

(19)



Евразийское
патентное
ведомство

(21) 202391644 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.10.02(51) Int. Cl. A61K 48/00 (2006.01)
A61K 31/395 (2006.01)
A61P 7/00 (2006.01)(22) Дата подачи заявки
2021.12.02

(54) КОМПОЗИЦИИ И СПОСОБЫ НАЦЕЛИВАНИЯ НА BCL11A

(31) 63/120,885

(32) 2020.12.03

(33) US

(86) PCT/US2021/061672

(87) WO 2022/120094 2022.06.09

(88) 2022.08.25

(71) Заявитель:

СКРАЙБ ТЕРАПЬЮТИКС ИНК. (US)

(72) Изобретатель:

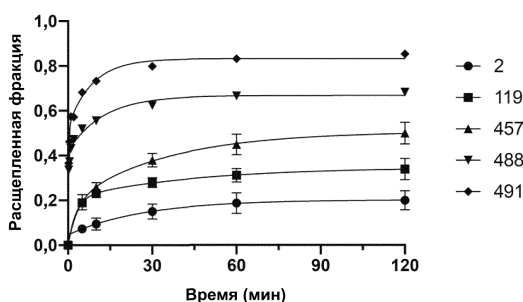
Оукс Бенджамин, Хиггинс Шон,
Денни Сара, Стаал Бретт Т., Колин
Изабел, Адил Маруф, Урнес Коул (US)

(74) Представитель:

Медведев В.Н. (RU)

(57) В настоящем документе предложены системы, содержащие полипептиды CRISPR класса 2, типа V, гидовые нуклеиновые кислоты (гНК) и необязательно нуклеиновые кислоты донорных матриц, используемые для модификации гена BCL11A. Системы также используются для модификации клеток у субъектов со связанным с гемоглобинопатией заболеванием. Также предложены способы лечения субъектов, имеющих связанное с гемоглобинопатией заболевание, посредством введения систем или нуклеиновых кислот, кодирующих такие системы, которые нацелены на ген BCL11A у таких субъектов.

Активные фракции вариантов белка



A1

202391644

202391644

A1

КОМПОЗИЦИИ И СПОСОБЫ НАЦЕЛИВАНИЯ НА BCL11A**ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА СМЕЖНЫЕ ЗАЯВКИ**

[0001] Настоящая заявка испрашивает приоритет по предварительной заявке на патент США № 63/120,885, поданной 3 декабря 2020 г., содержание которой полностью включено в настоящий документ путем ссылки.

ВКЛЮЧЕНИЕ ПЕРЕЧНЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПУТЕМ ССЫЛКИ

[0002] Настоящая заявка содержит перечень последовательностей, который был представлен в формате ASCII посредством EFS-WEB и полностью включен в настоящий документ путем ссылки. Копия упомянутого перечня в формате ASCII, созданная 1 декабря 2021 г., называется SCRB_030_01WO_SeqList_ST25.txt и имеет размер 8,78 Мб.

ПРЕДПОСЫЛКИ СОЗДАНИЯ ИЗОБРЕТЕНИЯ

[0003] Эмбриональный гемоглобин (также гемоглобин F, HbF или $\alpha_2\gamma_2$) представляет собой основной белок-носитель кислорода у эмбриона человека. HbF имеет состав, отличный от форм гемоглобина у взрослых, который позволяет ему сильнее связываться с кислородом, чем взрослая форма, что позволяет развивающемуся плоду извлекать кислород из кровотока матери. HbF представляет собой тетрамер из двух взрослых полипептидов α -глобина и двух эмбриональных β -подобных полипептидов γ -глобина. Во время беременности дублированные гены γ -глобина являются преобладающими генами, транскрибируемыми в кластере β -глобина. После рождения γ -глобин замещается взрослым β -глобином, что называется «эмбриональным переключением», и этот процесс включает в себя экспрессию BCL11A, регулятора сайленсинга HbF (Sankaran, V.G., et al. Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. *Science* 322(5909):1839-1842 (2008); Liu, N., et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell* 173(2):430 (2018)). У здоровых взрослых в состав гемоглобина входит гемоглобин A (~ 97%), гемоглобин A2 (2,2-3,5%) и гемоглобин F (<1%) (Thomas, C and Lumb, A.B. Physiology of haemoglobin. *Continuing Education in Anaesthesia Critical Care & Pain*. 12(5): 251-256 (2012)).

[0004] Гемоглобинопатии представляют собой наследственные моногенные расстройства, которые в большинстве случаев наследуются как аутосомные кодоминантные признаки. К распространенным гемоглобинопатиям относятся серповидноклеточная анемия и α - и β -талассемии. Гемоглобинопатии наиболее распространены среди населения Африки, Средиземноморского бассейна и Юго-Восточной Азии. Большинство гемоглобинопатий, включая серповидноклеточную анемию, являются просто структурными аномалиями в самих глобиновых белках. Серповидноклеточную анемию вызывает точечная мутация в структурном гене β -глобина *HBB*, что приводит к образованию аномального гемоглобина (HbS), и это ведет к снижению способности крови переносить кислород. С другой стороны, талассемии обычно приводят

к недостаточной продукции нормальных глобинов, зачастую вследствие мутации в регуляторных генах, что приводит к дефициту или отсутствию взрослого гемоглобина (HbA). При β -талассемии, при которой наблюдается дефицит β -глобина, увеличение экспрессии γ -глобина снижает дисбаланс между цепями α - и β -глобина, лежащего в основе патофизиологии анемии при этом заболевании (Liu, N., et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell* 173(2): 430 (2018)). И серповидноклеточная анемия, и талассемия могут приводить к анемии.

[0005] Белок В-клеточной лимфомы/лейкоза 11A (BCL11A) представляет собой белок, который у людей кодируется геном *BCL11A*. Во время дифференцировки гемопоэтических клеток этот ген подавлен, и было обнаружено, что он играет роль в подавлении продукции эмбрионального гемоглобина. BCL11A представляет собой основной репрессорный белок продукции гемоглобина F, действующий посредством связывания с геном, кодирующим субъединицу γ , в промоторной области (Sankaran VG, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322:1839 (2008)). Так как повышенный γ -глобин снижает клиническую тяжесть β -гемоглобинопатий, серповидноклеточной анемии и β -талассемии, вызванных мутацией или сниженной экспрессией β -глобина соответственно, редактирование гена *BCL11A* для повышения экспрессии γ -глобина за пределы остаточного $\sim 1\%$ эмбрионального гемоглобина предлагается в качестве привлекательной терапевтической стратегии у взрослых людей с гемоглобинопатиями (Smith, E.C., et al. Strict in vivo specificity of the Bcl11a erythroid enhancer. *Blood* 128(19):2338 (2016)).

[0006] Появление систем CRISPR/Cas и программируемый характер этих минимальных систем упрощают их применение в качестве универсальной технологии геномных манипуляций и конструирования. В настоящее время применение систем CRISPR/Cas для лечения гемоглобинопатий ограничено редактированием клеток *ex vivo* с последующей трансплантацией субъектам, страдающим базовой гемоглобинопатией. Таким образом, существует потребность в композициях и способах регуляции *BCL11A* для снижения прямого подавления промотора гена γ -глобина *in vivo* у субъектов с этими заболеваниями. В настоящем документе предложены композиции и способы нацеливания на ген *BCL11A* для удовлетворения этой потребности.

ИЗЛОЖЕНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

[0007] Настоящее описание относится к композициям модифицированных белков CRISPR класса 2, типа V и гидовых нуклеиновых кислот, применяемых для изменения в клетках целевой нуклеиновой кислоты, содержащей ген BCL11A. Белки CRISPR класса 2, типа V и гидовые нуклеиновые кислоты модифицированы для пассивного входа в клетки-мишени. Белки CRISPR класса 2, типа V и направляющие нуклеиновые кислоты используются в разнообразных способах модификации целевой нуклеиновой кислоты при связанных с BCL11A заболеваниях, и способы также предложены.

[0008] В одном аспекте настоящее описание относится к системам CasX:гидовая нуклеиновая кислота (системы CasX:гРНК) и к способам, применяемым для нокдауна или

нокаута гена BCL11A с целью уменьшения или устранения экспрессии продукта гена BCL11A у субъектов, имеющих связанное с β -гемоглобинопатией заболевание.

[0009] В некоторых вариантах осуществления гРНК в системе CasX:гРНК представляет собой гРНК или химеру РНК и ДНК и может представлять собой одиночную молекулу гРНК или двойную молекулу гРНК. В других вариантах осуществления гРНК в системе CasX:гРНК имеет нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, содержащей область внутри гена BCL11A. В некоторых вариантах осуществления нацеливающая последовательность гРНК выбрана из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789 или последовательности, идентичной ей на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 85%, по меньшей мере примерно 90% или по меньшей мере примерно 95%. гРНК может содержать нацеливающую последовательность, содержащую от 15 до 20 последовательных нуклеотидов. В других вариантах осуществления нацеливающая последовательность гРНК состоит из 20 нуклеотидов. В других вариантах осуществления нацеливающая последовательность состоит из 19 нуклеотидов. В других вариантах осуществления нацеливающая последовательность состоит из 18 нуклеотидов. В других вариантах осуществления нацеливающая последовательность состоит из 17 нуклеотидов. В других вариантах осуществления нацеливающая последовательность состоит из 16 нуклеотидов. В других вариантах осуществления нацеливающая последовательность состоит из 15 нуклеотидов. В других вариантах осуществления нацеливающая последовательность гРНК имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789. В других вариантах осуществления нацеливающая последовательность гРНК имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, с одним нуклеотидом, удаленным с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность состоит из 18 нуклеотидов, имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, с двумя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность состоит из 17 нуклеотидов, имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, с тремя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность состоит из 16 нуклеотидов, имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, с четырьмя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность состоит из 15 нуклеотидов, имеет последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, с пятью нуклеотидами, удаленными с 3'-конца последовательности.

[0010] В некоторых вариантах осуществления гРНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из последовательностей SEQ ID NO:

состоящей из SEQ ID NO: 132-148 и 26908-27154. В этих вариантах осуществления вариант CasX демонстрирует одну или более улучшенных характеристик относительно любого из референсных белков CasX с SEQ ID NO: 1-3. В некоторых вариантах осуществления вариант белка CasX обладает аффинностью связывания с последовательностью смежного с протоспейсером мотива (PAM), выбранной из группы, состоящей из TTC, ATC, GTC и CTC. В некоторых вариантах осуществления вариант белка CasX обладает аффинностью связывания с PAM-последовательностью, которая в по меньшей мере 1,5 раза выше по сравнению с аффинностью связывания любого из референсных белков CasX с SEQ ID NO: 1-3 в отношении PAM-последовательностей, выбранных из группы, состоящей из TTC, ATC, GTC и CTC.

[0012] В других вариантах осуществления системы CasX:гРНК молекула CasX и молекула гРНК связаны друг с другом в рибонуклеопротеиновый комплекс (RNP). В конкретном варианте осуществления RNP, содержащий вариант CasX и вариант гРНК, демонстрирует более высокую эффективность редактирования и/или связывания с целевой последовательностью в целевой ДНК, когда любая из PAM-последовательностей TTC, ATC, GTC или CTC расположена на расстоянии 1 нуклеотида к 5'-концу от последовательности нецелевой нити, имеющей идентичность с нацеливающей последовательностью гРНК в системе клеточного анализа по сравнению с эффективностью редактирования и/или связывания RNP, содержащего референсный белок CasX и референсную гРНК, в сопоставимой системе анализа.

[0013] В некоторых вариантах осуществления система CasX:гРНК дополнительно содержит донорную матрицу, содержащую нуклеиновую кислоту, включающую по меньшей мере участок гена BCL11A и имеющую по меньшей мере от 1 до примерно 5 мутаций относительно последовательности дикого типа, причем участок гена BCL11A выбран из группы, состоящей из экзона BCL11A, интрона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A или их комбинаций, при этом донорная матрица используется для нокдауна или нокаута гена BCL11A. В некоторых случаях донорная последовательность представляет собой одноцепочечную ДНК-матрицу или одноцепочечную РНК-матрицу. В других случаях донорная матрица представляет собой двухцепочечную ДНК-матрицу.

[0014] В других вариантах осуществления описание относится к нуклеиновым кислотам, кодирующим системы CasX:гРНК любого из вариантов осуществления, описанных в настоящем документе, а также векторов, содержащих нуклеиновые кислоты. В некоторых вариантах осуществления вектор выбран из группы, состоящей из ретровирусного вектора, лентивирусного вектора, аденовирусного вектора, аденоассоциированного вирусного вектора (AAV), вектора вируса простого герпеса (HSV), плазмиды, мини-кольца, наноплазмиды и РНК-вектора. В других вариантах осуществления вектор представляет собой частицу для доставки CasX (XDP), содержащую RNP из CasX и гРНК любого из вариантов осуществления, описанных в настоящем документе, и необязательно нуклеиновую кислоту донорной матрицы и нацеливающий фрагмент, такой

как вирусный гликопротеин.

[0015] В других вариантах осуществления в описании предложен способ модификации последовательности целевой нуклеиновой кислоты BCL11A в клетках популяции, включающий введение в клетку: а) системы CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе; b) нуклеиновой кислоты по любому из вариантов осуществления, описанных в настоящем документе; c) вектора по любому из вариантов осуществления, описанных в настоящем документе; d) XDP по любому из вариантов осуществления, описанных в настоящем документе; или e) комбинации вышеизложенного. В некоторых вариантах осуществления способа модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в последовательность целевой нуклеиновой кислоты по сравнению с последовательностью дикого типа. Целевой ген BCL11A включает в качестве регуляторного элемента сайт связывания эритроид-специфического энхансера GATA1 (GATA1). В некоторых вариантах осуществления способ модификации включает модификацию последовательности GATA1, причем ген BCL11A в результате модификации подвергается нокауту или нокдауну. В некоторых случаях способ дополнительно включает приведение целевой нуклеиновой кислоты в контакт с нуклеиновой кислотой донорной матрицы по любому из вариантов осуществления, описанных в настоящем документе. В некоторых вариантах осуществления способа донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок гена BCL11A, но с одной или более мутациями для нокаута или нокдауна гена BCL11A. В некоторых случаях модификация последовательности целевой нуклеиновой кислоты происходит *in vitro* или *ex vivo*. В некоторых случаях модификация последовательности целевой нуклеиновой кислоты происходит *in vivo*. В некоторых вариантах осуществления клетка представляет собой эукариотическую клетку, выбранную из группы, состоящей из клетки грызуна, клетки мыши, клетки крысы, клетки примата и клетки примата, не являющегося человеком. В некоторых вариантах осуществления клетка представляет собой клетку человека. В некоторых вариантах осуществления клетка выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта. В некоторых вариантах осуществления клетка представляет собой аутологичную клетку, полученную от субъекта со связанным с β -гемоглобинопатией заболеванием. В других вариантах осуществления клетка является аллогенной, но от того же вида, что и субъект, подлежащий лечению.

[0016] В других вариантах осуществления в описании предложены способы модификации последовательности целевой нуклеиновой кислоты гена BCL11A, причем клетки-мишени популяции приводят в контакт с векторами, кодирующими белок CasX и одну или более гРНК, содержащих нацеливающую последовательность, комплементарную гену BCL11A, и необязательно дополнительно содержащих донорную матрицу. В

некоторых случаях вектор представляет собой аденоассоциированный вирусный вектор (AAV), выбранный из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 или AAVRh10. В других случаях вектор представляет собой лентивирусный вектор. В других вариантах осуществления в описании предложены способы, в которых клетки-мишени приводят в контакт с применением вектора, и причем вектор представляет собой частицу для доставки CasX (XDP), содержащую RNP из CasX и гРНК любого из вариантов осуществления, описанных в настоящем документе, и необязательно нуклеиновую кислоту донорной матрицы. В некоторых вариантах осуществления способа вектор вводят субъекту в терапевтически эффективной дозе. Субъект может представлять собой мышь, крысу, свинью, примата, не являющегося человеком, или человека. Доза может быть введена способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0017] В других вариантах осуществления в описании предложен способ лечения связанного с β -гемоглобинопатией заболевания у нуждающегося в этом субъекта, включающий модификацию гена, кодирующего ген BCL11A, в клетке субъекта, причем модификация включает одно из приведения упомянутой клетки в контакт с: а) системой CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе; b) нуклеиновой кислоты по любому из вариантов осуществления, описанных в настоящем документе; c) вектора по любому из вариантов осуществления, описанных в настоящем документе; d) XDP по любому из вариантов осуществления, описанных в настоящем документе; или e) комбинации вышеизложенного. В некоторых вариантах осуществления связанное с β -гемоглобинопатией заболевание представляет собой серповидноклеточную анемию или бета-талассемию. В некоторых случаях способы лечения субъекта со связанным с β -гемоглобинопатией заболеванием приводят к улучшению по меньшей мере одного клинически значимого параметра. В других случаях способы лечения субъекта со связанным с β -гемоглобинопатией заболеванием приводят к улучшению по меньшей мере двух клинически значимых параметров.

[0018] В других вариантах осуществления в описании предложено применение систем CasX:гРНК, нуклеиновых кислот, векторов или XDP, описанных в настоящем документе, для лечения связанного с β -гемоглобинопатией заболевания у нуждающегося в этом субъекта. В некоторых вариантах осуществления применение включает модификацию гена, кодирующего ген BCL11A, в клетке субъекта, причем модификация включает одно из приведения упомянутой клетки в контакт с: а) системой CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе; b) нуклеиновой кислоты по любому из вариантов осуществления, описанных в настоящем документе; c) вектора по любому из вариантов осуществления, описанных в настоящем документе; d) XDP по любому из вариантов осуществления, описанных в настоящем документе; или e) комбинации вышеизложенного.

ВКЛЮЧЕНИЕ ПУТЕМ ССЫЛКИ

[0019] Все публикации, патенты и заявки на патенты, упомянутые в настоящем

описании, включены в настоящий документ путем ссылки в той степени, как если бы каждая отдельная публикация, патент или заявка на патент была конкретно и индивидуально обозначена как включенная путем ссылки. Содержания предварительных заявок на патент США 63/121,196, поданной 3 декабря 2020 г., 63/162,346, поданной 17 марта 2021 г., и 63/208,855, поданной 9 июня 2021 г., которые раскрывают варианты CasX и варианты гРНК, полностью включены в настоящий документ путем ссылки. Содержания опубликованных заявок на международный патент WO2020/247882, опубликованной 10 декабря 2020 г., WO2020/247883, опубликованной 10 декабря 2020 г., и WO2021/113772, опубликованной 10 июня 2021 г., полностью включены в настоящий документ путем ссылки.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0020] Новые признаки описания подробно описаны в приложенной формуле изобретения. Для лучшего понимания признаков и преимуществ настоящего описания рекомендуется изучить приведенное ниже подробное описание, в котором изложены наглядные примеры осуществления, в которых используются принципы описания, а также прилагаемые графические материалы, на которых:

[0021] На ФИГ. 1 представлен график результатов анализа по количественному определению активных фракций RNP, образованных огРНК 174 (SEQ ID NO: 2238) и вариантами CasX 119 (SEQ ID NO: 59), 457 (SEQ ID NO: 101), 488 (SEQ ID NO: 123) и 491 (SEQ ID NO: 126) CasX, описанными в примере 8. Эквимольные количества RNP и мишень совместно инкубировали и определяли количество расщепленной мишени в указанных временных точках. Для каждой временной точки показаны среднее и стандартное отклонение для трех независимых повторов. Показана аппроксимация объединенных повторов двухфазной моделью. «2» относится к референсному белку CasX с SEQ ID NO: 2.

[0022] На ФИГ. 2 представлено количественное определение активных фракций RNP, образованных CasX2 (референсный белок CasX с SEQ ID NO:2) и модифицированными огРНК, как описано в примере 8. Эквимольные количества RNP и мишень совместно инкубировали и определяли количество расщепленной мишени в указанных временных точках. Для каждой временной точки показаны среднее и стандартное отклонение для трех независимых повторов. Показана аппроксимация объединенных повторов двухфазной моделью.

[0023] На ФИГ. 3 представлено количественное определение активных фракций RNP, образованных CasX 491 и модифицированными огРНК в условиях ограничения гена, как описано в примере 8. Эквимольные количества RNP и мишень совместно инкубировали и определяли количество расщепленной мишени в указанных временных точках. Показана аппроксимация данных двухфазной моделью.

[0024] На ФИГ. 4 представлено количественное определение скоростей расщепления RNP, образованного огРНК174 и вариантами CasX, как описано в примере 8. Целевую ДНК инкубировали с 20-кратным избытком указанного RNP и определяли количество расщепленной мишени в указанных временных точках. Среднее и стандартное

отклонение трех независимых повторов показаны для каждой временной точки, за исключением 488 и 491, где показан один повтор. Показана однофазная аппроксимация объединенных повторов.

[0025] На ФИГ. 5 представлено количественное определение скоростей расщепления RNP, образованного CasX2 и указанными вариантами гРНК, как описано в примере 8. Целевую ДНК инкубировали с 20-кратным избытком указанного RNP и определяли количество расщепленной мишени в указанных временных точках. Для каждой временной точки показаны среднее и стандартное отклонение для трех независимых повторов. Показана однофазная аппроксимация объединенных повторов.

[0026] На ФИГ. 6 представлено количественное определение начальных скоростей для RNP, образованного CasX2 и вариантами гРНК, как описано в примере 8. Первые две временные точки из предыдущего эксперимента по расщеплению соответствовали аппроксимации линейной моделью для определения начальной скорости расщепления.

[0027] На ФИГ. 7 представлено количественное определение скоростей расщепления RNP, образованного CasX491 и вариантами гРНК, как описано в примере 8. Целевую ДНК инкубировали с 20-кратным избытком указанного RNP при 10 °C, и количество расщепленной мишени определяли в указанных временных точках. Показана однофазная аппроксимация временных точек.

[0028] На ФИГ. 8 представлено количественное определение компетентных фракций RNP из варианта 515 CasX (SEQ ID NO: 133) и 526 (SEQ ID NO: 143) в комплексе с вариантом 174 гРНК по сравнению с RNP из референсного CasX 2 в комплексе с гРНК 2 с применением эквимольных количеств указанного RNP и комплементарной мишени, как описано в примере 8. Показана аппроксимация двухфазной моделью для каждого периода времени или набора комбинированных повторов.

[0029] На ФИГ. 9 представлено количественное определение скоростей расщепления RNP из вариантов 515 и 526 CasX в комплексе с вариантом 174 гРНК по сравнению с RNP из референсного CasX 2 в комплексе с гРНК 2 с применением 20-кратного избытка указанного RNP, как описано в примере 8.

[0030] На ФИГ. 10А представлено количественное определение скоростей расщепления вариантов CasX на РАМ ТТС, как описано в примере 5. Целевые ДНК-субстраты с идентичными спейсерами и указанной РАМ-последовательностью инкубировали с 20-кратным избытком указанного RNP при 37°C и определяли количество расщепленной мишени в указанных временных точках. Показана однофазная аппроксимация одиночного повтора.

[0031] На ФИГ. 10В представлено количественное определение скоростей расщепления вариантов CasX на РАМ СТС, как описано в примере 5. Целевые ДНК-субстраты с идентичными спейсерами и указанной РАМ-последовательностью инкубировали с 20-кратным избытком указанного RNP при 37°C и определяли количество расщепленной мишени в указанных временных точках. Показана однофазная аппроксимация одиночного повтора.

[0032] На ФИГ. 10C представлено количественное определение скоростей расщепления вариантов CasX на РАМ GTC, как описано в примере 5. Целевые ДНК-субстраты с идентичными спейсерами и указанной РАМ-последовательностью инкубировали с 20-кратным избытком указанного RNP при 37 °C и определяли количество расщепленной мишени в указанных временных точках. Показана однофазная аппроксимация одиночного повтора.

[0033] На ФИГ. 10D представлено количественное определение скоростей расщепления вариантов CasX на РАМ ATC, как описано в примере 5. Целевые ДНК-субстраты с идентичными спейсерами и указанной РАМ-последовательностью инкубировали с 20-кратным избытком указанного RNP при 37 °C и определяли количество расщепленной мишени в указанных временных точках. Показана однофазная аппроксимация одиночного повтора.

[0034] На ФИГ. 11A представлено количественное определение скоростей расщепления RNP из варианта 491 CasX и гида 174 на РАМ NTC, как описано в примере 5. Временные точки были взяты в пределах 10 минут, и расщепленную фракцию графически изображали для каждой мишени и временной точки, но для ясности показаны только первые две минуты временного периода.

[0035] На ФИГ. 11B представлено количественное определение скоростей расщепления RNP из варианта 491 CasX и гида 174 на РАМ NTT, как описано в примере 5. Временные точки были взяты в пределах 10 минут, и расщепленную фракцию графически изображали для каждой мишени и временной точки.

[0036] На ФИГ. 12A представлено количественное определение расщепления с помощью RNP, образованного ogРНК174 и вариантом 515 CasX, с использованием спейсеров длиной 18, 19 или 20 нуклеотидов, как описано в примере 9. Целевую ДНК инкубировали с 20-кратным избытком указанного RNP и определяли количество расщепленной мишени в указанных временных точках. Для каждой временной точки показаны среднее и стандартное отклонение для трех независимых повторов. Показана однофазная аппроксимация объединенных повторов.

[0037] На ФИГ. 12B представлено количественное определение расщепления с помощью RNP, образованного ogРНК174 и вариантом 526 CasX, с использованием спейсеров длиной 18, 19 или 20 нуклеотидов, как описано в примере 9. Целевую ДНК инкубировали с 20-кратным избытком указанного RNP и определяли количество расщепленной мишени в указанных временных точках. Для каждой временной точки показаны среднее и стандартное отклонение для трех независимых повторов. Показана однофазная аппроксимация объединенных повторов.

[0038] На ФИГ. 13 представлено схематическое изображение примера белка CasX и каркасной последовательности ДНК для упаковки в аденоассоциированный вирус (AAV). Сегмент ДНК между инвертированными концевыми повторами (ITR) в AAV, состоящий из ДНК, кодирующей CasX, и ее промотора, и ДНК, кодирующей каркас, и ее промотора, упаковывают в капсид AAV во время получения AAV.

[0039] На ФИГ. 14 представлены результаты анализа редактирования, в котором сравниваются гРНК-каркасы 229-237 (см. таблицу 3 для соответствующих последовательностей и SEQ ID NO) с каркасом 174 в мышинных нейрональных клетках-предшественниках (mNPC), выделенных из трансгенных мышей Ai9-tdtomato. Клетки нуклеофицировали указанными дозами плазмид p59, кодирующих CasX 491, каркас и спейсер 11.30 (5' AAGGGGCUCCGCACCACGCC 3', SEQ ID NO: 27197), нацеленных на mRHO. Редактирование в локусе mRHO оценивали через 5 дней после трансфекции методом NGS, и редактирование с помощью конструкторов с каркасами 230, 231, 234 и 235, как продемонстрировано, было больше по сравнению с конструктами с каркасом 174 при обеих дозах.

[0040] На ФИГ. 15 представлены результаты анализа редактирования со сравнением гРНК-каркасов 229-237 с каркасом 174 в клетках mNPC. Клетки нуклеофицировали указанными дозами плазмид p59, кодирующих CasX 491, каркас и спейсер 12.7 (5' CUGCAUUCUAGUUGUGGUUU 3', SEQ ID NO: 27198), нацеленных на повторяющиеся элементы, предотвращающие экспрессию флуоресцентного белка tdTomato. Редактирование оценивали через 5 дней после трансфекции методом FACS для количественного определения доли tdTomato-положительных клеток. Клетки, нуклеофицированные каркасами 231-235, демонстрировали приблизительно на 35% большее редактирование по сравнению с конструктами с каркасом 174 в высокой дозе и приблизительно на 25% большее редактирование при низкой дозе.

[0041] На ФИГ. 16 представлены результаты анализа редактирования со сравнением CasX-нуклеаз 2, 119, 491, 515, 527, 528, 529, 530 и 531 (см. таблицу 4 для соответствующих последовательностей и SEQ ID NO) в специально разработанной линии клеток HEK293, PASS_V1.01. Клетки липофицировали 2 мкг плазмиды p67, кодирующей указанный белок CasX. Через пять дней извлекали геномную ДНК клеток. Проводили ПЦР-амплификацию и секвенирование нового поколения для выделения и количественного определения доли отредактированных клеток в специально разработанных сайтах редактирования на мишенях. Для каждой пробы редактирование оценивали в целевых сайтах (индивидуальных точках), состоящих из следующих РАМ-последовательностей: индивидуальные сайты 48 ТТС, 14 АТС, 22 СТС, 11 ГТС, и процентную долю редактирования нормализовали к контролю - несущей среде. Клетки, липофицированные любой нуклеазой, демонстрировали более высокое среднее редактирование в целевых сайтах РАМ ТТС (горизонтальная полоса), чем в случае нуклеазы дикого типа CasX 2, за исключением CasX 528. Относительное предпочтение любой данной нуклеазы к четырем разным РАМ-последовательностям также представлено на скрипичных графиках. В частности, CasX-нуклеазы 527, 528 и 529 демонстрируют по существу иные предпочтения РАМ, чем CasX 2 дикого типа.

[0042] На ФИГ. 17 представлены результаты анализа редактирования, где сравнивается улучшенная CasX-нуклеаза 491 с улучшенными нуклеазами 532 и 533 в специально разработанной линии клеток HEK293, PASS_V1.01. Клетки липофицировали в

двух повторностях 2 мкг плазмиды р67, кодирующей указанный белок CasX и ген резистентности к пурамицину, и выращивали в условиях пурамициновой селекции. Через три дня извлекали геномную ДНК клеток. Проводили ПЦР-амплификацию и секвенирование нового поколения для выделения и количественного определения доли отредактированных клеток в специально разработанных сайтах редактирования на мишенях. Для каждой пробы редактирование оценивали в целевых сайтах, состоящих из следующих РАМ-последовательностей: индивидуальные сайты 48 ТТС, 14 АТС, 22 СТС, 11 ГТС, и долю редактирования нормализовали к контролю - несущей среде. Клетки, липофицированные CasX 532 или 533, демонстрировали в среднем большее редактирование, чем Cas 491, в случае каждой из РАМ-последовательностей, за исключением CasX 533 в целевых сайтах ТТС РАМ. Планки погрешностей представляют стандартную ошибку среднего для n=2 биологических проб.

[0043] На ФИГ. 18 представлены результаты редактирования локуса эритроид-специфичного энхансера BCL11A в клетках НЕК293Т с помощью варианта 438 белка CasX с каркасом 174 по сравнению с системой Cas9, как описано в примере 13.

[0044] На ФИГ. 19 представлены результаты редактирования в области связывания GATA1 локуса эритроид-специфичного энхансера в BCL11A в клетках K562 с помощью варианта 491 белка CasX с каркасом 174 по сравнению с вариантом 119 белка CasX с каркасом 174, как описано в примере 14.

[0045] На ФИГ. 20 представлены результаты редактирования в области связывания GATA1 локуса эритроид-специфичного энхансера в BCL11A в клетках K562 с помощью варианта 491 белка CasX с каркасом 174, доставляемым различными дозами XDP, как описано в примере 14.

[0046] На ФИГ. 21 представлены результаты редактирования в области связывания GATA1 локуса эритроид-специфичного энхансера в BCL11A в клетках HSC с помощью варианта 491 белка CasX с каркасом 174 по сравнению с вариантом 119 белка CasX с каркасом 174, как описано в примере 15.

[0047] На ФИГ. 22 представлены результаты редактирования в области связывания GATA1 локуса эритроид-специфичного энхансера в BCL11A в клетках HSC с помощью варианта 491 белка CasX с каркасом 174, доставляемым различными дозами XDP, как описано в примере 15.

[0048] На ФИГ. 23 представлено схематическое изображение расположения спейсера 21.1 (SEQ ID NO: 22) относительно последовательности сайта связывания GATA1 в целевой нуклеиновой кислоте. Верхняя цепь: SEQ ID NO: 26790, нижняя цепь: SEQ ID NO: 26791.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0049] Несмотря на то что в настоящем документе показаны и описаны примеры осуществления, специалистам в данной области будет очевидно, что такие варианты осуществления предложены только в качестве примера. Специалистам в данной области будут понятны многочисленные вариации, изменения и замены без выхода за рамки

приведенной в документе формулы изобретения. Следует понимать, что для практического применения вариантов осуществления описания можно использовать различные альтернативы вариантам осуществления, описанным в настоящем документе. Предполагается, что пункты формулы изобретения определяют объем изобретения и что способы и структуры в рамках пунктов формулы, а также их эквиваленты охвачены формулой изобретения.

[0050] Если не указано иное, все технические и научные термины, применяемые в настоящем документе, имеют общепринятое значение, понятное любому рядовому специалисту в данной области, к которой относится настоящее изобретение. В настоящем документе описаны способы и материалы, хотя для проверки или анализа настоящих вариантов осуществления можно применять приемлемые способы и материалы, аналогичные или эквивалентные описанным ниже. В случае противоречий приоритет будет иметь настоящее описание патента, включая определения. Кроме того, материалы, способы и примеры приведены только для иллюстрации и не имеют ограничительного характера. Специалистам в данной области будут понятны многочисленные вариации, изменения и замены без выхода за рамки изобретения.

Определения

[0051] В настоящем документе термины «полинуклеотид» и «нуклеиновая кислота» применяются взаимозаменяемо и относятся к полимерной форме нуклеотидов любой длины - либо рибонуклеотидов, либо дезоксирибонуклеотидов. Таким образом, термины «полинуклеотид» и «нуклеиновая кислота» охватывают одноцепочечную ДНК; двухцепочечную ДНК; многоцепочечную ДНК; одноцепочечную РНК; двухцепочечную РНК; многоцепочечную РНК; геномную ДНК; кДНК; гибриды ДНК-РНК и полимер, содержащий пуриновые и пиримидиновые основания или другие природные, химически или биохимически модифицированные, неприродные или производные нуклеотидные основания.

[0052] Термины «гибридизируемый» или «комплементарный» применяются взаимозаменяемо для обозначения того, что нуклеиновая кислота (например, РНК, ДНК) содержит последовательность нуклеотидов, которая позволяет ей нековалентно связываться, т. е. образовывать пары оснований Уотсона - Крика и/или пары оснований G/U, «отжигаться», «гибридизоваться» с другой нуклеиновой кислотой специфичным для последовательности антипараллельным способом (т. е. нуклеиновая кислота специфически связывается с комплементарной нуклеиновой кислотой) в соответствующих условиях температуры и ионной силы раствора *in vitro* и/или *in vivo*. Следует понимать, что для обеспечения специфической гибридизации последовательность полинуклеотида необязательно должна быть на 100% комплементарна целевой нуклеиновой кислоте; он может быть идентичен последовательности на по меньшей мере примерно 70%, по меньшей мере примерно 80%, или по меньшей мере примерно 90%, или по меньшей мере примерно 95%, и по-прежнему гибридизоваться с целевой нуклеиновой кислотой. Более того, полинуклеотид может гибридизоваться на одном или более сегментах, так что

промежуточные или прилегающие сегменты не участвуют в событии гибридизации (например, структура петли или шпильки, «выпячивания», «пузыря» и т. п.).

[0053] Термин «ген» для целей настоящего описания включает область ДНК, кодирующую генный продукт (например, белок, РНК), а также все области ДНК, которые регулируют продукцию генного продукта, независимо от того, являются ли такие последовательности регуляторных элементов смежными с кодирующими и/или транскрибируемыми последовательностями. Соответственно, ген может включать регуляторные последовательности, включая, но не обязательно ограничиваясь ими, промоторные последовательности, терминаторы, последовательности регуляции трансляции, такие как сайты связывания рибосомы и внутренние сайты входа для рибосомы, энхансеры, сайленсеры, изоляторы, пограничные элементы, точки начала репликации, сайты присоединения матрицы и области управления локусом. Кодирующие последовательности кодируют генный продукт при транскрипции или транскрипции и трансляции; кодирующие последовательности описания могут содержать фрагменты и не обязательно содержат полноразмерную открытую рамку считывания. Ген может включать как транскрибируемую цепь, например, цепь, содержащую кодирующую последовательность, так и комплементарную цепь.

[0054] Термин «расположенный ниже» относится к нуклеотидной последовательности, которая расположена в 3'-направлении к исследуемой нуклеотидной последовательности. В определенных вариантах осуществления расположенные ниже нуклеотидные последовательности относятся к последовательностям, которые следуют за начальной точкой транскрипции. Например, кодон инициации трансляции гена расположен ниже стартового сайта транскрипции.

[0055] Термин «расположенный выше» относится к нуклеотидной последовательности, которая расположена в 5'-направлении к референсной нуклеотидной последовательности. В определенных вариантах осуществления расположенными выше нуклеотидными последовательностями называются последовательности, которые расположены с 5'-стороны от кодирующей области или начальной точки транскрипции. Например, большинство промоторов расположены выше от начальной точки транскрипции.

[0056] Термин «смежный» по отношению к полинуклеотидным или аминокислотным последовательностям относится к последовательностям, которые находятся рядом друг с другом или смежны друг с другом в полинуклеотиде или полипептиде. Специалисту в данной области будет понятно, что две последовательности могут считаться смежными друг с другом и все же включать в себя ограниченное количество промежуточной последовательности, например, 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 нуклеотидов или аминокислот.

[0057] Термин «регуляторный элемент» применяется в настоящем документе взаимозаменяемо с термином «регуляторная последовательность» и должен включать промоторы, энхансеры и другие элементы регуляции экспрессии (например, сигналы

терминации транскрипции, такие как сигналы полиаденилирования и последовательности поли-U). К примерам регуляторных элементов относятся промотор транскрипции, такой как, без ограничений, CMV, CMV+интрон A, SV40, RSV, HIV-Ltr, фактор элонгации 1 альфа (EF1 α), MMLV-ltr, внутренний сайт входа для рибосомы (IRES) или пептид P2A для обеспечения трансляции множества генов с одного транскрипта, металлотионеин, элемент-энхансер транскрипции, сигнал терминации транскрипции, последовательности полиаденилирования, последовательности для оптимизации инициации трансляции и последовательности терминации трансляции. Следует понимать, что выбор соответствующего регуляторного элемента будет зависеть от кодируемого компонента, который нужно экспрессировать (например, белка или РНК), или от того, содержит ли нуклеиновая кислота множество компонентов, которые требуют разных полимераз, или они не предназначены для экспрессии в виде слитного белка.

[0058] Термин «промотор» относится к последовательности ДНК, которая содержит сайт связывания РНК-полимеразы, сайт начала транскрипции, ТАТА-бокс и/или элемент распознавания В и помогает или способствует транскрипции и экспрессии связанной транскрибируемой полинуклеотидной последовательности и/или гена (или трансгена). Промотор может быть получен синтетически или может быть получен из известной или встречающейся в природе промоторной последовательности или другой промоторной последовательности. Промотор может быть проксимальным или дистальным по отношению к транскрибируемому гену. Промотор также может включать химерный промотор, содержащий комбинацию из двух или более гетерологичных последовательностей для придания определенных свойств. Промотор настоящего описания может включать варианты промоторных последовательностей, которые аналогичны по составу, но не идентичны другой (-им) промоторной (-ым) последовательности (-ям), известной или представленной в настоящем документе. Промотор может быть классифицирован по критериям, связанным с характером экспрессии связанной кодирующей или транскрибируемой последовательности или гена, функционально связанного с промотором, например, конститутивный, связанный с развитием, тканеспецифический, индуцибельный и т. д.

[0059] Термин «энхансер» относится к последовательностям ДНК регуляторных элементов, которые при связывании со специфическими белками, называемыми транскрипционными факторами, регулируют экспрессию ассоциированного гена. Энхансеры могут быть расположены в интроне гена или на 5'- или 3'-стороне от кодирующей последовательности гена. Энхансеры могут быть проксимальными к гену (т. е. располагаться в пределах нескольких десятков или сотен пар нуклеотидов (п. о.) от промотора) или могут быть расположены дистально по отношению к гену (т. е. на расстоянии тысячи п. о., сотни тысяч п. о. или даже миллионов п. о. от промотора). Один ген может регулироваться более чем одним энхансером, и все они предусмотрены в рамках объема настоящего описания.

[0060] В настоящем документе термин «посттранскрипционный регуляторный

элемент (PRE)», такой как PRE гепатита, относится к последовательности ДНК, которая при транскрибировании создает третичную структуру, способную демонстрировать посттранскрипционную активность для усиления или стимулирования экспрессии ассоциированного функционально связанного с ней гена.

[0061] Термин «сайт связывания GATA» относится к сайту связывания на ДНК с семейством транскрипционных факторов GATA. Транскрипционные факторы GATA, как правило, распознают сайт-мишень, соответствующий консенсусной последовательности WGATAR (где W=A или T и R=A или G).

[0062] В настоящем документе термин «рекомбинантный» означает, что конкретная нуклеиновая кислота (ДНК или РНК) представляет собой продукт различных комбинаций стадий клонирования, рестрикции и/или лигирования, что приводит к получению конструктора, имеющего структурную кодирующую или не кодирующую последовательность, отличную от эндогенных нуклеиновых кислот, встречающихся в природных системах. По существу, последовательности ДНК, кодирующие структурную кодирующую последовательность, могут быть собраны из фрагментов кДНК и коротких олигонуклеотидных линкеров или из ряда синтетических олигонуклеотидов для получения синтетической нуклеиновой кислоты, которая способна экспрессироваться из рекомбинантной единицы транскрипции, содержащейся в клетке или в бесклеточной системе транскрипции и трансляции. Такие последовательности могут быть представлены в форме открытой рамки считывания, без прерываний внутренними нетранслируемыми последовательностями или интронами, которые обычно присутствуют в эукариотических генах. Геномную ДНК, содержащую соответствующие последовательности, также можно применять при создании рекомбинантного гена или единицы транскрипции. Последовательности нетранслируемой ДНК могут присутствовать на 5'- или 3'-стороне от открытой рамки считывания, где такие последовательности не препятствуют манипуляциям или экспрессии кодирующих областей, и с их действием может быть фактически связано модулирование продукции желаемого продукта с помощью различных механизмов (см. «энхансеры» и «промоторы» выше).

[0063] Термин «рекомбинантный полинуклеотид» или «рекомбинантная нуклеиновая кислота» относится к таким, которые не встречаются в природе, например, получены посредством искусственного комбинирования двух отличных по иным параметрам сегментов последовательности посредством вмешательства человека. Такую искусственную комбинацию часто получают либо средствами химического синтеза, либо посредством искусственного манипулирования отдельными сегментами нуклеиновых кислот, например, методами геной инженерии. Эта операция может выполняться для замены кодона на избыточный кодон, кодирующий ту же или консервативную аминокислоту, при этом, как правило, вводят или удаляют сайт распознавания последовательности. Альтернативно эту операцию выполняют для объединения сегментов нуклеиновых кислот с желаемыми функциями для получения желаемой комбинации функций. Такую искусственную комбинацию часто получают либо средствами

химического синтеза, либо посредством искусственного манипулирования отдельными сегментами нуклеиновых кислот, например, методами генной инженерии.

[0064] Аналогичным образом, термин «рекомбинантный полипептид» или «рекомбинантный белок» относится к полипептиду или белку, который не встречается в природе, например, получен посредством искусственного комбинирования двух отличных по иным параметрам отдельных сегментов аминокислотной последовательности посредством вмешательства человека. Таким образом, например, белок, который содержит гетерологичную аминокислотную последовательность, является рекомбинантным.

[0065] В настоящем документе термин «приведение в контакт» означает формирование физического соединения между двумя или более объектами. Например, приведение последовательности целевой нуклеиновой кислоты в контакт с гидовой нуклеиновой кислотой означает, что обеспечена возможность общей физической связи последовательности целевой нуклеиновой кислоты и гидовой нуклеиновой кислоты; например, они могут гибридизоваться, если последовательности обладают сходством последовательностей.

[0066] Термины «константа диссоциации» или « K_d » применяются взаимозаменяемо и означают аффинность между лигандом «L» и белком «P»; т. е. то, насколько прочно лиганд связывается с конкретным белком. Ее можно рассчитать с применением формулы $K_d = [L][P]/[LP]$, где [P], [L] и [LP] представляют собой молярные концентрации белка, лиганда и комплекса соответственно.

[0067] В описании предложены композиции и способы, используемые для редактирования последовательности целевой нуклеиновой кислоты. В настоящем документе термин «редактирование» применяется взаимозаменяемо с термином «модификация» и включает в себя, без ограничений, расщепление, никирование, делецию, нокин, нокаут и т. п.

[0068] Термин «нокаут» относится к устранению гена или экспрессии гена. Например, ген может быть нокаутирован посредством либо делеции, либо добавления нуклеотидной последовательности, что приводит к нарушению рамки считывания. В качестве другого примера, ген может быть нокаутирован посредством замены части гена на нерелевантную последовательность. В настоящем документе термин «нокаут» относится к снижению экспрессии гена или его генного (-ых) продукта (-ов). В результате нокаута гена активность или функция белка могут быть ослаблены или уровни белка могут быть снижены или устранены.

[0069] В настоящем документе термин «направляемая гомологией репарация» (HDR) относится к форме репарации ДНК, которая происходит во время восстановления двухцепочечных разрывов в клетках. Этот процесс требует гомологии нуклеотидной последовательности, и в нем применяется донорная матрица для репарации или нокаута целевой ДНК, и он приводит к передаче генетической информации от донора (например, от донорной матрицы) к мишени. Направляемая гомологией репарация может приводить к изменению последовательности целевой нуклеиновой кислоты посредством вставки,

делеции или мутации, если донорная матрица отличается от целевой последовательности ДНК и если часть или вся последовательность донорной матрицы включается в целевую ДНК в правильном геномном локусе.

[0070] В настоящем документе термин «негомологичное соединение концов» (NHEJ) относится к репарации двухцепочечных разрывов в ДНК посредством прямого лигирования концов разрыва друг с другом без необходимости в гомологичной матрице (в отличие от направляемой гомологией репарации, которая требует гомологичной последовательности, направляющей репарацию). NHEJ часто приводит к инделам; потере (делеции) или вставке нуклеотидной последовательности вблизи сайта двухцепочечного разрыва.

[0071] В настоящем документе термин «соединение концов, обусловленное микрогомологией» (MMEJ), относится к мутагенному механизму репарации двухцепочечного разрыва (DSB), который всегда ассоциируется с делециями, фланкирующими сайты разрывов, без необходимости в гомологичной матрице (в отличие от направляемой гомологией репарации, которая требует того, чтобы гомологичная последовательность направляла репарацию). MMEJ часто приводит к потере (делеции) нуклеотидной последовательности рядом с сайтом двухцепочечного разрыва.

[0072] Полинуклеотид или полипептид (или белок) имеет определенную процентную долю «сходства последовательности» или «идентичности последовательности» с другим полинуклеотидом или полипептидом, и это означает, что при выравнивании процентные доли оснований или аминокислот двух сравниваемых последовательностей одинаковы, и они находятся в том же относительном положении. Сходство последовательности (иногда называемое процентным сходством, процентной идентичностью или гомологией) может быть определено несколькими разными способами. Для определения сходства последовательностей последовательности можно выравнивать с применением способов и компьютерных программ, известных в данной области, включая BLAST, доступ к которой во всемирной сети осуществляется по адресу ncbi.nlm.nih.gov/BLAST. Процент комплементарности между конкретными участками нуклеотидных последовательностей в пределах нуклеиновых кислот можно определять с применением любого удобного способа. Примеры способов включают программы BLAST (basic local alignment search tools) и программы PowerBLAST (Altschul et al., J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656) или с применением программ Gap (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, г. Мэдисон, штат Висконсин, США), например, с применением заданных по умолчанию параметров, в которых применяется алгоритм Смита - Ватермана (Adv. Appl. Math., 1981, 2, 482-489).

[0073] Термины «полипептид» и «белок» применяются в настоящем документе взаимозаменяемо и относятся к полимерной форме аминокислот любой длины, которая может включать кодированные и некодированные аминокислоты, химически или биохимически модифицированные или дериватизированные аминокислоты, а также

полипептиды, имеющие модифицированные пептидные основные цепи. Термин включает слитные белки, включая, без ограничений, слитные белки с гетерологичной аминокислотной последовательностью.

[0074] «Вектор» или «экспрессионный вектор» представляет собой репликон, такой как плаزمид, фаг, вирус, вирусоподобную частицу или космиду, к которой можно присоединить другой сегмент ДНК, т. е. «вставку», так чтобы вызвать репликацию или экспрессию присоединенного сегмента в клетке.

[0075] Термин «встречающийся в природе», или «немодифицированный», или «дикого типа», применяемый в настоящем документе применительно к нуклеиновой кислоте, полипептиду, клетке или организму, относится к нуклеиновой кислоте, полипептиду, клетке или организму, встречающемуся в природе.

[0076] В настоящем документе термин «мутация» относится к вставке, делеции, замене, дублированию или инверсии одной или более аминокислот или нуклеотидов по сравнению с референсной аминокислотной последовательностью или аминокислотной последовательностью дикого типа или с референсной нуклеотидной последовательностью или нуклеотидной последовательностью дикого типа.

[0077] В настоящем документе термин «выделенный» предназначен для описания полинуклеотида, полипептида или клетки, находящейся в среде, отличной от той, в которой полинуклеотид, полипептид или клетка встречается в природных условиях. Выделенная генетически модифицированная клетка-хозяин может присутствовать в смешанной популяции генетически модифицированных клеток-хозяев.

