	100	20	103532094	103532108	+	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr10	100	20	132186431	132186445	-	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr8	100	20	144600475	144600489	-	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr5	100	20	21926319	21926333	+	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTGGG	chr4	100	20	95491921	95491935	-	TCCACTGACCTGGG	0
LTRA	TCCACTGACCTTGG	ВИЧ	100	20	170	184	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	ВИЧ	100	20	9098	9112	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr16	100	20	86962569	86962583	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr11	100	20	68156214	68156228	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr6	100	20	98901054	98901068	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr5	100	20	72600080	72600094	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr5	100	20	136458169	136458183	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr4	100	20	25353030	25353044	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTGG	chr2	100	20	207833373	207833387	+	TCCACTGACCTTGG	0
LTRA	TCCACTGACCTTAAG	chr15	100	20	67850506	67850520	-	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr7	100	20	155511293	155511307	-	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr5	100	20	25142541	25142555	-	TCCACTGACCTTAAG	0

LTRA	TCCACTGACCTTAAG	chr3	100	20	116712579	116712593	+	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr1	100	20	163298514	163298528	+	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTCAG	chr20	100	20	22136764	22136778	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr19	100	20	50519462	50519476	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr18	100	20	74623621	74623635	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr16	100	20	71402733	71402747	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr14	100	20	24193180	24193194	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr11	100	20	133664063	133664077	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr9	100	20	140394271	140394285	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr6	100	20	47685256	47685270	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr5	100	20	152371290	152371304	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr4	100	20	110823169	110823183	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr3	100	20	46255327	46255341	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr3	100	20	186757301	186757315	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTGAG	chrX	100	20	74260261	74260275	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr11	100	20	76052171	76052185	-	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	33927660	33927674	-	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	71035331	71035345	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	95871690	95871704	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr7	100	20	137681847	137681861	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr6	100	20	32460609	32460623	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr3	100	20	42344237	42344251	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr2	100	20	64643586	64643600	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTAG	chr16	100	20	55133552	55133566	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTAG	chr15	100	20	90072212	90072226	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr12	100	20	69006300	69006314	+	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr3	100	20	170680338	170680352	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr2	100	20	215414950	215414964	-	TCCACTGACCTTTAG	0
LTRB	CAGCAGTTCTGAAGTACTCCGG		100	7,00E-04	9291	9313	-	CAGCAGTTCTGAAGTACTCCGG	0
LTRB	AGCAGTTCTGAAGTACTCCGG		100	0,003	9291	9312	-	AGCAGTTCTGAAGTACTCCGG	0
LTRB	GCAGTTCTGAAGTACTCCGG		100	0,009	9291	9311	-	GCAGTTCTGAAGTACTCCGG	0
LTRB	CAGTTCTGAAGTACTCCGG		100	0,033	9291	9310	-	CAGTTCTGAAGTACTCCGG	0
LTRB	AGTTCTGAAGTACTCCGG		100	0,12	9291	9309	-	AGTTCTGAAGTACTCCGG	0
LTRB	GTTCTTGAAAGTACTCCGG		100	0,42	9291	9308	-	GTTCTTGAAAGTACTCCGG	0
LTRB	TTCITGAAAGTACTCCGG		100	1,5	9291	9307	-	TTCITGAAAGTACTCCGG	0
LTRB	TCITGAAAGTACTCCGG		100	5,4	9291	9306	-	TCITGAAAGTACTCCGG	0
LTRB	TCTTGAAGTACTCTAG	chr11	100	5,4	91845834	91845849	-	TCTTGAAGTACTCTAG	0
LTRB	CTTGAAGTACTCGAG	chr19	100	20	45672789	45672803	-	CTTGAAGTACTCGAG	0
LTRB	CTTGAAGTACTCGAG	chr15	100	20	82132445	82132459	+	CTTGAAGTACTCGAG	0
LTRB	CTTGAAGTACTCGAG	chr11	100	20	94282411	94282425	+	CTTGAAGTACTCGAG	0
LTRB	CTTGAAGTACTCGAG	chr2	100	20	193312744	193312758	-	CTTGAAGTACTCGAG	0
LTRB	CTTGAAGTACTCCGG		100	20	9291	9305	-	CTTGAAGTACTCCGG	0
LTRB	CTTGAAGTACTCCGG	chr15	100	20	61274973	61274987	-	CTTGAAGTACTCCGG	0
LTRB	CTTGAAGTACTCAAG	chrX	100	20	36051764	36051778	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr15	100	20	31316465	31316479	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr13	100	20	23054474	23054488	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr9	100	20	83208046	83208060	+	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr8	100	20	13956368	13956382	+	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCCAG	chr16	100	20	57449025	57449039	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr15	100	20	41397831	41397845	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr11	100	20	70255488	70255502	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr3	100	20	134149643	134149657	+	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCTAG	chr11	100	20	91845834	91845848	-	CTTGAAGTACTCTAG	0
LTRB	CTTGAAGTACTCTAG	chr1	100	20	24520600	24520614	+	CTTGAAGTACTCTAG	0