[0078] Термин «клетка-хозяин», используемый в настоящем документе, обозначает эукариотическую клетку, прокариотическую клетку или клетку из многоклеточного организма (например, линию клеток), культивируемую в качестве одиночных объектов, причем эукариотические или прокариотические клетки применяются в качестве реципиентов нуклеиновой кислоты (например, экспрессионного вектора), и включают потомство исходной клетки, которая была генетически модифицирована нуклеиновой кислотой. Следует понимать, что потомство одной клетки не обязательно может быть полностью идентичным по морфологии, или геному, или общему комплементу ДНК исходному родителю вследствие естественной, случайной или преднамеренной мутации. «Рекомбинантная клетка-хозяин» (также называемая «генетически модифицированной клеткой-хозяином») представляет собой клетку-хозяина, в которую была введена гетерологичная нуклеиновая кислота, например, экспрессионный вектор.

[0079] Термин «консервативная аминокислотная замена» относится к взаимозаменяемости в белках аминокислотных остатков, имеющих сходные боковые цепи. Например, группа аминокислот, имеющих алифатические боковые цепи, состоит из глицина, аланина, валина, лейцина и изолейцина; группа аминокислот, имеющих алифатически-гидроксильные боковые цепи, состоит из серина и треонина; группа аминокислот, имеющих амидосодержащие боковые цепи, состоит из аспарагина и глутамина; группа аминокислот, имеющих ароматические боковые цепи, состоит из

фенилаланина, тирозина и триптофана; группа аминокислот, имеющих основные боковые цепи, состоит из лизина, аргинина и гистидина; а группа аминокислот, имеющих серосодержащие боковые цепи, состоит из цистеина и метионина. Примерами групп консервативных аминокислотных замен являются: валин-лейцин-изолейцин, фенилаланин-тирозин, лизин-аргинин, аланин-валин и аспарагин-глутамин.

[0080] В настоящем документе термины «лечение» или «терапия» применяются взаимозаменяемо в настоящем документе и относятся к подходу для получения благоприятных или желаемых результатов, включая, без ограничений, терапевтический благоприятный эффект и/или профилактический благоприятный эффект. Под терапевтическим благоприятным эффектом подразумевается устранение или облегчение базового расстройства или заболевания, подлежащего лечению. Терапевтический благоприятный эффект также может быть достигнут с помощью устранения или облегчения одного или более симптомов или улучшения одного или более клинических параметров, связанных с базовым заболеванием, так что у субъекта наблюдается улучшение, несмотря на то что субъект все еще может страдать базовым заболеванием.

[0081] Термины «терапевтически эффективное количество» и «терапевтически эффективная доза», используемые в настоящем документе, относятся к количеству лекарственного средства или биологического препарата, отдельно или в составе композиции, которое способно оказывать какой-либо обнаруживаемый благоприятный эффект в отношении любого симптома, аспекта, измеренного параметра или характеристики болезненного состояния или заболевания при введении одной дозы или повторяющихся доз субъекту, такому как человек или экспериментальное животное. Такой эффект не должен быть абсолютным, чтобы быть признанным благоприятным.

[0082] В настоящем документе термин «введение» означает способ предоставления субъекту дозировки композиции описания.

[0083] В настоящем документе термин «субъект» представляет собой млекопитающее. К млекопитающим относятся, без ограничений, одомашненные животные, приматы, приматы, не являющиеся человеком, люди, собаки, свиньи, кролики, мыши, крысы и другие грызуны.

[0084] Все публикации, патенты и заявки на патенты, упомянутые в настоящем описании, включены в настоящий документ путем ссылки в той степени, как если бы каждая отдельная публикация, патент или заявка на патент была конкретно и индивидуально обозначена как включенная путем ссылки.

I. Общие способы

[0085] При практическом применении настоящего изобретения, если не указано иное, используются традиционные методики иммунологии, биохимии, химии, молекулярной биологии, микробиологии, клеточной биологии, геномики и рекомбинантной ДНК, которые можно найти в таких стандартных справочных пособиях, как: *Molecular Cloning: A Laboratory Manual*, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001); *Short Protocols in Molecular Biology*, 4th Ed. (Ausubel et al. eds., John Wiley & Sons

1999); Protein Methods (Bollag et al., John Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner et al. eds., Academic Press 1999); Viral Vectors (Kaplift & Loewy eds., Academic Press 1995); Immunology Methods Manual (I. Lefkovits ed., Academic Press 1997) и Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998), содержимое которых включено в настоящий документ путем ссылки.

[0086] Если приведен диапазон значений, следует понимать, что в него входят конечные точки и каждое промежуточное значение до десятой доли нижнего предела, если контекстом явно не предусмотрено иное, между верхним и нижним пределами этого диапазона, а также любое другое указанное или промежуточное значение в этом указанном диапазоне. Верхний и нижний пределы этих меньших диапазонов могут быть независимо включены в меньшие диапазоны и также включены при наличии любого специально исключенного предела в указанном диапазоне. Если указанный диапазон включает один или оба предела, то диапазоны, исключающие любой из двух или оба таких предела сразу, также включены.

[0087] Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют общепринятое значение, понятное любому специалисту в области, к которой относится настоящее изобретение. Все публикации, упомянутые в настоящем документе, включены в настоящий документ путем ссылки для раскрытия и описания способов и/или материалов, в связи с которыми эти публикации были процитированы.

[0088] Необходимо отметить, что в настоящем документе и в прилагаемой формуле изобретения формы единственного числа также включают объекты во множественном числе, если из контекста явно не следует иное.

[0089] Следует понимать, что определенные признаки описания, которые для ясности приведены в контексте разных вариантов осуществления, также могут быть представлены в комбинации в одном варианте осуществления. В других случаях различные признаки описания, которые для краткости описаны в контексте одного варианта осуществления, также могут быть представлены отдельно или в любой приемлемой подкомбинации. Подразумевается, что все комбинации вариантов осуществления, относящихся к описанию, конкретно включены в настоящее описание и раскрыты в настоящем документе так, как если бы каждая комбинация была четко раскрыта в отдельности. Кроме того, все подкомбинации различных вариантов осуществления и их элементы также конкретно включены в настоящее описание и описаны в настоящем документе так, как если бы каждая такая подкомбинация была четко раскрыта в отдельности.

II. Системы генетического редактирования генов BCL11A

[0090] В первом аспекте в настоящем описании предложены системы, содержащие белок-нуклеазу CRISPR класса 2, типа V и одну или более гидовых нуклеиновых кислот (гРНК) для применения при модификации или редактировании гена BCL11A с целью снижения или устранения экспрессии продукта гена BCL11A. Примеры систем с белком-

нуклеазой CRISPR класса 2, типа V и гидовой нуклеиновой кислотой включают систему CasX:гРНК. Системы CasX:гРНК специально разработаны для модификации гена BCL11A в эукариотических клетках. В некоторых случаях системы CasX:гРНК предназначены для нокдауна или нокаута гена BCL11A. По существу, с применением программируемых композиций и способов, предложенных в настоящем документе, можно нацеливаться на любой участок гена BCL11A. В некоторых вариантах осуществления ген BCL11A, подлежащий модификации, представляет собой последовательность дикого типа, а участок, подлежащий модификации, выбран из группы, состоящей из интрона BCL11A, экзона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A и межгенной области, или модификация представляет собой делецию или мутацию в одном или более экзонах.

[0091] В настоящем документе термин «система», например системы, содержащие белок-нуклеазу CRISPR и одну или более гРНК описания, а также системы, содержащие нуклеиновые кислоты, кодирующие белки-нуклеазы CRISPR и гРНК, и системы, содержащие векторы, которые содержат нуклеиновые кислоты, или системы, содержащие белок-нуклеазу CRISPR и одну или более гРНК описания, применяются взаимозаменяемо с термином «композиция».

[0092] Человеческий ген *BCL11A* (HGNC: 13221) кодирует белок (Q9H165), имеющий последовательность

MSRRKQGKQPQLSKREFSPEPLEAILTDDEPDHGPLGAPEGDHDLLTCGQCQMN
FPLGDILIFIEHKRKQCNGSLCLEKAVDKPPSPSPIEMKKASNPVEVGIVTPEDDDCLSTS
SRGICPKQEHADKLLHWRGLSSPRSAHGALIPTPGMSAEYAPQGICKDEPSSYTCTTCK
QPFTSAWFLQHAQNTHGLRIYLESEHGSPLTPRVGIPSGLGAECPSQPPLHGIHIADNNP
FNLLRIPGSVSREASGLAEGFRPPTPPLFSPPPRHHLDPHRIERLGAEEMALATHHPSAFD
RVLRLNPMAMEPPAMDFSRRLRELAGENTSSPPLSPGRPSPMQRLLPFQPGSKPPFLATP
PLPPLQSAPPPSQPPVKSKEFCGKTFKFQSNLVVHRRSHTGEKPYKCNLCDHACTQAS
KLKRHMKTTHMHKSSPMTVKSDDGLSTASSPEPGTSDLVGSASSALKSVVAKFKSENDP
NLIPENGDEEEEEEDDEEEEEEEEEEEEEELTESERVDYGFGLSLEAARHHENSSRGAVVGV
GDESRLPDVMQGMVLSSMQHFSEAFHQVLGEKHKRGHLAEAEGRDTCDEDSVAGE
SDRIDDGTVNGRGCSPGESASGGLSKKLLLGSPSSLSPFSKRIKLEKEFDLPPAAMPNTEN
VYSQWLAGYAASRQLKDPFLSFGDSRQSPFASSEHSSENGSLRFSTPPGELDGGISGRS
GTGSGGSTPHISGPGPRPSSKEGRRSDTCEYCGKVFKNCSNLTVHRRSHTGERPYKCEL
CNYACAQSSKLTRHMKTHGQVGKDVYKCEICKMPFSVYSTLEKHKMKWHSDRVLNN
DIKTE (SEQ ID NO: 100). Ген BCL11A определяется как последовательность, которая охватывает chr2 60450520-60554467 (GRCh38/hg38 Ensembl 100) генома человека на хромосоме 2.

[0093] В некоторых вариантах осуществления в описании предложены системы, специально предназначенные для модификации гена BCL11A в эукариотических клетках; *in vitro*, *ex vivo* или *in vivo* у субъекта. По существу, можно нацеливаться на любой участок целевой нуклеиновой кислоты BCL11A с применением программируемых композиций и

способов, предложенных в настоящем документе. В некоторых вариантах осуществления нуклеаза CRISPR представляет собой нуклеазу класса 2, типа V. Несмотря на то что элементы систем CRISPR-Cas класса 2, типа V имеют различия, они также имеют некоторые общие характеристики, которые отличают их от систем Cas9. Во-первых, нуклеазы типа V обладают одиночным РНК-направляемым эффектором, содержащим домен RuvC, но не содержащим домен HNH, и они распознают богатый Т-нуклеотидами мотив PAM на 5'-стороне от целевой области на нецелевой цепи, и это отличается от систем Cas9, которые используют богатый G-нуклеотидами PAM на 3'-стороне от целевых последовательностей. Нуклеазы типа V создают ступенчатые двухцепочечные разрывы дистально от PAM-последовательности, в отличие от Cas9, которые создают тупой конец в проксимальном сайте вблизи PAM. Кроме того, нуклеазы типа V разрушают транс-активные оцДНК при активации целевой дцДНК или связывании с цис-активной оцДНК. В некоторых вариантах осуществления в описании предложена нуклеаза класса 2, типа V, выбранная из группы, состоящей из Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, CasZ и CasX. В некоторых вариантах осуществления в описании предложены системы, содержащие один или более белков CasX и одну или более гидовых нуклеиновых кислот (гРНК) в качестве системы CasX:гРНК. В других вариантах осуществления системы CasX:гРНК описания содержат один или более белков CasX, одну или более гидовых нуклеиновых кислот (гРНК) и одну или более нуклеиновую кислоту донорной матрицы, которая содержит нуклеиновую кислоту, кодирующую участок гена BCL11A, причем нуклеиновая кислота донорной матрицы содержит делецию, вставку или мутацию одного или более нуклеотидов по сравнению с геномной нуклеотидной последовательностью, кодирующей белок BCL11A. Каждый из этих компонентов и их применение при редактировании гена BCL11A описаны в настоящем документе ниже.

[0094] В некоторых вариантах осуществления в описании для редактирования генов предложены пары из CasX и гРНК любого из вариантов осуществления, описанных в настоящем документе, которые способны связываться друг с другом перед их применением для редактирования генов, и, таким образом, они являются «предварительно комплексованными» в рибонуклеопротеиновый комплекс (RNP). Применение предварительных комплексов RNP дает преимущества при доставке компонентов системы в клетку или последовательность целевой нуклеиновой кислоты для редактирования последовательности целевой нуклеиновой кислоты.

[0095] В некоторых вариантах осуществления функциональный RNP может быть доставлен *ex vivo* в клетку посредством электрофореза или химическими средствами. В других вариантах осуществления функциональный RNP может быть доставлен в своей функциональной форме либо *ex vivo*, либо *in vivo* с помощью вектора. В некоторых вариантах осуществления RNP может быть доставлен *in vivo* субъекту с применением частицы для доставки CasX (XDP). гРНК может придавать комплексу специфичность к мишени вследствие наличия нацеливающей последовательности (или «спейсера»), имеющей нуклеотидную последовательность, которая комплементарна

последовательности целевой нуклеиновой кислоты, тогда как вариант белка CasX из предварительного комплекса CasX:гРНК обеспечивает сайт-специфическую активность, такую как расщепление или никирование целевой последовательности, которая направлена на сайт-мишень (например, стабилизируется на сайте-мишени) в пределах последовательности целевой нуклеиновой кислоты за счет ее связи с гРНК.

[0096] Системы используются при лечении субъекта, имеющего связанное с гемоглобинопатией заболевание, такое как серповидноклеточная анемия или β -талассемия. Каждый из компонентов систем CasX:гРНК, их функции и их применение при редактировании целевых нуклеиновых кислот в клетках более подробно описаны ниже.

III. Гидовые нуклеиновые кислоты систем редактирования генов

[0097] В другом аспекте описание относится к специально разработанным гидовым рибонуклеиновым кислотам (гРНК), содержащим нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты гена BCL11A (и, следовательно, способную гибридизоваться с ней), которые могут находиться в комплексе с нуклеазой CRISPR для геномного редактирования целевой нуклеиновой кислоты BCL11A в клетке. Предполагается, что в некоторых вариантах осуществления в системах для модификации целевой нуклеиновой кислоты обеспечивается доставка множества гРНК. Например, пару гРНК с последовательностями, нацеливающимися на разные или перекрывающиеся области последовательности целевой нуклеиновой кислоты, можно применять, когда каждая из них находится в комплексе с нуклеазой CRISPR, чтобы связываться и выполнить расщепление в двух разных или перекрывающихся сайтах внутри гена, которые затем подвергаются редактированию методами негомологичного соединения концов (NHEJ), направляемой гомологией репарации (HDR), независимой от гомологии нацеленной интеграции (HITI), соединения концов, обусловленного микрогомологией (MMEJ), отжига одиночных цепей (SSA) или эксцизионной репарации оснований (BER).

[0098] В некоторых вариантах осуществления в описании предложены гРНК, используемые в системах CasX:гРНК, которые используются в геномном редактировании гена BCL11A в эукариотической клетке. В конкретном варианте осуществления гРНК систем способна образовывать комплекс с нуклеазой CasX. В настоящем описании предложены специально разработанные гРНК, в которых нацеливающая последовательность (или спейсер, описанный более подробно выше) гРНК комплементарна последовательностям целевых нуклеиновых кислот (и, следовательно, способна гибридизоваться с ними) при применении в качестве компонента систем для редактирования генов CasX:гРНК. Идентификаторы SEQ ID NO типичных, но не имеющих ограничительного характера примеров последовательностей, нацеливающихся на целевую нуклеиновую кислоту BCL11A, которые можно использовать в гРНК вариантов осуществления, представлены в таблице 1, которая более подробно описана ниже.

a. Референсная гРНК и варианты гРНК

[0099] В настоящем документе термин «референсная гРНК» относится к гидовой нуклеиновой кислоте для CRISPR, содержащей последовательность дикого типа

встречающейся в природе гРНК. В некоторых вариантах осуществления референсная гРНК описания может быть подвергнута одному или более способам мутагенеза, таким как способы мутагенеза, описанные в настоящем документе, которые могут включать глубокую мутационную эволюцию (DME), глубокое мутационное сканирование (DMS), подверженную ошибкам ПЦР, каскадный мутагенез, случайный мутагенез, ПЦР со ступенчатой достройкой, перетасовку генов или перестановку доменов, с целью создания одного или более вариантов гРНК с улучшенными или измененными свойствами относительно референсной гРНК. Варианты гРНК также включают варианты, содержащие одну или более экзогенных последовательностей, например слитых с 5'- или 3'-концом или вставленных внутрь. Активность референсных гРНК можно применять в качестве эталона, с которым сравнивают активность вариантов гРНК, измеряя таким образом улучшения функции или других характеристик вариантов гРНК. В других вариантах осуществления референсную гРНК можно подвергать одной или более произвольным специфически нацеленным мутациям для получения варианта гРНК, например, с помощью логически сконструированного варианта.

[0100] гРНК описания содержат два сегмента: нацеливающую последовательность и связывающийся с белком сегмент. Нацеливающий сегмент гРНК включает нуклеотидную последовательность (называемую взаимозаменяемо гидовой последовательностью, спейсером, нацеливателем или нацеливающей последовательностью), которая комплементарна (и, следовательно, гибридизуется с) конкретной последовательности (целевому сайту) в последовательности целевой нуклеиновой кислоты (например, целевой оцРНК, целевой оцДНК, цепи двухцепочечной целевой ДНК и т. д.), как более подробно описано ниже. Нацеливающая последовательность гРНК способна связываться с последовательностью целевой нуклеиновой кислоты, в том числе с кодирующей последовательностью, компонентом кодирующей последовательности, с некодирующей последовательностью и регуляторными элементами. Белок-связывающий сегмент (или «активатор», или «белок-связывающая последовательность») взаимодействует с (например, связывается) белком CasX в комплекс, образуя RNP (более подробно описан ниже). Белок-связывающий сегмент альтернативно называется в настоящем документе «каркасом», и он состоит из нескольких областей, более подробно описанных ниже.

[0101] В случае двойной гидовой РНК (дгРНК) каждый из нацеливающего и активаторного участка имеет дуплекс-образующий сегмент, где дуплекс-образующий сегмент нацеливателя и дуплекс-образующий сегмент активатора обладают комплементарностью друг к другу и гибридизуются друг с другом с образованием двухцепочечного дуплекса (дцРНК-дуплекс для гРНК). Если гРНК представляет собой гРНК, то термин «нацеливатель» или «нацеливающая РНК» применяется в настоящем документе для обозначения крРНК-подобной молекулы (крРНК: «CRISPR-РНК») двойной гидовой РНК CasX (и, следовательно, одинарной гидовой РНК CasX, если «активатор» и «нацеливатель» связаны друг с другом, например, посредством промежуточных нуклеотидов). крРНК имеет 5'-область, которая отжигается с тракрРНК, за которой

следуют нуклеотиды нацеливающей последовательности. Таким образом, например, гидовая РНК (дгРНК или огРНК) содержит гидовую последовательность и дуплекс-образующий сегмент крРНК, который также может называться крРНК-повтором. Соответствующая тракрРНК-подобная молекула (активирующий участок) содержит дуплекс-образующий участок нуклеотидов, который формирует другую половину дцРНК-дуплекса белок-связывающего сегмента гидовой РНК. Таким образом, нацеливатель и активатор как соответствующая пара гибридизуются с образованием двойной гидовой НК, называемой в настоящем документе «двойной гидовой НК», «двухмолекулярной гРНК», «дгРНК», «бимолекулярной гидовой НК» или «двухмолекулярной гидовой НК». Сайт-специфическое связывание и/или расщепление последовательности целевой нуклеиновой кислоты (например, геномной ДНК) белком CasX может происходить в одном или более местоположениях (например, в последовательности целевой нуклеиновой кислоты), определяемом комплементарностью связывания оснований между нацеливающей последовательностью гРНК и последовательностью целевой нуклеиновой кислоты. Таким образом, например, гРНК описания имеет комплементарность последовательности и, следовательно, может гибридизоваться с целевой нуклеиновой кислотой, которая смежна с последовательностью, комплементарной РАМ-мотиву ТС или РАМ-последовательности, такой как АТС, СТС, ГТС или ТТС. Так как нацеливающая последовательность гидовой последовательности гибридизуется с последовательностью целевой нуклеиновой кислоты, нацеливатель может быть модифицирован пользователем для гибридизации с конкретной последовательностью целевой нуклеиновой кислоты, при условии, что будет учитываться местоположение РАМ-последовательности. Таким образом, в некоторых случаях последовательность нацеливателя может представлять собой не встречающуюся в природе последовательность. В других случаях последовательность нацеливателя может представлять собой встречающуюся в природе последовательность, полученную из гена, подлежащего редактированию. В других вариантах осуществления активатор и нацеливатель в гРНК ковалентно связаны друг с другом (а не гибридизуются друг с другом) и представляют собой одну молекулу, которая в настоящем документе называется «одномолекулярной гРНК», «мономолекулярной гидовой НК», «одиночной гидовой НК», «одиночной гидовой РНК», «одномолекулярной гидовой РНК», «мономолекулярной гидовой РНК» или «огРНК». В некоторых вариантах осуществления огРНК включает «активатор» или «нацеливатель» и, таким образом, может представлять собой «активатор-РНК» и «нацеливатель-РНК» соответственно. В некоторых вариантах осуществления гРНК представляет собой молекулу рибонуклеиновой кислоты («гРНК»), а в других вариантах осуществления гРНК представляет собой химеру и содержит как ДНК, так и РНК. В настоящем документе термин «гРНК» охватывает встречающиеся в природе молекулы, а также варианты последовательностей.

[0102] В совокупности собранные гРНК описания содержат четыре различных области или домена: триплекс РНК, каркасный стебель, удлиненный стебель и нацеливающую последовательность, которая в вариантах осуществления описания

специфична для целевой нуклеиновой кислоты и расположена на 3'-конце гРНК. Триплекс РНК, каркасный стебель и удлинённый стебель в совокупности называются «каркасом» гРНК.

b. Триплекс РНК

[0103] В некоторых вариантах осуществления гидовых НК, предложенных в настоящем документе (включая референсные огРНК), присутствует триплекс РНК, и триплекс РНК содержит последовательность UUU--nX(~ 4-15)--UUU (SEQ ID NO: 226) стеблевой петли, которая заканчивается нуклеотидами AAAG после 2 промежуточных стеблевых петель (каркасной стеблевой петли и удлинительной стеблевой петли), образуя псевдоузел, который также может переходить из триплексного в дуплексный псевдоузел. Последовательность триплекса UU-UUU-AAA образуется как связующее звено между нацеливающей последовательностью, каркасным стеблем и удлинённым стеблем. В примерах огРНК для CasX сначала кодируется область UUU-петля-UUU, а затем каркасная стеблевая петля, после этого - удлинённая стеблевая петля, которая связана четверной петлей, а затем последовательность AAAG закрывает триплекс, после чего переходит в нацеливающую последовательность.

c. Каркасная стеблевая петля

[0104] В некоторых вариантах осуществления огРНК описания за триплексной областью следует каркасная стеблевая петля. Каркасная стеблевая петля представляет собой область гРНК, которая связывается с белком CasX (таким как референсный CasX или вариант белка). В некоторых вариантах осуществления каркасная стеблевая петля представляет собой достаточно короткую и стабильную стеблевую петлю. В некоторых случаях каркасная стеблевая петля не позволяет вносить много изменений, и для нее необходима некоторая форма РНК-пузыря. В некоторых вариантах осуществления каркасный стебель необходим для функционирования огРНК CasX. Хотя он, возможно, аналогичен связующему стеблю Cas9 в роли критической стеблевой петли, каркасный стебель огРНК CasX в некоторых вариантах осуществления имеет обязательное утолщение (РНК-пузырь), которое отличается от многих других стеблевых петель, обнаруживаемых в системах CRISPR/Cas. В некоторых вариантах осуществления присутствие этого утолщения консервативно для огРНК, которые взаимодействуют с разными белками CasX. Пример последовательности каркасной стеблевой петли гРНК содержит последовательность CCAGCGACUAUGUCGUAUGG (SEQ ID NO: 20).

d. Удлинённая стеблевая петля

[0105] В некоторых вариантах осуществления огРНК CasX описания за каркасной стеблевой петли следует удлинённая стеблевая петля. В некоторых вариантах осуществления удлинённый стебель содержит синтетический слитый конструктор из тракр- и крРНК, который по большей части не связывается с белком CasX. В некоторых вариантах осуществления удлинённая стеблевая петля может иметь высокую гибкость. В некоторых вариантах осуществления одиночная гидовая гРНК образована с четырехпетлевым линкером GAAA или линкером GAGAAA между тракр- и крРНК в удлинённой стеблевой

петле. В некоторых случаях нацеливатель и активатор в огРНК CasX связаны друг с другом промежуточными нуклеотидами, и линкер может иметь длину от 3 до 20 нуклеотидов. В некоторых вариантах осуществления в огРНК CasX описания удлиненный стебель представляет собой большую петлю 32 п. о., которая находится за пределами белка CasX в рибонуклеопротеиновом комплексе. Пример последовательности удлинительной стеблевой петли огРНК содержит последовательность GCGCUUAUUUAUCGGAGAGAAAUCCGAUAAUAAGAAGC (SEQ ID NO: 21). В некоторых вариантах осуществления удлиненная стеблевая петля содержит спейсерную последовательность GAGAAA.

е. Нацеливающая последовательность

[0106] В некоторых вариантах осуществления гРНК описания за удлинительной стеблевой петлей следует область, которая образует часть триплекса, а затем нацеливающая последовательность (или «спейсер») на 3'-конце гРНК. Нацеливающая последовательность нацеливает цельный рибонуклеопротеиновый комплекс CasX на конкретную область последовательности целевой нуклеиновой кислоты гена, подлежащего модификации. Таким образом, например, нацеливающие последовательности гРНК описания имеют комплементарность последовательностям и, следовательно, могут гибридизоваться с участком гена BCL11A в нуклеиновой кислоте в эукариотической клетке (например, эукариотической хромосоме, хромосомной последовательности, эукариотической РНК и т. д.) в качестве компонента RNP, если на расстоянии 1 нуклеотида с 5'-стороны от последовательности нецелевой цепи, комплементарной целевой последовательности, расположен РАМ-мотив ТС или любая из РАМ-последовательностей TTC, ATC, GTC или CTC. Нацеливающую последовательность гРНК можно модифицировать таким образом, чтобы гРНК могла нацеливаться на желаемую последовательность любой желаемой последовательности целевой нуклеиновой кислоты, при условии, что учитывается местоположение РАМ-последовательности. В некоторых вариантах осуществления каркас гРНК находится с 5'-стороны нацеливающей последовательности, а нацеливающая последовательность находится на 3'-конце гРНК. В некоторых вариантах осуществления последовательность РАМ-мотива, распознаваемая нуклеазой RNP, представляет собой ТС. В других вариантах осуществления РАМ-последовательность, распознаваемая нуклеазой RNP, представляет собой NTC.

[0107] В некоторых вариантах осуществления нацеливающая последовательность гРНК специфична для участка гена, кодирующего белок BCL11A. В некоторых вариантах осуществления нацеливающая последовательность гРНК специфична для экзона BCL11A. В некоторых вариантах осуществления нацеливающая последовательность гРНК специфична для интрона BCL11A. В некоторых вариантах осуществления нацеливающая последовательность гРНК специфична для соединения интрон-экзон BCL11A. В некоторых вариантах осуществления нацеливающая последовательность гРНК имеет последовательность, которая гибридизуется с регуляторным элементом BCL11A, кодирующей областью BCL11A, некодирующей областью BCL11A или их комбинациями

(например, пересечением двух областей). В некоторых вариантах осуществления регуляторный элемент содержит последовательность связывания с GATA. В некоторых вариантах осуществления нацеливающая последовательность гРНК комплементарна последовательности, содержащей один или более однонуклеотидных полиморфизмов (SNP) гена *BCL11A* или его комплемента. SNP, которые находятся в кодирующей последовательности *BCL11A* или в пределах некодирующей последовательности *BCL11A*, входят в объем настоящего описания. В других вариантах осуществления нацеливающая последовательность гРНК комплементарна последовательности межгенной области гена *BCL11A* или последовательности, комплементарной межгенной области гена *BCL11A*.

[0108] В некоторых вариантах осуществления нацеливающая последовательность гРНК выполнена специфичной для регуляторного элемента, который регулирует экспрессию продукта гена *BCL11A*. Такие регуляторные элементы включают, без ограничений, промоторные области, энхансерные области, межгенные области, 5'-нетранслируемые области (5'-UTR), 3'-нетранслируемые области (3'-UTR), консервативные элементы и области, содержащие цис-регуляторные элементы. Предполагается, что промоторная область охватывает нуклеотиды, находящиеся в пределах 5 т. п. о. от точки начала кодирующей последовательности или, в случае энхансерных элементов или консервативных элементов гена, находящиеся в пределах тысяч пар оснований (п. о.), сотен тысяч п. о. или даже миллионов п. о. от кодирующей последовательности гена целевой нуклеиновой кислоты. В конкретных вариантах осуществления нацеливающая последовательность гРНК гибридизируется с последовательностью, которая комплементарна регуляторному элементу *BCL11A*. В одном варианте осуществления нацеливающая последовательность гРНК представляет собой UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22), которая гибридизируется с последовательностью сайта связывания эритроид-специфического энхансера *BCL11A* GATA1, или идентична последовательности на по меньшей мере 90% или по меньшей мере 95% (см. ФИГ. 23). В другом варианте осуществления нацеливающая последовательность гРНК представляет собой UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23) или идентична последовательности на по меньшей мере 90% или по меньшей мере 95%. В другом варианте осуществления нацеливающая последовательность гРНК представляет собой UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23) или идентична последовательности на по меньшей мере 90% или по меньшей мере 95%. В других вариантах осуществления нацеливающая последовательность гРНК выбрана из группы, состоящей из CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949), GAGGCCAAACCCUCCUGGA (SEQ ID NO: 2948), AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747) и AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748).

[0109] У субъектов в период постнатального созревания связывание GATA1 усиливает экспрессию *BCL11A*, которая, в свою очередь, подавляет экспрессию гемоглобина F (HbF) в пользу гемоглобина гамма. Однако у субъектов с определенными гемоглобинопатиями было продемонстрировано, что подавление экспрессии *BCL11A*

обеспечивает возобновление экспрессии HbF, что может компенсировать у субъекта дефектный по иным параметрам гемоглобин.

[0110] В некоторых вариантах осуществления нацеливающая последовательность гРНК содержит от 14 до 35 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 14, 15, 16, 18, 18, 19, 20, 21, 22, 23 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 или 35 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 20 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 19 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 18 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 17 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 16 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность состоит из 15 последовательных нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 или 35 последовательных нуклеотидов, и нацеливающая последовательность может содержать от 0 до 5, от 0 до 4, от 0 до 3 или от 0 до 2 несовпадений относительно последовательности целевой нуклеиновой кислоты и сохранять достаточную специфичность связывания таким образом, чтобы RNP, содержащий гРНК, имел нацеливающую последовательность, способную образовывать комплементарную связь по отношению к целевой нуклеиновой кислоте.

[0111] Типичные, но не имеющие ограничительного характера примеры нацеливающихся последовательностей к последовательности целевой нуклеиновой кислоты, предусмотренные для применения в гРНК описания, представлены как SEQ ID NO: 272-2100 и 2286-26789 (см. таблицу 1). В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ АТС, содержащие последовательность, которая на по меньшей мере 50% идентична, на по меньшей мере 55% идентична, на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей мере 90% идентична, на по меньшей мере 95% идентична или на 100% идентична последовательности с SEQ ID NO: 272-2100 или 2286-5625. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ АТС, содержащие последовательность с SEQ ID NO: 272-2100 или 2286-5625. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ СТС, содержащие последовательность, которая на по меньшей мере 50% идентична, на по меньшей мере 55% идентична, на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей

мере 90% идентична, на по меньшей мере 95% идентична или на 100% идентична последовательности с SEQ ID NO: 5626-13616. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ CTC, содержащие последовательность с SEQ ID NO: 5626-13616. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ GTC, содержащие последовательность, которая на по меньшей мере 50% идентична, на по меньшей мере 55% идентична, на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей мере 90% идентична, на по меньшей мере 95% идентична или на 100% идентична последовательности с SEQ ID NO: 13617-17903. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ GTC, содержащие последовательность с SEQ ID NO: 13617-17903. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ TTC, содержащие последовательность, которая на по меньшей мере 50% идентична, на по меньшей мере 55% идентична, на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей мере 90% идентична, на по меньшей мере 95% идентична или на 100% идентична последовательности с SEQ ID NO: 17904-26789. В некоторых вариантах осуществления в описании предложены нацеливающие последовательности для РАМ TTC, содержащие последовательность с SEQ ID NO: 17904-26789. В некоторых вариантах осуществления нацеливающая последовательность, предусмотренная для применения в гРНК описания, гРНК содержит последовательность с SEQ ID NO: 272-2100 или 2286-26789 с одним нуклеотидом, удаленным с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность гРНК содержит последовательность с SEQ ID NO: 272-2100 или 2286-26789 с двумя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность гРНК содержит последовательность с SEQ ID NO: 272-2100 или 2286-26789 с тремя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность гРНК содержит последовательность с SEQ ID NO: 272-2100 или 2286-26789 с четырьмя нуклеотидами, удаленными с 3'-конца последовательности. В других вариантах осуществления нацеливающая последовательность гРНК содержит последовательность с SEQ ID NO: 272-2100 или 2286-26789 с пятью нуклеотидами, удаленными с 3'-конца последовательности. В вышеуказанных вариантах осуществления параграфа тиминового (Т) нуклеотида могут быть замещены одним или более или всеми урациловыми (U) нуклеотидами в любой из нацеливающих последовательностей таким образом, что нацеливающая последовательность гРНК может представлять собой гДНК или гРНК, или химеру РНК и ДНК, в тех случаях, когда кодирующая последовательность для спейсера включена в

экспрессионный вектор. В некоторых вариантах осуществления нацеливающая последовательность с SEQ ID NO: 272-2100 или 2286-26789 имеет замену по меньшей мере 1, 2, 3, 4, 5 или 6 или более тиминовых нуклеотидов на урациловые нуклеотиды.

Таблица 1. SEQ ID NO для нацеливающих последовательностей гРНК для гена BCL11A

Тип РАМ	SEQ ID NO
АТС	272-2100, 2286-5625
СТС	5626-13616
ГТС	13617-17903
ТТС	17904-26789

[0112] В некоторых вариантах осуществления система CasX:гРНК содержит первую гРНК и дополнительно содержит вторую (и необязательно третью, четвертую, пятую или более) гРНК, причем вторая гРНК или дополнительная гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку последовательности целевой нуклеиновой кислоты по сравнению с нацеливающей последовательностью первой гРНК, так что нацеливание происходит на множество точек в целевой нуклеиновой кислоте, и, например, в целевую нуклеиновую кислоту CasX вводит множество разрывов. Следует понимать, что в таких случаях вторая или дополнительная гРНК находится в комплексе с дополнительной копией белка CasX. Посредством выбора нацеливающих последовательностей гРНК можно модифицировать или редактировать определенные области последовательности целевой нуклеиновой кислоты, охватывающие конкретное местоположение в целевой нуклеиновой кислоте, применяя системы CasX:гРНК, описанные в настоящем документе, в том числе облегчать вставку донорной матрицы, содержащей мутацию гена BCL11A. В конкретном варианте осуществления вторая гРНК может содержать нацеливающую последовательность, комплементарную последовательности, находящейся на 5'- или 3'-стороне и примыкающей к сайту связывания GATA1, так что сайт связывания GATA1 нарушается.

f. Каркасы гРНК

[0113] Остальные компоненты гРНК, за исключением домена нацеливающей последовательности, называются в настоящем документе каркасом. В некоторых вариантах осуществления каркасы гРНК получены из встречающихся в природе последовательностей, описанных ниже в качестве референсной гРНК. В других вариантах осуществления каркасы гРНК представляют собой варианты референсной гРНК, в которую введены мутации, вставки, делеции или замены доменов для придания гРНК желательных или улучшенных свойств.

[0114] Термин «смежный» по отношению к полинуклеотидным или аминокислотным последовательностям относится к последовательностям, которые находятся рядом друг с другом или смежны друг с другом в полинуклеотиде или полипептиде. Специалисту в данной области будет понятно, что две последовательности

могут считаться смежными друг с другом и все же включать в себя ограниченное количество промежуточной последовательности, например, 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 нуклеотидов или аминокислот.

[0115] В таблице 2 представлены тракр- и каркасные последовательности референсной гРНК. В некоторых вариантах осуществления в описании предложены последовательности гРНК, в которых гРНК имеет каркас, содержащий последовательность с SEQ ID NO: 4-16, как указано в таблице 2, или последовательность, имеющую по меньшей мере одну нуклеотидную модификацию относительно референсной последовательности гРНК, имеющей последовательность по любой из SEQ ID NO: 4-16 таблицы 2. Следует понимать, что в тех вариантах осуществления, в которых вектор содержит ДНК, кодирующую последовательность для гРНК, или где гРНК представляет собой химеру РНК и ДНК, тиминовые (Т) основания могут быть заменены на урациловые (U) основания в любом из вариантов осуществления последовательности гРНК, описанных в настоящем документе.

Таблица 2. Тракр- и каркасные последовательности референсной гРНК

SEQ ID NO:	Нуклеотидная последовательность
4	ACAUCUGGCGCGUUUAUUC CAUACUUUGGAGCCAGUCCCAGCGAC UAUGUCGUAUGGACGAAGCGCUUAUUUAUCGGAGAGAAACCGAUA AGUAAAACGCAUCAAAAG
5	UACUGGCGCUUUUAUCUCAUACUUUGAGAGCCAUCACCAGCGACU AUGUCGUAUGGGUAAAGCGCUUAUUUAUCGGAGAGAAAUCCGAUA AAUAAGAAGCAUCAAAAG
6	ACAUCUGGCGCGUUUAUUC CAUACUUUGGAGCCAGUCCCAGCGAC UAUGUCGUAUGGACGAAGCGCUUAUUUAUCGGAGA
7	ACAUCUGGCGCGUUUAUUC CAUACUUUGGAGCCAGUCCCAGCGAC UAUGUCGUAUGGACGAAGCGCUUAUUUAUCGG
8	UACUGGCGCUUUUAUCUCAUACUUUGAGAGCCAUCACCAGCGACU AUGUCGUAUGGGUAAAGCGCUUAUUUAUCGGAGA
9	UACUGGCGCUUUUAUCUCAUACUUUGAGAGCCAUCACCAGCGACU AUGUCGUAUGGGUAAAGCGCUUAUUUAUCGG
10	GUUUACACACUCCCUCUCAUAGGGU
11	GUUUACACACUCCCUCUCAUGAGGU
12	UUUUACAUAACCCCUCAUGGGAU
13	GUUUACACACUCCCUCUCAUGGGGG
14	CCAGCGACUAUGUCGUAUGG

15	GCGCUUAUUUAUCGGAGAGAGAAAUCCGAUAAAUAAGAAGC
16	GGCGCUUUUAUCUCAUUACUUUGAGAGCCAUCACCAGCGACUAUGU CGUAUGGGUAAAGCGCUUAUUUAUCGGA

g. Варианты гРНК

[0116] В другом аспекте описание относится к вариантам гидовой нуклеиновой кислоты (альтернативно называемым в настоящем документе как «вариант гРНК» или «вариация гРНК»), которые содержат одну или более модификаций относительно каркаса референсной гРНК. В настоящем документе термин «каркас» относится ко всем частям гРНК, необходимым для функционирования гРНК, за исключением нацеливающей последовательности.

[0117] В некоторых вариантах осуществления вариант гРНК содержит одну или более нуклеотидных замен, вставок, делеций, перестановок или заменных областей относительно референсной последовательности гРНК описания. В некоторых вариантах осуществления мутация может осуществляться в любой области каркаса референсной гРНК с получением варианта гРНК. В некоторых вариантах осуществления каркас варианта последовательности гРНК на по меньшей мере 20%, по меньшей мере 30%, по меньшей мере 40%, по меньшей мере 50%, по меньшей мере 60% или по меньшей мере 70%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичен последовательности с SEQ ID NO: 4 или SEQ ID NO: 5.

[0118] В некоторых вариантах осуществления вариант гРНК содержит одно или более изменений нуклеотидов в пределах одной или более областей референсной гРНК, которые улучшают характеристику относительно референсной гРНК. Примеры областей включают триплекс РНК, псевдоузел, каркасную стеблевую петлю и удлинённую стеблевую петлю. В некоторых случаях вариант каркасного стебля дополнительно содержит пузырь. В других случаях вариант каркаса дополнительно содержит область петли триплекса. В других случаях вариант каркаса дополнительно содержит 5'-неструктурированную область. В одном варианте осуществления вариант каркаса гРНК содержит каркасную стеблевую петлю, которая на по меньшей мере 60% идентична последовательности с SEQ ID NO: 14. В другом варианте осуществления вариант гРНК содержит каркасную стеблевую петлю, имеющую последовательность CCAGCGACUAUGUCGUAGUGG (SEQ ID NO: 25). В другом варианте осуществления в описании предложен каркас гРНК, содержащий относительно SEQ ID NO:5 замену C18G, вставку G55, делецию U1 и модифицированную удлинённую стеблевую петлю, в которой исходная 6-нуклеотидная петля и 13 пар наиболее проксимальных к петле оснований (всего 32 нуклеотидов) заменены на шпильку Uvsx (4-нуклеотидная петля и 5 проксимальных к петле оснований; всего 14 нуклеотидов), и дистальное от петли основание удлинённого стебля преобразовано в стебель с полным спариванием оснований, смежный с новой

шпилькой Uvsx, посредством делеции A99 и замены G64U. В вышеуказанном варианте осуществления каркас гРНК содержит последовательность ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAG UGGGUAAGCUCUCCUCUUCGGAGGGAGCAUCAAAG (SEQ ID NO: 2238).

[0119] Все варианты гРНК, которые имеют одну или более улучшенных функций или характеристик или добавление одной или более новых функций при сравнении варианта гРНК с референсной гРНК, описанной в настоящем документе, предусмотрены в пределах объема описания. Типичный пример такого варианта гРНК представляет собой гид 174 (SEQ ID NO: 2238), конструкция которого (с обоснованием конструкции) описана в разделе «Примеры». В некоторых вариантах осуществления вариант гРНК добавляет новую функцию в RNP, содержащий вариант гРНК. В некоторых вариантах осуществления вариант гРНК имеет улучшенную характеристику, выбранную из: улучшенной стабильности; улучшенной растворимости; улучшенной транскрипции гРНК; улучшенной стойкости к нуклеазной активности; увеличенной скорости сворачивания гРНК; сниженного образования побочных продуктов при сворачивании; увеличенного продуктивного сворачивания; улучшенной аффинности связывания с белком CasX; улучшенной аффинности связывания с целевой ДНК в комплексе с белком CasX; улучшенного редактирования гена в комплексе с белком CasX; улучшенной специфичности редактирования в комплексе с белком CasX и улучшенной способности использовать больший спектр из одной или более PAM-последовательностей, включая ATC, CTC, GTC или TTC, при редактировании целевой ДНК в комплексе с белком CasX, или любая их комбинация. В некоторых случаях одна или более улучшенных характеристик варианта гРНК улучшена в по меньшей мере от примерно 1,1 до примерно 100 000 раз относительно референсной гРНК с SEQ ID NO: 4 или SEQ ID NO: 5. В других случаях одна или более улучшенных характеристик варианта гРНК улучшена в по меньшей мере примерно 1,1, в по меньшей мере примерно 10, в по меньшей мере примерно 100, в по меньшей мере примерно 1000, в по меньшей мере примерно 10 000, в по меньшей мере примерно 100 000 или более раз относительно референсной гРНК SEQ ID NO: 4 или SEQ ID NO: 5. В других случаях одна или более улучшенных характеристик варианта гРНК улучшена в примерно 1,1-100,00 раза, примерно 1,1-10,00 раза, примерно 1,1-1000 раз, примерно 1,1-500 раз, примерно 1,1-100 раз, примерно 1,1-50 раз, примерно 1,1-20 раз, примерно 10-100,00 раза, примерно 10-10,00 раза, примерно 10-1000 раз, примерно 10-500 раз, примерно 10-100 раз, примерно 10-50 раз, примерно 10-20 раз, примерно 2-70 раз, примерно 2-50 раз, примерно 2-30 раз, примерно 2-20 раз, примерно 2-10 раз, примерно 5-50 раз, примерно 5-30 раз, примерно 5-10 раз, примерно 100-100,00 раза, примерно 100-10,00 раза, примерно 100-1000 раз, примерно 100-500 раз, примерно 500-100,00 раза, примерно 500-10,00 раза, примерно 500-1000 раз, примерно 500-750 раз, примерно 1000-100,00 раза, примерно 10 000-100,00 раза, примерно 20-500 раз, примерно 20-250 раз, примерно 20-200 раз, примерно 20-100 раз, примерно 20-50 раз, примерно 50-10 000 раз, примерно 50-1000 раз, примерно 50-500 раз, примерно 50-200 раз или примерно 50-100 раз относительно референсной гРНК с SEQ ID

NO: 4 или SEQ ID NO: 5. В других случаях одна или более улучшенных характеристик варианта гРНК имеет улучшение в 1,1 раза, 1,2 раза, 1,3 раза, 1,4 раза, 1,5 раза, 1,6 раза, 1,7 раза, 1,8 раза, 1,9 раза, 2 раза, 3 раза, 4 раза, 5 раз, 6 раз, 7 раз, 8 раз, 9 раз, 10 раз, 11 раз, 12 раз, 13 раз, 14 раз, 15 раз, 16 раз, 17 раз, 18 раз, 19 раз, 20 раз, 25 раз, 30 раз, 40 раз, 45 раз, 50 раз, 55 раз, 60 раз, 70 раз, 80 раз, 90 раз, 100 раз, 110 раз, 120 раз, 130 раз, 140 раз, 150 раз, 160 раз, 170 раз, 180 раз, 190 раз, 200 раз, 210 раз, 220 раз, 230 раз, 240 раз, 250 раз, 260 раз, 270 раз, 280 раз, 290 раз, 300 раз, 310 раз, 320 раз, 330 раз, 340 раз, 350 раз, 360 раз, 370 раз, 380 раз, 390 раз, 400 раз, 425 раз, 450 раз, 475 раз или 500 раз относительно референсной гРНК с SEQ ID NO: 4 или SEQ ID NO: 5.

[0120] В некоторых вариантах осуществления вариант гРНК может быть создан посредством воздействия на референсную гРНК одним или более способами мутагенеза, такими как способы мутагенеза, описанные в настоящем документе ниже, которые могут включать в себя глубокую мутационную эволюцию (DME), глубокое мутационное сканирование (DMS), подверженную ошибкам ПЦР, каскадный мутагенез, случайный мутагенез, ПЦР со ступенчатой достройкой, перетасовку генов или перестановку доменов для получения вариантов гРНК описания. Активность референсных гРНК можно применять в качестве эталона, с которым сравнивают активность вариантов гРНК, измеряя таким образом улучшения функции вариантов гРНК по сравнению с референсной гРНК. В других вариантах осуществления референсная гРНК может быть подвергнута одной или более намеренным, нацеленным мутациям, заменам или перестановкам доменов для получения варианта гРНК, например, логически сконструированного варианта. Примеры вариантов гРНК, полученных такими способами, описаны в разделе «Примеры», а типичные последовательности каркасов гРНК представлены в таблице 3.

[0121] В некоторых вариантах осуществления вариант гРНК содержит одну или более модификаций по сравнению с референсной последовательностью каркаса гидовой нуклеиновой кислоты, причем одна или более модификаций выбраны из: по меньшей мере одной нуклеотидной замены в области варианта гРНК; по меньшей мере одной делеции нуклеотида в области варианта гРНК; по меньшей мере одной вставки нуклеотида в области варианта гРНК; полной или частичной замены области варианта гРНК; полной или частичной делеции области варианта гРНК или любой комбинации вышеизложенного. В некоторых случаях модификация представляет собой замену от 1 до 15 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях. В других случаях модификация представляет собой делецию от 1 до 10 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях. В других случаях модификация представляет собой вставку от 1 до 10 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях. В других случаях модификация представляет собой замену каркасной стеблевой петли или удлиненной стеблевой петли на последовательность стеблевой петли РНК из гетерологичного источника РНК с проксимальными 5'- и 3'-концами. В некоторых случаях вариант гРНК описания содержит две или более модификации в одной области. В других

случаях вариант гРНК описания содержит модификации в двух или более областях. В других случаях вариант гРНК содержит любую комбинацию вышеуказанных модификаций, описанных в данном параграфе.

[0122] В некоторых вариантах осуществления к варианту последовательности гРНК добавляют 5'-G для экспрессии *in vivo*, так как транскрипция с промотора U6 более эффективна и более единообразна для начального сайта, если нуклеотид +1 представляет собой G. В других вариантах осуществления к варианту последовательности гРНК добавляют два 5'-G для транскрипции *in vitro* с целью повышения эффективности производства, так как T7-полимераза имеет сильное предпочтение к G в положении +1 и пурины в положении +2. В некоторых случаях основания 5'-G добавляют к референсным каркасам из таблицы 2. В других случаях основания 5'-G добавляют к вариантам каркасов с SEQ ID NO: 2238-2285, 26794-26839 и 27219-2726 из таблицы 3.

[0123] В таблице 3 представлены примеры вариантов последовательностей каркаса гРНК описания. В некоторых вариантах осуществления вариант каркаса гРНК содержит любую из последовательностей, указанных в таблице 3, с SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99%. В некоторых вариантах осуществления вариант каркаса гРНК содержит любую из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265. В некоторых вариантах осуществления вариант каркаса гРНК содержит любую из SEQ ID NO: 2281-2285, 26794-26839 и 27219-27265 или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99%. В некоторых вариантах осуществления вариант каркаса гРНК содержит любую из SEQ ID NO: 2281-2285, 26794-26839 и 27219-27265. Следует понимать, что в тех вариантах осуществления, в которых вектор содержит ДНК, кодирующую последовательность для гРНК, или где гРНК представляет собой химеру РНК и ДНК, тиминовые (T) основания могут быть заменены на урациловые (U) основания в любом из вариантов осуществления последовательности гРНК, описанных в настоящем документе.

Таблица 3. Примеры последовательностей каркаса гРНК

SEQ ID NO:	НАЗВАНИЕ	НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ИЛИ ОПИСАНИЕ МОДИФИКАЦИИ
2101	НД	стабильная репликация фага
2102	НД	Слипающаяся петля b1
2103	НД	Слипающаяся петля a
2104	НД	32, шпилька uvsX

2105	НД	PP7
2106	НД	64, тройн. мут., усеченный удлинённый стебель
2107	НД	гиперстабильная четверная петля
2108	НД	C18G
2109	НД	U17G
2110	НД	петля CUUCGG
2111	НД	MS2
2112	НД	-1, A2G, -78, G77U
2113	НД	Q-бета
2114	НД	шпилька 45, 44
2115	НД	U1A
2116	НД	A14C, U17G
2117	НД	петля CUUCGG, модифицированная
2118	НД	Слипающаяся петля b2
2119	НД	-76:78, -83:87
2120	НД	-4
2121	НД	усеченный удлинённый стебель
2122	НД	C55
2123	НД	тройн. мут.
2124	НД	-76:78
2125	НД	-1:5
2126	НД	-83:87
2127	НД	=+G28, A82U, -84,
2128	НД	=+51U
2129	НД	-1:4, +G5A, +G86,
2130	НД	=+A94
2131	НД	=+G72
2132	НД	укороч. перед, модиф. петля CUUCGG. удлин. удлин.
2133	НД	A14C
2134	НД	-1:3, +G3
2135	НД	=+C45, +U46
2136	НД	модифицированная петля CUUCGG, улучш. начало
2137	НД	-93:94
2138	НД	=+U45
2139	НД	-69, -94
2140	НД	-94
2141	НД	модифицированная CUUCGG, минус U в 1-м триплексе
2142	НД	-1:4, +C4, A14C, U17G, +G72, -76:78, -83:87
2143	НД	U1C, -73
2144	НД	Каркас uuCG, стебель uuCG. Перестановка стебля, укорочение t
2145	НД	Каркас uuCG, стебель uuCG. Перестановка стебля
2146	НД	=+G60
2147	НД	без стебля Каркас uuCG
2148	НД	без стебля Каркас uuCG, улучш. начало
2149	НД	Каркас uuCG, стебель uuCG, улучш. начало
2150	НД	Псевдоузлы
2151	НД	Каркас uuCG, стебель uuCG
2152	НД	Каркас uuCG, стебель uuCG, без начала
2153	НД	Каркас uuCG
2154	НД	=+GCUC36

2155	НД	G-квадруплексная теломерная корзина+концы
2156	НД	G-квадруплекс M3q
2157	НД	G-квадруплексная теломерная корзина, без концов
2158	НД	шпилька 45, 44 (старая версия)
2159	НД	Сарцин-рициновая петля
2160	НД	uvsX, C18G
2161	НД	усеченная стеблевая петля, C18G, тройн. мут. (U10C)
2162	НД	корот. фаг. репл., C18G
2163	НД	петля фаг. репл., C18G
2164	НД	=+G18, наложенные на 64
2165	НД	усеченная стеблевая петля, C18G, -1 A2G
2166	НД	петля фаг. репл., C18G, тройн. мут. (U10C)
2167	НД	корот. фаг. репл., C18G, тройн. мут. (U10C)
2168	НД	uvsX, тройн. мут. (U10C)
2169	НД	усеченная стеблевая петля
2170	НД	=+A17, наложенные на 64
2171	НД	геномный рибозим 3' HDV
2172	НД	петля фаг. репл., тройн. мут. (U10C)
2173	НД	-79:80
2174	НД	корот. фаг. репл., тройн. мут. (U10C)
2175	НД	экстраусеченная стеблевая петля
2176	НД	U17G, C18G
2177	НД	корот. фаг. репл.
2178	НД	uvsX, C18G, -1 A2G
2179	НД	uvsX, C18G, тройн. мут. (U10C), -1 A2G, HDV -99 G65U
2180	НД	Антигеномный рибозим 3' HDV
2181	НД	uvsX, C18G, тройн. мут. (U10C), -1 A2G, HDV AA(98:99)C
2182	НД	3' рибозим HDV (Lior Nissim, Timothy Lu)
2183	НД	TAC(1:3)GA, наложенный на 64
2184	НД	uvsX, -1 A2G
2185	НД	усеченная стеблевая петля, C18G, тройн. мут. (U10C), -1 A2G, HDV -99 G65U
2186	НД	корот. фаг. репл., C18G, тройн. мут. (U10C), -1 A2G, HDV -99 G65U
2187	НД	вирусный рибозим Hammerhead дикого типа 3' sTRSV
2188	НД	корот. фаг. репл., C18G, -1 A2G
2189	НД	корот. фаг. репл., C18G, тройн. мут. (U10C), -1 A2G, 3' геномный HDV
2190	НД	петля фаг. репл., C18G, тройн. мут. (U10C), -1 A2G, HDV -99 G65U
2191	НД	3' рибозим HDV (Owen Ryan, Jamie Cate)
2192	НД	петля фаг. репл., C18G, -1 A2G
2193	НД	0,14
2194	НД	-78, G77U
2195	НД	НД
2196	НД	корот. фаг. репл., -1 A2G
2197	НД	усеченная стеблевая петля, C18G, тройн. мут. (U10C), -1 A2G
2198	НД	-1, A2G
2199	НД	усеченная стеблевая петля, тройн. мут. (U10C), -1 A2G
2200	НД	uvsX, C18G, тройн. мут. (U10C), -1 A2G
2201	НД	петля фаг. репл., -1 A2G
2202	НД	петля фаг. репл., тройн. мут. (U10C), -1 A2G

2203	НД	петля фаг. репл., C18G, тройн. мут. (U10C), -1 A2G
2204	НД	усеченная стеблевая петля, C18G
2205	НД	uvsX, тройн. мут. (U10C), -1 A2G
2206	НД	усеченная стеблевая петля, -1 A2G
2207	НД	корот. фаг. репл., тройн. мут. (U10C), -1 A2G
2208	НД	Рибозим 5' HDV (Owen Ryan, Jamie Cate)
2209	НД	Геномный рибозим 5' HDV
2210	НД	усеченная стеблевая петля, C18G, тройн. мут. (U10C), -1 A2G, HDV AA(98:99)C
2211	НД	Рибозим 5' env25 pistol (с добавленной петлей CUUCGG)
2212	НД	антигеномный рибозим 5' HDV
2213	НД	Рибозим 3' Hammerhead (Lior Nissim, Timothy Lu), гидовый каркас, scar
2214	НД	=+A27, наложенные на 64
2215	НД	Рибозим 5' Hammerhead (Lior Nissim, Timothy Lu), меньший scar
2216	НД	петля фаг. репл., C18G, тройн. мут. (U10C), -1 A2G, HDV AA(98:99)C
2217	НД	-27, наложенные на 64
2218	НД	3' Hatchet
2219	НД	Рибозим 3' Hammerhead (Lior Nissim, Timothy Lu)
2220	НД	5' Hatchet
2221	НД	Рибозим 5' HDV (Lior Nissim, Timothy Lu)
2222	НД	Рибозим 5' Hammerhead (Lior Nissim, Timothy Lu)
2223	НД	Минимальный рибозим Hammerhead 3' HH15
2224	НД	5'-рекрутирующий мотив RBMX
2225	НД	Рибозим 3' Hammerhead (Lior Nissim, Timothy Lu), меньший scar
2226	НД	Рибозим 3' env25 pistol (с добавленной петлей CUUCGG)
2227	НД	3' Env-9 Twister
2228	НД	=+AUUAUCUCAUACU25
2229	НД	5' Env-9 Twister
2230	НД	3' Twisted Sister 1
2231	НД	без стебля
2232	НД	Минимальный рибозим Hammerhead 5' HH15
2233	НД	Рибозим 5' Hammerhead (Lior Nissim, Timothy Lu), гидовый каркас, scar
2234	НД	5' Twisted Sister 1
2235	НД	Вирусный рибозим Hammerhead дикого типа 5' sTRSV
2236	НД	148, =+ G55, наложенный на 64
2237	НД	158, 103+148(+ G55) -99, G65U
2238	174	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2239	175	ACUGGCGCCUUUAUCUCAUUAUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2240	176	GCUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2241	177	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGCA UCAAG
2242	181	ACUGGCGCCUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC

		GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2243	182	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2244	183	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCUUACGGACUUCGGUCC GUAAGAAGCAUCAAAG
2245	184	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAUUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2246	185	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAUUGGGUAAAGCGCUUACGGACUUCGGUCC GUAAGAAGCAUCAAAG
2247	186	ACUGGCGCCUUUAUCAUCAUACUUUGAGAGCCAUCACCAG CGACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCC GUAAGAAGCAUCAAAG
2248	187	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCCCUCUUCGGAGGGAG CAUCAAAG
2249	188	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUCACCCA UGUGAGCAUCAAAG
2250	189	ACUGGCACUUUUACCUGAUUACUUUGAGAGCCAACACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2251	190	ACUGGCACUUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2252	191	ACUGGCCCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2253	192	ACUGGCGCUUUUACCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2254	193	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAACACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC AUCAAAG
2255	195	ACUGGCACCUUUACCUGAUUACUUUGAGAGCCAACACCAGC GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2256	196	ACUGGCACCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2257	197	ACUGGCCCCUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2258	198	ACUGGCGCCUUUAUCUGAUUACUUUGAGAGCCAACACCAGC GACUAUGUCGUAUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
2259	199	GCUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUCUUCGGAGGGAGC

		AUCAAAG
2260	200	GACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAG CGACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAG CAUCAAAG
2261	201	ACUGGCGCCUUUAUCUGAUUACUUUGGAGAGCCAUCACCAG CGACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAG CAUCAAAG
2262	202	ACUGGCGCAUUUAUCUGAUUACUUUGUGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2263	203	ACUGGCGCCUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2264	204	ACUGGCGCUUUUAUCUGAUUACUUUGGAGAGCCAUCACCAG CGACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAG CAUCAAAG
2265	205	ACUGGCGCAUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2266	206	ACUGGCGCUUUUAUCUGAUUACUUUGUGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2267	207	ACUGGCGCUUUUAUUCUGAUUACUUUGAGAGCCAUCACCAG CGACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAG CAUCAAAG
2268	208	ACGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCG ACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGCA UCAAG
2269	209	ACUGGCGCUUUUAUAUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2270	210	ACUGGCGCUUUUAUCUUGAUUACUUUGAGAGCCAUCACCAG CGACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAG CAUCAAAG
2271	211	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAGCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAAG
2272	212	ACUGGCGCUGUUAUCUGAUUACUUCGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCGAAG
2273	213	ACUGGCGCUCUUAUCUGAUUACUUCGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCGAAG
2274	214	ACUGGCGCUUGUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAGAG
2275	215	ACUGGCGCUUCUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAGAG
2276	216	ACUGGCGCUUGAUCUGAUUACCUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCCUCUUCGGAGGGAGC AUCAAGG

2277	217	ACUGGCGCUUUCaucUGAUUACCUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCAAGG
2278	218	ACUGGCGCUGUUAUCUGAUUACUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCAAAG
2279	219	ACUGGCGCUUUUAUCUGAUUACUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCGAAG
2280	220	ACUGGCGCUUUUAUCUGAUUACUUCGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCAAAG
2281	221	ACUGGCACUUCUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
2282	222	ACUGGCACUUCUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCAGAG
2283	223	ACUGGCACCUUUAUCUGAUUACUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAAAG
2284	224	ACUGGCACUUGUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
2285	225	ACUGGCACUUGUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGC AUCAGAG
27219	226	ACUGGCGCUUUUAUCUGAUUACUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCG UCAAUGACGCUGACGGUACAGGCCAGACAAUUAUUGUCUGG UAUAGUGCAGCAUCAAAG
26794	229	ACUGGCACUUUUAUCUGAUUACUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAAAG
26795	230	ACUGGCACUUCUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCUUACGGACUUCGGUCCG UAAGAAGCAUCAGAG
26796	231	ACUGGCGCUUCUAUCUGAUUACUCUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
26797	232	ACUGGCACUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
26798	233	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
26799	234	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCGCCUACGGACUUCGGUCC GUAAGGAGCAUCAGAG
26800	235	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUC CGUAAGAGGCAUCAGAG

26801	236	ACGGGACUUUCUAUCUGAUUACUCUGAAGUCCCUCACCAGC GACUAUGUCGUAUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
26802	237	ACCUGUAGUUCUAUCUGAUUACUCUGACUACAGUCACCAGC GACUAUGUCGUAUGGGUAAAGCCGCUUACGGACUUCGGUCC GUAAGAGGCAUCAGAG
26803	238	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGGGCGCAGCU UCGGCUGACGGUACACCGUGCAGCAUCAAAAG
26804	239	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGGGCGCAGCU UCGGCUGACGGUACACCGUGGGCGCAGCUUCGGCUGACGG UACACCGUGCAGCAUCAAAAG
26805	240	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGGGCGCAGCU UCGGCUGACGGUACACCGUGGGCGCAGCUUCGGCUGACGG UACACCGUGGGCGCAGCUUCGGCUGACGGUACACCGUGCA GCAUCAAAAG
26806	241	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGGGCGCAGCU UCGGCUGACGGUACACCGUGGGCGCAGCUUCGGCUGACGG UACACCGUGGGCGCAGCUUCGGCUGACGGUACACCGGUGG GCGCAGCUUCGGCUGACGGUACACCGUGCAGCAUCAAAAG
26807	242	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGGGCGCAGCU UCGGCUGACGGUACACCGUGGGCGCAGCUUCGGCUGACGG UACACCGUGGGCGCAGCUUCGGCUGACGGUACACCGGUGG GCGCAGCUUCGGCUGACGGUACACCGUGGGCGCAGCUUCG GUGACGGUACACCGUGCAGCAUCAAAAG
26808	243	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACCUAGCGGAGGCCUA GGUGCAGCAUCAAAAG
26809	244	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACCUAGCGGCUUGCUGA AGCGCGCACGGCAAGAGGGCGAGGUGCAGCAUCAAAAG
26810	245	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACCUAGCGGCUUGCUGA GGACUCGGCUUGCUGAAGCGCGCACGGCAAGAGGCGAGGGG CGGCGACUGGUGAGUACGCCAAAAAUUUUGACUAGCGGAGG CUAGAAGGAGAGAGGUGCAGCAUCAAAAG
26811	246	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGUGCCCGUCUGU UGUGUCGAGAGACGCCAAAAAUUUUGACUAGCGGAGGCUAG AAGGAGAGAGAUGGGUGCCGUGCAGCAUCAAAAG
26812	247	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACAUUGGAGAGGAGA UGUGCAGCAUCAAAAG
26813	248	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACAUUGGAGAUGUGCA GCAUCAAAAG
26814	249	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUUGGGCGCAGCGUCA AUG

		ACGCUGACGGUACAAGCAUCAAAAG
26815	250	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG UCA AUGACGCUGACGGUACAGGCCACAUGAGGAUCACCCA UGUGUAUAGUGCAGCAUCAAAAG
26816	251	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCU CAUGAGGAUCACCCAUGAGCUGACGGUACAGGCCACAUGAG GAUCACCCAUGUGGUUAUAGUGCAGCAUCAAAAG
26817	252	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG UCA AUGACGCUGACGGUACAGGCCACAUGGCAGUCGUAACG ACGCGGGUGGUUAUAGUGCAGCAUCAAAAG
26818	253	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA ACAUGGCAGUCCUAAGGACGCGGGUUUUGCUGACGGUACA GGCCACAUGGCAGUCGUAACGACGCGGGUGGUUAUAGUGCAG CAUCAAAAG
26819	254	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGAC AUGGCAGUCGUAACGACGCGGGUCUGACGGUACAGGCCACA UGAGGAUCACCCAUGUGGUUAUAGUGCAGCAUCAAAAG
26820	255	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAAGGAGUUUAU AUGGAAACCCUUAUGUGCAGCAUCAAAAG
26821	256	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCAGGAAGCACUAUGGGC GCAGCGUCAAUGACGCUGACGGUACAGGCCAGACAAUUAUU GUCUGGUUAUAGUGCAGCAGCAGAACAAUUGCUGAGGGCUA UUGAGGCGCAACAGCAUCUGUUGCAACUCACAGUCUGGGGC AUCAAGCAGCUCCAGGCAAGAAUCCUGAGCAUCAAAAG
26822	257	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGCCUGAAGAAGG GCGUGCAGCAUCAAAAG
26823	258	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGCUCGUGUAGCU CAUUAAGCUCCGAGCCGUGCAGCAUCAAAAG
26824	259	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACCCGUGUGCAUCCG CAGUGUCGGAUCCACGGGUGCAGCAUCAAAAG
26825	260	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACGGAAUCCAUUGCA CUCCGGAUUUCACUAGGUGCAGCAUCAAAAG
26826	261	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACAUUGCAUGUCUAAG ACAGCAUGUGCAGCAUCAAAAG
26827	262	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACAAAACAUAAGGAA AACCUAUGUUGUGCAGCAUCAAAAG
26828	263	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUAUGGGC GCAGCGUCAAUGACGCUGACGGUACAGGCCAGACAAUUAUU

		GUCUGGUAUAGUCCGUAAGAGGCAUCAGAG
26829	264	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGGUGGGCGCA GCGUCAAUAGACGCUGACGGUACAGGCCAGACAAUUAUUGUC UGGUACCCGUAAGAGGCAUCAGAG
26830	265	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGUAUGGGCGC AGCGUCAAUAGACGCUGACGGUACAGGCCACAUGAGGAUCAC CCAUGUGGUAUACCGUAAGAGGCAUCAGAG
26831	266	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUAUGGGGCGCAGCGU CAAUGACGCUGACGGUACAGGCCACAUGAGGAUCACCCAUG UGGUAUAGGGAGCAUCAAG
26832	267	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGUAUGGGCGC AGCUCAUGAGGAUCACCCAUGAGCUGACGGUACAGGCCACA UGAGGAUCACCCAUGUGGUAUACCGUAAGAGGCAUCAGAG
26833	268	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUAUGGGGCGCAGCUC AUGAGGAUCACCCAUGAGCUGACGGUACAGGCCACAUGAGG AUCACCCAUGUGGUAUAGGGAGCAUCAAG
26834	269	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGUAUGGGCGC AGCGUCAAUAGACGCUGACGGUACAGGCCACAUGGCAGUCGU AACGACGCGGGUGGUAUACCGUAAGAGGCAUCAGAG
26835	270	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUAUGGGGCGCAGCGU CAAUGACGCUGACGGUACAGGCCACAUGGCAGUCGUAACGA CGCGGGUGGUAUAGGGAGCAUCAAG
26836	271	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGUAUGGGCGC AGCAAACAUGGCAGUCCUAAGGACGCGGGUUUUGCUGACGG UACAGGCCACAUGGCAGUCGUAACGACGCGGGUGGUAUACC GUAAGAGGCAUCAGAG
26837	272	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUAUGGGGCGCAGCAA ACAUGGCAGUCCUAAGGACGCGGGUUUUGCUGACGGUACAG GCCACAUGGCAGUCGUAACGACGCGGGUGGUAUAGGGAGCA UCAAG
26838	273	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGUAUGGGCGC AGACAUGGCAGUCGUAACGACGCGGGUCUGACGGUACAGGC CACAUGAGGAUCACCCAUGUGGUAUACCGUAAGAGGCAUCA GAG
26839	274	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUAUGGGGCGCAGACA UGGCAGUCGUAACGACGCGGGUCUGACGGUACAGGCCACAU GAGGAUCACCCAUGUGGUAUAGGGAGCAUCAAG
27220	275	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CCUGAGGAUCACCCAGGUGCUGACGGUACAGGCCACCUGAG GAUCACCCAGGUGGUAUAGUGCAGCAUCAAG

27221	276	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG CAUGAGGAUACCCCAUGCGCUGACGGUACAGGCCGCAUGAG GAUCACCCCAUGCGGUUAUAGUGCAGCAUCAAAG
27222	277	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG CCUGAGGAUACCCAGGGCGCUGACGGUACAGGCCGCCUGAG GAUCACCCAGGGCGGUUAUAGUGCAGCAUCAAAG
27223	278	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG CCUGAGCAUCAGCCAGGGCGCUGACGGUACAGGCCGCCUGAG CAUCAGCCAGGGCGGUUAUAGUGCAGCAUCAAAG
27224	279	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUGAGCAUCAGCCAUGUGCUGACGGUACAGGCCACAUGAG CAUCAGCCAUGUGGUUAUAGUGCAGCAUCAAAG
27225	280	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUGAGUAUCAACCAUGUGCUGACGGUACAGGCCACAUGAG UAUCAACCAUGUGGUUAUAGUGCAGCAUCAAAG
27226	281	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUGAGAAUCAGCCAUGUGCUGACGGUACAGGCCACAUGAG AAUCAGCCAUGUGGUUAUAGUGCAGCAUCAAAG
27227	282	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCC CUUGAGGAUACCCCAUGUGCUGACGGUACAGGCCCCUUGAG GAUCACCCCAUGUGGUUAUAGUGCAGCAUCAAAG
27228	283	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CUUGAGGAUACCCCAUGUGCUGACGGUACAGGCCACUUGAG GAUCACCCCAUGUGGUUAUAGUGCAGCAUCAAAG
27229	284	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CCUGAGGAUACCCCAUGUGCUGACGGUACAGGCCACCUGAG GAUCACCCCAUGUGGUUAUAGUGCAGCAUCAAAG
27230	285	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUGAGGAUACCCUUAUGUGCUGACGGUACAGGCCACAUGAG GAUCACCUAUGUGGUUAUAGUGCAGCAUCAAAG
27231	286	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUUAGGAUACCCAAUGUGCUGACGGUACAGGCCACAUAUAG GAUCACCAUUGUGGUUAUAGUGCAGCAUCAAAG
27232	287	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUUAGGAUACCCGAUGUGCUGACGGUACAGGCCACAUAUAG GAUCACCGAUGUGGUUAUAGUGCAGCAUCAAAG
27233	288	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCAACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCA CAUUAGGAUACCCUUAUGUGCUGACGGUACAGGCCACAUAUAG GAUCACCUAUGUGGUUAUAGUGCAGCAUCAAAG

27234	289	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGAGGAUUAACCAUGUGCUGACGGUACAGGCCACAUGAG GAUUAACCAUGUGGUUAUAGUGCAGCAUCAAAG
27235	290	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGAGGAUAACCAUGUGCUGACGGUACAGGCCACAUGAG GAUAACCAUGUGGUUAUAGUGCAGCAUCAAAG
27236	291	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGAGGAUGACCAUGUGCUGACGGUACAGGCCACAUGAG GAUGACCAUGUGGUUAUAGUGCAGCAUCAAAG
27237	292	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGAGGACCACCAUGUGCUGACGGUACAGGCCACAUGAG GACCACCAUGUGGUUAUAGUGCAGCAUCAAAG
27238	293	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA GAUGAGGAUCACCAUGGGCUGACGGUACAGGCCAGAUGAG GAUCACCAUGGGGUUAUAGUGCAGCAUCAAAG
27239	294	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGGGGAUCACCAUGUGCUGACGGUACAGGCCACAUGGG GAUCACCAUGUGGUUAUAGUGCAGCAUCAAAG
27240	295	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGCGCAGCA CAUGAGGAUCACCAUGUGCUGACGGUACAGGCCACAUGAG GAUCACCAUGUGGUUAUAGUGCAGCAUCAAAG
27241	296	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACCUGAGGAUCACCA GGUGAGCAUCAAAG
27242	297	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCGCAUGAGGAUCACCA UGCGAGCAUCAAAG
27243	298	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCGCCUGAGGAUCACCA GGCGAGCAUCAAAG
27244	299	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCGCCUGAGCAUCAGCCA GGCGAGCAUCAAAG
27245	300	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGCAUCAGCCA UGUGAGCAUCAAAG
27246	301	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGUAUCAACCA UGUGAGCAUCAAAG
27247	302	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGAAUCAGCCA UGUGAGCAUCAAAG
27248	303	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUUGAGGAUCACCA UGUGAGCAUCAAAG

27249	304	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACUUGAGGAUCACCCA UGUGAGCAUCAAAG
27250	305	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACCUGAGGAUCACCCA UGUGAGCAUCAAAG
27251	306	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUCACCUA UGUGAGCAUCAAAG
27252	307	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUAUAGGAUCACCAA UGUGAGCAUCAAAG
27253	308	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUAUAGGAUCACCGA UGUGAGCAUCAAAG
27254	309	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUAUAGGAUCACCUA UGUGAGCAUCAAAG
27255	310	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUUAACCA UGUGAGCAUCAAAG
27256	311	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUAACCCA UGUGAGCAUCAAAG
27257	312	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUGACCCA UGUGAGCAUCAAAG
27258	313	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGACCACCCA UGUGAGCAUCAAAG
27259	314	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCAGAUGAGGAUCACCCA UGGGAGCAUCAAAG
27260	315	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGGGGAUCACCCA UGUGAGCAUCAAAG
27261	317	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCACAUGAGGAUCACCCA UGUGAGCAUCAGAG
27262	318	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCG UCAAUGACGCUGACGGUACAGGCCACAUGAGGAUCACCCA UGUGUAUAGUGCAGCAUCAGAG
27263	319	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGCU CAUGAGGAUCACCCAUGAGCUGACGGUACAGGCCACAUGAG GAUCACCCAUGUGGUUAUAGUGCAGCAUCAGAG
27264	320	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGCGCAGAC AUGGCAGUCGUAACGACGCGGGUCUGACGGUACAGGCCACA UGAGGAUCACCCAUGUGGUUAUAGUGCAGCAUCAGAG
27265	321	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC

		GACUAUGUCGUAGUGGGUAAAGCUGCACUAUGGGGGCCACAU GAGGAUCACCCAUGUGGUGUACAGCGCAGCGUCAAUAGACGC UGACGAUAGUGCAGCAUCAAAG
--	--	--

[0124] В некоторых вариантах осуществления вариант гРНК содержит одно или более дополнительных изменений в последовательности SEQ ID NO:2238, SEQ ID NO:2239, SEQ ID NO:2240, SEQ ID NO:2241, SEQ ID NO:2243, SEQ ID NO:2256, SEQ ID NO:2274, SEQ ID NO:2275, SEQ ID NO:2279, SEQ ID NO:2281, SEQ ID NO: 2285, SEQ ID NO: 26797 или SEQ ID NO: 26800 из таблицы 3.

[0125] В некоторых вариантах осуществления вариантов гРНК описания вариант гРНК содержит по меньшей мере одну модификацию, причем по меньшей мере одна модификация относительно каркаса референсной гидовой молекулы с SEQ ID NO:5 выбрана из одной или более из: (a) замены C18G в петле триплекса; (b) вставки G55 в пузыре стебля; (c) делеции U1; (d) модификации удлинительной стеблевой петли, в которой (i) петля из 6 нукл. и 13 проксимальных к петле пар оснований заменены шпилькой Uvsx и (ii) делеция A99 и замена G65U приводят к образованию дистального от петли основания, которое является полностью спаренным. В примерах осуществления вышеизложенного вариант гРНК содержит последовательность с любым из SEQ ID NO: 2238, 2241, 2244, 2248, 2249, 2256, 2259-2285, 26797 или 26800.

[0126] В некоторых вариантах осуществления вариант гРНК содержит экзогенную стеблевую петлю, имеющую длинную некодирующую РНК (днкРНК). В настоящем документе термин «днкРНК» относится к некодирующей РНК длиной более чем примерно 200 п. о. В некоторых вариантах осуществления 5'- и 3'-концы экзогенной стеблевой петли представляют собой спаренные основания; т. е. взаимодействуют с образованием области дуплексной РНК. В некоторых вариантах осуществления 5'- и 3'-концы экзогенной стеблевой петли представляют собой спаренные основания, и одна или более областей между 5'- и 3'-концами экзогенной стеблевой петли не являются спаренными основаниями.

[0127] В некоторых вариантах осуществления в описании предложены варианты гРНК с нуклеотидными модификациями относительно референсной гРНК, включающими: (a) замену 1-15 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях; (b) делецию от 1-10 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях; (c) вставку 1-10 последовательных или непоследовательных нуклеотидов в варианте гРНК в одной или более областях; (d) замену каркасной стеблевой петли или удлинительной стеблевой петли на последовательность стеблевой петли РНК из гетерологичного источника РНК с проксимальными 5'- и 3'-концами; или любую комбинацию (a)-(d). Любые из замен, вставок и делеций, описанных в настоящем документе, можно комбинировать для создания варианта гРНК описания. Например, вариант гРНК может содержать по меньшей мере одну замену и по меньшей мере одну делецию относительно референсной гРНК, по меньшей мере одну замену и по меньшей мере одну вставку относительно референсной гРНК, по меньшей мере одну вставку и по меньшей мере одну делецию относительно референсной гРНК или по меньшей мере одну замену, одну вставку и одну делецию относительно

референсной гРНК.

[0128] В некоторых вариантах осуществления вариант огРНК описания содержит одно или более дополнительных изменений в ранее созданном варианте, причем ранее созданный вариант сам по себе служит в качестве последовательности, подлежащей модификации. В некоторых вариантах осуществления вариант огРНК содержит одно или более дополнительных изменений последовательности SEQ ID NO: 2238, SEQ ID NO: 2239, SEQ ID NO: 2240, SEQ ID NO: 2241, SEQ ID NO: 2241, SEQ ID NO: 2274, SEQ ID NO: 2275, SEQ ID NO: 2279 или SEQ ID NO: 2285, SEQ ID NO: 26797 или SEQ ID NO: 26800.

[0129] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 174 (SEQ ID NO: 2238), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 174 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0130] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 175 (SEQ ID NO: 2239), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 174 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0131] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 215 (SEQ ID NO: 2275), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 215 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0132] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 221 (SEQ ID NO: 2281), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 221 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0133] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 225 (SEQ ID NO: 2285), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 225 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0134] В примерах осуществления вариант гРНК содержит одну или более модификаций относительно варианта каркаса гРНК 235 (SEQ ID NO: 26800), причем полученный вариант гРНК демонстрирует функциональное улучшение по сравнению с родительским 225 при оценке в анализе *in vitro* или *in vivo* в сопоставимых условиях.

[0135] В некоторых вариантах осуществления вариант гРНК содержит экзогенную удлиненную стеблевую петлю с такими отличиями от референсной гРНК, которые описаны ниже. В некоторых вариантах осуществления экзогенная удлиненная стеблевая петля имеет низкую идентичность или отсутствие идентичности с областями референсной стеблевой петли, описанными в настоящем документе (например, SEQ ID NO: 15). В некоторых вариантах осуществления экзогенная стеблевая петля содержит по меньшей мере 10 п. о., по меньшей мере 20 п. о., по меньшей мере 30 п. о., по меньшей мере 40 п. о., по меньшей мере 50 п. о., по меньшей мере 60 п. о., по меньшей мере 70 п. о., по меньшей мере 80 п. о.,

по меньшей мере 90 п. о., по меньшей мере 100 п. о., по меньшей мере 200 п. о., по меньшей мере 300 п. о., по меньшей мере 400 п. о., по меньшей мере 500 п. о., по меньшей мере 600 п. о., по меньшей мере 700 п. о., по меньшей мере 800 п. о., по меньшей мере 900 п. о., по меньшей мере 1000 п. о., по меньшей мере 2000 п. о., по меньшей мере 3000 п. о., по меньшей мере 4000 п. о., по меньшей мере 5000 п. о., по меньшей мере 6000 п. о., по меньшей мере 7000 п. о., по меньшей мере 8000 п. о., по меньшей мере 9000 п. о., по меньшей мере 10 000 п. о., по меньшей мере 12 000 п. о., по меньшей мере 15 000 п. о. или по меньшей мере 20 000 п. о. В некоторых вариантах осуществления вариант гРНК содержит область удлинительной стеблевой петли, содержащую по меньшей мере 10, по меньшей мере 100, по меньшей мере 500, по меньшей мере 1000 или по меньшей мере 10 000 нуклеотидов. В некоторых вариантах осуществления гетерологичная стеблевая петля повышает стабильность гРНК. В некоторых вариантах осуществления гетерологичная стеблевая петля РНК способна связывать белок, структуру РНК, последовательность ДНК или низкомолекулярное соединение. В некоторых вариантах осуществления область экзогенной стеблевой петли, заменяющая стеблевую петлю, содержит стеблевую петлю РНК или шпильку, в которой полученная гРНК обладает повышенной стабильностью и, в зависимости от выбора петли, может взаимодействовать с определенными клеточными белками или РНК. Такие экзогенные удлиненные стеблевые петли могут содержать, например, термостабильную РНК, такую как шпилька MS2 (ACAUGAGGAUCACCCAUGU (SEQ ID NO: 27)), Q β -шпилька (UGCAUGUCUAAGACAGCA (SEQ ID NO: 28)), шпилька II U1 (AAUCCAUUGCACUCCGGAUU (SEQ ID NO: 29)), U_{vsx} (CCUCUUCGGAGG (SEQ ID NO: 30)), шпилька PP7 (AGGAGUUUCUAUGGAAACCCU (SEQ ID NO: 31)), петля фаговой репликации (AGGUGGGACGACCUCUCGGUCGUCCUAUCU (SEQ ID NO: 32)), слипающаяся петля_a (UGCUCGCUCCGUUCGAGCA (SEQ ID NO: 33)), слипающаяся петля_b1 (UGCUCGACGCGUCCUCGAGCA (SEQ ID NO: 34)), слипающаяся петля_b2 (UGCUCGUUUGCGGCUACGAGCA (SEQ ID NO: 35)), G-квадруплекс M3q (AGGGAGGGAGGGAGAGG (SEQ ID NO: 149)), G-квадруплексная теломерная корзина (GGUUAGGGUUAGGGUUAGG (SEQ ID NO: 150)), сарцин-рициновая петля (CUGCUCAGUACGAGAGGAACCGCAG (SEQ ID NO: 151)) или псевдоузлы (UACACUGGGAUCGCUGAAUUAGAGAUCGGCGUCCUUUCAUUCUAUAUACUUUGGAGUUUUAAAAUGUCUCUAAGUACA (SEQ ID NO: 152)). В некоторых вариантах осуществления одна из вышеуказанных шпильчатых последовательностей встроена в стеблевую петлю для содействия встраиванию гРНК (и ассоциированного CasX в комплексе RNP) в отпочковывающуюся XDP (более подробно описано ниже).

[0136] В вариантах осуществления вариантов гРНК вариант гРНК дополнительно содержит спейсерную область (или нацеливающую последовательность), расположенную на 3'-конце гРНК, способную гибридизоваться с целевой нуклеиновой кислотой, специфичной к последовательности DMPK, описанной более подробно выше, которая содержит от по меньшей мере 14 до примерно 35 нуклеотидов, причем спейсер выполнен с

последовательностью, которая комплементарна целевой ДНК. В некоторых вариантах осуществления кодируемый вариант гРНК содержит нацеливающую последовательность из по меньшей мере 10-20 нуклеотидов, комплементарную целевой ДНК. В некоторых вариантах осуществления нацеливающая последовательность содержит 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 или 35 нуклеотидов. В некоторых вариантах осуществления кодируемый вариант гРНК содержит нацеливающую последовательность, имеющую 20 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 25 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 24 нуклеотида. В некоторых вариантах осуществления нацеливающая последовательность содержит 23 нуклеотида. В некоторых вариантах осуществления нацеливающая последовательность содержит 22 нуклеотида. В некоторых вариантах осуществления нацеливающая последовательность содержит 21 нуклеотид. В некоторых вариантах осуществления нацеливающая последовательность содержит 20 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 19 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 18 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 17 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 16 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 15 нуклеотидов. В некоторых вариантах осуществления нацеливающая последовательность содержит 14 нуклеотидов.

h. Образование комплекса с белком CasX

[0137] В некоторых вариантах осуществления при экспрессии вариант гРНК образует комплекс RNP с вариантом белка CasX, содержащим любую из последовательностей в таблице 4 (SEQ ID NO: 36-99, 101-148 и 26908-27154) или последовательность, которая идентична последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 85%, по меньшей мере примерно 90%, по меньшей мере примерно 91%, по меньшей мере примерно 92%, по меньшей мере примерно 93%, по меньшей мере примерно 94%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99%. В некоторых вариантах осуществления при экспрессии вариант гРНК образует комплекс RNP с вариантом белка CasX, содержащим любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154 или последовательность, которая идентична последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 85%, по меньшей мере примерно 90%, по меньшей мере примерно 91%, по меньшей мере примерно 92%, по меньшей мере примерно 93%, по меньшей мере примерно 94%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99%. В

некоторых вариантах осуществления при экспрессии вариант гРНК образует комплекс RNP с вариантом белка CasX, содержащим любую из SEQ ID NO:132-148 или 26908-27154 или последовательность, которая идентична последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 85%, по меньшей мере примерно 90%, по меньшей мере примерно 91%, по меньшей мере примерно 92%, по меньшей мере примерно 93%, по меньшей мере примерно 94%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99%.

[0138] В некоторых вариантах осуществления вариант гРНК имеет улучшенную способность образовывать комплекс с белком CasX (например, референсным CasX или вариантом CasX) по сравнению с референсной гРНК. В некоторых вариантах осуществления вариант гРНК имеет улучшенную аффинность к белку CasX (например, референсному белку или варианту белка) по сравнению с референсной гРНК, и таким образом улучшается ее способность образовывать рибонуклеопротеиновый (RNP) комплекс с белком CasX, как описано в разделе «Примеры». Улучшение образования рибонуклеопротеинового комплекса может в некоторых вариантах осуществления повышать эффективность, с которой собираются функциональные RNP. В некоторых вариантах осуществления более 90%, более 93%, более 95%, более 96%, более 97%, более 98% или более 99% RNP, содержащих вариант гРНК и ее нацеливающую последовательность, обладают компетентностью в отношении генетического редактирования целевой нуклеиновой кислоты.

[0139] Примеры нуклеотидных изменений, которые могут улучшать способность вариантов гРНК образовывать комплекс с белком CasX, могут в некоторых вариантах осуществления включать замену каркасного стебля на термостабильную стеблевую петлю. Без ограничений, накладываемых какой-либо теорией, замена каркасного стебля на термостабильную стеблевую петлю может повышать общую стабильность связывания варианта гРНК с белком CasX. Альтернативно или в дополнение удаление большого участка стеблевой петли может изменять кинетику сворачивания варианта гРНК и облегчать и ускорять структурную сборку функционально свернутой гРНК, например, посредством уменьшения степени «запутывания» вариант гРНК на самой себе. В некоторых вариантах осуществления выбор последовательности каркаса стеблевой петли может меняться при разных нацеливающих последовательностях, используемых для гРНК. В некоторых вариантах осуществления последовательность каркаса может быть адаптирована к нацеливающей последовательности и, следовательно, к целевой последовательности. Можно применять биохимические анализы для оценки аффинности связывания белка CasX с вариантом гРНК с образованием RNP, в том числе анализы из раздела «Примеры». Например, специалист в данной области может измерять изменения количества флуоресцентно-меченой гРНК, которая связана с иммобилизованным белком CasX, в ответ на увеличение концентраций дополнительной немеченой гРНК-«холодного

конкурента». Альтернативно или в дополнение флуоресцентный сигнал можно отслеживать или можно наблюдать, как он изменяется, при протекании разных количеств флуоресцентно-меченой гРНК над иммобилизованным белком CasX. Альтернативно возможность образования RNP можно оценивать с применением анализов расщепления *in vitro* относительно определенной последовательности целевой нуклеиновой кислоты.

IV. Белки для модификации целевой нуклеиновой кислоты

[0140] В настоящем описании предложены системы, содержащие нуклеазу CRISPR, которая может применяться при редактировании генома эукариотических клеток. В некоторых вариантах осуществления нуклеаза CRISPR, применяемая в системах редактирования генома, представляет собой нуклеазу класса 2, типа V. Хотя члены систем CRISPR-Cas класса 2, типа V имеют различия, они обладают некоторыми общими характеристиками, которые отличают их от систем Cas9. Во-первых, нуклеазы класса 2, типа V содержат единственный РНК-направляемый содержащий домен RuvC эффектор, не содержат домена HNH и распознают богатый Т-нуклеотидами РАМ-мотив на 5'-стороне от целевой области на нецелевой цепи, и этим отличаются от систем Cas9, которые используют богатый G-нуклеотидами РАМ-мотив на 3'-стороне от целевых последовательностей. Нуклеазы типа V создают ступенчатые двухцепочечные разрывы дистально от РАМ-последовательности, в отличие от Cas9, которые создают тупой конец в проксимальном сайте вблизи РАМ. Кроме того, нуклеазы типа V разрушают транскриптивные оцДНК при активации целевой дцДНК или связывании с цис-активной оцДНК. В некоторых вариантах осуществления нуклеазы типа V вариантов осуществления распознают РАМ-мотив 5'-ТС и образуют ступенчатые концы, расщепленные исключительно доменом RuvC. В некоторых вариантах осуществления нуклеаза типа V выбрана из группы, состоящей из Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, CasZ и CasX. В некоторых вариантах осуществления в настоящем описании предложены системы, содержащие белок CasX и одну или более гРНК (система CasX:гРНК), которые специально разработаны для модификации последовательности целевой нуклеиновой кислоты в эукариотических клетках.

[0141] Термин «белок CasX», используемый в настоящем документе, означает семейство белков и охватывает все встречающиеся в природе белки CasX, белки, которые на по меньшей мере 50% идентичны встречающимся в природе белкам CasX, а также варианты CasX, обладающие одной или более улучшенными характеристиками относительно встречающегося в природе референсного белка CasX.

[0142] Белки CasX описания содержат по меньшей мере один из следующих доменов: домен связывания с нецелевой цепью (NTSB), домен загрузки целевой цепи (TSL), спиральный домен I, спиральный домен II, олигонуклеотид-связывающий домен (OBD) и домен расщепления ДНК RuvC.

[0143] В некоторых вариантах осуществления белок CasX может связываться и/или модифицировать (например, никировать, катализировать двухцепочечный разрыв, метилировать, деметилировать и т. д.) целевую нуклеиновую кислоту в конкретной

последовательности, на которую нацелена ассоциированная гРНК, гибридизирующаяся с последовательностью в пределах последовательности целевой нуклеиновой кислоты.

a. Референсные белки CasX

[0144] В описании предложены встречающиеся в природе белки CasX (называемые в настоящем документе «референсным белком CasX»), которые подвергались последующей модификации для создания вариантов CasX описания. Например, референсные белки CasX могут быть выделены из встречающихся в природных условиях прокариотов, таких как виды *Deltaproteobacteria*, *Planctomycetes* или *Candidatus Sungbacteria*. Референсный белок CasX (взаимозаменяемо называемый в настоящем документе референсным полипептидом CasX) представляет собой эндонуклеазу CRISPR/Cas типа II, относящуюся к семейству белков CasX (взаимозаменяемо называемому Cas12e), которые взаимодействуют с гидовой РНК с образованием рибонуклеопротеинового комплекса (RNP).

[0145] В некоторых случаях референсный белок CasX выделяют или получают из *Deltaproteobacter*, и он имеет последовательность:

```

1 MEKRINKIRK KLSADNATKP VSRSGPMKTL LVRVMTDDLK KRLEKRRKKP EVMPQVISNN
61 AANNLRMLLD DYTKMKEAIL QVYWQEFKDD HVGLMCKFAQ PASKKIDQNK LKPEMDEKGN
121 LTTAGFACSQ CGQPLFVYKL EQVSEKGKAY TNYFGRCNVA EHEKLILLAQ LKPEKDSDEA
181 VTYSLGKFGQ RALDFYSIHV TKESTHPVKP LAQIAGNRYA SGPVGKALSD ACMGTIASFL
241 SKYQDIIEH QKVVKGNQKR LESLRELAGK ENLEYPSVTL PPQPHTKEGV DAYNEVIARV
301 RMWVNLNLWQ KLKLSRDDAK PLLRLKGFPF FPVVERRENE VDWWNTINEV KKLIDAKRDM
361 GRVFWSGVTA EKRNTILEGY NYLPNENDHK KREGSLENPK KPAKRQFGDL LLYLEKKYAG
421 DWGKVFDEAW ERIDKKIAGL TSHIEREEAR NAEDAQSKAV LTDWLRKAS FVLERLKEMD
481 EKEFYACEIQ LQKWYGDLRG NPFAVEAENR VVDISGFSIG SDGHSIQYRN LLAWKYLENG
541 KREFYLLMNY GKKGRIRFTD GTDIKSGKW QGLLYGGGKA KVIDLTFDPD DEQLILPLA
601 FGTRQGREFI WNDLLSLETG LIKLANGRVI EKTINYKKIG RDEPALFVAL TFERREVVDP
661 SNIKPVNLIG VDRGENIPAV IALTDPEGCP LPEFKDSSGG PTDILRIGEG YKEKQRAIQA
721 AKEVEQRRAG GYSRKFASKS RNLADDMVRN SARDLFYHAV THDAVLVFEN LSRGFGRQGG
781 RTFMTERQYT KMEDWLTAKL AYEGLTSKTY LSKTLAQYTS KTCSNCGFTI TTADYDGMLV
841 RLKKTSDGWA TTLNNKELKA EGQITYYNRY KRQTVEKELS AELDRLSEES GNNDISKWTK
901 GRRDEALFLL KKRFSHRPVQ EQFVCLDCGH EVHADEQAAL NIARSWLFLN SNSTEFKSYK
961 SGKQPFVGAW QAFYKRRLKE VWKPNA (SEQ ID NO: 1).
```

[0146] В некоторых случаях референсный белок CasX выделяют или получают из *Planctomycetes*, и он имеет последовательность:

1 MQEIKRINKI RRRLVKDSNT KKAGKTGPMK TLLVRVMTPD LRERLENLRK KPENIPQPIS
 61 NTSRANLNKL LTDYTEMKKA ILHVVWEEFQ KDPVGLMSRV AQPAPKNIDQ RKLIPVKDGN
 121 ERLTSSGFAC SQCCQPLYVY KLEQVNDKKGK PHTNYFGRCN VSEHERLILL SPHKPEANDE
 181 LVTYSLGKFG QRALDFYSIH VTRESNHPVK PLEQIGGNSC ASGPVGKALS DACMGAVASF
 241 LTKYQDIILE HQKVIKKNEK RLANLKDIA S ANGLAFPKIT LPPQPHTKEG IEAYNNVVAQ
 301 IVIWWNLNLW QKLKIGRDEA KPLQRLKGFP SFPLVERQAN EVDWWDVMVCN VKKLINEKKE
 361 DGKVFQWQNL GYKRQEALLP YLSSEEDRKK GKKFARYQFG DLLLHLEKKH GEDWGKVVYDE
 421 AWERIDKKVE GLSKHIKLEE ERRSEDAQSK AALTDWLRAK ASFVIEGLKE ADKDEFRCRCE
 481 LKLQKWYGD L RGKPF AIEAE NSILDISGFS KQYNCAFIWQ KDGVKKLNLY LIINYFKGGK
 541 LRFKKIKPEA FEANRFYTVI NKKSGEIVPM EVNFNFD DPN LIILPLAFGK RQGREFIWN D
 601 LLSLETGSLK LANGRVIEKT LYNRRTRQDE PALFVALTFE RREVL DSSNI KPMNLIGIDR
 661 GENIPAVIAL TDPEGCPLSR FKDSLGNPTH ILRIGESYKE KQRTIQAAKE VEQR RAGGYS
 721 RKYASKAKNL ADDMVRNTAR DLLYYAVTQD AMLIFENLSR GFGRQGKRTF MAERQYTRME
 781 DWLTAKLAYE GLPSKTYLSK TLAQYTSKTC SNCGFTITSA DYDRVLEK LK KTATGWM TTI
 841 NGKELKVEGQ ITYYNRYKRQ NVVKDLSVEL DRLSEESVNN DISSWTKGRS GEALSLLKKR
 901 FSHRPVQEFK VCLNCGFETH ADEQAALNIA RSWLFLRSQE YKKYQTNKTT GNTDKRAFVE
 961 TWQSFYRKKL KEVWKPAV (SEQ ID NO: 2).