Фиг. 15



Фиг. 16



Фиг. 17

LTR ВИЧ-1 (634 п.о.)

Вирус иммунодефицита человека 1 типа, NY5/BRU (LAV-1) рекомбинантный клон pNL4-3.
Номер доступа: M19921

TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTGATCTGTGGATCTACCACACACAA
GGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTGG
участок U3
ATGGTGCTACAAGCTAGTACCGAGTTGAGCAAGAGAAGGTAGAAGAAGCCAATGAAGGAGAGAACAA
CCCGCTTGTACACCCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTATTAGAGTGG
AGGTTTGACAGCCGCCAGCATTCATCACATGGCCCGAGAGCTGCATCCGGACTTCAAGAA
CTGCTGACATCGAGCTTGCTACAAGGGACTTCCGCTGGGACTTTCCAGGGAGGCCGTGGCTGG
GCAGGGACTGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTTGCTTGTACTGG
участок R
GTCTCTCTGGTTAGACCAGATCTGAGCCTGGAGCTCTGGCTAACTAGGAAACCCACTGCTTA
участок U5
AGCCTCAATAAAGCTGCCTTGAGTGCTCAAGTAGTGTGTGCCGTGTTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTAGTCAGTGTGGAAAATCTCTAGCA

Фиг. 18

SIVmm239-LTR (818 п.о.)

Вирус иммунодефицита обезьян (макак), изолят 239
Номер в Genbank: M33262

TGGAAGGGATTATTACAGTGCAAGAAGACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCAT
ACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTGGCTGGCTATGGAAATTAA
участок U3
GTCCCTGTAAATGTATCAGATGAGGCACAGGAGGATGAGGAGCATTATTAATGCATCCAGCTCAAACCTCCC
AGTGGGATGACCCCTGGGAGAGGTTCTAGCATGGAAGTTGATCCAACCTCTGGCCTACACTTATGAGGCATA
TGTTAGATACCCAGAAGAGTTGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCA
AGAGGCCTCTTAACATGGCTGACAAGAAGGAAACTCGCTGAAACAGCAGGGACTTCCACAAAGGGATGTTA
CGGGGAGGTACTGGGAGGGAGCCGGTCGGGAACGCCACTTCTTGATGTATAAATATCACTCCATTTCGCTC
TGTATTCACTCGCTCTGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAG
участок R
CCTGGGTGTTCCCTGCTAGACTCTCACCGACACTTGGCCGGTGCTGGCAGAGTGACTIONCCACGCTTGCTTGC
TAAAGCCCTCTCAATAAAGCTGCCATTAGAAGTAAGCTAGTGTGTGTTCCCATCTCCTAGCCGCC
участок U5
TGGTCAACTCGGTACTCAATAAAGAAGACCCTGGCTGTTAGGACCCCTTCTGCTTGGGAAACCGAAGCA
GGAAAATCCCTAGCA

Фиг. 19