[0147] В некоторых случаях референсный белок CasX выделяют или получают из *Candidatus Sungbacteria*, и он имеет последовательность

1 MDNANKPSTK SLVNTTRISD HFGVTPGQVT RVFSFGIIP T KRQYAIERW FAAVEAARER
 61 LYGMLYAHFQ ENPPAYLKEK FSYETFFKGR PVLNGLRDID PTIMTSAVFT ALRHKAE GAM
 121 AAFHTNHRRL FEEARKKMRE YAECLKANEA LLRGAADIDW DKIVNALRTR LNTCLAPEYD
 181 AVIADFGALC AFRALIAETN ALKGAYNHAL NQMLPALVKV DEPEEAESP RLRFFNGRIN
 241 DLPKFPVAER ETTPDTETII RQLEDMARVI PDTAEILGYI HRIRHKAARR KPGSAVPLPQ
 301 RVALYCAIRM ERNPEEDPST VAGHFLGEID RVCEKRRQGL VRTPFDSQIR ARYMDIISFR
 361 ATLAHPDRWT EIQLFRSNA SRRVRAETIS APFEGFSWTS NRTNPAPQYG MALAKDANAP
 421 ADAPELCICL SPSSAAFSVR EKG GDLIYMR PTGGRRGKDN PGKEITWVPG SFDEYPASGV
 481 ALKLRLYFGR SQARRMLTNK TWGLLSDNPR VFAANAELVG KKRNPQDRWK LFFHMVISGP
 541 PPVEYLD FSS DVRSRARTVI GINRGEVNPL AYAVVSVEDG QVLEEGLLGK KEYIDQLIET
 601 RRRISEYQSR EQTPPRDLRQ RVRHLQDTV L GSARAKIHS L IAFWKGILAI ERLDDQFHGR
 661 EQKIIPKKT Y LANKTGFMNA LSFSGAVRVD KKGNPWGGM I EIYPGGISRT CTQCGTVWLA
 721 RRPKNPGHRD AMVVIPDIVD DAAATGFDNV DCDAGTVDYG ELFTLSREWV RLTPRYSRVM
 781 RGTGLDLERA IRQGDDRKSR QMLELAL EPQ PQWGQFFCHR CGFNGQSDVL AATNLARRAI
 841 SLIRRLPDTD TPPTP (SEQ ID NO: 3).

b. Варианты белка CasX

[0148] В настоящем описании предложены варианты референсного белка CasX (взаимозаменяемо называемого в настоящем документе «вариант CasX» или «вариант белка CasX»), причем варианты CasX содержат по меньшей мере одну модификацию в по меньшей мере одном домене относительно референсного белка CasX, включая, без ограничений, последовательности с SEQ ID NO:1-3.

[0149] Варианты CasX описания имеют одну или более улучшенных характеристик по сравнению с референсными белками CasX. Примеры улучшенных характеристик вариантов осуществления вариантов CasX включают, без ограничений, улучшенное сворачивание варианта, улучшенную аффинность связывания с гРНК, улучшенную аффинность связывания с целевой нуклеиновой кислотой, улучшенную способность использовать более широкий спектр РАМ-последовательностей при редактировании и/или связывании целевой ДНК, улучшенное раскручивание целевой ДНК, улучшенную эффективность редактирования, улучшенную специфичность редактирования, повышенную долю эукариотического генома, который можно эффективно отредактировать, повышенную активность нуклеазы, повышенную загрузку целевой цепи для двухцепочечного расщепления, сниженную загрузку целевой цепи для одноцепочечного никирования, сниженное нецелевое расщепление, улучшенное связывание с нецелевой цепью ДНК, улучшенная стабильность белка, улучшенная стабильность комплекса белок:гРНК (RNP), улучшенная растворимость белка, улучшенная растворимость комплекса белок:гРНК (RNP), улучшенный выход белка, улучшенная экспрессия белка и улучшенные характеристики слияния, как более подробно описано ниже. Примеры улучшенных характеристик описаны в WO 2020/247882A1 и WO 2020/247883, включенных в настоящий документ путем ссылки. В вышеуказанных вариантах осуществления одна или более улучшенных характеристик варианта CasX улучшена в по меньшей мере от примерно 1,1 до примерно 100 000 раз относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 при анализе сопоставимым образом. В других вариантах осуществления улучшение составляет по меньшей мере примерно 1,1 раза, по меньшей мере примерно 2 раза, по меньшей мере примерно 5 раз, по меньшей мере примерно 10 раз, по меньшей мере примерно 50 раз, по меньшей мере примерно 100 раз, по меньшей мере примерно 500 раз, по меньшей мере примерно 1000 раз, по меньшей мере примерно 5000 раз, по меньшей мере примерно 10 000 раз или по меньшей мере примерно 100 000 раз по сравнению с референсным белком CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 при анализе сопоставимым образом. В других случаях одна или более улучшенных характеристик RNP из варианта CasX и варианта гРНК улучшена в по меньшей мере примерно 1,1, в по меньшей мере примерно 10, в по меньшей мере примерно 100, в по меньшей мере примерно 1000, в по меньшей мере примерно 10 000, в по меньшей мере примерно 100 000 или более раз относительно RNP с референсным белком CasX с SEQ ID NO:1, SEQ ID NO:2 или SEQ ID NO:3 и гРНК из таблицы 2. В других случаях одна или более улучшенных характеристик RNP из варианта CasX и варианта гРНК улучшены в от примерно 1,1 до 100,00 раза, от примерно 1,1 до 10,00 раза, от примерно 1,1 до 1000 раз, от примерно 1,1 до 500 раз, от примерно 1,1 до 100 раз, от примерно 1,1 до 50 раз, от примерно 1,1 до 20 раз, от примерно 10 до 100,00 раза, от примерно 10 до 10,00 раза, от примерно 10 до 1000 раз, от примерно 10 до 500 раз, от примерно 10 до 100 раз, от примерно 10 до 50 раз, от примерно 10 до 20 раз, от примерно 2 до 70 раз, от примерно 2 до 50 раз, от примерно 2 до 30 раз, от примерно 2 до 20 раз, от

примерно 2 до 10 раз, от примерно 5 до 50 раз, от примерно 5 до 30 раз, от примерно 5 до 10 раз, от примерно 100 до 100,00 раза, от примерно 100 до 10,00 раза, от примерно 100 до 1000 раз, от примерно 100 до 500 раз, от примерно 500 до 100,00 раза, от примерно 500 до 10,00 раза, от примерно 500 до 1000 раз, от примерно 500 до 750 раз, от примерно 10 000 до 100,00 раза, от примерно 1000 до 100,00 раза, от примерно 20 до 500 раз, от примерно 20 до 250 раз, от примерно 20 до 200 раз, от примерно 20 до 100 раз, от примерно 20 до 50 раз, от примерно 50 до 10 000 раз, от примерно 50 до 1000 раз, от примерно 50 до 500 раз, от примерно 50 до 200 раз или от примерно 50 до 100 раз относительно RNP с референсным белком CasX с SEQ ID NO:1, SEQ ID NO:2 или SEQ ID NO:3 и гРНК из таблицы 2 при анализе сопоставимым образом. В других случаях одна или более улучшенных характеристик RNP с вариантом CasX и вариантом гРНК имеет улучшение в примерно 1,1 раза, 1,2 раза, 1,3 раза, 1,4 раза, 1,5 раза, 1,6 раза, 1,7 раза, 1,8 раза, 1,9 раза, 2 раза, 3 раза, 4 раза, 5 раз, 6 раз, 7 раз, 8 раз, 9 раз, 10 раз, 11 раз, 12 раз, 13 раз, 14 раз, 15 раз, 16 раз, 17 раз, 18 раз, 19 раз, 20 раз, 25 раз, 30 раз, 40 раз, 45 раз, 50 раз, 55 раз, 60 раз, 70 раз, 80 раз, 90 раз, 100 раз, 110 раз, 120 раз, 130 раз, 140 раз, 150 раз, 160 раз, 170 раз, 180 раз, 190 раз, 200 раз, 210 раз, 220 раз, 230 раз, 240 раз, 250 раз, 260 раз, 270 раз, 280 раз, 290 раз, 300 раз, 310 раз, 320 раз, 330 раз, 340 раз, 350 раз, 360 раз, 370 раз, 380 раз, 390 раз, 400 раз, 425 раз, 450 раз, 475 раз или 500 раз относительно RNP с референсным белком CasX с SEQ ID NO:1, SEQ ID NO:2 или SEQ ID NO:3 и гРНК из таблицы 2 при анализе сопоставимым образом.

[0150] Термин «вариант CasX» включает варианты, которые представляют собой слитные белки; т. е. CasX является «слитным с» гетерологичной последовательностью. Он включает варианты CasX, содержащие последовательности вариантов CasX и N-концевые, C-концевые или внутренние слияния CasX с гетерологичным белком или его доменом.

[0151] В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в домене NTSB. В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в домене TSL. В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в спиральном домене I. В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в спиральном домене II. В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в домене OBD. В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию в домене расщепления ДНК RuvC. В некоторых вариантах осуществления по меньшей мере одна модификация в домене расщепления ДНК RuvC содержит аминокислотную замену одной или более из аминокислот K682, G695, A708, V711, D732, A739, D733, L742, V747, F755, M771, M779, W782, A788, G791, L792, P793, Y797, M799, Q804, S819 или Y857 или делецию аминокислоты P793 из SEQ ID NO:2.

[0152] В некоторых вариантах осуществления вариант белка CasX содержит по меньшей мере одну модификацию в по меньшей мере 1 домене, в по меньшей мере каждом из 2 доменов, в по меньшей мере каждом из 3 доменов, в по меньшей мере каждом из 4 доменов или в по меньшей мере каждом из 5 доменов референсного белка CasX,

включающего последовательности с SEQ ID NO: 1-3. В некоторых вариантах осуществления вариант белка CasX содержит две или более модификации в по меньшей мере одном домене референсного белка CasX. В некоторых вариантах осуществления вариант белка CasX содержит по меньшей мере две модификации в по меньшей мере одном домене референсного белка CasX, по меньшей мере три модификации в по меньшей мере одном домене референсного белка CasX или по меньшей мере четыре модификации в по меньшей мере одном домене референсного белка CasX. В некоторых вариантах осуществления, в которых вариант CasX содержит две или более модификации по сравнению с референсным белком CasX, каждая модификация осуществляется в домене, независимо выбранном из группы, состоящей из NTSBD, TSLD, спирального домена I, спирального домена II, OBD и домена расщепления ДНК RuvC. В некоторых вариантах осуществления по меньшей мере одна модификация варианта белка CasX включает делецию по меньшей мере участка одного домена референсного белка CasX с SEQ ID NO: 1-3. В некоторых вариантах осуществления делеция находится в NTSBD, TSLD, спиральном домене I, спиральном домене II, OBD или домене расщепления ДНК RuvC. В других вариантах осуществления в описании предложены варианты CasX, причем варианты CasX содержат по меньшей мере одну модификацию относительно другого варианта CasX; например, вариант 515 CasX представляет собой вариант 491 варианта CasX. Все варианты, которые улучшают одну или более функций или характеристик варианта белка CasX по сравнению с референсным белком CasX (или вариантом, из которого вариант был получен), описанным в настоящем документе, предусмотрены в пределах объема описания.

[0153] В некоторых вариантах осуществления модификация варианта CasX представляет собой мутацию одной или более аминокислот референсного CasX. В других вариантах осуществления модификация представляет собой замену одного или более доменов референсного CasX одним или более доменами из другого CasX. В некоторых вариантах осуществления вставка включает вставку части или всего домена из другого белка CasX. Мутации могут возникать в любом одном или более доменах референсного белка CasX и могут включать, например, полную или частичную делецию одного или более доменов или одну или более аминокислотных замен, делеций или вставок в любом домене референсного белка CasX. Домены белков CasX включают домен связывания с нецелевой цепью (NTSB), домен загрузки целевой цепи (TSL), спиральный домен I, спиральный домен II, олигонуклеотид-связывающий домен (OBD) и домен расщепления ДНК RuvC. Любое изменение в аминокислотной последовательности референсного белка CasX, которое приводит к улучшению характеристики белка CasX, считается вариантом белка CasX описания. Например, варианты CasX могут содержать одну или более аминокислотных замен, вставок, делеций или переставленных доменов или любые их комбинации относительно референсной последовательности белка CasX.

[0154] Приемлемые способы мутагенеза для создания вариантов белка CasX описания могут включать, например, глубокую мутационную эволюцию (DME), глубокое

мутационное сканирование (DMS), подверженную ошибкам ПЦР, кассетный мутагенез, случайный мутагенез, ПЦР со ступенчатой достройкой, перетасовку генов или перестановку доменов. В некоторых вариантах осуществления варианты CasX разработаны, например, посредством выбора одной или более желаемых мутаций в референсном CasX. В определенных вариантах осуществления активность референсного белка CasX применяется в качестве эталона, с которым сравнивают активность одного или более вариантов CasX, измеряя таким образом улучшение функции вариантов CasX.

[0155] В некоторых вариантах осуществления вариантов CasX, описанных в настоящем документе, по меньшей мере одна модификация содержит: (a) замену от 1 до 100 последовательных или непоследовательных аминокислот в варианте CasX по сравнению с референсным CasX с SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, вариантом 491 CasX или вариантом 515 CasX; (b) делецию от 1 до 100 последовательных или непоследовательных аминокислот в варианте CasX по сравнению с референсным CasX или вариантом, из которого он был получен; (c) вставку от 1 до 100 последовательных или непоследовательных аминокислот в CasX по сравнению с референсным CasX или вариантом, из которого он был получен; или (d) любую комбинацию (a)-(c). В некоторых вариантах осуществления по меньшей мере одна модификация включает: (a) замену 5-10 последовательных или непоследовательных аминокислот в варианте CasX по сравнению с референсным CasX с SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, CasX 491 или CasX 515; (b) делецию от 1 до 5 последовательных или непоследовательных аминокислот в варианте CasX по сравнению с референсным CasX или вариантом, из которого он был получен; (c) вставку от 1 до 5 последовательных или непоследовательных аминокислот в CasX по сравнению с референсным CasX или вариантом, из которого он был получен; или (d) любую комбинацию (a)-(c).

[0156] В некоторых вариантах осуществления вариант белка CasX содержит или состоит из последовательности, которая имеет по меньшей мере 1, по меньшей мере 2, по меньшей мере 3, по меньшей мере 4, по меньшей мере 5, по меньшей мере 6, по меньшей мере 7, по меньшей мере 8, по меньшей мере 9, по меньшей мере 10, по меньшей мере 20, по меньшей мере 30, по меньшей мере 40, по меньшей мере 50, по меньшей мере 60, по меньшей мере 70, по меньшей мере 80, по меньшей мере 90 или по меньшей мере 100 изменений относительно последовательности SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, CasX 491 (со ссылкой на таблицу 4) или CasX 515 (со ссылкой на таблицу 4). Эти изменения могут представлять собой аминокислотные вставки, делеции, замены или любые их комбинации. Изменения могут присутствовать в одном домене, или в любом домене, или в любой комбинации доменов варианта CasX. В заменах, описанных в настоящем документе, любая аминокислота может быть замещена на любую другую аминокислоту. Замещение может представлять собой консервативную замену (например, основную аминокислоту замещают другой основной аминокислотой). Замещение может представлять собой неконсервативную замену (например, основную аминокислоту замещают на кислотную аминокислоту, или наоборот). Например, пролин в референсном белке CasX может быть

замещен на любое из аргинина, гистидина, лизина, аспарагиновой кислоты, глутаминовой кислоты, серина, треонина, аспарагина, глутамина, цистеина, глицина, аланина, изолейцина, лейцина, метионина, фенилаланина, триптофана, тирозина или валина для создания варианта белка CasX описания.

[0157] Можно комбинировать любые сочетания вариантов осуществления с заменой, вставкой и делецией, описанных в настоящем документе, для создания варианта белка CasX описания. Например, вариант белка CasX может содержать по меньшей мере одну замену и по меньшей мере одну делецию относительно референсной последовательности белка CasX, по меньшей мере одну замену и по меньшей мере одну вставку относительно референсной последовательности белка CasX, по меньшей мере одну вставку и по меньшей мере одну делецию относительно референсной последовательности белка CasX или по меньшей мере одну замену, одну вставку и одну делецию относительно референсной последовательности белка CasX.

[0158] В некоторых вариантах осуществления вариант CasX содержит по меньшей мере одну модификацию по сравнению с референсной последовательностью CasX с SEQ ID NO:2, выбранную из одной или более из: (a) аминокислотной замены L379R; (b) аминокислотной замены A708K; (c) аминокислотной замены T620P; (d) аминокислотной замены E385P; (e) аминокислотной замены Y857R; (f) аминокислотной замены I658V; (g) аминокислотной замены F399L; (h) аминокислотной замены Q252K; (i) аминокислотной замены L404K и (j) аминокислотной делеции P793.

[0159] В некоторых вариантах осуществления вариант белка CasX содержит от 400 до 2000 аминокислот, от 500 до 1500 аминокислот, от 700 до 1200 аминокислот, от 800 до 1100 аминокислот или от 900 до 1000 аминокислот.

[0160] В некоторых вариантах осуществления вариант белка CasX содержит последовательность с SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, как указано в таблице 4. В некоторых вариантах осуществления вариант белка CasX состоит из последовательности с SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, как указано в таблице 4. В других вариантах осуществления вариант белка CasX содержит последовательность, которая на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична, на по меньшей мере 99,5% идентична последовательности с SEQ ID NO: 59, 72-99, 101-148 или

26908-27154, как указано в таблице 4. В некоторых вариантах осуществления вариант белка CasX содержит или состоит из последовательности с SEQ ID NO: 536-99, 101-148 или 26908-27154. В других вариантах осуществления вариант белка CasX содержит последовательность, которая на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична, на по меньшей мере 99,5% идентична последовательности с SEQ ID NO: 36-99, 101-148 или 26908-27154. В некоторых вариантах осуществления вариант белка CasX содержит или состоит из последовательности с SEQ ID NO: 132-148 или 26908-27154. В других вариантах осуществления вариант белка CasX содержит последовательность, которая на по меньшей мере 60% идентична, на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична, на по меньшей мере 99,5% идентична последовательности с SEQ ID NO: 132-148 или 26908-27154.

Таблица 4. Последовательности вариантов CasX

SEQ ID NO	Вариант	Описание варианта
36	НД	Домены TSL, спиральный I, спиральный II, OBD и RuvC из SEQ ID NO: 2 и домен NTSB из SEQ ID NO: 1
37	НД	Домены NTSB, спиральный I, спиральный II, OBD и RuvC из SEQ ID NO: 2 и домен TSL из SEQ ID NO: 1.
38	НД	Домены TSL, спиральный I, спиральный II, OBD и RuvC из SEQ ID NO: 1 и домен NTSB из SEQ ID NO: 2

39	НД	Домены NTSB, спиральный I, спиральный II, OBD и RuvC из SEQ ID NO: 1 и домен TSL из SEQ ID NO: 2.
40	НД	Домены NTSB, TSL, спиральный I, спиральный II и OBD из SEQ ID NO: 2 и экзогенный домен RuvC или его участок из второго белка CasX.
41	НД	НД
42	НД	Домены NTSB, TSL, спиральный II, OBD и RuvC из SEQ ID NO: 2 и домен спиральный I из SEQ ID NO: 1
43	НД	Домены NTSB, TSL, спиральный I, OBD и RuvC из SEQ ID NO: 2 и домен спиральный II из SEQ ID NO: 1
44	НД	Домены NTSB, TSL, спиральный I, спиральный II и RuvC из первого белка CasX и экзогенный OBD или его часть из второго белка CasX
45	НД	НД
46	НД	НД
47	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена T620P в SEQ ID NO: 2
48	НД	замена M771A в SEQ ID NO: 2.
49	НД	замена L379R, замена A708K, делеция Р в положении 793 и замена D732N в SEQ ID NO: 2.
50	НД	замена W782Q в SEQ ID NO: 2.
51	НД	замена M771Q в SEQ ID NO: 2
52	НД	замена R458I и замена A739V в SEQ ID NO: 2.
53	НД	L379R, замена A708K, делеция Р в положении 793 и замена M771N в SEQ ID NO: 2
54	НД	замена L379R, замена A708K, делеция Р в положении 793 и замена A739T в SEQ ID NO: 2
55	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена D489S в SEQ ID NO: 2.
56	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена D732N в SEQ ID NO: 2.
57	НД	замена V711K в SEQ ID NO: 2.
58	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена Y797L в SEQ ID NO: 2.
60	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена M771N в SEQ ID NO: 2.

61	НД	замена A708K, делеция Р в положении 793 и замена E386S в SEQ ID NO: 2.
62	НД	замена L379R, замена C477K, замена A708K и делеция Р в положении 793 в SEQ ID NO: 2.
63	НД	замена L792D в SEQ ID NO: 2.
64	НД	замена G791F в SEQ ID NO: 2.
65	НД	замена A708K, делеция Р в положении 793 и замена A739V в SEQ ID NO: 2.
66	НД	замена L379R, замена A708K, делеция Р в положении 793 и замена A739V в SEQ ID NO: 2.
67	НД	замена C477K, замена A708K и делеция Р в положении 793 в SEQ ID NO: 2.
68	НД	замена L249I и замена M771N в SEQ ID NO: 2.
69	НД	замена V747K в SEQ ID NO: 2.
70	НД	замена L379R, замена C477K, замена A708K, делеция Р в положении 793 и замена M779N в SEQ ID NO: 2.
71	НД	L379R, F755M
59	119	НД
72	429	НД
73	430	НД
74	431	НД
75	432	НД
76	433	НД
77	434	НД
78	435	НД
79	436	НД
80	437	НД
81	438	НД
82	439	НД
83	440	НД
84	441	НД
85	442	НД

86	443	НД
87	444	НД
88	445	НД
89	446	НД
90	447	НД
91	448	НД
92	449	НД
93	450	НД
94	451	НД
95	452	НД
96	453	НД
97	454	НД
98	455	НД
99	456	НД
101	457	НД
102	458	НД
103	459	НД
104	460	НД
105	278	НД
106	279	НД
107	280	НД
108	285	НД
109	286	НД
110	287	НД
111	288	НД
112	290	НД
113	291	НД
114	293	НД
115	300	НД
116	492	НД

117	493	НД
118	387	НД
119	395	НД
120	485	НД
121	486	НД
122	487	НД
123	488	НД
124	489	НД
125	490	НД
126	491	НД
127	494	НД
128	328	НД
129	388	НД
130	389	НД
131	390	НД
132	514	НД
133	515	НД
134	516	НД
135	517	НД
136	518	НД
137	519	НД
138	520	НД
139	522	НД
140	523	НД
141	524	НД
142	525	НД
143	526	НД
144	527	НД
145	528	НД
146	529	НД

147	530	НД
148	531	НД
26908	532	НД
26909	533	НД
26910	534	НД
26911	535	НД
26912	536	НД
26913	537	НД
26914	538	НД
26915	539	НД
26916	540	НД
26917	541	НД
26918	542	НД
26919	543	НД
26920	544	НД
26921	545	НД
26922	546	НД
26923	547	НД
26924	548	НД
26925	550	НД
26926	551	НД
26927	552	НД
26928	553	НД
26929	554	НД
26930	555	НД
26931	556	НД
26932	557	НД
26933	558	НД
26934	559	НД
26935	560	НД

26936	561	НД
26937	562	НД
26938	563	НД
26939	564	НД
26940	565	НД
26941	566	НД
26942	567	НД
26943	568	НД
26944	569	НД
26945	570	НД
26946	571	НД
26947	572	НД
26948	573	НД
26949	574	НД
26950	575	НД
26951	576	НД
26952	577	НД
26953	578	НД
26954	579	НД
26955	580	НД
26956	581	НД
26957	582	НД
26958	583	НД
26959	584	НД
26960	585	НД
26961	586	НД
26962	587	НД
26963	588	НД
26964	589	НД
26965	590	НД

26966	591	НД
26967	592	НД
26968	593	НД
26969	594	НД
26970	595	НД
26971	596	НД
26972	597	НД
26973	598	НД
26974	599	НД
26975	600	НД
26976	601	НД
26977	602	НД
26978	603	НД
26979	604	НД
26980	605	НД
26981	606	НД
26982	607	НД
26983	608	НД
26984	609	НД
26985	610	НД
26986	611	НД
26987	612	НД
26988	613	НД
26989	614	НД
26990	615	НД
26991	616	НД
26992	617	НД
26993	618	НД
26994	619	НД
26995	620	НД

26996	621	НД
26997	622	НД
26998	623	НД
26999	624	НД
27000	625	НД
27001	626	НД
27002	627	НД
27003	628	НД
27004	629	НД
27005	630	НД
27006	631	НД
27007	632	НД
27008	633	НД
27009	634	НД
27010	635	НД
27011	636	НД
27012	637	НД
27013	638	НД
27014	639	НД
27015	640	НД
27016	641	НД
27017	642	НД
27018	643	НД
27019	644	НД
27020	645	НД
27021	646	НД
27022	647	НД
27023	648	НД
27024	649	НД
27025	650	НД

27026	651	НД
27027	652	НД
27028	653	НД
27029	654	НД
27030	655	НД
27031	656	НД
27032	657	НД
27033	658	НД
27034	659	НД
27035	660	НД
27036	661	НД
27037	662	НД
27038	663	НД
27039	664	НД
27040	665	НД
27041	666	НД
27042	667	НД
27043	668	НД
27044	669	НД
27154	670	НД
27045	671	НД
27046	672	НД
27047	673	НД
27048	674	НД
27049	675	НД
27050	676	НД
27051	677	НД
27052	678	НД
27053	679	НД
27054	680	НД

27055	681	НД
27056	682	НД
27057	683	НД
27058	684	НД
27059	685	НД
27060	686	НД
27061	687	НД
27062	688	НД
27063	689	НД
27064	690	НД
27065	691	НД
27066	692	НД
27067	693	НД
27068	694	НД
27069	701	НД
27070	702	НД
27071	703	НД
27072	704	НД
27073	705	НД
27074	706	НД
27075	707	НД
27076	708	НД
27077	709	НД
27078	710	НД
27079	711	НД
27080	712	НД
27081	713	НД
27082	714	НД
27083	715	НД
27084	716	НД

27085	717	НД
27086	718	НД
27087	719	НД
27088	720	НД
27089	721	НД
27090	722	НД
27091	723	НД
27092	724	НД
27093	725	НД
27094	726	НД
27095	727	НД
27096	728	НД
27097	729	НД
27098	730	НД
27099	731	НД
27100	732	НД
27101	733	НД
27102	734	НД
27103	735	НД
27104	736	НД
27105	737	НД
27106	738	НД
27107	739	НД
27108	740	НД
27109	741	НД
27110	742	НД
27111	743	НД
27112	744	НД
27113	745	НД
27114	746	НД

27115	747	НД
27116	748	НД
27117	749	НД
27118	750	НД
27119	751	НД
27120	752	НД
27121	753	НД
27122	754	НД
27123	755	НД
27124	756	НД
27125	757	НД
27126	758	НД
27127	759	НД
27128	760	НД
27129	761	НД
27130	762	НД
27131	763	НД
27132	764	НД
27133	765	НД
27134	766	НД
27135	767	НД
27136	768	НД
27137	769	НД
27138	770	НД
27139	777	НД
27140	778	НД
27141	779	НД
27142	780	НД
27143	781	НД
27144	782	НД

27145	783	НД
27146	784	НД
27147	785	НД
27148	786	НД
27149	787	НД
27150	788	НД
27151	789	НД
27152	790	НД
27153	791	НД

с. Варианты белков CasX с доменами из множества исходных белков

[0161] В определенных вариантах осуществления в описании предложен химерный белок CasX, содержащий белковые домены из двух или более разных белков CasX, таких как два или более встречающихся в природе белков CasX, или две или более последовательностей вариантов белков CasX, описанных в настоящем документе. В настоящем документе термин «химерный белок CasX» относится к CasX, содержащему по меньшей мере два домена, выделенных или полученных из разных источников, таких как два встречающихся в природе белка, которые в некоторых вариантах осуществления могут быть выделены из разных биологических видов. Например, в некоторых вариантах осуществления химерный белок CasX содержит первый домен из первого белка CasX и второй домена из второго, другого белка CasX. В некоторых вариантах осуществления первый домен может быть выбран из группы, состоящей из доменов NTSB, TSL, спирального I, спирального II, OBD и RuvC. В некоторых вариантах осуществления второй домен выбран из группы, состоящей из NTSB, TSL, спирального I, спирального II, OBD и RuvC, причем второй домен отличается от вышеуказанного первого домена. В случае разделенных или несмежных доменов, таких как спиральный I, RuvC и OBD, участок несмежного домена может быть заменен на соответствующий участок из любого другого источника. Например, спиральный домен I-I (иногда называемый спиральный I-a) в SEQ ID NO: 2 можно заменить на соответствующую последовательность спирального домена I-I из SEQ ID NO: 1 и т. п. Последовательности доменов из референсных белков CasX и их координаты представлены в таблице 5. Типичные примеры химерных белков CasX включают варианты CasX 472-483, 485-491 и 515, последовательности которых приведены в таблице 4.

Таблица 5. Координаты доменов в референсных белках CasX

Доменное имя	Координаты в SEQ ID NO: 1	Координаты в SEQ ID NO: 2
OBD a	1-55	1-57
спиральный I a	56-99	58-101
NTSB	100-190	102-191
спиральный I b	191-331	192-332

спиральный II	332-508	333-500
OBD b	509-659	501-646
RuvC a	660-823	647-810
TSL	824-933	811-920
RuvC b	934-986	921-978

* OBD a и b, спиральный I a и b, a также RuvC a и b также называются в настоящем документе OBD I и II, спиральные I-I и I-II и RuvC I и II.

d. Аффинность белка к гРНК

[0162] В некоторых вариантах осуществления вариант белка CasX имеет улучшенную аффинность к гРНК относительно референсного белка CasX, что приводит к образованию рибонуклеопротеинового комплекса (RNP). Повышенная аффинность варианта белка CasX к гРНК может, например, приводить к более низкой K_d образования комплекса RNP, что в некоторых случаях может приводить к образованию более стабильного рибонуклеопротеинового комплекса. В некоторых вариантах осуществления повышенная аффинность варианта белка CasX к гРНК приводит к повышению стабильности рибонуклеопротеинового комплекса при доставке в клетки человека. Эта повышенная стабильность может влиять на функциональность и пригодность комплекса в клетках субъекта, а также приводить к улучшенным фармакокинетическим свойствам в крови при доставке субъекту. В некоторых вариантах осуществления повышенная аффинность варианта белка CasX и полученная повышенная стабильность рибонуклеопротеинового комплекса позволяют доставлять более низкую дозу варианта белка CasX субъекту или клеткам с сохранением желаемой активности, например редактирования гена *in vivo* или *in vitro*. В некоторых вариантах осуществления более высокая аффинность (более плотное связывание) варианта белка CasX с гРНК обеспечивает большее число событий редактирования, когда и вариант белка CasX, и гРНК остаются в комплексе RNP. Увеличение числа событий редактирования можно оценивать с применением анализов редактирования, таких как анализы редактирования tdTom, описанные в настоящем документе. В некоторых вариантах осуществления K_d связи варианта белка CasX с гРНК увеличивается относительно референсного белка CasX с коэффициентом по меньшей мере примерно 1,1, по меньшей мере примерно 1,2, по меньшей мере примерно 1,3, по меньшей мере примерно 1,4, по меньшей мере примерно 1,5, по меньшей мере примерно 1,6, по меньшей мере примерно 1,7, по меньшей мере примерно 1,8, по меньшей мере примерно 1,9, по меньшей мере примерно 2, по меньшей мере примерно 3, по меньшей мере примерно 4, по меньшей мере примерно 5, по меньшей мере примерно 6, по меньшей мере примерно 7, по меньшей мере примерно 8, по меньшей мере примерно 9, по меньшей мере примерно 10, по меньшей мере примерно 15, по меньшей мере примерно 20, по меньшей мере примерно 25, по меньшей мере примерно 30, по меньшей мере примерно 35, по меньшей мере примерно 40, по меньшей мере примерно 45, по меньшей мере примерно 50, по меньшей мере примерно 60, по меньшей мере примерно 70, по меньшей мере примерно 80, по меньшей мере примерно 90 или по меньшей мере примерно 100. В некоторых вариантах осуществления вариант CasX имеет

увеличенную в примерно 1,1 до примерно 10 раз аффинность связывания с гРНК по сравнению с референсным белком CasX с SEQ ID NO: 2.

[0163] В некоторых вариантах осуществления повышение аффинности варианта белка CasX к гРНК приводит к повышению стабильности рибонуклеопротеинового комплекса при доставке в клетки млекопитающих, в том числе при доставке *in vivo* субъекту. Эта повышенная стабильность может влиять на функциональность и пригодность комплекса в клетках субъекта, а также приводить к улучшенным фармакокинетическим свойствам в крови при доставке субъекту. В некоторых вариантах осуществления повышенная аффинность варианта белка CasX и полученная повышенная стабильность рибонуклеопротеинового комплекса позволяют доставлять более низкую дозу варианта белка CasX субъекту или в клетки с сохранением желаемой активности; например, редактирования генов *in vivo* или *in vitro*. Повышенную способность к образованию RNP и поддержанию их в стабильной форме можно оценивать с применением анализов, таких как анализы расщепления *in vitro*, описанные в разделе «Примеры» в настоящем документе. В некоторых вариантах осуществления RNP, содержащий варианты CasX описания, способен достигать константы скорости $k_{\text{расщепл.}}$ в комплексе в виде RNP, которая в по меньшей мере 2 раза, по меньшей мере 5 раз или по меньшей мере 10 раз выше по сравнению с таковой для RNP, содержащего референсный CasX с SEQ ID NO: 1-3.

[0164] В некоторых вариантах осуществления более высокая аффинность (более плотное связывание) варианта белка CasX с гРНК обеспечивает большее число событий редактирования, когда и вариант белка CasX, и гРНК остаются в комплексе RNP. Увеличение числа событий редактирования можно оценивать с применением анализов редактирования, таких как анализы, описанные в настоящем документе.

[0165] Без ограничений, накладываемых какой-либо теорией, в некоторых вариантах осуществления аминокислотные изменения в домене спиральный I могут повышать аффинность связывания варианта белка CasX с нацеливающей последовательностью гРНК, тогда как изменения в домене спиральный II могут повышать аффинность связывания варианта белка CasX с каркасной стеблевой петлей гРНК, а изменения в олигонуклеотид-связывающем домене (OBD) повышают аффинность связывания варианта белка CasX с триплексом гРНК.

[0166] Способы измерения аффинности связывания белка CasX с гРНК включают способы *in vitro* с применением очищенного белка CasX и гРНК. Аффинность связывания для референсного CasX и вариантов белков может быть измерена посредством флуоресцентной поляризации, если гРНК или белок CasX помечен флуорофором. Альтернативно или в дополнение к этому аффинность связывания может быть измерена посредством интерферометрии биослоев, анализа методом сдвига электрофоретической подвижности (EMSA) или связывания на фильтрах. Дополнительные стандартные методики количественного определения абсолютных аффинностей связывающих РНК белков, таких как референсный CasX и варианты белков описания для конкретных гРНК, таких как референсные гРНК и их варианты, включают, без ограничений, изотермическую

калориметрию (ИТС) и поверхностный плазмонный резонанс (SPR), а также способы из раздела «Примеры».

е. Аффинность к целевой нуклеиновой кислоте

[0167] В некоторых вариантах осуществления вариант белка CasX имеет улучшенную аффинность связывания с последовательностью целевой нуклеиновой кислоты относительно аффинности референсного белка CasX с последовательностью целевой нуклеиновой кислоты. В некоторых вариантах осуществления аффинность варианта белка CasX описания к молекуле целевой нуклеиновой кислоты увеличивается относительно референсного белка CasX с коэффициентом по меньшей мере примерно 1,1, по меньшей мере примерно 1,2, по меньшей мере примерно 1,3, по меньшей мере примерно 1,4, по меньшей мере примерно 1,5, по меньшей мере примерно 1,6, по меньшей мере примерно 1,7, по меньшей мере примерно 1,8, по меньшей мере примерно 1,9, по меньшей мере примерно 2, по меньшей мере примерно 3, по меньшей мере примерно 4, по меньшей мере примерно 5, по меньшей мере примерно 6, по меньшей мере примерно 7, по меньшей мере примерно 8, по меньшей мере примерно 9, по меньшей мере примерно 10, по меньшей мере примерно 15, по меньшей мере примерно 20, по меньшей мере примерно 25, по меньшей мере примерно 30, по меньшей мере примерно 35, по меньшей мере примерно 40, по меньшей мере примерно 45, по меньшей мере примерно 50, по меньшей мере примерно 60, по меньшей мере примерно 70, по меньшей мере примерно 80, по меньшей мере примерно 90 или по меньшей мере примерно 100.

[0168] Варианты CasX с более высокой аффинностью к своей целевой нуклеиновой кислоте могут в некоторых вариантах осуществления расщеплять последовательность целевой нуклеиновой кислоты быстрее, чем референсный белок CasX, который не имеет повышенной аффинности к целевой нуклеиновой кислоте. В некоторых вариантах осуществления улучшенная аффинность к последовательности целевой нуклеиновой кислоты включает улучшенную аффинность к последовательности целевой нуклеиновой кислоты, улучшенную аффинность связывания с более широким спектром РАМ-последовательностей, улучшенную способность поиска в ДНК последовательности целевой нуклеиновой кислоты или любые их комбинации, что приводит к улучшению способности модифицировать целевую нуклеиновую кислоту. В некоторых вариантах осуществления вариант белка CasX с улучшенной аффинностью к целевой нуклеиновой кислоте имеет повышенную аффинность к специфическим РАМ-последовательностям, отличным от канонической РАМ ТТС, распознаваемой референсным белком CasX SEQ ID NO: 2, включая аффинность связывания к РАМ-последовательностям, выбранным из группы, состоящей из ТТС, АТС, ГТС и СТС. Более высокая общая аффинность к ДНК также в некоторых вариантах осуществления может увеличивать частоту, с которой белок CasX может эффективно начать и завершить стадию связывания и раскручивания, облегчая таким образом внедрение в целевую цепь, образование R-петли и, в конечном итоге, расщепление последовательности целевой нуклеиновой кислоты.

[0169] В некоторых вариантах осуществления вариант белка CasX обладает

улучшенной аффинностью связывания с нецелевой цепью целевой нуклеиновой кислоты. В настоящем документе термин «нецелевая цепь» относится к цепи ДНК последовательности целевой нуклеиновой кислоты, которая не образует пары оснований Уотсона и Крика с нацеливающей последовательностью в гРНК и является комплементарной целевой цепи ДНК. В некоторых вариантах осуществления вариант белка CasX имеет увеличенную в от примерно 1,1 до примерно 100 раз аффинность связывания с нецелевой цепью целевой нуклеиновой кислоты по сравнению с референсным белком с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3.

[0170] Способы измерения аффинности варианта белка CasX к молекуле целевой нуклеиновой кислоты могут включать анализы методом сдвига электрофоретической подвижности (EMSA), связывание на фильтрах, изотермическую калориметрию (ИТС), а также поверхностный плазмонный резонанс (SPR), флуоресцентную поляризацию и интерферометрию биослоев (BLI). Дополнительные способы измерения аффинности белка CasX к мишени включают биохимические анализы *in vitro*, которые измеряют количество событий расщепления ДНК с течением времени; например определение константы скорости $k_{\text{расщепл.}}$, как описано в разделе «Примеры».

f. Улучшенная специфичность к целевому сайту

[0171] В некоторых вариантах осуществления вариант белка CasX имеет улучшенную специфичность к последовательности целевой нуклеиновой кислоты относительно референсного белка CasX. В настоящем документе термин «специфичность» используется взаимозаменяемо с термином «специфичность мишени» и означает уровень расщепления рибонуклеопротеиновым комплексом CRISPR/Cas нецелевых последовательностей, которые являются сходными, но не идентичными последовательности целевой нуклеиновой кислоты; например, RNP с вариантом CasX с более высокой степенью специфичности будет демонстрировать меньшее нецелевое расщепление последовательностей относительно референсного белка CasX. Специфичность и уменьшение потенциально вредных побочных эффектов системы CRISPR/Cas могут быть очень важными для достижения приемлемого терапевтического индекса при применении у субъектов-млекопитающих.

[0172] В некоторых вариантах осуществления вариант белка CasX имеет улучшенную специфичность к целевому сайту в целевой последовательности, который комплементарен нацеливающей последовательности гРНК, относительно референсного белка CasX с SEQ ID NO: 1-3. Без ограничений, накладываемых какой-либо теорией, существует возможность того, что аминокислотные изменения в спиральных доменах I и II, которые увеличивают специфичность варианта белка CasX к цепи целевой нуклеиновой кислоты, могут повышать специфичность варианта белка CasX к целевой нуклеиновой кислоте в целом. В некоторых вариантах осуществления аминокислотные изменения, которые повышают специфичность вариантов белков CasX к целевой нуклеиновой кислоте, также могут приводить к снижению аффинности вариантов белков CasX к ДНК.

[0173] Способы тестирования специфичности белка CasX (такого как вариант или

референсный белок) к мишени могут включать направление и циркуляризацию для регистрации *in vitro* эффектов расщепления посредством секвенирования (CIRCLE-seq) или аналогичные способы. Вкратце, в методиках CIRCLE-seq геномную ДНК расщепляют и циркуляризуют посредством лигирования адаптеров петля-стебель, которые никируют в областях петля-стебель с открытием 4-нуклеотидных палиндромных липких концов. После этого осуществляется внутримолекулярное лигирование и разложение оставшейся линейной ДНК. Затем циклические молекулы ДНК, содержащие сайт расщепления CasX, линейаризуют с помощью CasX, и адаптеры адаптера лигируют к открытым концам, после чего проводят высокопроизводительное секвенирование для создания прочтений спаренных концов, которые содержат информацию о нецелевом сайте. Дополнительные анализы, которые можно применять для обнаружения нецелевых событий, и, следовательно, специфичности белка CasX, включают анализы, применяемые для обнаружения и количественного определения инделов (вставок и делеций), образованных в этих выбранных нецелевых сайтах, например, анализы нуклеазами обнаружения несовпадений и секвенирование нового поколения (NGS). Примеры анализов обнаружения несовпадений включают нуклеазные анализы, в которых геномную ДНК из клеток, обработанных CasX и огРНК, подвергают ПЦР амплификации, денатурированию и регибридизации с образованием гетеродуплексной ДНК, содержащей одну цепь дикого типа и одну цепь с инделом. Несовпадения распознаются и расщепляются нуклеазами обнаружения несовпадений, такими как нуклеаза Surveyor или T7 эндонуклеаза I.

g. Протоспейсер и РАМ-последовательности

[0174] В настоящем документе протоспейсер определяется как последовательность ДНК, комплементарная нацеливающей последовательности гидовой РНК, и ДНК, комплементарная этой последовательности, называемые целевой цепью и нецелевой цепью соответственно. В настоящем документе РАМ представляет собой нуклеотидную последовательность, расположенную на расстоянии 1 нуклеотида с 5'-стороны от последовательности нецелевой цепи, которая комплементарна последовательности целевой нуклеиновой кислоты в целевой цепи целевой нуклеиновой кислоты, которая в сочетании с нацеливающей последовательностью гРНК способствует ориентированию и позиционированию CasX для потенциального расщепления цепи (-ей) протоспейсера.

РАМ-последовательности могут быть вырожденными, и конкретные конструкторы RNP могут иметь разные предпочтительные и допустимые РАМ-последовательности, которые поддерживают разные эффективности расщепления. В соответствии с соглашением, если не указано иное, описание относится как к РАМ, так и к последовательности протоспейсера и их направленности в соответствии с ориентацией нецелевой цепи. Это не подразумевает, что РАМ-последовательность нецелевой цепи, а не целевая цепь является определяющей для расщепления или механистически вовлечена в распознавание мишени. Например, если речь идет о РАМ TTC, то она фактически может быть комплементарна последовательности GAA, которая необходима для расщепления мишени, или может представлять собой некоторую комбинацию нуклеотидов из обеих

цепей. В случае белков CasX, описанных в настоящем документе, РАМ находится с 5'-стороны от протоспейсера, и один нуклеотид отделяет РАМ от первого нуклеотида протоспейсера. Таким образом, в случае референсного CasX следует понимать, что РАМ TTC означает последовательность, соответствующую формуле 5'-...NN TTCN(протоспейсер)NNNNNN...3', где «N» представляет собой любой ДНК-нуклеотид, а «(протоспейсер)» представляет собой последовательность ДНК, имеющую идентичность с нацеливающей последовательностью гидовой РНК. В случае варианта CasX с расширенным распознаванием РАМ следует понимать, что РАМ TTC, CTC, GTC или ATC означают последовательности, соответствующие формулам: 5'-...NN TTCN(протоспейсер)NNNNNN...3';

5'-...NN CTCN(протоспейсер)NNNNNN...3';

5'-...NN GTCN(протоспейсер)NNNNNN...3' или

5'-...NN ATCN(протоспейсер)NNNNNN...3'.

[0175] Альтернативно следует понимать, что РАМ TC означает последовательность, соответствующую формуле: 5'-...NN TTCN(протоспейсер)NNNNNN...3'. Кроме того, варианты белков CasX описания обладают улучшенной способностью эффективно редактировать и/или связывать целевую ДНК в комплексе с гРНК в виде RNP, с использованием РАМ-мотива TC, включая РАМ-последовательности, выбранные из TTC, ATC, GTC или CTC (в ориентации от 5' к 3'), по сравнению с RNP из референсного белка CasX и референсной гРНК. В вышеизложенном РАМ-последовательность расположена на расстоянии по меньшей мере 1 нуклеотида с 5'-стороны от нецелевой цепи протоспейсера, имеющей идентичность с нацеливающей последовательностью гРНК, в системе анализа по сравнению с эффективностью редактирования и/или связывания RNP, содержащего референсный белок CasX и референсную гРНК, в сопоставимой системе анализа. В одном варианте осуществления RNP из варианта CasX и варианта гРНК демонстрирует более высокую эффективность редактирования и/или связывания целевой последовательности в целевой ДНК по сравнению с RNP, содержащим референсный белок CasX и референсную гРНК, в сопоставимой системе анализа, причем РАМ-последовательность целевой ДНК представляет собой TTC. В другом варианте осуществления RNP из варианта CasX и варианта гРНК демонстрирует более высокую эффективность редактирования и/или связывания целевой последовательности в целевой ДНК по сравнению с RNP, содержащим референсный белок CasX и референсную гРНК, в сопоставимой системе анализа, причем РАМ-последовательность целевой ДНК представляет собой ATC. В одном варианте осуществления RNP из варианта CasX и варианта гРНК демонстрирует более высокую эффективность редактирования и/или связывания целевой последовательности в целевой ДНК по сравнению с RNP, содержащим референсный белок CasX и референсную гРНК, в сопоставимой системе анализа, причем РАМ-последовательность целевой ДНК представляет собой CTC. В одном варианте осуществления RNP из варианта CasX и варианта гРНК демонстрирует более высокую эффективность редактирования и/или связывания целевой последовательности в целевой ДНК по сравнению с RNP, содержащим

референсный белок CasX и референсную гРНК, в сопоставимой системе анализа, причем РАМ-последовательность целевой ДНК представляет собой GTC. В вышеуказанных вариантах осуществления повышение эффективности редактирования и/или аффинности связывания для одной или более РАМ-последовательностей составляет по меньшей мере 1,5 раза или более по сравнению с эффективностью редактирования и/или аффинностью связывания RNP из любого из белков CasX с SEQ ID NO:1-3 и гРНК из таблицы 2 для РАМ-последовательностей. Примеры анализов, демонстрирующие улучшенное редактирование, описаны в настоящем документе в разделе «Примеры».

h. Раскручивание ДНК

[0176] В некоторых вариантах осуществления вариант белка CasX обладает улучшенной способностью к раскручиванию ДНК относительно референсного белка CasX. Ранее было показано, что плохое раскручивание дцДНК нарушает или препятствует способности белков системы CRISPR/Cas AnaCas9 или Cas14s расщеплять ДНК. Таким образом, без ограничений, накладываемых какой-либо теорией, вероятно, что увеличение активности по расщеплению ДНК некоторыми вариантами белков CasX описания по меньшей мере частично обусловлено повышенной способностью находить и раскручивать дцДНК в целевом сайте.

[0177] Без ограничений, накладываемых какой-либо теорией, считается, что изменения аминокислот в домене NTSB могут давать варианты белков CasX с повышенными характеристиками раскручивания ДНК. Альтернативно или дополнительно аминокислотные изменения в OBD или областях спирального домена, которые взаимодействуют с РАМ, могут также давать варианты белков CasX с повышенными характеристиками раскручивания ДНК.

[0178] Способы измерения способности белков CasX (таких как вариант или референсный белок) к раскручиванию ДНК включают, без ограничений, анализы *in vitro*, которые оценивают повышение скоростей для дцДНК-мишеней при флуоресцентной поляризации или интерферометрии биослоев.

i. Каталитическая активность

[0179] Рибонуклеопротеиновый комплекс описанных в настоящем документе систем CasX:гРНК содержит вариант CasX, который связывается с последовательностью целевой нуклеиновой кислоты и расщепляет последовательность целевой нуклеиновой кислоты. В некоторых вариантах осуществления вариант белка CasX обладает улучшенной каталитической активностью относительно референсного белка CasX. Без ограничений, накладываемых какой-либо теорией, считается, что в некоторых случаях расщепление целевой цепи может быть ограничивающим фактором для Cas12-подобных молекул при создании разрыва в дцДНК. В некоторых вариантах осуществления варианты белков CasX улучшают изгибание целевой цепи ДНК и расщепление этой цепи, что приводит к улучшению общей эффективности расщепления дцДНК рибонуклеопротеиновым комплексом CasX.

[0180] В некоторых вариантах осуществления вариант белка CasX обладает

повышенной нуклеазной активностью по сравнению с референсным белком CasX. Варианты с повышенной нуклеазной активностью могут быть созданы, например, посредством аминокислотных изменений в нуклеазном домене RuvC. В некоторых вариантах осуществления вариант CasX содержит нуклеазный домен RuvC, обладающий никазой активностью. В вышеизложенном никаза CasX системы CasX:гРНК создает одноцепочечный разрыв в пределах 10-18 нуклеотидов в сторону 3' от сайта РАМ в нецелевой цепи. В других вариантах осуществления вариант CasX содержит нуклеазный домен RuvC, обладающий активностью по расщеплению двух цепей. В вышеизложенном CasX системы CasX:гРНК создает двухцепочечный разрыв в пределах 18-26 нуклеотидов в сторону 5' от сайта РАМ на целевой цепи и 10-18 нуклеотидов в сторону 3' на нецелевой цепи. Нуклеазную активность можно анализировать различными способами, включая приведенные в разделе «Примеры». В некоторых вариантах осуществления вариант CasX имеет константу $k_{\text{расщепл.}}$, которая в по меньшей мере 2 раза, или в по меньшей мере 3 раза, или в по меньшей мере 4 раза, или в по меньшей мере 5 раз, или в по меньшей мере 6 раз, или в по меньшей мере 7 раз, или в по меньшей мере 8 раз, или в по меньшей мере 9 раз, или в по меньшей мере 10 раз выше по сравнению с референсным CasX.

[0181] В некоторых вариантах осуществления вариант белка CasX обладает улучшенной характеристикой образования RNP с гРНК, которая приводит к более высокой процентной доли компетентных в отношении расщепления RNP по сравнению с RNP референсного белка CasX SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, описанной в примерах. Под фразой «компетентный в расщеплении» подразумевается, что образованный RNP обладает способностью расщеплять целевую нуклеиновую кислоту. В некоторых вариантах осуществления RNP из варианта CasX и гРНК демонстрирует по меньшей мере 2-кратную, или по меньшей мере 3-кратную, или по меньшей мере 4-кратную, или по меньшей мере 5-кратную, или по меньшей мере 10-кратную скорость расщепления по сравнению с RNP из референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК из таблицы 2. В вышеуказанном варианте осуществления улучшенный уровень компетентности может быть продемонстрирован в анализе *in vitro*, таком как описанный в разделе «Примеры».

[0182] В некоторых вариантах осуществления вариант белка CasX имеет повышенную загрузку целевой цепи при двухцепочечном расщеплении по сравнению с референсным CasX. Варианты с повышенной активностью загрузки целевой цепи могут быть созданы, например, посредством аминокислотных изменений в домене TLS. Без ограничений, накладываемых какой-либо теорией, аминокислотные изменения в домене TSL могут приводить к получению вариантов CasX с улучшенной каталитической активностью. Альтернативно или дополнительно аминокислотные изменения вокруг канала связывания для дуплекса РНК:ДНК также могут улучшать каталитическую активность варианта белка CasX. В некоторых вариантах осуществления вариант белка CasX имеет повышенную побочную активность расщепления по сравнению с референсным белком CasX. В настоящем документе «побочная активность расщепления» относится к

дополнительному нецелевому расщеплению нуклеиновых кислот после распознавания и расщепления последовательности целевой нуклеиновой кислоты. В некоторых вариантах осуществления вариант белка CasX имеет сниженную побочную активность расщепления по сравнению с референсным белком CasX.

[0183] В некоторых вариантах осуществления, например, в тех вариантах осуществления, которые охватывают области применения, где расщепление последовательности целевой нуклеиновой кислоты не является желательным результатом, улучшение каталитической активности варианта белка CasX включает изменение, уменьшение или устранение каталитической активности варианта белка CasX. В некоторых вариантах осуществления рибонуклеопротеиновый комплекс, содержащий вариант белка dCasX, связывается с последовательностью целевой нуклеиновой кислоты и не расщепляет целевую нуклеиновую кислоту.

[0184] В некоторых вариантах осуществления рибонуклеопротеиновый комплекс CasX, содержащий вариант белка CasX, связывает целевую ДНК, но создает одноцепочечный разрыв в целевой ДНК. В некоторых вариантах осуществления, в частности в тех вариантах осуществления, где белок CasX представляет собой нуклеазу, вариант белка CasX имеет сниженную загрузку целевой цепи для одноцепочечного никирования. Варианты со сниженной загрузкой целевой цепи могут быть созданы, например, посредством аминокислотных замен в домене TSL.

[0185] Примеры способов характеристики каталитической активности белков CasX могут включать, без ограничений, анализы расщепления *in vitro*, включая описанные в разделе «Примеры». В некоторых вариантах осуществления кинетику расщепления цепи можно оценивать посредством электрофореза продуктов ДНК на агарозных гелях.

j. Слитные белки CasX

[0186] В некоторых вариантах осуществления в описании предложены белки CasX, содержащие гетерологичный белок, слитный с CasX. В некоторых случаях CasX представляет собой референсный белок CasX. В других случаях CasX представляет собой вариант CasX по любому из вариантов осуществления, описанных в настоящем документе.

[0187] В некоторых вариантах осуществления вариант белка CasX содержит любую из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154 из таблицы 4, слитую с одним или более белками или их доменами, которые имеют другую интересующую активность, что приводит к получению слитного белка. В некоторых вариантах осуществления вариант белка CasX содержит любую из SEQ ID NO: 36-99, 101-148, 26908-27154, слитую с одним или более белками или их доменами. В некоторых вариантах осуществления вариант белка CasX содержит любую из SEQ ID NO: 132-148, 26908-2715, слитую с одним или более белками или их доменами. Например, в некоторых вариантах осуществления вариант белка CasX слит с белком (или его доменом), который ингибирует транскрипцию, изменяет последовательность целевой нуклеиновой кислоты или модифицирует полипептид, ассоциированный с нуклеиновой кислотой (например, модификация гистона).

[0188] В некоторых вариантах осуществления гетерологичный полипептид (или

гетерологичную аминокислоту, например, остаток цистеина или неприродная аминокислота) можно вставить в одно или более положений в белке CasX с созданием слитного белка CasX. В других вариантах осуществления остаток цистеина может быть вставлен в одно или более положений в белке CasX с последующей конъюгацией гетерологичного полипептида, описанного ниже. В некоторых альтернативных вариантах осуществления гетерологичный полипептид или гетерологичную аминокислоту можно добавлять на N- или C-конце белка CasX. В других вариантах осуществления гетерологичный полипептид или гетерологичную аминокислоту можно вставлять внутрь последовательности белка CasX.

[0189] В некоторых вариантах осуществления слитный вариант белка CasX сохраняет РНК-направляемую специфичную к последовательности активность по связыванию с целевой нуклеиновой кислотой и расщеплению. В некоторых случаях слитный вариант белка CasX имеет (сохраняет) 50% или более активности (например, по расщеплению и/или связыванию) относительно соответствующего варианта белка CasX, который не имеет вставки гетерологичного белка. В некоторых случаях слитный вариант белка CasX сохраняет по меньшей мере примерно 60%, или по меньшей мере примерно 70% или более, по меньшей мере примерно 80%, или по меньшей мере примерно 90%, или по меньшей мере примерно 92%, или по меньшей мере примерно 95%, или по меньшей мере примерно 98%, или по меньшей мере примерно 100% активности (например, активности по расщеплению и/или связыванию) относительно соответствующего белка CasX, который не имеет вставки гетерологичного белка.

[0190] В некоторых случаях слитный вариант белка CasX сохраняет (имеет) активность по связыванию с целевой нуклеиновой кислотой относительно активности белка CasX без вставленной гетерологичной аминокислоты или гетерологичного полипептида. В некоторых случаях слитный вариант белка CasX сохраняет по меньшей мере примерно 60%, или по меньшей мере примерно 70% или более, по меньшей мере примерно 80%, или по меньшей мере примерно 90%, или по меньшей мере примерно 92%, или по меньшей мере примерно 95%, или по меньшей мере примерно 98%, или по меньшей мере примерно 100% активности по связыванию относительно соответствующего белка CasX, который не имеет вставки гетерологичного белка.

[0191] В некоторых случаях слитный вариант белка CasX сохраняет (имеет) активность по связыванию и/или расщеплению целевой нуклеиновой кислоты относительно активности исходного белка CasX без вставленной гетерологичной аминокислоты или гетерологичного полипептида. Например, в некоторых случаях слитный вариант белка CasX имеет (сохраняет) 50% или более активности связывания и/или расщепления относительно соответствующего исходного белка CasX (белка CasX, который не имеет вставки). Например, в некоторых случаях слитный вариант белка CasX имеет (сохраняет) 60% или более (70% или более, 80% или более, 90% или более, 92% или более, 95% или более, 98% или более, или 100%) активности по связыванию и/или расщеплению относительно соответствующего исходного белка CasX (белка CasX, который не имеет

вставки). Способы измерения активности по расщеплению и/или связыванию белка CasX и/или слитного белка CasX известны специалисту в данной области, и можно применять любой удобный способ.

[0192] Для включения в референсный слитный белок CasX или слитный вариант белка CasX описания подходят разнообразные гетерологичные полипептиды. В некоторых случаях партнер по слиянию может модулировать транскрипцию (например, ингибировать транскрипцию, увеличивать транскрипцию) целевой ДНК. Например, в некоторых случаях партнер по слиянию представляет собой белок (или домен из белка), который ингибирует транскрипцию (например, репрессор транскрипции, белок, который функционирует посредством рекрутирования белков-ингибиторов транскрипции, модификации целевой ДНК, такой как метилирование, рекрутирования модификатора ДНК, модуляции гистонов, связанных с целевой ДНК, рекрутирования модификатора гистонов, например, изменяющего ацетилирование и/или метилирование гистонов, и т. п.).

[0193] В некоторых случаях партнер по слиянию представляет собой белок (или домен из белка), который усиливает транскрипцию (например, активатор транскрипции, белок, который функционирует посредством рекрутирования белков-активаторов транскрипции, модификации целевой ДНК, такой как деметилирование, рекрутирования модификатора ДНК, модуляции гистонов, связанных с целевой ДНК, рекрутирования модификатора гистонов, например, изменяющего ацетилирование и/или метилирование гистонов, и т. п.). В некоторых случаях партнер по слиянию обладает ферментативной активностью, которая изменяет последовательность целевой нуклеиновой кислоты; например, нуклеазной активностью, метилтрансферазной активностью, деметилазной активностью, активностью репарации ДНК, активностью повреждения ДНК, активностью деаминарования, дисмутазной активностью, активностью алкилирования, активностью депуринизации, оксидазной активностью, активностью образования пиримидинового димера, интегразной активностью, транспозазной активностью, лигазной активностью, геликазной активностью, фотолиазной активностью или гликозилазной активностью. В некоторых вариантах осуществления вариант CasX содержит любую из SEQ ID NO: 36-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 132-148 или 26908-27154, и полипептид с активностью метилтрансферазы, активностью деметилазы, активностью ацетилтрансферазы, активностью деацетилазы, активностью киназы, активностью фосфатазы, активностью убиквитинлигазы, деубиквитирующей активностью, аденилирующей активностью, деаденилирующей активностью, SUMOилирующей активностью, деSUMOилирующей активностью, рибозилирующей активностью, дерибозилирующей активностью, миристоилирующей активностью или демиристоилирующей активностью.

[0194] В некоторых вариантах осуществления вариант CasX содержит любую из SEQ ID NO: 36-99, 101-148 и 26908-27154, или любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 132-148 или 26908-27154, и партнера по слиянию, обладающего ферментативной активностью, которая модифицирует полипептид

(например, гистон), ассоциированный с целевой нуклеиновой кислотой (например, полипептид с активностью метилтрансферазы, активностью деметилазы, активностью ацетилтрансферазы, активностью деацетилазы, активностью киназы, активностью фосфатазы, активностью убиквитинлигазы, деубиквитирующей активностью, аденилирующей активностью, деаденилирующей активностью, SUMOилирующей активностью, деSUMOилирующей активностью, рибозилирующей активностью, дерибозилирующей активностью, миристоилирующей активностью или демиристоилирующей активностью). Примеры белков (или их фрагментов), которые можно применять в качестве партнера по слиянию для усиления транскрипции, включают, без ограничений: активаторы транскрипции, такие как VP16, VP64, VP48, VP160, субдомен р65 (например, из NFκB) и активационный домен EDLL и/или активационный домен TAL (например, для активности в растениях); гистоновые лизинметилтрансферазы, такие как SET1A, SET1B, MLL1-5, ASH1, SYMD2, NSD1 и т. п.; гистоновые лизиндеметилазы, такие как JHDM2a/b, UTX, JMJD3 и т. п.; гистоновые ацетилтрансферазы, такие как GCN5, PCAF, CBP, p300, TAF1, TIP60/PLIP, MOZ/MYST3, MORF/MYST4, SRC1, ACTR, P160, CLOCK и т. п.; и ДНК-деметилазы, такие как Ten-Eleven Translocation (TET) dioxygenase 1 (TET1CD), TET1, DME, DML1, DML2, ROS1 и т. п.

[0195] Примеры белков (или их фрагментов), которые можно применять в качестве партнера по слиянию для снижения транскрипции, включают, без ограничений: репрессоры транскрипции, такие как Kruppel-ассоциированный бокс (KRAB или SKD); репрессивный домен KOX1; домен взаимодействия mSIN3 (SID) из Mad; репрессорный домен ERF (ERD), репрессорный домен SRDX (например, для репрессии в растениях) и т. п.; гистоновые лизинметилтрансферазы, такие как Pr-SET7/8, SUV4-20H1, RIZ1 и т. п.; гистоновые лизиндеметилазы, такие как JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, JARID1A/RBP2, JARID1B/PLU-1, JARID1C/SMCX, JARID1D/SMCY и т. п.; гистоновые лизиндеацетилазы, такие как HDAC1, HDAC2, HDAC3, HDAC8, HDAC4, HDAC5, HDAC7, HDAC9, SIRT1, SIRT2, HDAC11 и т. п.; ДНК-метидазы, такие как HhaI ДНК m5c-метилтрансфераза (M.HhaI), (M.HhaI), ДНК метилтрансфераза 1 (DNMT1), ДНК метилтрансфераза 3a (DNMT3a), ДНК метилтрансфераза 3b (DNMT3b), MET1, DRM3 (растения), ZMET2, CMT1, CMT2 (растения) и т. п.; и периферические рекрутирующие элементы, такие как lamin A, lamin B и т. п.

[0196] В некоторых случаях партнер по слиянию с вариантом CasX обладает ферментативной активностью, которая модифицирует последовательность целевой нуклеиновой кислоты (например, оцРНК, дцРНК, оцДНК, дцДНК). Примеры ферментативной активности, которую может обеспечивать партнер по слиянию, включают, без ограничений: нуклеазную активность, например, обеспечиваемую рестрикционным ферментом (например, нуклеазой FokI), метилтрансферазную активность, например, обеспечиваемую метилтрансферазой (например, HhaI ДНК m5c-метилтрансферазой (M.HhaI), ДНК метилтрансферазой 1 (DNMT1), ДНК метилтрансферазой 3a (DNMT3a), ДНК метилтрансферазой 3b (DNMT3b), MET1, DRM3 (растения), ZMET2, CMT1, CMT2

(растения) и т. п.); деметилазную активность, например, обеспечиваемую деметилазой (например, Ten-Eleven Translocation (TET) dioxygenase 1 (TET 1 CD), TET1, DME, DML1, DML2, ROS1 и т. п.), активность репарации ДНК, активность повреждения ДНК, деаминирующую активность, например, обеспечиваемую деаминазой (например, ферментом цитозиндеаминазой, например, белком APOBEC, таким как APOBEC1 крысы), дисмутазную активность, алкилирующую активность, активность депуринизации, оксидазную активность, активность образования пиримидинового димера, интегразную активность, например, обеспечиваемую интегразой и/или резольвазой (например, Gin-инвертазой, такой как гиперактивный мутант Gin-инвертазы, GinH106Y; интегразой вируса иммунодефицита человека типа 1 (IN); Tn3-резольвазой; и т. п.), транспозазную активность, рекомбиназную активность, такую как обеспечиваемая рекомбиназой (например, каталитическим доменом рекомбиназы Gin), полимеразную активность, лигазную активность, геликазную активность, фотолиазную активность и гликозилазную активность).

[0197] В некоторых случаях вариант белка CasX настоящего описания слит с полипептидом, выбранным из домена для усиления транскрипции (например, домена VP16, домена VP64), домена для ослабления транскрипции (например, домена KRAB, например, из белка Kox1), основного каталитического домена гистоновой ацетилтрансферазы (например, гистоновой ацетилтрансферазы p300), белка/домена, который обеспечивает обнаруживаемый сигнал (например, флуоресцентного белка, такого как GFP), нуклеазного домена (например, нуклеазы FokI) или редактирующего основания фермента (например, цитидиндеаминазы, такой как APOBEC1).

[0198] В некоторых вариантах осуществления вариант CasX содержит любую из SEQ ID NO: 36-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154 или любую из SEQ ID NO: 132-148 или 26908-27154, и партнера по слиянию, имеющего ферментативную активность, которая модифицирует белок, ассоциированный с целевой нуклеиновой кислотой (например, оцРНК, дцРНК, оцДНК, дцДНК) (например, гистон, РНК-связывающий белок, ДНК-связывающий белок и т. п.). Примеры ферментативной активности (которая модифицируют белок, ассоциированный с целевой нуклеиновой кислотой), которую может обеспечивать партнер по слиянию, включают, без ограничений: метилтрансферазную активность, например, обеспечиваемую гистоновой метилтрансферазой (HMT) (например, гомологом 1 супрессора мозаичности 3-9 (SUV39H1, также известен как KMT1A), эухроматической гистоновой лизинметилтрансферазой 2 (G9A, также известна как KMT1C и EHMT2), SUV39H2, ESET/SETDB1 и т. п., SET1A, SET1B, MLL1-5, ASH1, SYMD2, NSD1, DOT1L, Pr-SET7/8, SUV4-20H1, EZH2, RIZ1), деметилазную активность, например, обеспечиваемую гистоновой деметилазой (например, лизиндеметилазой 1A (KDM1A также известна как LSD1), JHDM2a/b, JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, JARID1A/RBP2, JARID1B/PLU-1, JARID1C/SMCX, JARID1D/SMCY, UTX, JMJD3 и т. п.), ацетилтрансферазную активность, например, обеспечиваемую гистоновой ацетилазой-

трансферазой (например, каталитическое ядро/фрагмент ацетилтрансферазы p300, GCN5, PCAF, CBP, TAF1, TIP60/PLIP, MOZ/MYST3, MORF/MYST4, HB01/MYST2, HMOF/MYST1, SRC1, ACTR, P160, CLOCK и т. п.), деацетилазную активность, например, обеспечиваемую гистоновой деацетилазой (например, HDAC1, HDAC2, HDAC3, HDAC8, HDAC4, HDAC5, HDAC7, HDAC9, SIRT1, SIRT2, HDAC11 и т. п.), киназную активность, фосфатазную активность, активность убиквитинлигазы, деубиквитинирующую активность, аденилирующую активность, деаденилирующую активность, SUMOилирующую активность, деSUMOилирующую активность, рибозилирующую активность, дерибозилирующую активность, миристоилирующую активность или демиристоилирующую активность.

[0199] Дополнительные примеры приемлемых партнеров по слиянию для варианта CasX представляют собой (i) домен дестабилизации дигидрофолатредуктазы (DHFR) (например, для получения химически контролируемого направляемого РНК субъекта полипептида или условно активного РНК-направляемого полипептида) и (ii) пептида транзита к хлоропластам. В некоторых вариантах осуществления вариант CasX содержит любую из SEQ ID NO: 36-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 132-148 или 26908-27154, или последовательность из таблицы 4, и пептид транзита к хлоропластам, включая, без ограничений:

MASMISSSAVTTVSRASRGQSAAMAPFGGLKSMTGFPVRKVNTDITSITSNGGR
VKCMQVWPPIGKKKFETLSYLPPLTRDSRA (SEQ ID NO: 154);

MASMISSSAVTTVSRASRGQSAAMAPFGGLKSMTGFPVRKVNTDITSITSNGGRVKS
(SEQ ID NO: 155);

MASSMLSSATMVASPAQATMVAPFNGLKSSAAFPATRKANNDITSITSNGGRVNCMQV
WPPIEKKKFETLSYLPDLTDSGGRVNC (SEQ ID NO: 156);

MAQVSRICNGVQNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWGLKKSGMTLIG
SELRPLKVMSSVSTAC (SEQ ID NO: 157);

MAQVSRICNGVWNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWGLKKSGMTLIG
SELRPLKVMSSVSTAC (SEQ ID NO: 158);

MAQINNMAQGIQTLNPNSNFHKPQVPKSSSFLVFGSKKLKNSANSMLVLKKDSIFMQLF
CSFRISASVATAC (SEQ ID NO: 159);

MAALVTSQLATSGTVLSVTDRFRRPGFQGLRPRNPADAALGMRTVGASAAPKQSRKPH
RFDRRCLSMVV (SEQ ID NO: 160);

MAALTTSQLATSATGFGIADRSAPSSLLRHGFQGLKPRSPAGGDATSLSVTTSARATPKQ
QRSVQRGRSRRFPSVVVC (SEQ ID NO: 161);

MASSVLSSAAVATRSNVAQANMVAFTGLKSAASFPVSRKQNLDTIASNGGRVQC
(SEQ ID NO: 162);

MESLAATSVFAPSRVAVPAARALVRAGTVVPTRRTSSTSGTSGVKCSAAVTPQASPVIS
RSAAAA (SEQ ID NO: 163) и

MGAAATSMQSLKFSNRLVPPSRRLSPVPNNVTCNNLPKSAAPVRTVKCCASSWNSTING

AAATTNGASAASS (SEQ ID NO: 164).

[0200] В некоторых случаях вариант белка CasX настоящего описания может включать пептид избегания эндосом. В некоторых случаях полипептид избегания эндосом содержит аминокислотную последовательность GLFXALLXLLXSLWXLLLXA (SEQ ID NO: 165), причем каждый X независимо выбран из лизина, гистидина и аргинина. В некоторых случаях полипептид избегания эндосом содержит аминокислотную последовательность GLFHALLHLLHSLWHLLLHA (SEQ ID NO: 166) или НННННННННН (SEQ ID NO: 167).

[0201] Не имеющие ограничительного характера примеры партнеров по слиянию для применения с вариантами белков CasX при нацеливании на последовательности целевых нуклеиновых кислот оцРНК включают (без ограничений): факторы сплайсинга (например, домены RS); компоненты трансляции белка (например, факторы инициации трансляции, элонгации и/или высвобождения; например, eIF4G); РНК-метилазы; ферменты редактирования РНК (например, РНК-деаминазы, например, аденозиндеаминаза, действующая на РНК (ADAR), в том числе ферменты редактирования А в I и/или С в U); геликазы; РНК-связывающие белки и т. п. Следует понимать, что гетерологичный полипептид может включать весь белок или в некоторых случаях может включать фрагмент белка (например, функциональный домен).

[0202] В некоторых вариантах осуществления вариант CasX содержит любую из SEQ ID NO: 36-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 59, 72-99, 101-148 или 26908-27154, или любую из SEQ ID NO: 132-148 или 26908-27154 и партнера по слиянию - любого домена, способного взаимодействовать с оцРНК (к которому для целей настоящего описания относятся внутримолекулярные и/или межмолекулярные вторичные структуры, например, двухцепочечные дуплексы РНК, такие как шпильки, стеблевые петли и т. д.), как временно, так и необратимо, прямо или косвенно, включая, без ограничений, эффекторный домен, выбранный из группы, включающей: эндонуклеазы (например, РНКазы III, домен CRR22 DYW, домены Dicer и PIN (N-конец PilT) из белков, таких как SMG5 и SMG6); белки и белковые домены, ответственные за стимуляцию расщепления РНК (например, CPSF, CstF, CFIm и CFIIIm); экзонуклеазы (например, XRN-1 или экзонуклеаза T); деаденилазы (например, HNT3); белки и белковые домены, ответственные за нонсенс-опосредованный распад РНК (например, UPF1, UPF2, UPF3, UPF3b, RNP SI, Y14, DEK, REF2 и SRm160); белки и белковые домены, ответственные за стабилизацию РНК (например, PABP); белки и белковые домены, ответственные за подавление трансляции (например, Ago2 и Ago4); белки и белковые домены, ответственные за стимуляцию трансляции (например, Staufen); белки и белковые домены, ответственные за (например, способные к) модуляцию трансляции (например, факторы трансляции, такие как факторы инициации, факторы элонгации, факторы высвобождения и т. д., например, eIF4G); белки и белковые домены, ответственные за полиаденилирование РНК (например, PAP1, GLD-2 и Star-PAP); белки и белковые домены, ответственные за полиуридилирование РНК (например, CI D1 и концевая уридилаттрансфераза); белки и

белковые домены, ответственные за локализацию РНК (например, IMP1, ZBP1, She2p, She3p и Bicaudal-D); белки и белковые домены, ответственные за задержку РНК в ядре (например, Rrp6); белки и белковые домены, ответственные за ядерный экспорт РНК (например, TAP, NXF1, THO, TREX, REF и Aly); белки и белковые домены, ответственные за подавление сплайсинга РНК (например, PTB, Sam68 и hnRNP A1); белки и белковые домены, ответственные за стимуляцию сплайсинга РНК (например, домены, богатые серином/аргинином (SR)); белки и белковые домены, ответственные за снижение эффективности транскрипции (например, FUS (TLS)); и белки и белковые домены, ответственные за стимуляцию транскрипции (например, CDK7 и Tat). Альтернативно эффекторный домен может быть выбран из группы, содержащей эндонуклеазы; белки и белковые домены, способные стимулировать расщепление РНК; экзонуклеазы; деаденилазы; белки и белковые домены, имеющие нонсенс-опосредованную активность по распаду РНК; белки и белковые домены, способные стабилизировать РНК; белки и белковые домены, способные подавлять трансляцию; белки и белковые домены, способные стимулировать трансляцию; белки и белковые домены, способные модулировать трансляцию (например, факторы трансляции, такие как факторы инициации, факторы элонгации, факторы высвобождения и т. д., например, eIF4G); белки и белковые домены, способные к полиаденилированию РНК; белки и белковые домены, способные к полиуридилированию РНК; белки и белковые домены, обладающие активностью по локализации РНК; белки и белковые домены, способные к задерживанию РНК в ядре; белки и белковые домены, обладающие активностью ядерного экспорта РНК; белки и белковые домены, способные подавлять сплайсинг РНК; белки и белковые домены, способные стимулировать сплайсинг РНК; белки и белковые домены, способные снижать эффективность транскрипции; и белки и белковые домены, способные стимулировать транскрипцию. Другой приемлемый гетерологичный полипептид представляет собой РНК-связывающий домен PUF, который более подробно описан в WO2012068627, полностью включенной в настоящий документ путем ссылки.

[0203] Некоторые факторы сплайсинга РНК, которые можно применять (полностью или в виде его фрагментов) в качестве партнера по слиянию с вариантом CasX, имеют модульную организацию с отдельными специфичными для последовательности РНК-связывающими модулями и эффекторными доменами сплайсинга. Например, члены семейства богатых серином/аргинином (SR) белков содержат N-концевые мотивы распознавания РНК (RRM), которые связываются с экзонными энхансерами сплайсинга (ESE) в пре-мРНК и С-концевыми доменами RS, которые способствуют включению экзона. В качестве другого примера, гЯRNP-белок hnRNP A1 связывается с экзонными сайленсерами сплайсинга (ESS) своими доменами RRM и ингибирует включение экзона с помощью С-концевого богатого глицином домена. Некоторые факторы сплайсинга могут регулировать альтернативное использование сайта сплайсинга (ss) посредством связывания с регуляторными последовательностями между двумя альтернативными сайтами. Например, ASF/SF2 может распознавать ESE и способствовать применению

проксимальных к интрону сайтов, тогда как hRNP A1 может связываться с ESS и сдвигать сплайсинг к применению дистальных к интрону сайтов. Одним из вариантов применения таких факторов является создание ESF, которые модулируют альтернативный сплайсинг эндогенных генов, в частности генов, связанных с заболеванием. Например, пре-mРНК Bcl-x дает две сплайсинговые изоформы с двумя альтернативными 5'-сайтами сплайсинга для кодирования белков с противоположными функциями. Длинная сплайсинговая изоформа Bcl-xL представляет собой мощный ингибитор апоптоза, экспрессируемый в долгоживущих постмитотических клетках, и она имеет повышенную экспрессию во многих раковых клетках, защищая клетки против апоптотических сигналов. Короткая изоформа Bcl-xS представляет собой проапоптотическую изоформу и экспрессируется с высокими уровнями в клетках с высокой скоростью оборота (например, в развивающихся лимфоцитах). Соотношение двух сплайсинговых изоформ Bcl-x регулируется множеством cis-элементов, которые расположены либо в основной области экзона, либо в удлинительной области экзона (т. е. между двумя альтернативными 5'-сайтами сплайсинга). См. дополнительные примеры в WO2010075303, полностью включенной в настоящий документ путем ссылки.

[0204] Дополнительные приемлемые партнеры по слиянию для применения с вариантом CasX включают, без ограничений, белки (или их фрагменты), которые представляют собой граничные элементы (например, CTCF), белки и их фрагменты, которые обеспечивают рекрутирование периферии (например, lamin A, lamin B и т. д.) и стыковочные элементы белков (например, FKBP/FRB, Pill/Abyl и т. д.).

[0205] В некоторых случаях гетерологичный полипептид (партнер по слиянию) для применения с вариантом CasX обеспечивает субклеточную локализацию, т. е. гетерологичный полипептид содержит последовательность субклеточной локализации (например, сигнал ядерной локализации (NLS) для нацеливания на ядро, последовательность для удержания слитного белка вне ядра, например последовательность экспорта из ядра (NES), последовательность для удержания слитного белка в цитоплазме, сигнал митохондриальной локализации для нацеливания на митохондрию, сигнал локализации в хлоропластах для нацеливания на хлоропласт, сигнал удерживания в эндоплазматическом ретикулуме (ER) и т. п.). В некоторых вариантах осуществления указанный РНК-направляемый полипептид или условно активный РНК-направляемый полипептид и/или рассматриваемый слитый белок CasX не включает NLS, так что белок не нацелен на ядро (что может быть предпочтительно, например, если последовательность целевой нуклеиновой кислоты представляет собой РНК, которая находится в цитозоле). В некоторых вариантах осуществления партнер по слиянию может обеспечивать метку (т. е. гетерологичный полипептид представляет собой детектируемую метку) для простоты отслеживания и/или очистки (например, флуоресцентный белок, например, зеленый флуоресцентный белок (GFP), желтый флуоресцентный белок (YFP), красный флуоресцентный белок (RFP), голубой флуоресцентный белок (CFP), mCherry, tdTomato и т. п.; гистидиновую метку, например метку 6XHis; гемагглютининовую (HA) метку; метку

FLAG; метку Мус и т. п.).

[0206] В некоторых случаях не имеющие ограничительного характера примеры NLS, приемлемых для применения с вариантом CasX, включают последовательности, на по меньшей мере примерно 80%, по меньшей мере примерно 90%, или по меньшей мере примерно 95% идентичные или полностью идентичные последовательностям, полученным из: NLS большого Т-антигена вируса SV40, имеющего аминокислотную последовательность PKKKRKV (SEQ ID NO: 168); NLS из нуклеоплазмينا (например, нуклеоплазмينا двухкомпонентного NLS с последовательностью KRPAATKKAGQAKKKK (SEQ ID NO: 169); NLS, с-мус, имеющий аминокислотную последовательность PAAKRVKLD (SEQ ID NO: 170) или RQRRNELKRSP (SEQ ID NO: 171); NLS hRNPAl M9, имеющий последовательность NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 172); последовательность RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 173) домена IBB из импортина-альфа; последовательность VSRKRPRP (SEQ ID NO: 174) и PPKKARED (SEQ ID NO: 175) из Т-белка миомы; последовательность PQPKKKPL (SEQ ID NO: 176) из человеческого p53; последовательность SALIKKKKKMAP (SEQ ID NO: 177) из мышинового c-abl IV; последовательности DRLRR (SEQ ID NO: 178) и PKQKKRK (SEQ ID NO: 179) из вируса гриппа NS1; последовательность RKLKKKIKKL (SEQ ID NO: 180) из дельта-антигена вируса гепатита; последовательность REKKKFLKRR (SEQ ID NO: 181) из мышинового белка Mxl; последовательность KRKGDEV DGVDEVAKKKSKK (SEQ ID NO: 182) из поли(АДФ-рибоза) полимеразы человека; последовательность RKCLQAGMNLEARKTKK (SEQ ID NO: 183) из рецепторов стероидных гормонов (человеческих), глюкокортикоидных; последовательность PRPRKIPR (SEQ ID NO: 184) из Р-белка вируса болезни Борна (BDV-P1); последовательность PPRKKRTVV (SEQ ID NO: 185) из неструктурного белка вируса гепатита С (HCV-NS5A); последовательность NLSKKKKRKREK (SEQ ID NO: 186) из LEF1; последовательность RRPSRPFRKP (SEQ ID NO: 187) из ORF57 обезьян; последовательность KRPRSPSS (SEQ ID NO: 188) из EBV LANA; последовательность KRGINDRNFWRGENERKTR (SEQ ID NO: 189) из белка гриппа А; последовательность PRPPKMARYDN (SEQ ID NO: 190) из РНК-хеликазы А человека (RHA); последовательность KRSFSKAF (SEQ ID NO: 191) из ядрышковой РНК-геликазы II; последовательность KLKIKRPVK (SEQ ID NO: 192) из TUS-белка; последовательность PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 193), ассоциированная с импортином-альфа; последовательность PKTRRRPRRSQRKRPT (SEQ ID NO: 26792) из Rex-белка HTLV-1; последовательность SRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 194) из белка EGL-13 *Caenorhabditis elegans* и последовательности KTRRRPRRSQRKRPT (SEQ ID NO: 195), RRKKRRPRRKRR (SEQ ID NO: 196), PKKKS RKP KKS R K (SEQ ID NO: 197), HKKKHPDASVNFSEFSK (SEQ ID NO: 198), QRPGPYDRPQRPGPYDRP (SEQ ID NO: 199), LSPSLSPLLSPSLSP (SEQ ID NO: 200), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 201), PKRGRGRPKRGRGR (SEQ ID NO: 202), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 203),

PKKKRKVPPPPKKKKRKV (SEQ ID NO: 204), PAKRARRGYKC (SEQ ID NO: 27199), KLGPRKATGRW (SEQ ID NO: 27200), PRRKREE (SEQ ID NO: 27201), PYRGRKE (SEQ ID NO: 27202), PLRKPRR (SEQ ID NO: 27203), PLRKPRRGSPLRKRPRR (SEQ ID NO: 27204), PAAKRVKLDGGKRTADGSEFESPKKKKRKV (SEQ ID NO: 27205), PAAKRVKLDGGKRTADGSEFESPKKKKRKVGIHGVPAA (SEQ ID NO: 27206), PAAKRVKLDGGKRTADGSEFESPKKKKRKVAEAAAKEAAAKEAAKA (SEQ ID NO: 207), PAAKRVKLDGGKRTADGSEFESPKKKKRKVPG (SEQ ID NO: 27208), KRKGSPERGERKRHW (SEQ ID NO: 27209), KRTADSQHSTPPKTKRKVEFEPKKKKRKV (SEQ ID NO: 27210) и PKKKRKVGGSKRTADSQHSTPPKTKRKVEFEPKKKKRKV (SEQ ID NO: 27211). В некоторых вариантах осуществления один или более NLS связаны с CRISPR-белком или со смежным NLS линкерным пептидом, причем линкерный пептид выбран из группы, состоящей из RS, (G)_n (SEQ ID NO: 27212), (GS)_n (SEQ ID NO: 27213), (GSGGS)_n (SEQ ID NO: 214), (GGSGGS)_n (SEQ ID NO: 215), (GGGS)_n (SEQ ID NO: 216), GGSG (SEQ ID NO: 217), GGSGG (SEQ ID NO: 218), GSGSG (SEQ ID NO: 219), GSGGG (SEQ ID NO: 220), GGGSG (SEQ ID NO: 221), GSSSG (SEQ ID NO: 222), GPGP (SEQ ID NO: 223), GGP, PPP, PPAPPA (SEQ ID NO: 224), PPPG (SEQ ID NO: 27214), PPPGPPP (SEQ ID NO: 225), PPP(GGGS)_n (SEQ ID NO: 27215), (GGGS)_nPPP (SEQ ID NO: 27216), AEAAAKEAAAKEAAKA (SEQ ID NO: 27217) и TPPKTKRKVEFE (SEQ ID NO: 27218), где n составляет от 1 до 5. Как правило, NLS (или множество NLS) обладают достаточной силой, чтобы обеспечивать накопление слитного варианта белка CasX в ядре эукариотической клетки. Обнаружение накопления в ядре может осуществляться любой приемлемой методикой. Например, обнаруживаемый маркер может быть слит со слитным вариантом белка CasX, так что местоположение внутри клетки может быть визуализировано. Также можно выделить клеточные ядра из клеток и проанализировать их содержание любым приемлемым способом обнаружения белка, например, методом иммуногистохимии, вестерн-блоттинга или анализом ферментативной активности. Накопление в ядре также может быть определено опосредованно.

[0207] Как правило, NLS (или множество NLS) обладает достаточной силой, чтобы обеспечить накопление экспрессированного слитного варианта белка CasX в ядре эукариотической клетки. Обнаружение накопления в ядре может осуществляться любой приемлемой методикой. Например, обнаруживаемый маркер может быть слит со слитным вариантом белка CasX, так что местоположение внутри клетки может быть визуализировано. Также можно выделить клеточные ядра из клеток и проанализировать их содержание любым приемлемым способом обнаружения белка, например, методом иммуногистохимии, вестерн-блоттинга или анализом ферментативной активности. Накопление в ядре также может быть определено опосредованно.

[0208] В некоторых случаях слитный вариант белка CasX включает «домен трансдукции белка», или PTD (также известен как CPP - проникающий в клетку пептид), который относится к белку, полинуклеотиду, углеводу или органическому или неорганическому соединению, которое облегчает прохождение через липидный бислой,

мицеллу, клеточную мембрану, мембрану органеллы или везикулярную мембрану. PTD, присоединенный к другой молекуле, которая может находиться в диапазоне от малой полярной молекулы до крупной макромолекулы и/или наночастицы, способствует прохождению молекулы через мембрану, например, из внеклеточного пространства во внутриклеточное пространство или из цитозоля внутрь органеллы. В некоторых вариантах осуществления PTD ковалентно связан с аминоконцом слитного варианта белка CasX. В некоторых вариантах осуществления PTD ковалентно связан с карбоксильным концом слитного варианта белка CasX. В некоторых случаях PTD вставлен внутрь последовательности слитного варианта белка CasX в приемлемом для вставки месте. В некоторых случаях слитный вариант белка CasX включает (конъюгирован, слит с) один или более PTD (например, два или более, три или более, четыре или более PTD). В некоторых случаях PTD включает один или более сигналов ядерной локализации (NLS). Примеры PTD включают, без ограничений, домен трансдукции пептида HIV TAT, содержащий YGRKKRRQRRR (SEQ ID NO: 205), RKKRRQRR (SEQ ID NO: 206); YARAAARQARA (SEQ ID NO: 207); THRLPRRRRRR (SEQ ID NO: 208) и GGRRARRRRRRR (SEQ ID NO: 209); полиаргининовую последовательность, содержащую ряд остатков аргинина, достаточный для прямого входа в клетку (например, 3, 4, 5, 6, 7, 8, 9, 10 или 10-50 аргининовых остатков (SEQ ID NO: 26793); домен VP22 (Zender et al. (2002) *Cancer Gene Ther.* 9(6):489-96); домен трансдукции белка *Drosophila Antennapedia* (Noguchi et al. (2003) *Diabetes* 52(7): 1732-1737); усеченный пептид кальцитонина человека (Trehin et al. (2004) *Pharm. Research* 21:1248-1256); полилизин (Wender et al. (2000) *Proc. Natl. Acad. Sci. USA* 97: 13003-13008); RRQRRTSKLMKR (SEQ ID NO: 210); транспортан GWTLNSAGYLLGKINLKALAALAKKIL (SEQ ID NO: 211); KALAWREAKLAKALAKALAKHLAKALAKALKCEA (SEQ ID NO: 212) и RQIKIWFQNRRMKWKK (SEQ ID NO: 213). В некоторых вариантах осуществления PTD представляет собой активируемый CPP (ACPP) (Aguilera et al. (2009) *Integr Biol (Camb)* June; 1(5-6): 371-381). ACPP содержат поликатионный CPP (например, Arg9 или «R9»), соединенный посредством расщепляемого линкера с соответствующим полианионом (например, Glu9 или «E9»), что снижает чистый заряд почти до нуля и таким образом ингибирует адгезию и поглощение в клетки. После расщепления линкера полианион высвобождается, локально демаскируя полиаргинин и присущую ему адгезионную способность, таким образом «активируя» ACPP для прохождения через мембрану.

[0209] В некоторых вариантах осуществления слитный вариант белка CasX для применения в системах может включать белок CasX, который связан с вставленной внутрь гетерологичной аминокислотой или гетерологичным полипептидом (гетерологичной аминокислотной последовательностью) посредством линкерного полипептида (например, одного или более линкерных полипептидов). В некоторых вариантах осуществления слитный вариант белка CasX может быть связан на С-конце и/или N-конце с гетерологичным полипептидом (партнером по слиянию) посредством линкерного полипептида (например, одного или более линкерных полипептидов). Линкерный

полипептид может содержать любую из разнообразных аминокислотных последовательностей. Белки могут быть соединены спейсерным пептидом, как правило, гибким, хотя не исключаются другие химические связи. Приемлемые линкеры включают полипептиды длиной от 4 аминокислот до 40 аминокислот или длиной от 4 аминокислот до 25 аминокислот. Эти линкеры по существу получают с применением синтетических кодирующих линкер олигонуклеотидов для соединения белков. Могут применяться пептидные линкеры с определенной степенью гибкости. Линкерные пептиды могут иметь практически любую аминокислотную последовательность, учитывая, что предпочтительные линкеры будут иметь последовательность, которая приводит к получению по существу гибкого пептида. При создании гибкого пептида полезно применять небольшие аминокислоты, такие как глицин и аланин. Создание таких последовательностей является обычным для специалистов в данной области. Различные линкеры доступны в продаже и считаются приемлемыми для применения. Примеры линкерных полипептидов включают пептиды, выбранные из группы, состоящей из RS, (G)_n (SEQ ID NO: 27212), (GS)_n (SEQ ID NO: 27213), (GSGGS)_n (SEQ ID NO: 214), (GGSGGS)_n (SEQ ID NO: 215), (GGGS)_n (SEQ ID NO: 216), где n представляет собой целое число от 1 до 5, GGSG (SEQ ID NO: 217), GGSGG (SEQ ID NO: 218), GSGSG (SEQ ID NO: 219), GSGGG (SEQ ID NO: 220), GGGSG (SEQ ID NO: 221), GSSSG (SEQ ID NO: 222), GPGP (SEQ ID NO: 223), GGP, PPP, PPAPPA (SEQ ID NO: 224), PPPG (SEQ ID NO: 27214), PPPGPPP (SEQ ID NO: 225), PPP(GGGS)_n (SEQ ID NO: 27215), (GGGS)_nPPP (SEQ ID NO: 27216), AEAAAKEAAAKEAAKA (SEQ ID NO: 27217) и TPPKTKRKVEFE (SEQ ID NO: 27218), где n составляет от 1 до 5 и т. п. Специалисту в данной области будет понятно, что конструкция пептида, конъюгированного с любыми описанными выше элементами, может включать линкеры, которые являются полностью или частично гибкими, так что линкер может включать гибкий линкер, а также один или более участков, которые придают ему менее гибкую структуру.

V. Системы и способы модификации генов BCL11A

[0210] Белки CRISPR, гидовые нуклеиновые кислоты и их варианты, предложенные в настоящем документе, можно использовать в различных областях применения, включая терапевтические, диагностические и исследовательские. В некоторых вариантах осуществления для реализации способов редактирования генов описания в настоящем документе предложены программируемые системы CasX:гРНК. Программируемый характер систем, предложенных в настоящем документе, позволяет точно нацеливать их на достижение желаемой модификации в одной или более заданных интересующих областях в целевой нуклеиновой кислоте гена BCL11A. Для модификации последовательности целевой нуклеиновой кислоты в клетке с применением систем, предложенных в настоящем документе, можно использовать различные стратегии и способы. В настоящем документе термин «модификация» включает, без ограничений, расщепление, никирование, редактирование, делецию, нокаут, нокдаун, мутирование, коррекцию, пропуск экзона и т. п. В зависимости от используемых компонентов системы событие редактирования может

представлять собой событие расщепления с последующим введением случайных вставок или делеций (инделов) или других мутаций (например, замены, дублирования или инверсии одного или более нуклеотидов), например, посредством использования неточного соединения нехомологичных концов ДНК (NHEJ), которое может создать, например, мутацию сдвига рамки. Альтернативно событие редактирования может представлять собой событие расщепления с последующей направляемой гомологией репарацией (HDR), независимой от гомологии нацеленной интеграцией (HITI), соединением концов, обусловленных микрогомологией (MMEJ), отжигом одиночных цепей (SSA) или эксцизионной репарацией оснований (BER), что приводит к модификации последовательности целевой нуклеиновой кислоты.

[0211] В некоторых вариантах осуществления способа ген BCL11A, подлежащий модификации, содержит последовательность, соответствующую полинуклеотиду, кодирующему всю или часть последовательности SEQ ID NO: 100, или содержит полинуклеотидную последовательность, которая охватывает весь или часть chr2 60450520-60554467 (GRCh38/hg38 Ensembl 100) генома человека на хромосоме 2. В других вариантах осуществления способа последовательность целевой нуклеиновой кислоты, подлежащая модификации, включает области гена BCL11A, кодирующие белок BCL11A, регуляторный элемент BCL11A, некодирующую область гена BCL11A или их перекрывающиеся участки. В конкретном варианте осуществления способа последовательность целевой нуклеиновой кислоты, подлежащая модификации, содержит последовательность связывания с GATA1 или ее комплемент.

[0212] В некоторых вариантах осуществления в описании предложены способы модификации целевой нуклеиновой кислоты BCL11A в клетке, причем способ включает введение в клетку системы CRISPR класса 2, типа V. В некоторых вариантах осуществления способов клетки, подлежащие модификации, являются аутологичными по отношению к субъекту, которому вводят упомянутую (-ые) клетку (-и). В других вариантах осуществления клетки, подлежащие модификации, являются аллогенными по отношению к субъекту, которому вводят упомянутую (-ые) клетку (-и). Таким образом, системы и способы, описанные в настоящем документе, могут применяться для конструирования различных клеток, в которых в гене β -глобина присутствуют мутации, связанные с заболеванием, например, гемоглобинопатиями, включая серповидно-клеточную болезнь и α - и β -талассемию. Таким образом, этот подход можно применять для модификации клеток для применения у субъекта со связанным с гемоглобинопатией заболеванием, таким как, без ограничений, серповидно-клеточная болезнь и α - и β -талассемию.

[0213] В некоторых вариантах осуществления в описании предложены способы модификации целевой нуклеиновой кислоты BCL11A в клетке, причем способ включает введение в клетку: i) системы CasX:гРНК, содержащей CasX и гРНК любого из вариантов осуществления, описанных в настоящем документе; ii) системы CasX:гРНК, содержащей CasX, гРНК и донорную матрицу по любому из вариантов осуществления, описанных в настоящем документе; iii) нуклеиновой кислоты, кодирующей CasX и гРНК и

необязательно содержащей донорную матрицу; iv) вектора, содержащего нуклеиновую кислоту по п. (iii), выше; v) XDP, содержащей систему CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе; или vi) комбинаций двух или более из пп. (i)-(v), причем последовательность целевой нуклеиновой кислоты клеток модифицирована белком CasX и необязательно донорной матрицей. В некоторых вариантах осуществления вектор представляет собой вектор AAV. В некоторых вариантах осуществления в описании предложены системы CasX:гРНК для применения в способах модификации гена BCL11A в клетке, причем система содержит вариант CasX, выбранный из группы, состоящей из SEQ ID NO: 36-99, 101-148 и 26908-27154, или вариант CasX, выбранный из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, или вариант CasX, выбранный из группы, состоящей из SEQ ID NO: 132-148 и 26908-27154, или последовательность варианта на по меньшей мере 60% идентична, на по меньшей мере 70% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична или на по меньшей мере 99,5% идентична с ней, каркас гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 2101-2285, 26794-26839 и 27219-27265, как указано в таблице 3, или из группы, состоящей из SEQ ID NO: 2281-2285, 26794-26839 и 27219-27265, или последовательности, которая на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична, на по меньшей мере 99,5% идентична с ней, и гРНК содержит нацеливающую последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 или 2286-26789, или последовательности, которая на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей мере 90% или на

по меньшей мере 95% идентична с ней и содержит от 15 до 20 нуклеотидов. В конкретных вариантах осуществления нацеливающая последовательность гРНК комплементарна и, следовательно, способна гибридизоваться с последовательностью в пределах последовательности мотива связывания с GATA1 или находится с 5'- или 3'-стороны от последовательности мотива связывания с GATA1. В одном варианте осуществления нацеливающая последовательность гРНК представляет собой UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22), которая гибридизуется с последовательностью сайта связывания эритроид-специфического энхансера GATA1 в BCL11A или представляет собой последовательность, идентичную последовательности на по меньшей мере 90% или на по меньшей мере 95%. В другом варианте осуществления нацеливающая последовательность гРНК представляет собой UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23), которая гибридизуется с последовательностью, комплементарной обратному комплементу последовательности сайта связывания эритроид-специфического энхансера GATA1 в BCL11A или представляет собой последовательность, идентичную последовательности на по меньшей мере 90% или на по меньшей мере 95%. В другом конкретном варианте осуществления нацеливающая последовательность гРНК комплементарна и, следовательно, способна гибридизоваться с последовательностью внутри промотора гена BCL11A. В одном варианте осуществления способа CasX и гРНК связаны друг с другом в рибонуклеопротеиновый комплекс (RNP). В некоторых вариантах осуществления способа модификации последовательности целевой нуклеиновой кислоты BCL11A в клетке модификация включает введение одноцепочечного разрыва в последовательность целевой нуклеиновой кислоты. В других вариантах осуществления способа модификация включает введение двухцепочечного разрыва в последовательность целевой нуклеиновой кислоты. В некоторых вариантах осуществления способа модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в последовательность целевой нуклеиновой кислоты. Как описано в настоящем документе, вариант CasX, вводящий расщепление двухцепочечных молекул в целевую нуклеиновую кислоту, создает двухцепочечный разрыв в пределах 18-26 нуклеотидов с 5'-стороны от сайта РАМ на целевой цепи и 10-18 нуклеотидов с 3'-стороны на нецелевой цепи. Таким образом, в некоторых вариантах осуществления полученная в результате осуществления способа модификация может приводить к случайным вставкам или делециям (инделам) или замене, дублированию или инверсии одного или более нуклеотидов в этой области с помощью репарационных механизмов соединения негомологичных концов ДНК (NHEJ).

[0214] В других вариантах осуществления способа модификации последовательности целевой нуклеиновой кислоты BCL11A в клетке способ включает приведение последовательности целевой нуклеиновой кислоты в контакт с системой CasX:гРНК с первой и второй или множеством гРНК, нацеленных на разные или перекрывающиеся участки гена BCL11A (например, при котором нацеливающая последовательность второй гРНК комплементарна последовательности, которая находится

на 5'- или 3'-стороне от сайта связывания с GATA1), причем белок CasX вводит множество разрывов в целевую нуклеиновую кислоту, что приводит к устойчивому инделу или мутации в целевой нуклеиновой кислоте, как описано в настоящем документе, или вырезанию последовательности мотива связывания с GATA1 с соответствующей модуляцией экспрессии или изменением функции продукта гена BCL11A, создавая таким образом отредактированную клетку. В некоторых вариантах вышеизложенного множество гРНК нацеливаются на положения с 5'- и 3'-стороны от последовательности мотива связывания с GATA1 в гене BCL11A так, что последовательность мотива связывания с GATA1 вырезается из целевого гена между двумя сайтами разрезов, на которые нацелены две гРНК. Следует понимать, что вышеуказанные варианты осуществления способа также могут быть реализованы посредством применения кодирования нуклеиновых кислот, векторов, содержащих кодирующие кислоты, или XDP, содержащих компоненты системы CasX:гРНК.

[0215] В некоторых вариантах осуществления в способах описания предложены пары из белка CasX и гРНК, которые создают сайт-специфичные двухцепочечные разрывы (DSB) или одноцепочечные разрывы (SSB) (например, когда белок CasX представляет собой нуклеазу, которая может расщеплять только одну цепь целевой нуклеиновой кислоты) в пределах 18-24 нуклеотидов с 3'-стороны от сайта PAM, которые затем могут подвергаться репарации либо посредством нехомологичного соединения концов (NHEJ), либо посредством направляемой гомологией репарации (HDR), независимой от гомологии нацеленной интеграции (HTI), соединения концов, обусловленного микрогомологией (MMEJ), отжига одиночных цепей (SSA) или эксцизионной репарации оснований (BER), в результате чего модификация гена BCL11A включает введение вставки, делеции, инверсии или мутации дублирования одного или более нуклеотидов по сравнению с последовательностью дикого типа с соответствующей модуляцией экспрессии или изменением продукта гена BCL11A, в результате чего создается отредактированная клетка.

[0216] В некоторых случаях система CasX:гРНК для применения в способах модификации гена BCL11A дополнительно содержит нуклеиновую кислоту-донорную матрицу по любому из вариантов осуществления, описанных в настоящем документе, причем донорная матрица может быть вставлена посредством механизмов направляемой гомологией репарации (HDR) или независимой от гомологии нацеленной интеграции (HTI) клетки-хозяина. Таким образом, в некоторых случаях способы, предложенные в настоящем документе, включают приведение гена BCL11A в контакт с донорной матрицей посредством введения донорной матрицы (либо *in vitro* внутрь клетки, либо *in vivo* внутрь клетки), причем донорная матрица, участок донорной матрицы, копия донорной матрицы или участок копии донорной матрицы интегрируется в ген BCL11A, заменяя участок гена BCL11A. Донорная матрица может представлять собой короткий одноцепочечный или двухцепочечный олигонуклеотид или длинный одноцепочечный или двухцепочечный олигонуклеотид. В некоторых вариантах осуществления донорная матрица содержит по меньшей мере участок гена BCL11A, причем участок гена BCL11A выбран из группы,

состоящей из экзона BCL11A, интрона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A или их комбинации. В некоторых вариантах осуществления в описании предложены донорные матрицы для применения в нацеливании или в нарушении работы сайта связывания с активатором транскрипции GATA1 в целевой последовательности BCL11A, причем донорная матрица содержит последовательности, которые являются негомологичными областям ДНК в сайте GATA1 или вблизи него в гене BCL11A, фланкированы двумя областями гомологии («гомологичные плечи») с 5'- и 3'-сторон от сайта (-ов) разрыва, так что работа механизмов репарации между целевой областью ДНК и двумя фланкирующими последовательностями приводит к вставке донорной матрицы в целевую область и облегчению вставки посредством HDR. Донорная матрица может содержать одну или более из одиночных замен оснований, вставок, делеций, инверсий или перестроений по отношению к геномной последовательности, при условии, что существует достаточная гомология с последовательностью целевой нуклеиновой кислоты для обеспечения ее интеграции в целевую нуклеиновую кислоту, что может приводить к мутации сдвига рамки или другой мутации, так что белок BCL11A не будет экспрессироваться (нокаут) или будет экспрессироваться на более низком уровне (нокдаун). Экзогенная донорная матрица, вставленная методом НПТ, может иметь любую длину, например быть относительно короткой последовательностью длиной от 10 до 50 нуклеотидов или более длинной последовательностью длиной более 50-1000 нуклеотидов. Отсутствие гомологии может иметь, например, идентичность последовательности не более чем 20-50% и/или отсутствие специфической гибридизации в условиях сниженной жесткости. В других случаях отсутствие гомологии может дополнительно включать критерий идентичности не более 5, 6, 7, 8 или 9 п. о. В некоторых вариантах осуществления полинуклеотид донорной матрицы содержит по меньшей мере примерно 10, по меньшей мере примерно 50, по меньшей мере примерно 100, или по меньшей мере примерно 200, или по меньшей мере примерно 300, или по меньшей мере примерно 400, или по меньшей мере примерно 500, или по меньшей мере примерно 600, или по меньшей мере примерно 700, или по меньшей мере примерно 800, или по меньшей мере примерно 900, или по меньшей мере примерно 1000, или по меньшей мере примерно 10 000, или по меньшей мере примерно 15 000 нуклеотидов. В других вариантах осуществления донорная матрица содержит от по меньшей мере примерно 10 до примерно 15 000 нуклеотидов, или от по меньшей мере примерно 100 до примерно 10 000 нуклеотидов, или от по меньшей мере примерно 400 до примерно 8000 нуклеотидов, или от по меньшей мере примерно 600 до примерно 5000 нуклеотидов, или от по меньшей мере примерно 1000 до примерно 2000 нуклеотидов. Последовательность донорной матрицы может содержать определенные различия в последовательности по сравнению с геномной последовательностью, например сайты рестрикции, нуклеотидные полиморфизмы, маркеры селекции (например, гены резистентности к лекарственному средству, флуоресцентные белки, ферменты и т. д.) и т. д., которые можно применять для оценки успешной вставки донорной нуклеиновой кислоты в сайт расщепления или, в некоторых случаях, применять для других целей

(например, чтобы обозначить экспрессию в целевом геномном локусе). Альтернативно эти различия в последовательности могут включать фланкирующие последовательности рекомбинации, такие как FLP, последовательности loxP или т. п., которые могут быть активированы позднее для удаления маркерной последовательности.

[0217] В некоторых вариантах осуществления способов модификации целевой нуклеиновой кислоты BCL11A в клетке *in vitro* или *ex vivo* для индукции расщепления или любой желаемой модификации целевой нуклеиновой кислоты гРНК и/или белок CasX настоящего описания и необязательно последовательность донорной матрицы, независимо от того, вводят ли их в виде нуклеиновых кислот или полипептидов, комплексов RNP, векторов или XDP, вводят в клетки в течение периода от примерно 30 минут до примерно 24 часов или по меньшей мере примерно 1 часа, 1,5 часа, 2 часов, 2,5 часа, 3 часа, 3,5 часа, 4 часа, 5 часов, 6 часов, 7 часов, 8 часов, 12 часов, 16 часов, 18 часов, 20 часов, или в течение любого периода от примерно 30 минут до примерно 24 часов, и введение можно повторять с частотой от примерно раза в день до примерно раза в 4 дня, например, раз в 1,5 дня, раз в 2 дня, раз в 3 дня или с любой частотой от примерно раза в день до примерно раза в четыре дня. Агент (-ы) могут подаваться в клетки субъекта один или более раз, например, один раз, два, три или более трех раз, и клеткам позволяют инкубироваться с агентом (-ами) в течение некоторого периода времени после каждого события приведения в контакт, например, от 30 минут до примерно 24 часов. В случае способов на основе *in vitro* после периода инкубации с CasX и гРНК (и необязательно донорной матрицей) среду заменяют свежей средой, и клетки культивируют дополнительно.

[0218] В некоторых вариантах осуществления способов модификации целевой нуклеиновой кислоты BCL11A в клетке способы дополнительно включают приведение в контакт последовательности целевой нуклеиновой кислоты клетки с: а) дополнительной нуклеазой CRISPR и гРНК, нацеленной на другой или перекрывающийся участок целевой нуклеиновой кислоты BCL11A относительно первой гРНК; б) полинуклеотидом, кодирующим дополнительную нуклеазу CRISPR и гРНК по п. (а); с) вектором, содержащим полинуклеотид по п. (б); или d) XDP, содержащей дополнительную нуклеазу CRISPR и гРНК по п. (а), причем приведение в контакт приводит к модификации целевой нуклеиновой кислоты BCL11A в другом месте в последовательности по сравнению с первой гРНК. В некоторых случаях дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий последовательность, отличную от белка CasX по любому из предшествующих пунктов. В других случаях дополнительная нуклеаза CRISPR не представляет собой белок CasX и выбрана из группы, состоящей из Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, Cas13a, Cas13b, Cas13c, Cas13d, CasY, Cas14, Cpf1, C2cl, Csn2, Cas Phi и вариантов их последовательностей.

[0219] В тех случаях, когда модификация приводит к нокдауну гена BCL11A, экспрессия белка BCL11A снижается на по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно

70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% по сравнению с клетками, которые не были модифицированы. В других случаях, в которых модификация приводит к нокауту гена BCL11A, целевая нуклеиновая кислота клеток популяции модифицирована таким образом, что экспрессия белка BCL11A не может быть обнаружена. Экспрессию белка BCL11A можно измерять посредством проточной цитометрии, твердофазного ИФА, клеточными анализами, вестерн-блоттингом, количественной ОТ-ПЦР или другими способами, известными в данной области или описанными в разделе «Примеры».

[0220] В некоторых вариантах осуществления в описании предложены способы модификации целевой нуклеиновой кислоты BCL11A в популяции клеток *in vivo* у субъекта. В некоторых вариантах осуществления модификация последовательности целевой нуклеиновой кислоты осуществляется *ex vivo* в эукариотической клетке, причем эукариотическая клетка выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта. В вышеуказанном варианте осуществления популяцию модифицированных клеток можно использовать в способе лечения субъекта, причем модифицированные клетки вводят нуждающемуся в этом субъекту, при этом субъект выбран из группы, состоящей из мыши, крысы, свиньи, примата, не являющегося человеком, и человека. В некоторых случаях клетка *ex vivo* является аутологичной и выделена из костного мозга субъекта или из периферической крови. В других случаях клетка *ex vivo* является аллогенной и выделена из костного мозга или из периферической крови другого субъекта. В способах лечения модифицированную клетку можно вводить субъекту способом введения, выбранным из интрапаренхиматозного, внутривенного, внутриартериального, внутримышечного, субкутикулярного, внутрисуставного, внутрисердечного, внутриперикардального, интравитреального, субкапсулярного, или посредством подкожной инъекции, и она может быть введена субъекту посредством трансплантации, местной инъекции, системной инфузии или их комбинаций. В вышеуказанном варианте осуществления способ приводит к сохранению модифицированной клетки или ее потомства в течение по меньшей мере примерно 1 месяца, по меньшей мере примерно 2 месяцев, по меньшей мере примерно 3 месяцев, по меньшей мере примерно 4 месяцев, по меньшей мере примерно 6 месяцев, по меньшей мере примерно 7 месяцев, по меньшей мере примерно 8 месяцев, по меньшей мере примерно 9 месяцев, по меньшей мере примерно 10 месяцев, по меньшей мере примерно 11 месяцев, по меньшей мере примерно 12 месяцев, по меньшей мере примерно 18 месяцев, по меньшей мере примерно 2 лет, по меньшей мере примерно 3 лет, по меньшей мере примерно 4 лет или по меньшей мере примерно 5 лет.

[0221] В некоторых вариантах осуществления способов модификации последовательности целевой нуклеиновой кислоты модификация гена BCL11A включает

связывание комплекса CasX:гРНК с последовательностью целевой нуклеиновой кислоты и введение в клетки в виде RNP. В некоторых вариантах осуществления CasX представляет собой каталитически неактивный белок CasX (dCasX), который сохраняет способность связываться с гРНК и последовательностью целевой нуклеиновой кислоты. Например, последовательность целевой нуклеиновой кислоты содержит последовательность BCL11A, содержащую последовательность, комплементарную последовательности мотива связывания с GATA1, и связывание комплекса dCasX:гРНК с целевой последовательностью препятствует транскрипции аллеля BCL11A или подавляет ее. В некоторых вариантах осуществления dCasX содержит мутацию в остатках D672, E769 и/или D935, соответствующих белку CasX SEQ ID NO: 1, или D659, E756 и/или D922, соответствующих белку CasX SEQ ID NO: 2. В некоторых вариантах осуществления вышеизложенного мутация в варианте белка CasX представляет собой замену данного остатка на аланин или глицин и может использоваться в любом из вариантов, описанных в настоящем документе.

[0222] Введение рекомбинантных экспрессионных векторов, содержащих компоненты или нуклеиновые кислоты, кодирующие компоненты вариантов осуществления системы, в клетку-мишень можно осуществлять *in vivo*, *in vitro* или *ex vivo*. В некоторых вариантах осуществления способа векторы могут быть введены непосредственно в целевую клетку-хозяин. Способы введения нуклеиновой кислоты (например, нуклеиновой кислоты, содержащей донорную полинуклеотидную последовательность, одну или более нуклеиновых кислот (ДНК или РНК), кодирующих белок CasX и/или гРНК или содержащий их вектор) в клетку известны в данной области, и для введения нуклеиновой кислоты (например, экспрессионного конструкта) в клетку можно применять любой удобный способ. Приемлемые способы включают, например, вирусную инфекцию, трансфекцию, липофекцию, электропорацию, преципитацию фосфатом кальция, опосредованную полиэтиленгликолем (PEG) трансфекцию, опосредованную DEAE-декстраном трансфекцию, опосредованную липосомами трансфекцию, технологию генной пушки, нуклеофекцию, электропорацию, прямое добавление проникающих в клетку белков CasX, которые слиты или рекрутируют донорную ДНК, сжатие клеток, преципитацию фосфатом кальция, прямую микроинъекцию, опосредованную наночастицами доставку нуклеиновой кислоты и т. п. Нуклеиновые кислоты могут быть введены в клетки с применением хорошо разработанных коммерчески доступных методик трансфекции, таких как применение реагентов TransMessenger® от Qiagen, наборов для трансфекции Stemfect™ RNA Transfection Kit от Stemgent и TransIT®-mRNA Transfection Kit от Mirus Bio LLC, нуклеофекции Lonza, электропорации Maxagen и т. п. Введение рекомбинантных экспрессионных векторов, содержащих последовательности, кодирующие системы CasX:гРНК (и необязательно донорные последовательности) описания, в клетки в условиях *in vitro* может происходить в любой приемлемой культуральной среде и в любых приемлемых условиях культивирования, которые способствуют выживанию клеток. Например, клетки можно приводить в контакт с векторами, содержащими нуклеиновые кислоты субъекта (например,

рекомбинантные экспрессионные векторы, содержащие последовательность донорной матрицы и нуклеиновую кислоту, кодирующую CasX и гРНК), так чтобы векторы поглощались клетками. Векторы, применяемые для обеспечения нуклеиновых кислот, кодирующих гРНК и/или белки CasX для целевой клетки-хозяина, могут включать приемлемые промоторы для осуществления экспрессии, то есть активации транскрипции представляющей интерес нуклеиновой кислоты. В некоторых случаях представляющая интерес нуклеиновая кислота функционально связана с промотором. К ним могут относиться промоторы универсального действия, например, промотор CMV-бета-актина или индуцибельные промоторы, такие как промоторы, активные в конкретных популяциях клеток, или которые отвечают на присутствие лекарственных средств, таких как тетрациклин или канамицин. Под активацией транскрипции подразумевается, что транскрипция будет повышена выше исходных уровней в целевой клетке-хозяине, содержащей вектор, в по меньшей мере примерно 10 раз, в по меньшей мере примерно 100 раз, более типично в по меньшей мере примерно 1000 раз. Кроме того, векторы, применяемые для введения в клетку нуклеиновой кислоты, кодирующей гРНК и/или белок CasX, могут включать нуклеотидные последовательности, которые кодируют маркеры селекции в клетках-мишенях, чтобы можно было идентифицировать клетки, которые поглотили белок CasX и/или гРНК.

[0223] Для доставки вирусным вектором клетки можно приводить в контакт с вирусными частицами, содержащими рассматриваемые вирусные экспрессионные векторы, и нуклеиновую кислоту, кодирующую CasX и гРНК, и необязательно донорную матрицу. В некоторых вариантах осуществления вектор представляет собой аденоассоциированный вирусный (AAV) вектор, причем AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV 44.9, AAV-Rh74 или AAVRh10. В других случаях AAV выбран из AAV1, AAV2, AAV5, AAV6, AAV7, AAV8 и AAV9, которые эффективны для трансдукции в мышцы (Gruntman AM, et al. Gene transfer in skeletal and cardiac muscle using recombinant adeno-associated virus. *Curr Protoc Microbiol.* 14(14D):3 (2013). Варианты осуществления векторов AAV более полно описаны ниже. В некоторых вариантах осуществления вектор представляет собой лентивирусный вектор. Ретровирусы, например лентивирусы, могут быть приемлемы для применения в способах настоящего описания. Обычно применяемые ретровирусные векторы являются «дефектными», например, они неспособны продуцировать вирусные белки, необходимые для продуктивной инфекции. Вместо этого для репликации вектора требуется рост в пакующей линии клеток. Для создания вирусных частиц, содержащих представляющие интерес нуклеиновые кислоты, ретровирусные нуклеиновые кислоты, содержащие нуклеиновую кислоту, упаковывают в вирусные капсиды с помощью пакующей линии клеток. Разные пакующие линии клеток обеспечивают разные белки оболочки (экотропный, амфотропный или ксенотропный), которые должны быть включены в капсид, и белок оболочки определяет специфичность или тропизм вирусной частицы к клеткам (экотропный - для мышей и крыс; Амфотропный - для большинства типов клеток

млекопитающих, включая человека, собаку и мышь; и ксенотропный - для большинства типов клеток млекопитающих, за исключением мышинных клеток). Подходящую паковочную линию клеток можно применять для обеспечения того, что упакованные вирусные частицы нацелены на клетки. Способы введения рассматриваемых экспрессионных векторов в паковочные линии клеток и сбора вирусных частиц, которые создаются паковочными линиями, хорошо известны в данной области, например, из патента США № 5,173,414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); and Samulski et al., *J. Virol.* 63:03822-3828 (1989). Нуклеиновые кислоты также можно вводить посредством прямой микроинъекции (например, инъекции РНК).

[0224] В других вариантах осуществления способов модификации гена BCL11A в способе используются частицы для доставки CasX (XDP) для нацеленной доставки RNP в клетки субъекта. XDP представляет собой частицы, которые очень похожи на вирусы, но не содержат вирусного генетического материала и, следовательно, не являются инфекционными. В некоторых вариантах осуществления XDP содержит CasX и гРНК в комплексе в виде RNP и необязательно донорную матрицу, которая содержит, полностью или частично, ген BCL11A для нокаута или нокдауна гена BCL11A или участка гена путем вставки посредством механизмов HDR или NHEJ. Варианты осуществления XDP более подробно описаны ниже.

VI. Полинуклеотиды и векторы

[0225] В другом аспекте настоящее описание относится к полинуклеотидам, кодирующим нуклеазы класса 2, типа V и гРНК, которые можно использовать при редактировании гена BCL11A. В некоторых вариантах осуществления в описании предложены полинуклеотиды, кодирующие белки CasX и полинуклеотиды гРНК по любому из вариантов осуществления системы CasX:гРНК, описанных в настоящем документе. В дополнительных вариантах осуществления в описании предложены полинуклеотиды донорной матрицы, кодирующие участки или полностью ген BCL11A. В некоторых случаях донорная матрица содержит мутацию или гетерологичную последовательность для нокдауна или нокаута гена BCL11A после ее вставки в целевую нуклеиновую кислоту. В дополнительных вариантах осуществления в описании предложены векторы, содержащие полинуклеотиды, кодирующие белки CasX и гРНК для CasX, описанные в настоящем документе, а также донорные матрицы по вариантам осуществления.

[0226] В некоторых вариантах осуществления в описании предложена полинуклеотидная последовательность, кодирующая варианты CasX по любому из вариантов осуществления, описанных в настоящем документе, включая варианты белка CasX с SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, как описано в таблице 4, или последовательности, на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по

меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичные последовательности с SEQ ID NO: 59, 72-99, 101-148 и 26908-27154 из таблицы 4. В некоторых вариантах осуществления в описании предложена полинуклеотидная последовательность, кодирующая варианты CasX по любому из вариантов осуществления, описанных в настоящем документе, включая варианты белка CasX с SEQ ID NO: 36-99, 101-148 и 26908-27154, или описанные в таблице 4, или последовательности, на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99% идентичные последовательности с SEQ ID NO: 36-99, 101-148 и 26908-27154 из таблицы 4. В некоторых вариантах осуществления в описании предложена изолированная полинуклеотидная последовательность, кодирующая последовательность гРНК по любому из вариантов осуществления, описанных в настоящем документе, включая последовательности SEQ ID NO: 4-16, 2238-2285, 26794-26839 или 27219-27265 из таблиц 2 и 3, вместе с нацеливающими последовательностями SEQ ID NO: 272-2100 или 2286-26789. В некоторых вариантах осуществления в описании предложена изолированная полинуклеотидная последовательность, кодирующая последовательность гРНК по любому из вариантов осуществления, описанных в настоящем документе, включая последовательности SEQ ID NO: 2101-2285, 26794-26839 и 27219-27265, вместе с нацеливающими последовательностями SEQ ID NO: 272-2100 или 2286-26789. В некоторых вариантах осуществления в описании предложена изолированная полинуклеотидная последовательность, кодирующая последовательность гРНК по любому из вариантов осуществления, описанных в настоящем документе, включая последовательности SEQ ID NO: 2281-2285, 26794-26839 или 27219-27265, вместе с нацеливающими последовательностями SEQ ID NO: 272-2100 или 2286-26789. В некоторых вариантах осуществления последовательности, кодирующие белок CasX, кодон-оптимизированы для экспрессии в эукариотической клетке.

[0227] В некоторых вариантах осуществления в описании предложен полинуклеотид, кодирующий последовательность каркаса гРНК с SEQ ID NO: 4-16, 2238-2285, 26794-26839 или 27219-27265, или как указано в таблице 2 или таблице 3, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99%. В других вариантах осуществления в описании предложен полинуклеотид нацеливающей последовательности из таблицы 1 или последовательность, которая на по меньшей мере примерно 65%, по меньшей мере примерно 75%, по меньшей мере примерно 85% или по меньшей мере примерно 95% идентична последовательности с SEQ ID NO: 272-2100 или 2286-26789. В

некоторых вариантах осуществления полинуклеотид нацеливающей последовательности, в свою очередь, связан с 3'-концом каркасной последовательности гРНК; либо в виде огРНК, либо в виде дгРНК. В других вариантах осуществления в описании предложены гРНК, содержащие полинуклеотиды нацеливающей последовательности, имеющие один или более однонуклеотидных полиморфизмов (SNP) относительно последовательности с SEQ ID NO: 272-2100 или 2286-26789.

[0228] В других вариантах осуществления в описании предложена выделенная полинуклеотидная последовательность, кодирующая гРНК, содержащую нацеливающую последовательность, которая комплементарна и, следовательно, способна гибридизоваться с геном BCL11A. В некоторых вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с экзоном BCL11A. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с интроном BCL11A. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с соединением интрон-экзон BCL11A. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с межгенной областью гена BCL11A. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с регуляторным элементом BCL11A. В некоторых случаях регуляторный элемент BCL11A представляет собой промотор или энхансер BCL11A. В некоторых случаях регуляторный элемент BCL11A расположен с 5'-стороны от сайта начала транскрипции BCL11A, с 3'-стороны от сайта начала транскрипции BCL11A или в интроне BCL11A. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с последовательностью, расположенной с 5'-стороны от последовательности мотива связывания с GATA1. В других вариантах осуществления полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность, которая гибридизуется с последовательностью, перекрывающейся с последовательностью мотива связывания с GATA1. В конкретном варианте осуществления вышеизложенного полинуклеотидная последовательность кодирует гРНК, содержащую нацеливающую последовательность с SEQ ID NO: 22. В некоторых случаях регуляторный элемент BCL11A находится в интроне гена BCL11A. В других случаях регуляторный элемент BCL11A содержит 5'-UTR гена BCL11A. В других случаях регуляторный элемент BCL11A содержит 3'-UTR гена BCL11A.

[0229] В других вариантах осуществления в описании предложены нуклеиновые кислоты донорных матриц, причем донорная матрица содержит нуклеотидную последовательность, имеющую гомологию с последовательностью целевой нуклеиновой кислоты BCL11A. В некоторых вариантах осуществления донорная матрица BCL11A

предназначена для редактирования генов в сочетании с системой CasX:гРНК и содержит по меньшей мере участок гена BCL11A. В других вариантах осуществления донорная последовательность BCL11A содержит последовательность, которая кодирует по меньшей мере участок экзона BCL11A. В других вариантах осуществления донорная матрица BCL11A содержит последовательность, которая кодирует по меньшей мере участок интрона BCL11A. В других вариантах осуществления донорная матрица BCL11A содержит последовательность, которая кодирует по меньшей мере участок места соединения интрон-экзон BCL11A. В других вариантах осуществления донорная матрица BCL11A содержит последовательность, которая кодирует по меньшей мере участок межгенной области гена BCL11A. В других вариантах осуществления донорная матрица BCL11A содержит последовательность, которая кодирует по меньшей мере участок регуляторного элемента BCL11A. В некоторых случаях донорная матрица BCL11A представляет собой последовательность дикого типа, которая кодирует по меньшей мере участок SEQ ID NO: 100. В других случаях последовательность донорной матрицы BCL11A содержит одну или более мутаций относительно гена BCL11A дикого типа. В конкретном варианте осуществления донорная матрица содержит последовательность, которая кодирует, частично или полностью, последовательность мотива связывания с GATA1, но с по меньшей мере 1-5 мутациями относительно последовательности дикого типа. В вышеуказанных вариантах осуществления донорная матрица содержит по меньшей мере 10 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 200 нуклеотидов, по меньшей мере 300 нуклеотидов, по меньшей мере 400 нуклеотидов, по меньшей мере 500 нуклеотидов, по меньшей мере 600 нуклеотидов, по меньшей мере 700 нуклеотидов, по меньшей мере 800 нуклеотидов, по меньшей мере 900 нуклеотидов, по меньшей мере 1000 нуклеотидов, по меньшей мере 2000 нуклеотидов, по меньшей мере 3000 нуклеотидов, по меньшей мере 4000 нуклеотидов, по меньшей мере 5000 нуклеотидов, по меньшей мере 6000 нуклеотидов, по меньшей мере 7000 нуклеотидов, по меньшей мере 8000 нуклеотидов, по меньшей мере 9000 нуклеотидов, по меньшей мере 10 000 нуклеотидов, по меньшей мере 12 000 нуклеотидов или по меньшей мере 15 000 нуклеотидов. В некоторых вариантах осуществления донорная матрица содержит по меньшей мере от примерно 10 до примерно 15 000 нуклеотидов. В некоторых вариантах осуществления донорная матрица представляет собой одноцепочечную ДНК-матрицу. В других вариантах осуществления донорная матрица представляет собой одноцепочечную РНК-матрицу. В других вариантах осуществления донорная матрица представляет собой двухцепочечную ДНК-матрицу. В некоторых вариантах осуществления донорная матрица может быть предоставлена в виде оголенной нуклеиновой кислоты в системах для редактирования гена BCL11A, и ее не нужно включать в вектор. В других вариантах осуществления донорная матрица может быть включена в вектор для облегчения ее доставки в клетку; например, в вирусный вектор.

[0230] В других аспектах описание относится к способам получения полинуклеотидных последовательностей, кодирующих варианты CasX или гРНК по любому из вариантов осуществления, описанных в настоящем документе, включая их

гомологичные варианты, а также способам экспрессии белков, в которых с полинуклеотидных последовательностей экспрессируются белки или транскрибируется РНК. В целом, способы включают получение полинуклеотидной последовательности, кодирующей варианты CasX или гРНК по любому из вариантов осуществления, описанных в настоящем документе, и включение кодирующего гена в экспрессионный вектор, подходящий для клетки-хозяина. Для получения полинуклеотидов и экспрессионных векторов настоящего описания можно применять стандартные рекомбинантные методики молекулярной биологии. Для получения кодируемого референсного CasX, вариантов CasX или гРНК по любому из вариантов осуществления, описанных в настоящем документе, способы включают трансформирование соответствующей клетки-хозяина экспрессионным вектором, содержащим кодирующий полинуклеотид, и культивирование клетки-хозяина в условиях, приводящих или способствующих получению референсного CasX, вариантов CasX или гРНК по любому из вариантов осуществления, описанных в настоящем документе, подлежащих экспрессированию или транскрибированию в трансформированной клетке-хозяине, с получением таким образом вариантов CasX или гРНК, которые извлекают способами, описанными в настоящем документе, или с помощью стандартных способов очистки, известных в данной области или описанных в разделе «Примеры».

[0231] В соответствии с описанием последовательности нуклеиновых кислот, которые кодируют варианты CasX или гРНК по любому из вариантов осуществления, описанных в настоящем документе (или их комплемент), применяются для создания рекомбинантных молекул ДНК, которые направляют экспрессию в соответствующих клетках-хозяевах. Несколько стратегий клонирования подходят для реализации настоящего описания, многие из которых применяются для создания конструкта, содержащего ген, кодирующий композицию настоящего описания, или его комплемент. В некоторых вариантах осуществления стратегия клонирования применяется для создания гена, который кодирует конструкт, содержащий нуклеотиды, кодирующие варианты CasX, или гРНК, которая применяется для трансформации клетки-хозяина и экспрессии композиции.

[0232] В некоторых подходах сначала получают конструкт, содержащий последовательность ДНК, кодирующую вариант CasX или гРНК. Примеры способов получения таких конструктов описаны в разделе «Примеры». Затем конструкт применяют для создания экспрессионного вектора, приемлемого для трансформации клетки-хозяина, такой как прокариотическая или эукариотическая клетка-хозяин, для экспрессии и извлечения белкового конструкта, в случае CasX или гРНК. При необходимости клетка-хозяин представляет собой *E. coli*. В других вариантах осуществления клетка-хозяин представляет собой эукариотическую клетку. Эукариотическая клетка-хозяин может быть выбрана из клеток-фибробластов почки новорожденного хомяка (ВНК), клеток эмбриональной почки человека 293 (HEK293), клеток эмбриональной почки человека 293Т (HEK293Т), клеток NS0, клеток SP2/0, клеток миеломы YO, мышинных клеток миеломы P3X63, клеток PER, клеток PER.C6, клеток гибридомы, клеток NIH3T3, клеток,

происходящих от CV-1 (обезьяньих), с генетическим материалом SV40 (COS), клеток HeLa, клеток яичника китайского хомячка (CHO), или дрожжевых клеток, или других эукариотических клеток, известных в данной области, приемлемых для получения рекомбинантных продуктов. Примеры способов создания экспрессионных векторов, трансформации клеток-хозяев и экспрессии и извлечения вариантов CasX или гРНК описаны в разделе «Примеры».

[0233] Ген, кодирующий вариант CasX или конструкт гРНК, может быть получен в одну или более стадий, либо полностью синтетически, либо посредством синтеза в комбинации с ферментативными процессами, такими как опосредованное рестрикционным ферментом клонирование, ПЦР с перекрывающимися праймерами, включая способы, более подробно описанные в разделе «Примеры». Способы, описанные в настоящем документе, можно применять, например, для лигирования к последовательностям полинуклеотидов, кодирующих гены различных компонентов (например, CasX и гРНК) желаемой последовательности. Гены, кодирующие полипептидные композиции, собирают из олигонуклеотидов с применением стандартных методик синтеза генов.

[0234] В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая белок CasX, кодон-оптимизирована для предполагаемой клетки-хозяина. Этот тип оптимизации может предполагать мутацию кодирующей нуклеотидной последовательности для имитации предпочтительных кодонов предполагаемого организма или клетки-хозяина при кодировании того же белка CasX. Таким образом, кодоны можно изменять, но кодируемый белок или гРНК остается неизменным. Например, если предполагаемая клетка-мишень для белка CasX представляла собой клетку человека, то можно было применять человеческую кодон-оптимизированную кодирующую CasX нуклеотидную последовательность. В качестве другого не имеющего ограничительного характера примера, если предполагаемая клетка-хозяин представляла собой клетку мыши, то можно было создать мышиную кодон-оптимизированную кодирующую CasX нуклеотидную последовательность. Конструирование гена может быть выполнено с применением алгоритмов, которые оптимизируют применение кодонов и аминокислотный состав, делая их подходящими для клетки-хозяина, используемой при получении референсного CasX или вариантов CasX. В одном способе описания создают и затем собирают библиотеку полинуклеотидов, кодирующих компоненты конструктов, описанных выше. Затем полученные гены собирают и готовые гены применяют для трансформации клетки-хозяина, продуцируют и извлекают варианты CasX или композиции гРНК для оценки их свойств, как описано в настоящем документе.

[0235] Описание предусматривает применение плазмидных экспрессионных векторов, содержащих репликационные и управляющие последовательности, которые совместимы с, распознаются клеткой-хозяином и функционально связаны с геном, кодирующим полипептид, для управляемой экспрессии полипептида или транскрипции РНК. Такие последовательности векторов хорошо известны для разнообразных бактерий, дрожжей и вирусов. Используемые экспрессионные векторы, подходящие для применения,

включают, например, сегменты хромосомных, нехромосомных и синтетических последовательностей ДНК. Термин «экспрессионный вектор» относится к конструкту ДНК, содержащему последовательность ДНК, которая функционально связана с приемлемой управляющей последовательностью, и способна осуществлять экспрессию ДНК, кодирующей полипептид, в приемлемом хозяине. Требования заключаются в том, чтобы векторы были способны к репликации и жизнеспособны в выбранной клетке-хозяине. При необходимости можно применять низкокопийные или высококопийные векторы. Управляющие последовательности вектора включают промотор для осуществления транскрипции, необязательную операторную последовательность для контроля такой транскрипции, последовательность, кодирующую приемлемые сайты связывания мРНК с рибосомой и последовательности, которые контролируют прекращение транскрипции и трансляции. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая гРНК, функционально связана с управляющим элементом, например, элементом управления транскрипцией, таким как промотор. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая белок CasX, функционально связана с управляющим элементом, например, элементом управления транскрипцией, таким как промотор. В других случаях нуклеотидная последовательность, кодирующая CasX и гРНК, связана и функционально связана с одним элементом управления. Промотор может представлять собой любую последовательность ДНК, которая демонстрирует транскрипционную активность в выбранной клетке-хозяине и может быть получена из генов, кодирующих белки, либо гомологичные, либо гетерологичные для клетки-хозяина. Примеры регуляторных элементов включают промотор транскрипции, элемент-энхансер транскрипции, сигнал терминации транскрипции, внутренний сайт входа рибосомы (IRES) или пептид P2A для обеспечения трансляции множества генов с одного транскрипта, последовательности полиаденилирования для активации последующей терминации транскрипции, последовательности для оптимизации инициации трансляции и последовательности терминации трансляции. В некоторых случаях промотор представляет собой конститутивно активный промотор. В некоторых случаях промотор представляет собой регулируемый промотор. В одном варианте осуществления промотор представляет собой индуцибельный промотор. В некоторых случаях промотор представляет собой тканеспецифичный промотор. В некоторых случаях промотор представляет собой специфичный для типа клеток промотор. В некоторых случаях элемент управления транскрипцией (например, промотор) является функциональным в типе клеток-мишеней или популяции клеток-мишеней. Например, в некоторых случаях элемент управления транскрипцией может быть функциональным в эукариотических клетках, например, пакующих клетках для вирусных или XDP-векторов, гематопозитических стволовых клетках (HSC), гематопозитических клетках-предшественниках (HPC), клетках CD34+, мезенхимальных стволовых клетках (MSC), эмбриональных стволовых (ES) клетках, индуцированных плюрипотентных стволовых клетках (iPSC), общих миелоидных клетках-предшественниках, клетках-

проэритроблестах и клетках-эритроблестах.

[0236] Не имеющие ограничительного характера примеры промоторов pol II включают, без ограничений, EF-1альфа, основной промотор EF-1альфа, Jens Tornoe (JeT), промоторы из цитомегаловируса (CMV), немедленный-ранний промотор CMV (CMVIE), энхансер CMV, тимидинкиназу вируса простого герпеса (HSV), ранний и поздний промотор обезьяньего вируса 40 (SV40), энхансер SV40, промотор с длинными концевыми повторами (LTR) ретровируса, промотор металлотионеина-I мыши, главный поздний промотор аденовируса (Ad MLP), полноразмерный промотор CMV, минимальный промотор CMV, промотор α -актина цыпленка (CBA), гибридный CBA (CBh), промотор α -актина цыпленка с энхансером цитомегаловируса (CB7), слитый промотор бета-актина цыпленка и акцепторный сайт сплайса бета-глобина кролика (CAG), промотор вируса саркомы Рауса (RSV), промотор ВИЧ-Ltr, промотор hPGK, промотор ТК HSV, промотор 7SK, промотор Mini-TK, промотор синапсина I человека (SYN), который придает нейрон-специфическую экспрессию, промотор бета-актина, суперкоровый промотор 1 (SCP1), промотор Меср2 для селективной экспрессии в нейронах, минимальный промотор IL-2, энхансер/промотор (единый) вируса саркомы Рауса, промотор с длинными концевыми повторами (LTR) вируса некроза селезенки, промотор TBG, промотор из гена тироксинсвязывающего глобулина человека (специфический для печени), промотор PGK, промотор убиквитина С человека (UBC), промотор UC0E (промотор HNRPA2B1-CBX3), синтетический промотор CAG, промотор гистона H2, промотор гистона H3, промотор малой ядерной РНК U1a1 (226 нукл.), промотор малой ядерной РНК U1a1 (226 нукл.), промотор малой ядерной РНК U1b2 (246 нукл.) 26, промотор GUSB, промотор CBh, промотор родопсина (Rho), склонный к сайленсингу промотор вируса некроза селезенки (SFFV), промотор H1 человека (H1), промотор POL1, минимальный энхансер/промотор TTR, промотор b-кинезина, промотор с длинными терминальными повторами (LTR) вируса опухоли молочной железы мыши, промотор эукариотического фактора 4A инициации человека (EIF4A1), промотор ROSA26, промотор глицеральдегид-3-фосфатдегидрогеназы (GAPDH), промоторы tРНК, а также усеченные версии и варианты последовательности вышеизложенного. В конкретном варианте осуществления промотор pol II представляет собой EF-1альфа, причем промотор усиливает эффективность трансфекции, транскрипцию или экспрессию нуклеазы CRISPR, долю экспрессионно-положительных клонов и число копий эписомального вектора в долгосрочной культуре.

[0237] Не имеющие ограничительного характера примеры промоторов pol III включают, без ограничений, U6, мини-U6, U6-усеченные промоторы, 7SK, варианты H1, BiH1 (двусторонний промотор H1), BiU6, Bi7SK, BiH1 (двусторонние промоторы U6, 7SK и H1), U6 гориллы, U6 резуса, 7SK человека, H1 человека и варианты их последовательностей. В вышеуказанном варианте осуществления промотор pol III усиливает транскрипцию гРНК.

[0238] Выбор соответствующего вектора и промотора знаком рядовому специалисту в данной области, так как связан с контролем экспрессии, например, для модификации гена

BCL11A. Экспрессионный вектор также может содержать сайт связывания рибосомы для инициации трансляции и терминатор транскрипции. Экспрессионный вектор также может включать подходящие последовательности для усиления экспрессии. Экспрессионный вектор также может включать нуклеотидные последовательности, кодирующие белковые метки (например, метка 6xHis, гемагглютининовая метка, флуоресцентный белок и т. д.), которые могут быть слиты с белком CasX, что приводит к получению химерного белка CasX, который применяют для очистки или обнаружения.

[0239] Рекомбинантные экспрессионные векторы описания могут также содержать элементы, которые способствуют устойчивой экспрессии белков CasX и гРНК описания. Например, рекомбинантные экспрессионные векторы могут включать одно или более из сигнала полиаденилирования (поли(A)), интронной последовательности или посттранскрипционного регуляторного элемента, такого как посттранскрипционный регуляторный элемент вируса гепатита лесного сурка (WPRE). Типичные последовательности поли(A) включают сигнал поли(A) hGH (короткий), сигнал поли(A) HSV TK, синтетические сигналы полиаденилирования, сигнал поли(A) SV40, сигнал поли(A) β -глобина и т. п. Специалист в данной области сможет выбрать приемлемые элементы для включения в рекомбинантные экспрессионные векторы, описанные в настоящем документе.

[0240] В некоторых вариантах осуществления в настоящем документе предложены один или более рекомбинантных экспрессионных векторов, содержащих одно или более из: (i) нуклеотидной последовательности нуклеиновой кислоты-донорной матрицы, где донорная матрица содержит нуклеотидную последовательность, имеющую гомологию с последовательностью целевого локуса BCL11A целевой нуклеиновой кислоты (например, целевого генома); (ii) нуклеотидной последовательности, кодирующей гРНК, которая гибридизируется с целевой последовательностью локуса BCL11A целевого генома (например, выполненную в виде одинарной или двойной гидовой РНК), функционально связанной с промотором, который способен функционировать в клетке-мишени, такой как эукариотическая клетка; и (iii) нуклеотидной последовательности, кодирующей белок CasX, функционально связанный с промотором, который способен функционировать в клетке-мишени, такой как эукариотическая клетка. В некоторых вариантах осуществления последовательности, кодирующие донорную матрицу, гРНК и белок CasX, находятся в разных рекомбинантных экспрессионных векторах, а в других вариантах осуществления одна или более полинуклеотидных последовательностей (донорной матрицы, CasX и гРНК) находятся в одном рекомбинантном экспрессионном векторе. В других случаях CasX и гРНК доставляют в клетку-мишень в виде RNP (например, посредством электропорации или химическими средствами), а донорную матрицу доставляют вектором.

[0241] Полинуклеотидную (-ые) последовательность (-и) вводят в вектор различными процедурами. В целом, ДНК вставляют в соответствующий (-ие) сайт (-ы) рестрикционной эндонуклеазы с применением методик, известных в данной области. Компоненты вектора, как правило, включают, без ограничений, одно или более из

сигнальной последовательности, точки начала репликации, одного или более маркерных генов, энхансерного элемента, промотора и последовательности терминации транскрипции. При конструировании приемлемых векторов, содержащих один или более из этих компонентов, используют стандартные методики лигирования, известные специалистам в данной области. Такие методики хорошо известны в данной области и хорошо описаны в научной и патентной литературе. Различные векторы являются общедоступными. Вектор может, например, быть в форме плазмиды, космиды, вирусной частицы или фага, который можно удобно подвергать процедурам получения рекомбинантной ДНК, и выбор вектора часто зависит от клетки-хозяина, в которую его необходимо ввести. Таким образом, вектор может представлять собой автономно реплицирующийся вектор, т. е. вектор, который существует как внехромосомный объект, репликация которого не зависит от хромосомной репликации, например, плазмиду. Альтернативно вектор может представлять собой вектор, который при введении в клетку-хозяина интегрируется в геном клетки-хозяина и реплицируется вместе с хромосомой (-ами), в которую он интегрирован. После введения в приемлемую клетку-хозяина экспрессию белка, участвующего в процессинге антигена, презентации антигена, распознавания антигена и/или реакции на антиген, можно определять с применением любого анализа нуклеиновых кислот или белков, известного в данной области. Например, наличие транскрибируемой мРНК референсного CasX или вариантов CasX может быть обнаружено и/или количественно определено обычными гибридизационными анализами (например, анализом нозерн-блот), процедурами амплификации (например, ОТ-ПЦР), серийным анализом генной экспрессии (SAGE) (патент США № 5,695,937) и технологиями на основе матриц (см., например, патенты США № 5,405,783, 5,412,087 и 5,445,934) с применением зондов, комплементарных любой области полинуклеотида.

[0242] Полинуклеотиды и рекомбинантные экспрессионные векторы могут быть доставлены в целевые клетки-хозяева различными способами. Такие способы включают, без ограничений, вирусную инфекцию, трансфекцию, липофекцию, электропорацию, преципитацию фосфатом кальция, опосредованную полиэтиленимином (PEI) трансфекцию, опосредованную DEAE-декстраном трансфекцию, микроинъекцию, опосредованную липосомами трансфекцию, технологию генной пушки, нуклеофекцию, прямое добавление проникающих в клетки белков CasX, которые слиты или рекрутируют донорную ДНК, сжатие клеток, преципитацию фосфатом кальция, прямую микроинъекцию, опосредованную наночастицами доставку нуклеиновых кислот, а также применение коммерческих наборов реагентов TransMessenger® от Qiagen, наборов для трансфекции Stemfect™ RNA Transfection Kit от Stemgent и TransIT®-mRNA Transfection Kit от Mirus Bio LLC, нуклеофекция Lonza, электропорация Maxagen и т. п.

[0243] Последовательность рекомбинантного экспрессионного вектора может быть упакована в вирус или вирусоподобную частицу (также называемую в настоящем документе «частица» или «вирион») для последующей инфекции и трансформации клетки *ex vivo*, *in vitro* или *in vivo*. Такие частицы или вирионы, как правило, включают белки,

которые заключают в капсид или упаковывают векторный геном. Приемлемые экспрессионные векторы могут включать вирусные экспрессионные векторы на основе вируса осповакцины; полиовируса; аденовируса; ретровирусный вектор (например, вирус лейкоза мыши), вирус некроза селезенки и векторы, производные от ретровирусов, таких как вирус саркомы Рауса, вирус саркомы Харви, вирус птичьего лейкоза, лентивирус, вирус иммунодефицита человека, вирус миелопролиферативной саркомы и вирус опухоли молочной железы; и т. п. В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный аденоассоциированный вирусный (AAV) вектор. В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный лентивирусный вектор. В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный ретровирусный вектор.

[0244] В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный аденоассоциированный вирусный (AAV) вектор. В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный лентивирусный вектор. В некоторых вариантах осуществления рекомбинантный экспрессионный вектор настоящего описания представляет собой рекомбинантный ретровирусный вектор.

[0245] AAV представляет собой небольшой (20 нм) непатогенный вирус, который можно использовать для лечения заболеваний человека в ситуациях, когда вирусный вектор используется для доставки в клетку, такую как эукариотическая клетка, *in vivo* или *ex vivo* с целью подготовки клеток для введения субъекту. Создают конструкт, например конструкт, кодирующий любой из белков CasX и/или вариантов осуществления CasX и гРНК, как описано в настоящем документе, и фланкируют последовательностями инвертированных концевых повторов AAV (ITR), таким образом обеспечивая упаковку вектора AAV в вирусную частицу AAV.

[0246] Вектор «AAV» может относиться к встречающемуся в природе вирусу дикого типа или его производным. Термин охватывает все подтипы, серотипы и псевдотипы, и как природные, так и рекомбинантные формы, за исключением случаев с другими требованиями. В настоящем документе термин «серотип» относится к AAV, который идентифицируется и отличается от других AAV на основе реактивности белка капсида с определенной антисывороткой, например существует множество известных серотипов AAV приматов. В некоторых вариантах осуществления вектор AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV 44.9, AAV-Rh74 (AAV, полученный от макака-резуса) и AAVRh10 и модифицированных капсидов этих серотипов. Например, серотип AAV-2 означает AAV, который содержит капсидные белки, кодируемые геном Cap AAV-2, и геном содержит последовательности 5'- и 3'-ITR из того же серотипа AAV-2. Псевдотипом AAV называется AAV, который

содержит капсидные белки из одного серотипа, а вирусный геном, включающий 5'-3' ITR, - от второго серотипа. Ожидается, что псевдотипичный rAAV будет иметь свойства связывания с клеточной поверхностью, характерные для серотипа капсида, и генетические свойства, соответствующие серотипу ITR. Псевдотипичный рекомбинантный AAV (rAAV) получают с применением стандартных методик, описанных в данной области. В настоящем документе, например, термин «rAAV1» может применяться для обозначения AAV, имеющего как белки капсида, так и 5'-3' ITR из одного серотипа, или он может относиться к AAV, имеющему белки капсида из серотипа 1 и 5'-3' ITR из другого серотипа AAV, например AAV серотипа 2. В каждом примере, представленном в настоящем документе, в описании конструкции и продукции вектора приводится серотип капсида и последовательностей 5'-3' ITR.

[0247] Термин «вирус AAV» или «вирусная частица AAV» относится к вирусной частице, состоящей из по меньшей мере одного капсидного белка AAV (предпочтительно всех капсидных белков AAV дикого типа) и инкапсулированного полинуклеотида. Если частица дополнительно содержит гетерологичный полинуклеотид (т. е. полинуклеотид, отличный от генома AAV дикого типа, который должен быть доставлен в клетку млекопитающего), она, как правило, называется «rAAV». Пример гетерологичного полинуклеотида представляет собой полинуклеотид, содержащий белок CasX и/или оgРНК, и необязательно донорную матрицу по любому из вариантов осуществления, описанных в настоящем документе.

[0248] Под «инвертированными концевыми повторами аденоассоциированного вируса», или «ITR AAV», понимаются области, известные в данной области, находящиеся на каждом конце генома AAV, которые в совокупности функционируют как цис-активные в качестве точки начала репликации ДНК и в качестве сигналов паковки вируса. ITR AAV вместе с кодирующей областью гер AAV обеспечивают эффективное вырезание и извлечение, а также интеграцию нуклеотидной последовательности, расположенной между двумя фланкирующими ITR, в геном клетки млекопитающего. Нуклеотидные последовательности областей ITR AAV известны. См., например, Kotin, R.M. (1994) *Human Gene Therapy* 5:793-801; Berns, K. I. "Parvoviridae and their Replication" in *Fundamental Virology*, 2nd Edition, (B. N. Fields and D. M. Knipe, eds.). В настоящем документе ITR AAV не обязательно должен иметь изображенную нуклеотидную последовательность дикого типа, она может быть изменена, например, посредством вставки, делеции или замены нуклеотидов. Кроме того, ITR AAV могут быть получены из любого из нескольких серотипов AAV, включая, без ограничений, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 и AAVRh10, и модифицированных капсидов этих серотипов. Более того, 5' и 3'-ITR, которые фланкируют выбранную нуклеотидную последовательность в векторе AAV, не обязательно должны быть идентичными или полученными из одного и того же серотипа или изолята AAV, при условии, что они функционируют так, как ожидается, т. е. обеспечивают вырезание и выделение представляющей интерес последовательности из генома или вектора клетки-хозяина и

обеспечивают интеграцию гетерологичной последовательности в геном клетки-реципиента, когда в клетке присутствуют продукты гена Rep AAV. Применение серотипов AAV для интеграции гетерологичных последовательностей в клетку-хозяина известно в данной области (см., например, публикации WO2018195555A1 и US20180258424A1, включенные в настоящий документ путем ссылки).

[0249] Под термином «кодирующая область AAV Rep» подразумевается область генома AAV, кодирующая репликационные белки Rep 78, Rep 68, Rep 52 и Rep 40. Было показано, что продукты экспрессии Rep обладают многими функциями, включая распознавание, связывание и никирование точки начала репликации ДНК AAV, ДНК-геликазной активностью и модуляцией транскрипции с промоторов AAV (или других гетерологичных промоторов). Продукты экспрессии Rep в совокупности требуются для репликации генома AAV. Под термином «область кодирования AAV Cap» подразумевается область генома AAV, кодирующая капсидные белки VP1, VP2 и VP3 или их функциональные гомологи. Эти продукты экспрессии гена Cap обеспечивают функции упаковки, которые в совокупности необходимы для упаковки вирусного генома.

[0250] В некоторых вариантах осуществления капсиды AAV, используемые для доставки кодирующих последовательностей для CasX и гРНК и необязательно донорных матриц DMPK в клетку-хозяина, могут быть получены из любого из нескольких серотипов AAV, включая, без ограничений, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV 44.9, AAV-Rh74 (AAV из макака-резуса) и AAVRh10, а AAV ITR получают из AAV серотипа 2. В конкретном варианте осуществления AAV1, AAV7, AAV6, AAV8 или AAV9 используют для доставки CasX, гРНК и необязательно донорных матричных нуклеотидов в мышечную клетку-хозяина.

[0251] Для получения вирусных частиц гAAV экспрессионный вектор AAV вводят в приемлемую клетку-хозяина с применением известных методик, таких как трансфекция. Для формирования частиц вируса обычно применяют пакующие клетки; к таким клеткам относятся клетки HEK293 (и другие клетки, известные в данной области), которые пакут аденовирус. Ряд методик трансфекции общеизвестен в данной области; см., например, Sambrook et al. (1989) *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratories, New York. Особенно подходящие способы трансфекции включают соперецитацию с фосфатом кальция, прямую микроинъекцию в культивируемые клетки, электропорацию, опосредованный липосомами перенос гена, опосредованную липидами трансдукцию и доставку нуклеиновой кислоты с применением бомбардировки микрочастицами.

[0252] В качестве преимущества конструкторов гAAV настоящего описания меньший размер нуклеаз CRISPR типа V; например, CasX по вариантам осуществления позволяет включать все необходимые компоненты для редактирования и вспомогательные компоненты для экспрессии в трансген таким образом, чтобы одна частица гAAV могла доставлять и трансдуцировать эти компоненты в клетку-мишень в форме, которая приводит к экспрессии нуклеазы CRISPR и гРНК, способным эффективно модифицировать целевую нуклеиновую кислоту клетки-мишени. Типичная схема такого конструктора представлена на

ФИГ. 13. Этим она заметно отличается от других систем CRISPR, таких как Cas9, где обычно используется система с двумя частицами для доставки необходимых компонентов для редактирования в целевую клетку. Таким образом, в некоторых вариантах осуществления систем гAAV в описании предложены: i) первая плаزمид, содержащая ITR, последовательности, кодирующие вариант CasX, последовательности, кодирующие одну или более гРНК, первый промотор, функционально связанный с CasX, и второй промотор, функционально связанный с гРНК, и необязательно один или более энхансерных элементов; ii) вторая плазмид, содержащая гены гер и сар; и iii) третья плазмид, содержащая хелперные гены, причем при трансфекции соответствующей пакующей клетки клетка способна продуцировать гAAV, обладающий способностью доставлять в клетку-мишень в одной частице последовательности, способные экспрессировать нуклеазу CasX и гРНК, имеющие способность редактировать целевую нуклеиновую кислоту клетки-мишени. В некоторых вариантах осуществления систем гAAV последовательность, кодирующая белок CRISPR, и последовательность, кодирующая по меньшей мере первую гРНК, имеют длину менее чем примерно 3100, менее чем примерно 3090, менее чем примерно 3080, менее чем примерно 3070, менее чем примерно 3060, менее чем примерно 3050 или менее чем примерно 3040 нуклеотидов в длину, так что последовательности, кодирующие первый и второй промотор, и необязательно один или более энхансерных элементов могут иметь в общей сложности длину по меньшей мере примерно 1300, по меньшей мере примерно 1350, по меньшей мере примерно 1360, по меньшей мере примерно 1370, по меньшей мере примерно 1380, по меньшей мере примерно 1390 нуклеотидов, по меньшей мере примерно 1400, по меньшей мере примерно 1500, по меньшей мере примерно 1600 нуклеотидов, по меньшей мере примерно 1650, по меньшей мере примерно 1700, по меньшей мере примерно 1750, по меньшей мере примерно 1800, по меньшей мере примерно 1850 или по меньшей мере примерно 1900 нуклеотидов. В некоторых вариантах осуществления систем гAAV последовательность, кодирующая первый промотор и по меньшей мере один вспомогательный элемент, имеет в общей сложности длину более чем по меньшей мере примерно 1300, по меньшей мере примерно 1350, по меньшей мере примерно 1360, по меньшей мере примерно 1370, по меньшей мере примерно 1380, по меньшей мере примерно 1390, по меньшей мере примерно 1400, по меньшей мере примерно 1500, по меньшей мере примерно 1600 нуклеотидов, по меньшей мере примерно 1650, по меньшей мере примерно 1700, по меньшей мере примерно 1750, по меньшей мере примерно 1800, по меньшей мере примерно 1850 или по меньшей мере примерно 1900 нуклеотидов. В некоторых вариантах осуществления систем гAAV последовательность, кодирующая первый и второй промоторы и по меньшей мере один вспомогательный элемент, имеет в общей сложности длину более чем по меньшей мере примерно 1300, по меньшей мере примерно 1350, по меньшей мере примерно 1360, по меньшей мере примерно 1370, по меньшей мере примерно 1380, по меньшей мере примерно 1390, по меньшей мере примерно 1400, по меньшей мере примерно 1500, по меньшей мере примерно 1600 нуклеотидов, по меньшей мере примерно 1650, по меньшей мере примерно 1700, по меньшей мере примерно 1750, по меньшей мере примерно 1800, по меньшей мере примерно 1850, по меньшей мере примерно 1900 нуклеотидов.

1800, по меньшей мере примерно 1850 или по меньшей мере примерно 1900 нуклеотидов.

[0253] В некоторых вариантах осуществления клетки-хозяева, трансфицированные вышеописанными экспрессионными векторами AAV, становятся способными обеспечивать хелперные функции AAV для репликации и заключения в капсид нуклеотидных последовательностей, фланкированных ITR AAV, с получением вирусных частиц rAAV. Хелперные функции AAV обычно представляют собой кодирующие последовательности, полученные из AAV, которые могут быть экспрессированы с получением генных продуктов AAV, которые, в свою очередь, функционируют как транс-активные, обеспечивая продуктивную репликацию AAV. В настоящем документе хелперные функции AAV применяются для дополнения необходимых функций AAV, которые отсутствуют в экспрессионных векторах AAV. Таким образом, хелперные функции AAV включают одну или обе из основных ORF (открытых рамок считывания) AAV, которые кодируют кодирующие области *тер* и *сар*, или их функциональные гомологи. Вспомогательные функции могут быть введены в клетки-хозяева и затем экспрессированы способами, известными специалистам в данной области. Обычно вспомогательные функции обеспечиваются посредством инфицирования клеток-хозяев неродственным хелперным вирусом. В некоторых вариантах осуществления вспомогательные функции обеспечиваются с применением вспомогательного функционального вектора. В зависимости от используемой системы хозяин/вектор в экспрессионном векторе можно применять любое число приемлемых элементов управления транскрипцией и трансляцией, включая конститутивные и индуцибельные промоторы, транскрипционные энхансерные элементы, терминаторы транскрипции и т. д. В некоторых вариантах осуществления в описании предложены клетки-хозяева, содержащие векторы AAV по вариантам осуществления, описанным в настоящем документе.

[0254] В других вариантах осуществления приемлемые векторы могут включать вирусоподобные частицы (VLP). Вирусоподобные частицы (VLP) представляют собой частицы, которые очень похожи на вирусы, но не содержат вирусного генетического материала и, следовательно, не являются инфекционными. В некоторых вариантах осуществления VLP содержат полинуклеотид, кодирующий интересующий трансген, например любой из белка CasX и/или вариантов осуществления гРНК, и необязательно полинуклеотиды донорных матриц, описанные в настоящем документе, упакованные с одним или более вирусными структурными белками. В других вариантах осуществления в описании предложены XDP, полученные *in vitro*, которые содержат RNP-комплекс CasX:гРНК и необязательно донорную матрицу. Для создания XDP можно применять комбинации структурных белков из разных вирусов, включая компоненты из семейств вирусов, включая Parvoviridae (например, аденоассоциированный вирус), Retroviridae (например, альфа-ретровирус, бета-ретровирус, гамма-ретровирус, дельта-ретровирус, эпсилон-ретровирус или лентивирус), Flaviviridae (например, вирус гепатита С), Paramyxoviridae (например, вирус Нипах) и бактериофаги (например, Q β , AP205). В некоторых вариантах осуществления в описании предложены системы XDP, разработанные

с применением компонентов ретровируса, включая лентивирусы (такие как ВИЧ) и альфа-ретровирус, бета-ретровирус, гамма-ретровирус, дельта-ретровирус, эпсилон-ретровирус, в которых отдельные плазмиды, содержащие полинуклеотиды, кодирующие различные компоненты, вводят в пакующую клетку, которая, в свою очередь, продуцирует XDP. В некоторых вариантах осуществления в описании предложена XDP, содержащая один или более компонентов из i) протеазы, ii) сайта расщепления протеазой, iii) одного или более компонентов полипротеина gag, выбранных из матриксного белка (МА), нуклеокапсидного белка (NC), капсидного белка (CA), пептида p1, пептида p6, пептида P2A, пептида P2B, пептида P10, пептида p12, пептида PP21/24, пептида P12/P3/P8 и пептида P20; v) CasX; vi) гРНК и vi) нацеливающих гликопротеинов или фрагментов антител, причем полученная XDP-частица включает в капсид RNP-комплекс CasX:гРНК. Полинуклеотиды, кодирующие Gag, CasX и гРНК, могут дополнительно содержать парные компоненты, предназначенные для содействия перемещению компонентов из ядра клетки-хозяина и в отпочковывающуюся XDP. Не имеющие ограничительного характера примеры таких транспортировочных компонентов включают шпильчатую РНК, такую как шпилька MS2, шпилька PP7, шпилька Q β и U1-шпилька II, которые имеют аффинность связывания к белку оболочки MS2, белку оболочки PP7, белку оболочки Q β и частице распознавания сигнала U1A соответственно. В других вариантах осуществления гРНК может содержать rev-отвечающий элемент (RRE) или его участки, которые имеют аффинность связывания с Rev, который может быть связан с полипротеином Gag. В других вариантах осуществления гРНК может содержать один или более RRE и одну или более последовательностей шпильки MS2. В других вариантах осуществления гРНК может содержать rev-отвечающий элемент (RRE) или его участки, которые имеют аффинность связывания с Rev, который может быть связан с полипротеином Gag. RRE может быть выбран из группы, состоящей из стебля IIВ rev-отвечающего элемента (RRE), стебля II-V RRE, стебля II RRE, rev-связывающего элемента (RBE) стебля IIВ и полноразмерного RRE. В вышеуказанном варианте осуществления компоненты включают последовательности UGGGCGCAGCGUCA AUGACGCUGACGGUACA (стебель IIВ; SEQ ID NO: 27266), GCACUAUGGGGCGCAGCGUCA AUGACGCUGACGGUACAGGCCAGACAAUUAUUGUC UGGUAUAGUGC (стебель II; SEQ ID NO: 27267), GCUGACGGUACAGGC (RBE, SEQ ID NO: 27268), CAGGAAGCACUAUGGGGCGCAGCGUCA AUGACGCUGACGGUACAGGCCAGACAAUUAUUGUCUGGUUAUAGUGCAGCAGCAGACAACAAUUAUUGUCUGAGGGGCUAUUGAGGCGCAACAGCAUCUGUUGCAACUCACAGUCUGGGGCAUCAAGCAGCUCCAGGCAAGAAUC CUG (стебель II-V; SEQ ID NO: 27269) и AGGAGCUUUGUCCUUGGGUUCUUGGGAGCAGCAGGAAGCACUAUGGGGCGCAGCGUCA AUGACGCUGACGGUACAGGCCAGACAAUUAUUGUCUGGUUAUAGUGCAGCAGCAGACAACAAUUAUUGUCUGAGGGGCUAUUGAGGCGCAACAGCAUCUGUUGCAACUCACAGUCUGGGGCAUCAAGCAGCUCCAGGCAAGAAUCCUGGCUGUGGAAAGAUACCUAAAGGAUCAACAGCUCCU (полноразмерный RRE; SEQ ID NO: 27270). В других

вариантах осуществления гРНК может содержать один или более RRE и одну или более последовательностей шпильки MS2. В конкретном варианте осуществления гРНК содержит вариант шпильки MS2, который оптимизирован для повышения аффинности связывания с белком оболочки MS2, улучшая таким образом встраивание гРНК и ассоциированного CasX в отпочковывающийся XDP. Варианты гРНК, содержащие варианты шпильки MS2, включают варианты гРНК 275-315 и 317-320 (SEQ ID NO: 2722-27264).

[0255] Нацеливающие гликопротеины или фрагменты антител на поверхности, которые обеспечивает тропизм XDP к клетке-мишени, причем при введении и входе в клетку-мишень молекула RNP свободно транспортируется в ядро клетки. Гликопротеин оболочки может быть получен из любых оболочечных вирусов, которые, как известно в данной области, придают XDP тропизм, включая, без ограничений, группу, состоящую из вируса аргентинской геморрагической лихорадки, вируса австралийских летучих мышей, вируса множественного ядерного полиэндроза *Autographa californica*, вируса птичьего лейкоза, эндогенного вируса бабуина, вируса боливийской геморрагической лихорадки, вируса болезни Борна, вируса Бреда, вируса Буньямвера, вируса Чандипура, вируса Чикунгунья, вируса конго-крымской геморрагической лихорадки, вируса лихорадки Денге, вируса Дювенажа, вируса восточного энцефалита лошадей, вируса геморрагической лихорадки Эбола, вируса Эбола Заира, кишечного аденовируса, эфемеровируса, вируса Эпштейна - Барр (EBV), вируса 1 европейских летучих мышей, вируса 2 европейских летучих мышей, продукта синтетического слияния Fug и gP, вируса лейкоза обезьян-гиббонов, хантавируса, вируса Хендра, вируса гепатита А, вируса гепатита В, вируса гепатита С, вируса гепатита D, вируса гепатита Е, вируса гепатита G (вируса GB C), вируса простого герпеса типа 1, вируса простого герпеса типа 2, цитомегаловируса человека (HHV5), пенящего вируса человека, герпесвируса человека (HHV), герпесвируса человека типа 7, герпесвируса человека типа 6, герпесвируса человека типа 8, вируса иммунодефицита человека 1 (ВИЧ-1), метапневмовируса человека, Т-лимфотропного вируса 1 человека, вируса гриппа А, гриппа В, гриппа С, вируса японского энцефалита, вируса герпеса, ассоциированного с саркомой Капоши (HHV8), вируса Кьясанурской лесной болезни, вируса Ла-Кросс, вируса лагосских летучих мышей, вируса лихорадки Ласса, вируса лимфоцитарного хориоменингита (LCMV), вируса Мачупо, вируса геморрагической лихорадки Марбурга, вируса кори, коронавируса, связанного с ближневосточным респираторным синдромом, вируса Мокола, вируса мышиноного лейкоза Молони, вируса оспы обезьян, вируса опухоли молочной железы мыши, вируса эпидемического паротита, мышиноного гамма-герпесвируса, вируса болезни Ньюкасла, вируса Нипах, вируса Нипах, норовируса, вируса омской геморрагической лихорадки, вируса папилломы, парвовируса, вируса псевдобешенства, вируса Quagranfil, вируса бешенства, эндогенного ретровируса кошачьих RD114, респираторно-синцитиального вируса (RSV), вируса лихорадки долины Рифт, вируса реки Росс, ротавируса, вируса саркомы Рауса, вируса краснухи, вируса геморрагической лихорадки, ассоциированной с

вирусом Сабиа, SARS-ассоциированного коронавируса (SARS-CoV), вируса Сендай, вируса Такарибе, тогтовируса, вируса, вызывающего клещевой энцефалит, вируса ветряной оспы (HHV3), вируса ветряной оспы (HHV3), вируса натуральной оспы, вируса малой оспы, вируса венесуэльского энцефалита лошадей, вируса венесуэльской геморрагической лихорадки, вируса везикулярного стоматита (VSV), VSV-G, везикуловируса, вируса Западного Нила, вируса западного энцефалита лошадей и вируса Зика.

[0256] В других вариантах осуществления в описании предложена XDP согласно вышеизложенному, дополнительно содержащая один или более компонентов полипротеина pol (например, протеазу) и необязательно второй CasX или донорную матрицу. В описании предусмотрено множество конфигураций расположения закодированных компонентов, включая дубликаты некоторых закодированных компонентов. Вышеизложенное обеспечивает преимущества по сравнению с другими векторами в данной области тем, что вирусная трансдукция в делящиеся и неделящиеся клетки является эффективной и что XDP доставляет эффективный и короткий RNP, который избегает механизмов иммунного надзора субъекта, которые в противном случае обнаружили бы чужеродный белок. Не имеющие ограничительного характера примеры систем XDP описаны в PCT/US20/63488 и WO2021113772A1, включенных в настоящий документ путем ссылки. В некоторых вариантах осуществления в описании предложены клетки-хозяева, содержащие полинуклеотиды или векторы, кодирующие любые из вышеуказанных вариантов осуществления XDP.

[0257] После получения и извлечения XDP, содержащей RNP CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе, XDP можно применять в способах редактирования клеток-мишеней субъектов посредством введения такого XDP, как более подробно описано ниже.

VII. Клетки

[0258] В другом аспекте в настоящем документе предложены популяции клеток, содержащих ген BCL11A, модифицированный *ex vivo* в соответствии с вариантами осуществления любых из систем или способов, описанных в настоящем документе. В некоторых вариантах осуществления клетки, которые были генетически модифицированы таким образом, можно вводить субъекту для таких целей, как генная терапия; например, в способах лечения связанного с гемоглобинопатией заболевания, такого как серповидноклеточная анемия или бета-талассемия, в которых введение приводит к повышению экспрессии γ -глобина и повышению эмбрионального гемоглобина (HbF) у субъекта. В других вариантах осуществления в описании предложены композиции модифицированных клеток для применения в качестве лекарственного средства при лечении связанного с гемоглобинопатией заболевания.

[0259] Клетки настоящего описания, приемлемые для модификации *ex vivo* гена BCL11A посредством нуклеазы Cas класса 2, типа V и одной или более гидовых РНК, нацеленных на целевую нуклеиновую кислоту BCL11A, могут представлять собой

гематопозитическую клетку-предшественника (HPC), гематопозитическую стволовую клетку (HSC), клетку CD34+, мезенхимальную стволовую клетку (MSC), индуцированную плюрипотентную стволовую клетку (iPSC), общую миелоидную клетку-предшественника, клетку-проэритробласт или клетку-эритробласт. В некоторых вариантах осуществления популяция модифицированных клеток представляет собой клетки животного происхождения; например, полученные из грызуна, крысы, мыши, кролика, собаки, или клетки примата, не являющегося человеком; например, клетка обезьяны яванского макака. В некоторых вариантах осуществления клетка представляет собой клетку человека. В некоторых случаях клетки, подлежащие модификации, являются аутологичными по отношению к субъекту, которому вводят клетки. В других случаях клетки являются аллогенными по отношению к субъекту, которому вводят клетки. В некоторых случаях клетка выделена *ex vivo* из костного мозга или из периферической крови субъекта.

[0260] В некоторых вариантах осуществления в описании предложены способы и популяции клеток, модифицированные посредством введения в каждую клетку популяции: i) системы CasX:гРНК, содержащей CasX и гРНК по любому из вариантов осуществления, описанных в настоящем документе; ii) системы CasX:гРНК, содержащей CasX, гРНК и донорную матрицу по любому из вариантов осуществления, описанных в настоящем документе; iii) нуклеиновой кислоты, кодирующей CasX и гРНК и необязательно содержащей донорную матрицу; iv) вектора, содержащего нуклеиновую кислоту по п. (iii) выше, который может представлять собой AAV по любому из вариантов осуществления, описанных в настоящем документе; v) XDP, содержащей систему CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе; или vi) комбинаций двух или более из пп. (i)-(v), причем последовательность целевой нуклеиновой кислоты BCL11A клеток, на которую нацелена гРНК, модифицирована белком CasX и необязательно донорной матрицей. В некоторых вариантах осуществления способ дополнительно включает введение второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку последовательности целевой нуклеиновой кислоты по сравнению с первой гРНК. В некоторых случаях CasX и гРНК доставляют в клетки популяции в виде RNP (варианты осуществления которых описаны в настоящем документе выше) и необязательно донорной матрицы. В некоторых вариантах осуществления в описании предложена популяция клеток, модифицированных вышеуказанными способами, причем клетки были модифицированы таким образом, что по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не демонстрируют обнаружимой экспрессии белка BCL11A. В других вариантах осуществления в описании предложена популяция клеток, в которой клетки были модифицированы таким образом, что экспрессия белка BCL11A снижена на по меньшей мере 70%, на по меньшей мере 75%, на по меньшей мере 80%, на по меньшей мере 85%, на по меньшей мере 90% или на по меньшей мере 95% по сравнению с клетками, которые не

были модифицированы. В других вариантах осуществления в описании предложена популяция клеток, в которых экспрессия белка BCL11A не может быть обнаружена в модифицированных клетках популяции. Эффекты модификации можно оценивать методами вестерн-блоттинга, проточной цитометрии, твердофазной ИФА, клеточных анализов, кОТ-ПЦР, электрохемилюминесцентных анализов, смысловые транскрипты можно анализировать с помощью анализа флуоресцентной гибридизации РНК *in situ* (FISH) или другими способами, известными в данной области, или так, как описано в разделе «Примеры».

[0261] В некоторых вариантах осуществления в описании предложены способы модификации целевой нуклеиновой кислоты BCL11A в популяции клеток способами *in vitro* или *ex vivo*. В способе предусмотрено, что клетки могут быть получены от субъекта с применением любого числа методик, известных специалисту в данной области; например, посредством биопсии костного мозга или получения пробы периферической крови. Желаемые клетки можно отделять от остальной части пробы, отмывать для удаления жидкостей и остатков и необязательно помещать в подходящий буфер или среду для последующих стадий обработки. Способ может включать одну или более стадий из i) введения в клетки компонентов системы CasX:гРНК для редактирования целевых нуклеиновых кислот; ii) введения в клетки нуклеиновой кислоты или вектора, кодирующего компоненты системы CasX:гРНК; iii) размножения клеток в подходящей среде в условиях, приемлемых для их размножения, и iv) криоконсервации клеток для последующего введения субъекту. Таким образом, системы и способы CasX:гРНК, описанные в настоящем документе, можно применять для модификации различных клеток, связанных с гемоглобинопатией, для получения популяций клеток, в которых экспрессия белка BCL11A снижена или устранена, а HbF повышен. Таким образом, этот подход может применяться для способов лечения субъекта с гемоглобинопатией, такого как, помимо прочего, серповидноклеточная анемия или бета-талассемии. В некоторых случаях клетки приводят в контакт с CasX и гРНК, причем гРНК представляет собой гидовую РНК (гРНК). В других случаях клетки приводят в контакт с CasX и гРНК, причем гРНК представляет собой химеру, содержащую ДНК и РНК. Как описано в настоящем документе, в вариантах осуществления любой из комбинаций каждая из упомянутых молекул гРНК (комбинация каркасной и нацеливающей последовательности, которая может быть выполнена в виде огРНК или дгРНК) может быть представлена в виде RNP с вариантом осуществления CasX, описанным в настоящем документе, для включения в клетки, подлежащие модификации. В одном варианте осуществления целевую нуклеиновую кислоту клетки модифицируют посредством приведения клеток в контакт с белком CasX, гидовой нуклеиновой кислотой (гРНК), содержащей нацеливающую последовательность, комплементарную целевой нуклеиновой кислоте BCL11A, и донорной матрицей, причем донорная матрица вставляется в участок последовательности целевой нуклеиновой кислоты клетки или заменяет его таким образом, что белок BCL11A не экспрессируется или экспрессируется на сниженном уровне. В других случаях CasX и гРНК доставляют в клетки популяции в

векторе (варианты осуществления которых описаны в настоящем документе выше), причем целевая нуклеиновая кислота модифицирована таким образом, что белок BCL11A не экспрессируется или экспрессируется на сниженном уровне.

[0262] В некоторых вариантах осуществления клетки популяции приводят в контакт с вариантом CasX, содержащим последовательность из таблицы 4, или последовательность, которая на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична или на по меньшей мере 99,5% идентична ей, каркас гРНК содержит последовательность из таблицы 3, или последовательность, которая на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 81% идентична, на по меньшей мере 82% идентична, на по меньшей мере 83% идентична, на по меньшей мере 84% идентична, на по меньшей мере 85% идентична, на по меньшей мере 86% идентична, на по меньшей мере 86% идентична, на по меньшей мере 87% идентична, на по меньшей мере 88% идентична, на по меньшей мере 89% идентична, на по меньшей мере 89% идентична, на по меньшей мере 90% идентична, на по меньшей мере 91% идентична, на по меньшей мере 92% идентична, на по меньшей мере 93% идентична, на по меньшей мере 94% идентична, на по меньшей мере 95% идентична, на по меньшей мере 96% идентична, на по меньшей мере 97% идентична, на по меньшей мере 98% идентична, на по меньшей мере 99% идентична, на по меньшей мере 99,5% идентична с ней, и гРНК содержит нацеливающую последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789 из таблицы 1 или последовательности, которая на по меньшей мере 65% идентична, на по меньшей мере 70% идентична, на по меньшей мере 75% идентична, на по меньшей мере 80% идентична, на по меньшей мере 85% идентична, на по меньшей мере 90% или на по меньшей мере 95% идентична ей и содержит от 15 до 20 аминокислот. В других случаях CasX и одну или более гРНК вводят в популяцию клеток в виде кодирующих полинуклеотидов с помощью вектора, варианты осуществления которого описаны в настоящем документе. Дополнительные способы модификации клеток с применением компонентов системы CasX:гРНК включают вирусную инфекцию, трансфекцию, конъюгирование, слияние протопластов, технологию генной пушки, преципитацию фосфатом кальция, прямую микроинъекцию и т. п. Выбор способа по существу зависит от типа трансформируемой клетки и обстоятельств, при

которых происходит трансформация; например, *in vitro*, *ex vivo* или *in vivo*. Общее описание этих способов можно найти в Ausubel, et al, Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons, 1995.

[0263] При гибридизации с целевой нуклеиновой кислотой CasX и гРНК CasX вводит один или более одноцепочечных разрывов или двухцепочечных разрывов в ген BCL11A, что приводит к модификации целевой нуклеиновой кислоты, такой как постоянный индел (делеция или вставка) или другая мутация (например, замена, дублирование или инверсия) в целевой нуклеиновой кислоте, в связи с механизмами репарации клетки-хозяина, что приводит к соответствующему снижению или устранению экспрессии функционального белка BCL11A, создавая таким образом модифицированную популяцию клеток. Как описано в настоящем документе, вариант CasX, вводящий расщепление двухцепочечных молекул в целевую нуклеиновую кислоту, создает двухцепочечный разрыв в пределах 18-26 нуклеотидов с 5'-стороны от сайта PAM на целевой цепи и 10-18 нуклеотидов с 3'-стороны на нецелевой цепи. Таким образом, в некоторых вариантах осуществления полученная в результате осуществления способа модификация может приводить к случайным вставкам или делециям (инделам) или замене, дублированию или инверсии одного или более нуклеотидов в этой области с помощью репарационных механизмов соединения нехомологичных концов ДНК (NHEJ).

[0264] В некоторых вариантах осуществления способа модификации популяции клеток первая гРНК содержит нацеливающую последовательность, комплементарную последовательности, находящейся проксимально к или в пределах любого из экзонов гена BCL11A. В одном варианте осуществления первая гРНК содержит нацеливающую последовательность, комплементарную последовательности, находящейся проксимально к или смежной с любым из регуляторных элементов гена BCL11A. В конкретном варианте осуществления первая гРНК содержит нацеливающую последовательность, комплементарную последовательности внутри или смежную с 5'-стороной последовательности мотива связывания с GATA1 гена BCL11A. В конкретном варианте осуществления нацеливающая последовательность представляет собой SEQ ID NO: 22. Таким образом, нарушение последовательности целевой нуклеиновой кислоты приводит к модификации гена BCL11A, так что экспрессия функционального белка BCL11A в модифицированных клетках популяции снижается или устраняется.

[0265] В некоторых вариантах осуществления способа целевая нуклеиновая кислота клеток популяции модифицирована с помощью множества гРНК (например, двух, трех, четырех или более), нацеленных на разные или перекрывающиеся участки гена BCL11A, причем белок CasX вводит множество разрывов в последовательность целевой нуклеиновой кислоты, которые приводят к созданию постоянного индела (делеции или вставки) или других мутаций (например, замены, дублирования или инверсии одного или более нуклеотидов), так что в модифицированных клетках популяции экспрессия функционального белка BCL11A снижается или устраняется.

[0266] В других вариантах осуществления в описании предложены популяции

клеток, модифицированные посредством приведения клетки в контакт с белком CasX, одной или более гРНК, содержащими нацеливающую последовательность, и донорной матрицей по любому из вариантов осуществления, описанных в настоящем документе, причем донорная матрица вставляется в сайты разрыва, введенные нуклеазой, заменяя полностью или частично последовательность целевой нуклеиновой кислоты гена BCL11A, подлежащего модификации. В одном варианте осуществления вышеизложенного донорная матрица содержит по меньшей мере участок экзона BCL11A и одну или более мутаций, и модификация клетки приводит к модификации гена, так что в модифицированных клетках популяции экспрессия функционального белка BCL11A снижается или устраняется. В другом варианте осуществления вышеизложенного донорная матрица содержит последовательность, находящуюся в пределах или смежно с 5'-стороны с последовательностью мотива связывания с GATA1, но имеет одну или более мутаций относительно последовательности дикого типа, и модификация клетки приводит к снижению или устранению экспрессии функционального белка BCL11A в модифицированных клетках популяции. Следует понимать, что в таких случаях заменяющие донорные матрицы в 5'- и 3'-направлении больше, чем позиции сайтов расщепления, вводимых нуклеазой в конкретных участках целевой нуклеиновой кислоты, подлежащей замене, и могут дополнительно содержать гомологичные плечи с 5'- и 3'-сторон от сайтов расщепления, введенных нуклеазой, для облегчения ее вставки. В некоторых случаях донорная матрица представляет собой одноцепочечную ДНК-матрицу или одноцепочечную РНК-матрицу. В других случаях донорная матрица представляет собой двухцепочечную ДНК-матрицу. Вставка донорной матрицы в целевую область, которая может быть опосредована направляемой гомологией репарацией (HDR, как описано выше) или независимой от гомологии нацеленной интеграцией (НТИ). Экзогенная последовательность, вставленная посредством НТИ, может иметь любую длину, например, относительно короткая последовательность длиной от 10 до 50 нуклеотидов или более длинная последовательность длиной примерно 50-1000 нуклеотидов. Последовательность донорной матрицы может содержать определенные различия последовательности по сравнению с геномной последовательностью, например сайты рестрикции, нуклеотидные полиморфизмы, штрихкоды, маркеры селекции (например, гены резистентности к лекарственному средству, флуоресцентные белки, ферменты и т. д.) и т. д., которые можно применять для оценки успешной вставки донорной нуклеиновой кислоты в сайт расщепления или, в некоторых случаях, можно применять для других целей (например, чтобы обозначить экспрессию в целевом геномном локусе). Альтернативно эти различия в последовательности могут включать фланкирующие последовательности рекомбинации, такие как FLP, последовательности loxP или т. п., которые могут быть активированы позднее для удаления маркерной последовательности.

[0267] В некоторых вариантах осуществления способа модификации популяции клеток способ дополнительно включает приведение в контакт последовательности целевой нуклеиновой кислоты BCL11A популяции клеток с: i) дополнительной нуклеазой CRISPR

и гРНК, нацеленной на другой или перекрывающийся участок целевой нуклеиновой кислоты BCL11A по сравнению с первой гРНК; ii) полинуклеотидом, кодирующим дополнительную нуклеазу CRISPR и гРНК по п. (i); iii) вектором, содержащим полинуклеотид по п. (ii); или iv) XDP, содержащей дополнительную нуклеазу CRISPR и гРНК по п. (a), причем приведение в контакт приводит к модификации гена BCL11A в другом положении в последовательности по сравнению с последовательностью, на которую нацелена первая гРНК. В одном варианте осуществления вышеизложенного дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий последовательность, отличную от белка CasX по предыдущим вариантам осуществления. В другом варианте осуществления вышеизложенного дополнительная нуклеаза CRISPR не представляет собой белок CasX и выбрана из группы, состоящей из Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas13a, Cas13b, Cas13c, Cas13d, Cas14, Cpf1, C2cl, Csn2 и вариантов их последовательностей.

[0268] В других вариантах осуществления в описании предложены способы модификации целевой нуклеиновой кислоты BCL11A в популяции клеток *in vivo* у субъекта. В одном варианте осуществления способа модификации *in vivo* способ включает введение субъекту вектора по вариантам осуществления, описанным в настоящем документе, в терапевтически эффективной дозе. В некоторых вариантах осуществления вектор вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг. В других вариантах осуществления вектор вводят субъекту в дозе от по меньшей мере примерно 1×10^5 вг/кг до по меньшей мере примерно 1×10^{16} вг/кг, или от по меньшей мере примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг, или от по меньшей мере примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг, или от по меньшей мере примерно 1×10^8 вг/кг до примерно 1×10^{14} вг/кг. В другом варианте осуществления способа модификации *in vivo* способ включает введение XDP субъекту в терапевтически эффективной дозе, причем XDP содержит CasX и гРНК в комплексе RNP и необязательно донорную матрицу (описанную более подробно выше), причем XDP имеет тропизм к клеткам-мишеням и способна доставлять RNP для редактирования гена BCL11A, как описано в настоящем документе. Варианты осуществления XDP, используемые в вышеуказанном способе редактирования, описаны в настоящем документе. В одном варианте осуществления XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14}

частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг, или от по меньшей мере примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или от по меньшей мере примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг. В вышеуказанных вариантах осуществления параграфа вектор или XDP вводят субъекту способом введения, выбранным из интрапаренохиматозного, внутривенного, внутриаартериального, внутрибрюшинного, внутрикапсульного, подкожного, внутримышечного, интраабдоминального или их комбинаций, причем способ введения представляет собой инъекцию, трансфузию или имплантацию.

VIII. Терапевтические способы

[0269] В другом аспекте настоящее описание относится к способам лечения связанного с гемоглобинопатией заболевания у нуждающегося в этом субъекта, включая, без ограничений, серповидноклеточную анемию или бета-талассемию, при которых подавление или устранение экспрессии белка BCL11A посредством модификации гена BCL11A в клетках-мишенях субъекта облегчает признаки, симптомы или эффекты заболевания, несмотря на то что субъект все еще может страдать базовым заболеванием.

[0270] Для разработки композиций для применения в способах лечения субъекта, страдающего связанным с гемоглобинопатией заболеванием, применяли ряд терапевтических стратегий. В некоторых вариантах осуществления способ включает введение субъекту с гемоглобинопатией (например, серповидноклеточной анемией или бета-талассемией) терапевтически эффективной дозы нуклеазы CRISPR класса 2, типа V и гидовой РНК, описанной в настоящем документе. В некоторых вариантах осуществления способ лечения включает введение субъекту терапевтически эффективной дозы: i) системы CasX:гРНК, содержащей первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте; ii) системы CasX:гРНК, содержащей первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте, и донорной матрицей; iii) нуклеиновой кислоты, кодирующей систему CasX:гРНК, из п. (i) или п. (ii); iv) вектора, содержащего нуклеиновую кислоту по п. (iii), который может представлять собой AAV по любому из вариантов осуществления, описанных в настоящем документе; v) XDP, содержащей систему CasX:гРНК из п. (i) или п. (ii); или vi) комбинаций двух или более из пп. (i)-(v), причем 1) ген BCL11A клеток субъекта, на который нацелена первая гРНК, модифицирован (например, проведен нокаут или нокдаун) белком CasX и необязательно донорной матрицей; и 2) это приводит к повышению продукции гемоглобина F (HbF) у субъекта. В некоторых вариантах осуществления способ лечения дополнительно включает введение второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку последовательности целевой нуклеиновой кислоты по сравнению с первой гРНК. В некоторых случаях клетки, являющиеся мишенями модификации, выбраны из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников

(NPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов. В некоторых вариантах осуществления субъект, подлежащий лечению, выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком. В другом варианте осуществления субъект представляет собой человека.

[0271] В некоторых вариантах осуществления способа лечения вектор представляет собой вектор AAV, кодирующий систему CasX:гРНК, и вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг. В других вариантах осуществления способа вектор AAV вводят субъекту в дозе от по меньшей мере примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, от по меньшей мере примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или от по меньшей мере примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг. В других вариантах осуществления способ лечения включает введение субъекту терапевтически эффективной дозы XDP, содержащей систему CasX:гРНК. В одном варианте осуществления XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг. В другом варианте осуществления XDP вводят субъекту в дозе от по меньшей мере примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или от по меньшей мере примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или от по меньшей мере примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг. В вышеуказанных вариантах осуществления параграфа вектор или XDP вводят субъекту способом введения, выбранным из интрапаренохиматозного, внутривенного, внутриартериального, внутрибрюшинного, внутрикапсульного, подкожного, внутримышечного, интраабдоминального или их комбинаций, причем способ введения представляет собой инъекцию, трансфузию или имплантацию. Введение может осуществляться один раз, два раза или может осуществляться множество раз по схеме: еженедельно, раз в две недели, ежемесячно, раз в четыре месяца или раз в шесть месяцев.

[0272] В некоторых вариантах осуществления способ лечения включает введение вектора, содержащего полинуклеотид, кодирующий CasX, и множество гРНК, нацеленных на разные или перекрывающиеся области гена BCL11A, субъекту, причем введение приводит к приведению последовательности целевой нуклеиновой кислоты в контакт с

продуктом (-ами) экспрессии векторов в клетке субъекта, и при этом ген BCL11A в клетке субъекта модифицируется. В других вариантах осуществления способов лечения способы включают введение субъекту вектора, кодирующего белок CasX и гРНК и дополнительно содержащего донорную матрицу, причем упомянутое введение приводит к модификации последовательности целевой нуклеиновой кислоты в клетке субъекта посредством расщепления белком CasX и вставки донорной матрицы в целевую нуклеиновую кислоту. В других вариантах осуществления способы включают введение субъекту первого вектора, содержащего полинуклеотид, кодирующий CasX, и множества гРНК, нацеленных на разные или перекрывающиеся последовательности гена BCL11A, и второго вектора, содержащего полинуклеотид донорной матрицы, кодирующий по меньшей мере участок или весь ген BCL11A, причем введение векторов приводит к приведению последовательности целевой нуклеиновой кислоты субъекта в клетке субъекта с продуктом (-ами) экспрессии CasX и гРНК-векторов и донорной матрицы, при этом ген BCL11A в клетке субъекта модифицируется, как описано в настоящем документе. В некоторых вариантах осуществления способов лечения вектор, вводимый субъекту, представляет собой вектор AAV, как описано в настоящем документе. В вышеизложенном вектор AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 или AAVRh10. В некоторых вариантах осуществления способов лечения вектор, вводимый субъекту, представляет собой XDP, описанную в настоящем документе, содержащую RNP системы CasX:гРНК.

[0273] В некоторых вариантах осуществления способа модификация включает введение одноцепочечного разрыва в ген BCL11A клеток популяции. В других случаях модификация включает введение двухцепочечного разрыва в ген BCL11A клеток популяции. В некоторых вариантах осуществления модификация вводит одну или более мутаций в целевой нуклеиновой кислоте BCL11A, например, вставку, делецию, замену, дублирование или инверсию одного или более нуклеотидов в гене BCL11A, причем экспрессия белка BCL11A в клетках субъекта снижается на по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% по сравнению с клеткой, которая не была модифицирована. В некоторых случаях ген BCL11A клеток субъекта модифицирован таким образом, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют обнаруживаемый уровень белка BCL11A. В других случаях способа лечения модификация приводит к увеличению продукции HbF в кровотоке субъекта на по меньшей мере примерно 5%, по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40% или по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения. В других вариантах осуществления способ приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке

субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0. В других вариантах осуществления способ приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта. В вышеуказанных вариантах осуществления субъект выбран из группы, состоящей из мыши, крысы, свиньи, примата, не являющегося человеком, и человека. Способы получения проб от получавших лечение субъектов для анализа с целью определения эффективности лечения, такого как жидкости или ткани тела, и способы подготовки проб для анализа хорошо известны специалистам в данной области. Способы анализа уровней РНК и белка описаны выше и хорошо известны специалистам в данной области. Влияние лечения также можно оценивать посредством измерения биомаркеров, связанных с экспрессией целевого гена в вышеупомянутых жидкостях, тканях или органах, собранных у животного, приведенного в контакт с одним или более соединениями согласно настоящему изобретению, с помощью стандартных клинических способов, известных в данной области. Биомаркеры связанных с гемоглобинопатией заболеваний включают, без ограничений, процентную долю серповидных клеток в кровотоке, уровни BCL11A, РНК BCL11A, уровни гемоглобина S, уровни гемоглобина-гамма и уровни гемоглобина F.

[0274] В некоторых случаях способ лечения гемоглобинопатии у субъекта дополнительно включает введение терапевтически эффективной дозы дополнительной нуклеазы CRISPR или полинуклеотида, кодирующего дополнительную нуклеазу CRISPR. В одном варианте осуществления дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий последовательность, отличную от первого CasX. В другом варианте осуществления дополнительная нуклеаза CRISPR не является белком CasX; т. е. представляет собой Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, Cas13a, Cas13b, Cas13c, Cas13d, Cas14, Cpf1, C2c1, Csn2 или вариант их последовательности. В некоторых вариантах осуществления способ лечения гемоглобинопатии у субъекта дополнительно предусматривает введение химиотерапевтического препарата.

[0275] В других вариантах осуществления в описании предложены способы лечения связанного с гемоглобинопатией заболевания у нуждающегося в этом субъекта посредством введения субъекту терапевтически эффективного количества популяции клеток, модифицированной *in vitro* или *ex vivo* композициями системы CasX:гРНК по вариантам осуществления, описанным в настоящем документе, включая i) систему CasX:гРНК, содержащую первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте; ii) систему CasX:гРНК, содержащую первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте, и донорную матрицу; iii) нуклеиновую кислоту, кодирующую систему CasX:гРНК, из п. (i) или п. (ii);

iv) вектор, содержащий нуклеиновую кислоту по п. (iii), который может представлять собой AAV по любому из вариантов осуществления, описанных в настоящем документе; v) XDP, содержащую систему CasX:гРНК из п. (i) или п. (ii); или vi) комбинации двух или более из пп. (i)-(v). В одном варианте осуществления способ лечения включает: i) выделение индуцированных плюрипотентных стволовых клеток (iPSC) или гематопоэтических стволовых клеток (HSC) у субъекта; ii) модификацию целевой нуклеиновой кислоты BCL11A в iPSC или HSC способами по любому из вариантов осуществления, описанных в настоящем документе; iii) дифференцировку модифицированных iPSC или HSC в гемопоэтическую клетку-предшественник и iv) имплантацию гемопоэтических клеток-предшественников субъекту с гемоглобинопатией, причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40% или по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения. В некоторых случаях клетки являются аутологичными по отношению к субъекту, которому вводят клетки, и выделены из костного мозга или периферической крови субъекта. В других случаях клетки являются аллогенными по отношению к субъекту, которому вводят клетки, и выделены из костного мозга или из периферической крови другого субъекта. Модифицированные клетки могут быть имплантированы субъекту посредством трансплантации, местной инъекции, системной инфузии или их комбинаций. Способы модификации клеток для введения субъекту описаны в настоящем документе, но, вкратце, модификация включает приведение клеток в контакт с: i) системой CasX:гРНК, содержащей первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте; ii) системой CasX:гРНК, содержащей первый белок CasX и первую гРНК с нацеливающей последовательностью, комплементарной целевой нуклеиновой кислоте, и донорной матрицей; iii) нуклеиновой кислотой, кодирующей систему CasX:гРНК, из п. (i) или п. (ii); iv) вектором, содержащим нуклеиновую кислоту по п. (iii), который может представлять собой AAV по любому из вариантов осуществления, описанных в настоящем документе; v) XDP, содержащей систему CasX:гРНК из п. (i) или п. (ii); или vi) комбинацией двух или более из пп. (i)-(v), причем экспрессия белка BCL11A понижается или клетка не экспрессирует обнаруживаемый уровень белка BCL11A. В некоторых вариантах осуществления способ дополнительно включает введение второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку последовательности целевой нуклеиновой кислоты по сравнению с первой гРНК. В некоторых случаях CasX и гРНК доставляют в клетки популяции в виде RNP (варианты осуществления которых описаны в настоящем документе выше) и необязательно донорной матрицы, причем целевая нуклеиновая кислота модифицирована таким образом, что белок BCL11A не экспрессируется или экспрессируется на сниженном уровне. В других случаях CasX и гРНК доставляют в клетки популяции в векторе (варианты осуществления которых

описаны в настоящем документе выше), причем целевая нуклеиновая кислота модифицирована таким образом, что белок BCL11A не экспрессируется или экспрессируется на сниженном уровне. В некоторых вариантах осуществления клетки популяции, подлежащие модификации посредством введения композиций, представляют собой эукариотические клетки, выбранные из группы, состоящей из клеток грызунов, клеток мыши, клеток крысы и клеток приматов, не являющихся человеком. В некоторых вариантах осуществления эукариотические клетки представляют собой человеческие клетки. В некоторых вариантах осуществления эукариотическая клетка выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта. В некоторых вариантах осуществления способа клетки или их потомство, вводимое субъекту, сохраняются у субъекта в течение по меньшей мере одного месяца, двух месяцев, трех месяцев, четырех месяцев, пяти месяцев, шести месяцев, семи месяцев, восьми месяцев, девяти месяцев, десяти месяцев, одиннадцать месяцев, двенадцати месяцев, тринадцати месяцев, четырнадцати месяцев, пятнадцати месяцев, шестнадцати месяцев, семнадцати месяцев, восемнадцати месяцев, девятнадцати месяцев, двадцати месяцев, двадцати одного месяца, двадцати двух месяцев, двадцати трех месяцев, двух лет, трех лет, четырех лет или пяти лет после введения субъекту. В некоторых вариантах осуществления способы лечения по описанию приводят к повышенным уровням гемоглобина F (HbF) в кровотоке на по меньшей мере примерно 5%, по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40% или по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения. В других вариантах осуществления способ приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0. В других вариантах осуществления способ приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0276] В других вариантах осуществления в описании предложены способы повышения эмбрионального гемоглобина (HbF) у субъекта, имеющего гемоглобинопатию, посредством редактирования генома, причем способ включает: i) введение субъекту эффективной дозы вектора или XDP по варианту осуществления, описанному в настоящем документе, причем вектор или XDP доставляет систему CasX:гРНК в клетки субъекта; ii) целевую нуклеиновую кислоту BCL11A клеток субъекта редактируют с помощью CasX, нацеленного первой гРНК; iii) редактирование включает введение вставки, делеции,

замены, дублирования или инверсии одного или более нуклеотидов в последовательности целевой нуклеиновой кислоты, так что экспрессия белка BCL11A снижается или устраняется, причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40% или по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения. В вышеизложенном клетки выбраны из группы, состоящей из гемопоэтических стволовых клеток (HSC), гемопоэтических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов. В одном варианте осуществления способа целевая нуклеиновая кислота клеток отредактирована таким образом, что экспрессия белка BCL11A снижена на по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% по сравнению с целевой нуклеиновой кислотой клеток, которые не были отредактированы. В некоторых случаях субъект выбран из группы, состоящей из мыши, крысы, свиньи и примата, не являющегося человеком. В других случаях субъект представляет собой человека.

[0277] В некоторых вариантах осуществления способа лечения гемоглобинопатии у субъекта способ приводит к улучшению по меньшей мере одного клинически значимого параметра, выбранного из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзивных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни. В других вариантах осуществления способа лечения гемоглобинопатии у субъекта способ приводит к улучшению по меньшей мере двух клинически значимых параметров, выбранных из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения

вазоокклюзивных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

[0278] В некоторых вариантах осуществления способ лечения включает введение субъекту липосомы или липидной наночастицы, содержащей белок CasX и гРНК. В некоторых вариантах осуществления липосома или липидная наночастица дополнительно содержит донорную матрицу по любому из вариантов осуществления, описанных в настоящем документе.

[0279] В некоторых вариантах осуществления в описании предложен способ лечения субъекта, имеющего связанное с гемоглобинопатией заболевание, причем способ включает введение субъекту композиции CasX:гРНК, или вектора, или XDP, содержащую RNP с композицией CasX:гРНК по любому из вариантов осуществления, описанных в настоящем документе, в соответствии со схемой лечения, содержащей одну или более последовательных доз, с применением терапевтически эффективной дозы. В некоторых вариантах осуществления схемы лечения терапевтически эффективную дозу композиции или вектора вводят в виде однократной дозы. В других вариантах осуществления схемы лечения терапевтически эффективную дозу вводят субъекту в виде двух или более доз в течение по меньшей мере двух недель, или по меньшей мере одного месяца, или по меньшей мере двух месяцев, или по меньшей мере трех месяцев, или по меньшей мере четырех месяцев, или по меньшей мере пяти месяцев, или по меньшей мере шести месяцев. В некоторых вариантах осуществления схемы лечения эффективные дозы вводят путем, выбранным из группы, состоящей из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0280] В некоторых вариантах осуществления способы лечения дополнительно включают введение химиотерапевтического препарата, причем препарат эффективно облегчает признаки или симптомы, ассоциированные со связанным с гемоглобинопатией заболеванием, включая, без ограничений, гидроксимочевину, порошок L-глутамин для перорального введения, вокселотор и анальгетики.

[0281] В некоторых вариантах осуществления в настоящем описании предложена композиция CasX:гРНК, нуклеиновая кислота, кодирующая композицию CasX:гРНК, вектор, содержащий нуклеиновую кислоту, или XDP, содержащую RNP CasX:гРНК, для применения в качестве лекарственного средства для лечения гемоглобинопатии, включая серповидно-клеточную болезнь или бета-талассемию.

XIV. Наборы и композиции

[0282] В других вариантах осуществления в настоящем документе предложены наборы, содержащие белок CasX, одну или множество гРНК по любому из вариантов осуществления описания, имеющих нацеливающую последовательность, специфичную для гена BCL11A, и приемлемый контейнер (например, пробирку, флакон или планшет). В некоторых вариантах осуществления набор дополнительно содержит буфер, ингибитор нуклеазы, ингибитор протеазы, липосому, терапевтический агент, метку, реагент для

визуализации метки или любую комбинацию вышеизложенного. В некоторых вариантах осуществления набор дополнительно содержит фармацевтически приемлемый носитель, разбавитель или эксципиент. В некоторых вариантах осуществления набор содержит соответствующие контрольные композиции для применения при модификации генов и инструкции по применению. В некоторых вариантах осуществления набор содержит вектор, содержащий последовательность, кодирующую белок CasX описания, гРНК описания, необязательно донорную матрицу или их комбинацию.

[0283] В других вариантах осуществления наборов описания набор содержит композицию для лечения гемоглобинопатии у субъекта посредством модификации целевой нуклеиновой кислоты BCL11A в выделенных клетках субъекта, причем модификация включает приведение последовательности целевой нуклеиновой кислоты клеток в контакт с вариантом осуществления, описанным в настоящем документе: i) системой CasX:гРНК; ii) нуклеиновой кислотой, кодирующей компоненты системы CasX:гРНК; iii) вектором, содержащим нуклеиновую кислоту; iv) XDP, содержащей белок CasX и гидовую нуклеиновую кислоту (гРНК); или v) комбинацией любого из пп. (i)-(iv), причем i) упомянутое приведение в контакт приводит к модификации последовательности целевой нуклеиновой кислоты BCL11A белком CasX; ii) снижению экспрессии белка BCL11A и iii) повышенной продукции гемоглобина F (HbF) при созревании клеток. В некоторых случаях клетка представляет собой индуцированную плюрипотентную стволовую клетку (iPSC). В других случаях клетка представляет собой гематопоетическую стволовую клетку (HSC). В одном варианте осуществления применение композиции приводит к снижению экспрессии белка BCL11A зрелыми клетками на по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% по сравнению с целевой нуклеиновой кислотой, которая не была модифицирована. В другом варианте осуществления экспрессия белка BCL11A зрелыми клетками не обнаруживается.

[0284] В некоторых вариантах осуществления набор содержит множество клеток, отредактированных с применением систем CasX:гРНК, описанных в настоящем документе.

[0285] В настоящем описании представлены многочисленные примеры конфигураций, способов, параметров и т. п. Однако следует понимать, что такое описание не предназначено для ограничения объема настоящего описания, а, напротив, представлено в качестве описания примеров осуществления.

ВАРИАНТЫ ОСУЩЕСТВЛЕНИЯ В ВИДЕ НУМЕРОВАННОГО СПИСКА

[0286] Изобретение может быть определено путем ссылки на следующие иллюстративные варианты осуществления, перечисленные ниже в виде нумерованного списка.

Набор I

[0287] Вариант осуществления 1. Композиция, содержащая белок CRISPR класса 2, типа V и первую гидовую нуклеиновую кислоту (гНК), причем гНК содержит

нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты гена связывающего с полипиримидиновым трактом белка 1 (BCL11A).

[0288] Вариант осуществления 2. Композиция по варианту осуществления 1, причем гНК содержит нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, выбранной из группы, состоящей из:

- a. интрона BCL11A;
- b. экзона BCL11A;
- c. соединения интрон-экзон BCL11A;
- d. регуляторного элемента BCL11A; и
- e. межгенной области.

[0289] Вариант осуществления 3. Композиция по варианту осуществления 1, причем ген BCL11A содержит последовательность дикого типа.

[0290] Вариант осуществления 4. Композиция по любому из вариантов осуществления 1-3, причем гНК представляет собой гидовую РНК (гРНК).

[0291] Вариант осуществления 5. Композиция по любому из вариантов осуществления 1-3, причем гНК представляет собой гидовую ДНК (гДНК).

[0292] Вариант осуществления 6. Композиция по любому из вариантов осуществления 1-3, причем гНК представляет собой химеру, содержащую ДНК и РНК.

[0293] Вариант осуществления Композиция по любому из вариантов осуществления 1-6, причем гНК представляет собой одномолекулярную гНК (огНК).

[0294] Вариант осуществления 8. Композиция по любому из вариантов осуществления 1-6, причем гНК представляет собой двухмолекулярную гНК (дгНК).

[0295] Вариант осуществления 9. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, или последовательность, идентичную ей на по меньшей мере примерно 65%, по меньшей мере примерно 75%, по меньшей мере примерно 85% или по меньшей мере примерно 95%.

[0296] Вариант осуществления 10. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789.

[0297] Вариант осуществления 11. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность с SEQ ID NO: 272-2100 и 2286-26789, с одним нуклеотидом, удаленным с 3'-конца последовательности.

[0298] Вариант осуществления 12. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность с SEQ ID NO: 272-2100 и 2286-26789, с двумя нуклеотидами,

удаленными с 3'-конца последовательности.

[0299] Вариант осуществления 13. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность с SEQ ID NO: 272-2100 и 2286-26789, с тремя нуклеотидами, удаленными с 3'-конца последовательности.

[0300] Вариант осуществления 14. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность с SEQ ID NO: 272-2100 и 2286-26789, с четырьмя нуклеотидами, удаленными с 3'-конца последовательности.

[0301] Вариант осуществления 15. Композиция по любому из вариантов осуществления 1-8, причем нацеливающая последовательность гНК содержит последовательность с SEQ ID NO: 272-2100 и 2286-26789, с пятью нуклеотидами, удаленными с 3'-конца последовательности.

[0302] Вариант осуществления 16. Композиция по любому из вариантов осуществления 1-15, причем нацеливающая последовательность гНК комплементарна последовательности экзона BCL11A.

[0303] Вариант осуществления 17. Композиция по варианту осуществления 16, причем нацеливающая последовательность гНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A, последовательности экзона 3 BCL11A, последовательности экзона 4 BCL11A, последовательности экзона 5 BCL11A, последовательности экзона 6 BCL11A, последовательности экзона 7 BCL11A, последовательности экзона 8 BCL11A и последовательности экзона 9 BCL11A.

[0304] Вариант осуществления 18. Композиция по варианту осуществления 17, причем нацеливающая последовательность гНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A и последовательности экзона 3 BCL11A.

[0305] Вариант осуществления 19. Композиция по любому из вариантов осуществления 1-15, причем нацеливающая последовательность гНК комплементарна последовательности регуляторного элемента BCL11A.

[0306] Вариант осуществления 20. Композиция по варианту осуществления 19, причем нацеливающая последовательность гНК комплементарна последовательности промотора гена BCL11A.

[0307] Вариант осуществления 21. Композиция по варианту осуществления 19, причем нацеливающая последовательность гНК комплементарна последовательности энхансерного регуляторного элемента.

[0308] Вариант осуществления 22. Композиция по варианту осуществления 21, причем нацеливающая последовательность гНК комплементарна последовательности, которая содержит сайт связывания эритроид-специфического энхансера GATA1 (GATA1) гена BCL11A.

[0309] Вариант осуществления 23. Композиция по варианту осуществления 22, причем нацеливающая последовательность гНК имеет последовательность UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0310] Вариант осуществления 24. Композиция по варианту осуществления 22, причем нацеливающая последовательность гНК состоит из последовательности UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22).

[0311] Вариант осуществления 25. Композиция по варианту осуществления 21, причем нацеливающая последовательность гНК имеет последовательность UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0312] Вариант осуществления 26. Композиция по варианту осуществления 21, причем нацеливающая последовательность гНК состоит из последовательности UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23).

[0313] Вариант осуществления 27. Композиция по любому из вариантов осуществления 1-26, дополнительно содержащая вторую гНК, причем вторая гНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты BCL11A по сравнению с нацеливающей последовательностью гНК первой гНК.

[0314] Вариант осуществления 28. Композиция по варианту осуществления 27, причем нацеливающая последовательность второй гНК комплементарна последовательности целевой нуклеиновой кислоты, которая находится с 5'- или 3'-стороны от последовательности сайта связывания GATA1.

[0315] Вариант осуществления 29. Композиция по варианту осуществления 27, причем вторая гНК имеет нацеливающую последовательность, комплементарную тому же экзону, на который нацелена первая гНК.

[0316] Вариант осуществления 30. Композиция по любому из вариантов осуществления 1-29, причем первая или вторая гНК имеет каркас, содержащий последовательность, на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99% идентичную последовательности, выбранной из группы, состоящей из SEQ ID NO: 4-16 и 2101-2285, как указано в таблицах 1 и 2.

[0317] Вариант осуществления 31. Композиция по любому из вариантов осуществления 1-30, причем первая или вторая гНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 2101-2285.

[0318] Вариант осуществления 32. Композиция по любому из вариантов осуществления 1-30, причем первая или вторая гНК имеет каркас, состоящий из последовательности, выбранной из группы, состоящей из SEQ ID NO: 2101-2285.

[0319] Вариант осуществления 33. Композиция по любому из вариантов осуществления 1-30, причем первый или второй каркас гНК содержит последовательность, имеющую по меньшей мере одну модификацию относительно последовательности референсной гНК, выбранной из группы, состоящей из SEQ ID NO: 4-16.

[0320] Вариант осуществления 34. Композиция по варианту осуществления 33, причем по меньшей мере одна модификация референсной гНК содержит по меньшей мере одну замену, делецию или замену нуклеотида последовательности референсной гНК.

[0321] Вариант осуществления 35. Композиция по любому из вариантов осуществления 1-34, причем первый или второй вариант гНК содержит нацеливающую последовательность UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22).

[0322] Вариант осуществления 36. Композиция по любому из вариантов осуществления 1-35, причем первая или вторая гНК химически модифицирована.

[0323] Вариант осуществления 37. Композиция по любому из вариантов осуществления 1-36, причем белок CRISPR класса 2, типа V представляет собой референсный белок CasX, имеющий последовательность по любой из SEQ ID NO: 1-3, вариант белка CasX, имеющий последовательность SEQ ID NO: 36-99 или 101-148, как указано в таблице 4, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, или по меньшей мере примерно 95%, или по меньшей мере примерно 95%, или по меньшей мере примерно 96%, или по меньшей мере примерно 97%, или по меньшей мере примерно 98%, или по меньшей мере примерно 99%.

[0324] Вариант осуществления 38. Композиция по варианту осуществления 37, причем белок CRISPR класса 2, типа V представляет собой вариант белка CasX, имеющий последовательность SEQ ID NO: 36-99 или 101-148.

[0325] Вариант осуществления 39. Композиция по варианту осуществления 37, причем вариант белка CasX состоит из последовательности SEQ ID NO: 36-99 или 101-148.

[0326] Вариант осуществления 40. Композиция по варианту осуществления 37, причем вариант белка CasX содержит по меньшей мере одну модификацию относительно референсного белка CasX, имеющего последовательность, выбранную из SEQ ID NO: 1-3.

[0327] Вариант осуществления 41. Композиция по варианту осуществления 40, причем по меньшей мере одна модификация содержит по меньшей мере одну аминокислотную замену, делецию или замену в домене варианта белка CasX относительно референсного белка CasX.

[0328] Вариант осуществления 42. Композиция по варианту осуществления 41, причем домен выбран из группы, состоящей из домена связывания с нецелевой цепью (NTSB), домена загрузки целевой цепи (TSL), спирального домена I, спирального домена II, олигонуклеотид-связывающего домена (OBD) и домена расщепления ДНК RuvC.

[0329] Вариант осуществления 43. Композиция по любому из вариантов осуществления 37-42, причем белок CasX дополнительно содержит один или более

сигналов ядерной локализации (NLS).

[0330] Вариант осуществления 44. Композиция по варианту осуществления 43, причем один или более NLS выбраны из группы последовательностей, состоящей из PKKKRKV (SEQ ID NO: 168), KRPAATKKAGQAKKKK (SEQ ID NO: 169), PAAKRVKLD (SEQ ID NO: 170), RQRRNELKRSP (SEQ ID NO: 171), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 172), RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 173), VSRKRPRP (SEQ ID NO: 174), PPKKARED (SEQ ID NO: 175), PPKKKKPL (SEQ ID NO: 176), SALIKKKKKMAP (SEQ ID NO: 177), DRLRR (SEQ ID NO: 178), PKQKKRK (SEQ ID NO: 179), RKLKKKIKKL (SEQ ID NO: 180), REKKKFLKRR (SEQ ID NO: 181), KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 182), RKCLQAGMNLEARKTKK (SEQ ID NO: 183), PRPRKIPR (SEQ ID NO: 184), PPRKKRTVV (SEQ ID NO: 185), NLSKKKKRKREK (SEQ ID NO: 186), RRPSRPFRKP (SEQ ID NO: 187), KRPRSPSS (SEQ ID NO: 188), KRGINDRNFWRGENERKTR (SEQ ID NO: 189), PRPPKMARYDN (SEQ ID NO: 190), KRSFSKAF (SEQ ID NO: 191), KLKIKRPVK (SEQ ID NO: 192), PKTRRRPRRSQRKRPPT (SEQ ID NO: 26792), RRKKRRPRRKKRR (SEQ ID NO: 196), PKKKSRRPKKKSRK (SEQ ID NO: 197), HKKKHPDASVNFSEFSK (SEQ ID NO: 198), QRP GPYDRPQRP GPYDRP (SEQ ID NO: 199), LSPSLSPLLSPSLSPL (SEQ ID NO: 200), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 201), PKRGRGRPKRGRGR (SEQ ID NO: 202), MSRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 194), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 193) и PKKKRKVPPPPKKKKRKV (SEQ ID NO: 204).

[0331] Вариант осуществления 45. Композиция по варианту осуществления 43 или варианту осуществления 44, причем один или более NLS экспрессируются на С-конце белка CasX или рядом с ним.

[0332] Вариант осуществления 46. Композиция по варианту осуществления 43 или варианту осуществления 44, причем один или более NLS экспрессируются на N-конце белка CasX или рядом с ним.

[0333] Вариант осуществления 47. Композиция по варианту осуществления 43 или варианту осуществления 44, содержащая один или более NLS, расположенных на N-конце и на С-конце белка CasX или рядом с ними.

[0334] Вариант осуществления 48. Композиция по любому из вариантов осуществления 37-47, причем вариант CasX способен образовывать комплекс рибонуклеопротеиновый комплекс (RNP) с гидовой нуклеиновой кислотой (гНК).

[0335] Вариант осуществления 49. Композиция по варианту осуществления 48, причем RNP из варианта белка CasX и варианта гНК демонстрирует по меньшей мере одну или более улучшенных характеристик по сравнению с RNP из референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3, и гНК, содержащей последовательность SEQ ID NO: 4-16.

[0336] Вариант осуществления 50. Композиция по варианту осуществления 49, причем улучшенная характеристика выбрана из одного или более из группы, состоящей из

улучшенного сворачивания варианта CasX; улучшенной аффинности связывания с гидовой нуклеиновой кислотой (гНК); улучшенной аффинности связывания с целевой ДНК; улучшенной способности использовать более широкий спектр из одной или более РАМ-последовательностей, включая АТС, СТС, ГТС или ТТС, при редактировании целевой ДНК; улучшенного раскручивания целевой ДНК; повышенной активности редактирования; улучшенной эффективности редактирования; улучшенной специфичности редактирования; повышенной нуклеазной активности; повышенной загрузки целевой цепи для двухцепочечного расщепления; сниженной загрузки целевой цепи для одноцепочечного никирования; сниженного нецелевого расщепления; улучшенного связывания с нецелевой цепью ДНК; улучшенной стабильности белка; улучшенной растворимости белка; улучшенной стабильности комплекса белок:гНК (RNP); улучшенной растворимости комплекса белок:гНК; улучшенного выхода белка; улучшенной экспрессии белка; и улучшенных характеристик слияния.

[0337] Вариант осуществления 51. Композиция по варианту осуществления 49 или варианту осуществления 50, причем кратность улучшения характеристики RNP из варианта белка CasX и варианта гНК составляет от по меньшей мере примерно 1,1 до примерно 100 раз или более относительно RNP из референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3, и гНК, содержащей последовательность SEQ ID NO: 4-16.

[0338] Вариант осуществления 52. Композиция по варианту осуществления 49 или варианту осуществления 50, причем кратность улучшения характеристики варианта белка CasX составляет по меньшей мере примерно 1,1, по меньшей мере примерно 2, по меньшей мере примерно 10, по меньшей мере примерно 100 раз или более относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3, и гНК, содержащей последовательность SEQ ID NO: 4-16.

[0339] Вариант осуществления 53. Композиция по любому из вариантов осуществления 49-52, в которой улучшенная характеристика включает эффективность редактирования, и RNP варианта белка CasX и варианта гНК характеризуется кратностью улучшения эффективности редактирования от 1,1 до 100 раз по сравнению с RNP референсного белка CasX с SEQ ID NO: 2 и гНК с SEQ ID NO: 4-16.

[0340] Вариант осуществления 54. Композиция по любому из вариантов осуществления 48-53, причем RNP, содержащий вариант CasX и вариант гНК, демонстрирует более высокую эффективность редактирования и/или связывания с целевой последовательностью в целевой ДНК, если любая из РАМ-последовательностей ТТС, АТС, ГТС или СТС расположена на расстоянии 1 нуклеотида к 5'-концу от протоспейсера нецелевой цепи, имеющего идентичность с нацеливающей последовательностью гНК в системе клеточного анализа, по сравнению с эффективностью редактирования и/или связывания RNP, содержащего референсный белок CasX и референсную гНК, в сопоставимой системе анализа.

[0341] Вариант осуществления 55. Композиция по варианту осуществления 54, причем РАМ-последовательность представляет собой ТТС.

[0342] Вариант осуществления 56. Композиция по варианту осуществления 54, причем РАМ-последовательность представляет собой АТС.

[0343] Вариант осуществления 57. Композиция по варианту осуществления 54, причем РАМ-последовательность представляет собой СТС.

[0344] Вариант осуществления 58. Композиция по варианту осуществления 54, причем РАМ-последовательность представляет собой ГТС.

[0345] Вариант осуществления 59. Композиция по любому из вариантов осуществления 54-58, причем повышенная аффинность связывания с одной или более РАМ-последовательностями в по меньшей мере 1,5 раза выше по сравнению с аффинностью связывания любого из белков CasX с SEQ ID NO: 1-3 для РАМ-последовательностей.

[0346] Вариант осуществления 60. Композиция по любому из вариантов осуществления 48-59, причем RNP имеет на по меньшей мере 5%, по меньшей мере 10%, по меньшей мере 15% или по меньшей мере 20% более высокую процентную долю компетентных в отношении расщепления RNP по сравнению с RNP с референсным CasX и гНК с SEQ ID NO: 4-16.

[0347] Вариант осуществления 61. Композиция по любому из вариантов осуществления 37-60, причем вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий нуклеазной активностью.

[0348] Вариант осуществления 62. Композиция по любому из вариантов осуществления 37-60, причем вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий активностью двухцепочечного расщепления.

[0349] Вариант осуществления 63. Композиция по любому из вариантов осуществления 1-48, причем белок CasX представляет собой каталитически неактивный белок CasX (dCasX), и при этом dCasX и гНК сохраняют способность связываться с целевой нуклеиновой кислотой BCL11A.

Вариант осуществления 64. Композиция по варианту осуществления 63, причем dCasX содержит мутацию в остатках:

а. D672, E769 и/или D935, соответствующим белку CasX с SEQ ID NO:1; или

[0350] б. D659, E756 и/или D922, соответствующим белку CasX с SEQ ID NO: 2.

[0351] Вариант осуществления 65. Композиция по варианту осуществления 64, причем мутация представляет собой замену остатка на аланин.

[0352] Вариант осуществления 66. Композиция по любому из вариантов осуществления 1-62, дополнительно содержащая нуклеиновую кислоту донорной матрицы.

[0353] Вариант осуществления 67. Композиция по варианту осуществления 66, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок гена BCL11A, выбранный из группы, состоящей из экзона BCL11A, интрона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A и последовательности сайта связывания GATA1.

[0354] Вариант осуществления 68. Композиция по варианту осуществления 67, в

которой последовательность донорной матрицы содержит одну или более мутаций относительно соответствующего участка гена BCL11A дикого типа.

[0355] Вариант осуществления 69. Композиция по варианту осуществления 67 или варианту осуществления 68, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A, экзона 3 BCL11A, экзона 4 BCL11A, экзона 5 BCL11A, экзона 6 BCL11A, экзона 7 BCL11A, экзона 8 BCL11A и экзона 9 BCL11A.

[0356] Вариант осуществления 70. Композиция по варианту осуществления 69, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A и экзона 3 BCL11A.

[0357] Вариант осуществления 71. Композиция по любому из вариантов осуществления 66-70, в которой размер донорной матрицы находится в диапазоне от 10 до 15 000 нуклеотидов.

[0358] Вариант осуществления 72. Композиция по любому из вариантов осуществления 66-71, в которой донорная матрица представляет собой одноцепочечную ДНК-матрицу или одноцепочечную РНК-матрицу.

[0359] Вариант осуществления 73. Композиция по любому из вариантов осуществления 66-71, в которой донорная матрица представляет собой двухцепочечную ДНК-матрицу.

[0360] Вариант осуществления 74. Композиция по любому из вариантов осуществления 66-73, в которой донорная матрица содержит гомологичные плечи на 5'- и 3'-концах донорной матрицы, которые комплементарны последовательностям, фланкирующим сайты расщепления, в целевой нуклеиновой кислоте BCL11A, вводимой белком CRISPR класса 2, типа V, или рядом с ними.

[0361] Вариант осуществления 75. Нуклеиновая кислота, содержащая донорную матрицу по любому из вариантов осуществления 66-74.

[0362] Вариант осуществления 76. Нуклеиновая кислота, содержащая последовательность, которая кодирует CasX по любому из вариантов осуществления 37-65.

[0363] Вариант осуществления 77. Нуклеиновая кислота, содержащая последовательность, которая кодирует гНК по любому из вариантов осуществления 1-36.

[0364] Вариант осуществления 78. Нуклеиновая кислота по варианту осуществления 76, в которой последовательность, которая кодирует белок CasX, является кодон-оптимизированной для экспрессии в эукариотической клетке.

[0365] Вариант осуществления 79. Вектор, содержащий гНК по любому из вариантов осуществления 1-36, белок CasX по любому из вариантов осуществления 37-65 или нуклеиновую кислоту по любому из вариантов осуществления 75-78.

[0366] Вариант осуществления Вектор по варианту осуществления 79, который дополнительно содержит промотор.

[0367] Вариант осуществления 81. Вектор по варианту осуществления 79 или

варианту осуществления 80, который выбран из группы, состоящей из ретровирусного вектора, лентивирусного вектора, аденовирусного вектора, аденоассоциированного вирусного (AAV) вектора, вектора на основе вируса простого герпеса (HSV), вирусоподобной частицы (VLP), плазмиды, миникольца, наноплазмиды, ДНК-вектора и РНК-вектора.

[0368] Вариант осуществления 82. Вектор по варианту осуществления 81, который представляет собой вектор AAV.

[0369] Вариант осуществления 83. Вектор по варианту осуществления 82, причем вектор AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 или AAVRh10.

[0370] Вариант осуществления 84. Вектор по варианту осуществления 81, который представляет собой ретровирусный вектор.

[0371] Вариант осуществления 85. Вектор по варианту осуществления 81, который представляет собой VLP, содержащую один или более компонентов полибелка gag.

[0372] Вариант осуществления 86. Вектор по варианту осуществления 85, в котором один или более компонентов полибелка gag выбраны из группы, состоящей из матричного белка (МА), нуклеокапсидного белка (NC), капсидного белка (CA) и белка p1-р6.

[0373] Вариант осуществления 87. Вектор по варианту осуществления 85 или варианту осуществления 86, содержащий белок CasX и гНК.

[0374] Вариант осуществления 88. Вектор по варианту осуществления 87, в котором белок CasX и гНК связаны друг с другом в RNP.

[0375] Вариант осуществления 89. VLP по любому из вариантов осуществления 85-88, дополнительно содержащая псевдотипирующий гликопротеин вирусной оболочки или фрагмент антитела, который обеспечивает связывание и слияние VLP с клеткой-мишенью.

[0376] Вариант осуществления 90. VLP по варианту осуществления 89, причем клетка-мишень выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), эмбриональной стволовой (ES) клетки, индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

[0377] Вариант осуществления 91. Вектор по любому из вариантов осуществления 85-90, дополнительно содержащий донорную матрицу.

[0378] Вариант осуществления 92. Клетка-хозяин, содержащая вектор по любому из вариантов осуществления 79-91.

[0379] Вариант осуществления 93. Клетка-хозяин по варианту осуществления 92, которая выбрана из группы, состоящей из ВНК, HEK293, HEK293T, NS0, SP2/0, клеток миеломы YO, мышинных клеток миеломы P3X63, PER, PER.C6, NIH3T3, COS, HeLa, CHO и дрожжевых клеток.

[0380] Вариант осуществления 94. Способ модификации последовательности целевой нуклеиновой кислоты BCL11A в популяции клеток, включающий введение в

клетки популяции:

- a. композиции по любому из вариантов осуществления 1-74;
- b. нуклеиновой кислоты по любому из вариантов осуществления 75-78;
- c. вектора по любому из вариантов осуществления 79-84;
- d. VLP по любому из вариантов осуществления 85-91; или
- e. комбинаций двух или более из пп. (a)-(d),

причем последовательность целевой нуклеиновой кислоты гена BCL11A клеток, на которую нацелена первая гНК, модифицирована белком CasX.

[0381] Вариант осуществления 95. Способ по варианту осуществления 94, в котором модификация включает введение одноцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0382] Вариант осуществления 96. Способ по варианту осуществления 94, в котором модификация включает введение двухцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0383] Вариант осуществления 97. Способ по любому из вариантов осуществления 94-96, дополнительно включающий введение в клетки популяции второй гНК или нуклеиновой кислоты, кодирующей вторую гНК, причем вторая гНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты гена BCL11A по сравнению с первой гНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток популяции.

[0384] Вариант осуществления 98. Способ по любому из вариантов осуществления 94-97, в котором модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток популяции.

[0385] Вариант осуществления 99. Способ по варианту осуществления 94-98, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

[0386] Вариант осуществления 100. Способ по любому из вариантов осуществления 94-97, который включает вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0387] Вариант осуществления 101. Способ по варианту осуществления 98, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (НТИ).

[0388] Вариант осуществления 102. Способ по варианту осуществления 100 или варианту осуществления 101, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

[0389] Вариант осуществления 103. Способ по любому из вариантов осуществления 100-102, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в клетках популяции.

[0390] Вариант осуществления 104. Способ по любому из вариантов осуществления 94-103, в котором ген BCL11A клеток популяции модифицирован так, что экспрессия белка

BCL11A понижается на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, в которых ген BCL11A не модифицирован.

[0391] Вариант осуществления 105. Способ по любому из вариантов осуществления 94-103, в котором ген BCL11A клеток популяции модифицирован так, что по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0392] Вариант осуществления 106. Способ по любому из вариантов осуществления 94-105, в котором клетки являются эукариотическими.

[0393] Вариант осуществления 107. Способ по варианту осуществления 106, в котором эукариотические клетки выбраны из группы, состоящей из клеток грызунов, клеток мыши, клеток крысы и клеток примата, не являющегося человеком.

[0394] Вариант осуществления 108. Способ по варианту осуществления 106, в котором эукариотические клетки представляют собой клетки человека.

[0395] Вариант осуществления 109. Способ по любому из вариантов осуществления 106-108, в котором эукариотическая клетка выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

[0396] Вариант осуществления 110. Способ по любому из вариантов осуществления 94-109, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vitro* или *ex vivo*.

[0397] Вариант осуществления 111. Способ по любому из вариантов осуществления 94-109, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vivo* у субъекта.

[0398] Вариант осуществления 112. Способ по варианту осуществления 111, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0399] Вариант осуществления 113. Способ по варианту осуществления 111, в котором субъект является человеком.

[0400] Вариант осуществления 114. Способ по любому из вариантов осуществления 111-113, включающий введение субъекту терапевтически эффективной дозы вектора AAV.

[0401] Вариант осуществления 115. Способ по варианту осуществления 114, в котором вектор AAV вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно

1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0402] Вариант осуществления 116. Способ по варианту осуществления 114, в котором вектор AAV вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

[0403] Вариант осуществления 117. Способ по любому из вариантов осуществления 111-113, который включает введение терапевтически эффективной дозы VLP субъекту.

[0404] Вариант осуществления 118. Способ по варианту осуществления 117, в котором VLP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

[0405] Вариант осуществления 119. Способ по варианту осуществления 117, в котором VLP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг.

[0406] Вариант осуществления 120. Способ по любому из вариантов осуществления 112-119, в котором вектор или VLP вводят субъекту способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0407] Вариант осуществления 121. Способ по любому из вариантов осуществления 94-120, дополнительно включающий приведение последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток в контакт с:

a. дополнительной нуклеазой CRISPR и гНК, нацеленной на другой или перекрывающийся участок целевой нуклеиновой кислоты BCL11A по сравнению с первой гНК;

b. полинуклеотидом, кодирующим дополнительную нуклеазу CRISPR и гНК по п. (a);

c. вектором, содержащим полинуклеотид по п. (b); или

d. VLP, содержащей дополнительную нуклеазу CRISPR и гНК по п. (a),

причем приведение в контакт приводит к модификации гена BCL11A в другом положении в последовательности по сравнению с последовательностью, на которую нацелена первая гНК.

[0408] Вариант осуществления 122. Способ по варианту осуществления 121, причем дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий последовательность, отличную от белка CasX по любому из предшествующих вариантов осуществления.

[0409] Вариант осуществления 123. Способ по варианту осуществления 121, причем дополнительная нуклеаза CRISPR не представляет собой белок CasX.

[0410] Вариант осуществления 124. Способ по варианту осуществления 123, причем дополнительная нуклеаза CRISPR выбрана из группы, состоящей из Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12J, Cas13a, Cas13b, Cas13c, Cas13d, CasX, CasY, Cas14, Cpf1, C2cl, Csn2 и вариантов их последовательностей.

[0411] Вариант осуществления 125. Популяция клеток, модифицированная способом по любому из вариантов осуществления 94-124, причем клетки были модифицированы так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют обнаруживаемый уровень белка BCL11A.

[0412] Вариант осуществления 126. Популяция клеток, модифицированная способом по любому из вариантов осуществления 94-124, причем клетки были модифицированы так, что экспрессия белка BCL11A снижена на по меньшей мере 70%, на по меньшей мере 75%, на по меньшей мере 80%, на по меньшей мере 85%, на по меньшей мере 90% или на по меньшей мере 95% по сравнению с клетками, в которых ген BCL11A не был модифицирован.

[0413] Вариант осуществления 127. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий введение субъекту терапевтически эффективного количества клеток по варианту осуществления 125 или варианту осуществления 126.

[0414] Вариант осуществления 128. Способ по варианту осуществления 127, причем гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0415] Вариант осуществления 129. Способ по любому из вариантов осуществления 127 или варианту осуществления 128, в котором клетки являются аутологичными по отношению к субъекту, которому должны быть введены клетки.

[0416] Вариант осуществления 130. Способ по любому из вариантов осуществления 127 или варианту осуществления 128, в котором клетки являются аллогенными по отношению к субъекту, которому должны быть введены клетки.

[0417] Вариант осуществления 131. Способ по любому из вариантов осуществления 127-130, в котором клетки или их потомство сохраняются у субъекта в течение по меньшей мере одного месяца, двух месяцев, трех месяцев, четырех месяцев, пяти месяцев, шести месяцев, семи месяцев, восьми месяцев, девяти месяцев, десяти месяцев, одиннадцать месяцев, двенадцати месяцев, тринадцати месяцев, четырнадцати месяцев, пятнадцати месяцев, шестнадцати месяцев, семнадцати месяцев, восемнадцати месяцев, девятнадцати

месяцев, двадцати месяцев, двадцати месяцев, двадцати двух месяцев, двадцати трех месяцев, двух лет, трех лет, четырех лет или пяти лет после введения модифицированных клеток субъекту.

[0418] Вариант осуществления 132. Способ по любому из вариантов осуществления 127-131, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0419] Вариант осуществления 133. Способ по любому из вариантов осуществления 127-131, который приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0420] Вариант осуществления 134. Способ по любому из вариантов осуществления 127-131, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0421] Вариант осуществления 135. Способ по любому из вариантов осуществления 127-134, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0422] Вариант осуществления 136. Способ по любому из вариантов осуществления 127-134, в котором субъект является человеком.

[0423] Вариант осуществления 137. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий модификацию гена BCL11A в клетках субъекта, причем модификация включает приведение упомянутых клеток в контакт с терапевтически эффективной дозой:

- a. композиции по любому из вариантов осуществления 1-74;
- b. нуклеиновой кислоты по любому из вариантов осуществления 75-78;
- c. вектора по любому из вариантов осуществления 79-84;
- d. VLP по любому из вариантов осуществления 85-88; или
- e. комбинаций двух или более из пп. (a)-(d),

причем ген BCL11A клеток, на который нацелена первая гНК, модифицирован белком CasX.

[0424] Вариант осуществления 138. Способ по варианту осуществления 137, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0425] Вариант осуществления 139. Способ по любому из вариантов осуществления 137 или варианта осуществления 138, в котором клетка выбрана из группы, состоящей из

гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

[0426] Вариант осуществления 140. Способ по любому из вариантов осуществления 137-139, в котором модификация включает введение одноцепочечного разрыва в ген BCL11A клеток.

[0427] Вариант осуществления 141. Способ по любому из вариантов осуществления 137-139, в котором модификация включает введение двухцепочечного разрыва в ген BCL11A клеток.

[0428] Вариант осуществления 142. Способ по любому из вариантов осуществления 137-141, дополнительно включающий введение в клетки субъекта второй гНК или нуклеиновой кислоты, кодирующей вторую гНК, причем вторая гНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты по сравнению с первой гНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток субъекта.

[0429] Вариант осуществления 143. Способ по любому из вариантов осуществления 137-142, в котором модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток.

[0430] Вариант осуществления 144. Способ по варианту осуществления 143, в котором модификация приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

[0431] Вариант осуществления 145. Способ по любому из вариантов осуществления 137-144, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, которые не были модифицированы.

[0432] Вариант осуществления 146. Способ по любому из вариантов осуществления 137-144, в котором ген BCL11A клеток субъекта модифицирован так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0433] Вариант осуществления 147. Способ по любому из вариантов осуществления 137-146, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0434] Вариант осуществления 148. Способ по любому из вариантов осуществления 137-147, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере, 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0435] Вариант осуществления 149. Способ по любому из вариантов осуществления 137-146, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0436] Вариант осуществления 150. Способ по любому из вариантов осуществления 137-142, который включает вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток.

[0437] Вариант осуществления 151. Способ по варианту осуществления 149, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (НТИ).

[0438] Вариант осуществления 152. Способ по варианту осуществления 149 или варианту осуществления 151, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

[0439] Вариант осуществления 153. Способ по любому из вариантов осуществления 147-152, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, которые не были модифицированы.

[0440] Вариант осуществления 154. Способ по любому из вариантов осуществления 147-152, в котором ген BCL11A клеток субъекта модифицирован так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0441] Вариант осуществления 155. Способ по любому из вариантов осуществления 147-154, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0442] Вариант осуществления 156. Способ по любому из вариантов осуществления 147-154, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке

субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере, 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0443] Вариант осуществления 157. Способ по любому из вариантов осуществления 147-154, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0444] Вариант осуществления 158. Способ по любому из вариантов осуществления 137-156, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0445] Вариант осуществления 159. Способ по любому из вариантов осуществления 137-156, в котором субъект является человеком.

[0446] Вариант осуществления 160. Способ по любому из вариантов осуществления 137-159, в котором вектор представляет собой AAV, который вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0447] Вариант осуществления 161. Способ по любому из вариантов осуществления 137-159, в котором вектор представляет собой AAV, который вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

[0448] Вариант осуществления 162. Способ по любому из вариантов осуществления 137-159, в котором VLP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

[0449] Вариант осуществления 163. Способ по любому из вариантов осуществления 137-159, в котором VLP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг.

[0450] Вариант осуществления 164. Способ по любому из вариантов осуществления 137-163, в котором вектор или VLP вводят субъекту способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0451] Вариант осуществления 165. Способ по любому из вариантов осуществления 137-164, который приводит к улучшению по меньшей мере одной клинически значимой конечной точки у субъекта.

[0452] Вариант осуществления 166. Способ по варианту осуществления 165, который приводит к улучшению по меньшей мере одного клинически значимого параметра, выбранного из группы, состоящей из развития заболевания органа-мишени, альбинурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

[0453] Вариант осуществления 167. Способ по варианту осуществления 165, который приводит к улучшению по меньшей мере двух клинически значимых параметров, выбранных из группы, состоящей из развития заболевания органа-мишени, альбинурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

[0454] Вариант осуществления 168. Способ лечения субъекта с гемоглобинопатией, включающий:

- a. выделение индуцированных плюрипотентных стволовых клеток (iPSC) или гематопозитических стволовых клеток (HSC) у субъекта;
- b. модификацию целевой нуклеиновой кислоты BCL11A в iPSC или HSC способом по любому из вариантов осуществления 94-110;
- c. дифференцировку модифицированных iPSC или HSC в гематопозитическую клетку-предшественника; и
- d. имплантацию гематопозитической клетки-предшественника субъекту с гемоглобинопатией,

причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0455] Вариант осуществления 169. Способ по варианту осуществления 168, в котором iPSC или HSC является аутологичной и выделена из костного мозга или из периферической крови субъекта.

[0456] Вариант осуществления 170. Способ по варианту осуществления 168, в котором iPSC или HSC является аллогенной и выделена из костного мозга или из периферической крови другого субъекта.

[0457] Вариант осуществления 171. Способ по любому из вариантов осуществления 168-170, в котором имплантация включает введение гематопозитической клетки-предшественника субъекту посредством трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0458] Вариант осуществления 172. Способ по любому из вариантов осуществления 168-171, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0459] Вариант осуществления 173. Способ повышения уровня эмбрионального гемоглобина (HbF) у субъекта посредством редактирования генома, включающий: а. введение субъекту эффективной дозы вектора по любому из вариантов осуществления 79-84 или VLP по любому из вариантов осуществления 85-90, причем вектор или VLP доставляет систему CasX:гНК в клетки субъекта;

б. редактирование целевой нуклеиновой кислоты BCL11A клеток субъекта с помощью CasX, на который нацелена первая гНК;

с. редактирование включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в последовательности целевой нуклеиновой кислоты так, что экспрессия белка BCL11A снижается или устраняется,

причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0460] Вариант осуществления 174. Способ по варианту осуществления 173, в котором клетки выбраны из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

[0461] Вариант осуществления 175. Способ по варианту осуществления 173 или

варианту осуществления 174, в котором целевая нуклеиновая кислота клеток отредактирована так, что экспрессия белка BCL11A снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с целевой нуклеиновой кислотой клеток, которые не были отредактированы.

[0462] Вариант осуществления 176. Способ по любому из вариантов осуществления 173-175, в котором субъект выбран из группы, состоящей из мыши, крысы, свиньи и примата, не являющегося человеком.

[0463] Вариант осуществления 177. Способ по любому из вариантов осуществления 173-175, в котором субъект является человеком.

[0464] Вариант осуществления 178. Способ по любому из вариантов осуществления 173-177, в котором вектор вводят в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0465] Вариант осуществления 179. Способ по любому из вариантов осуществления 173-177, в котором VLP вводят в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

[0466] Вариант осуществления 180. Способ по любому из вариантов осуществления 173-179, в котором вектор или VLP вводят способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0467] Вариант осуществления 181. Композиция по любому из вариантов осуществления 1-74, нуклеиновая кислота по любому из вариантов осуществления 75-78, вектор по любому из вариантов осуществления 79-84, VLP по любому из вариантов осуществления 85-88, клетка-хозяин по варианту осуществления 92 или варианту осуществления 93 или популяция клеток по варианту осуществления 125 или варианту осуществления 126 для применения в качестве лекарственного средства для лечения гемоглобинопатии.

[0468] Вариант осуществления 182. Композиция по варианту осуществления 1, в которой последовательность целевой нуклеиновой кислоты комплементарна

последовательности нецелевой цепи, расположенной на расстоянии 1 нуклеотида от 3'-конца последовательности смежного с протоспейсером мотива (PAM).

[0469] Вариант осуществления 183. Композиция по варианту осуществления 182, в которой PAM-последовательность содержит мотив TC.

[0470] Вариант осуществления 184. Композиция по варианту осуществления 183, в которой PAM-последовательность содержит ATC, GTC, CTC или TTC.

[0471] Вариант осуществления 185. Композиция по любому из вариантов осуществления 182-184, в которой белок CRISPR класса 2, типа V содержит домен RuvC.

[0472] Вариант осуществления 186. Композиция по варианту осуществления 185, в которой домен RuvC создает смещенный двухцепочечный разрыв в последовательности целевой нуклеиновой кислоты.

[0473] Вариант осуществления 187. Композиция по любому из вариантов осуществления 182-186, в которой белок CRISPR класса 2, типа V не содержит нуклеазного домена HNH.

[0474] Набор II

[0475] Вариант осуществления 1. Система, содержащая белок CRISPR класса 2, типа V и первую гидовую рибонуклеиновую кислоту (гРНК), причем гРНК содержит нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, содержащей ген связывающегося с полипиримидиновым трактом белка 1 (BCL11A).

[0476] Вариант осуществления 2. Система по варианту осуществления 1, в которой гРНК содержит нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, выбранной из группы, состоящей из:

- a. интрона BCL11A;
- b. экзона BCL11A;
- c. соединения интрон-экзон BCL11A;
- d. регуляторного элемента BCL11A; и
- e. межгенной области.

[0477] Вариант осуществления 3. Система по варианту осуществления 1 или варианту осуществления 2, в которой ген BCL11A содержит последовательность дикого типа.

[0478] Вариант осуществления 4. Система по любому из вариантов осуществления 1-3, в которой гРНК представляет собой одномолекулярную гРНК (огРНК).

[0479] Вариант осуществления 5. Система по любому из вариантов осуществления 1-4, в которой гРНК представляет собой двухмолекулярную гРНК (дгРНК).

[0480] Вариант осуществления 6. Система по любому из вариантов осуществления 1-5, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, или последовательность, идентичную ей на по меньшей мере примерно 65%, по меньшей мере примерно 75%, по меньшей мере примерно 85% или по меньшей мере примерно 95%.

[0481] Вариант осуществления 7. Система по любому из вариантов осуществления 1-5, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789.

[0482] Вариант осуществления 8. Система по варианту осуществления 7, в которой у нацеливающей последовательности удален один нуклеотид с 3'-конца последовательности.

[0483] Вариант осуществления 9. Система по варианту осуществления 7, в которой у нацеливающей последовательности удалены два нуклеотида с 3'-конца последовательности.

[0484] Вариант осуществления 10. Система по варианту осуществления 7, в которой у нацеливающей последовательности удалены три нуклеотида с 3'-конца последовательности.

[0485] Вариант осуществления 11. Система по варианту осуществления 7, в которой у нацеливающей последовательности удалены четыре нуклеотида с 3'-конца последовательности.

[0486] Вариант осуществления 12. Система по варианту осуществления 7, в которой у нацеливающей последовательности удалены пять нуклеотидов с 3'-конца последовательности.

[0487] Вариант осуществления 13. Система по любому из вариантов осуществления 1-12, в которой нацеливающая последовательность гРНК комплементарна последовательности экзона BCL11A.

[0488] Вариант осуществления 14. Система по варианту осуществления 13, в которой нацеливающая последовательность гРНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A, последовательности экзона 3 BCL11A, последовательности экзона 4 BCL11A, последовательности экзона 5 BCL11A, последовательности экзона 6 BCL11A, последовательности экзона 7 BCL11A, последовательности экзона 8 BCL11A и последовательности экзона 9 BCL11A.

[0489] Вариант осуществления 15. Система по варианту осуществления 14, в которой нацеливающая последовательность гРНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A и последовательности экзона 3 BCL11A.

[0490] Вариант осуществления 16. Система по любому из вариантов осуществления 1-12, в которой нацеливающая последовательность гРНК комплементарна последовательности регуляторного элемента BCL11A.

[0491] Вариант осуществления 17. Система по варианту осуществления 16, в которой нацеливающая последовательность гРНК комплементарна последовательности промотора гена BCL11A.

[0492] Вариант осуществления 18. Система по варианту осуществления 16, в которой нацеливающая последовательность гРНК комплементарна последовательности

энхансерного регуляторного элемента.

[0493] Вариант осуществления 19. Система по варианту осуществления 18, в которой нацеливающая последовательность гРНК комплементарна последовательности, которая содержит эритроид-специфичный энхансерный сайт связывания фактора GATA1 (GATA1) гена BCL11A.

[0494] Вариант осуществления 20. Система по варианту осуществления 16, в которой нацеливающая последовательность гРНК комплементарна последовательности, которая расположена с 5'-конца от сайта связывания GATA1 гена BCL11A.

[0495] Вариант осуществления 21. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК содержит последовательность UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0496] Вариант осуществления 22. Система по варианту осуществления 19, в которой нацеливающая последовательность гРНК состоит из последовательности UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22).

[0497] Вариант осуществления 23. Система по варианту осуществления 18, в которой нацеливающая последовательность гРНК содержит последовательность UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0498] Вариант осуществления 24. Система по варианту осуществления 18, в которой нацеливающая последовательность гРНК состоит из последовательности UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23).

[0499] Вариант осуществления 25. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК содержит последовательность CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0500] Вариант осуществления 26. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК состоит из последовательности CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949).

[0501] Вариант осуществления 27. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК содержит последовательность GAGGCCAAACCCUCCUGGA (SEQ ID NO: 2948) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0502] Вариант осуществления 28. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК состоит из последовательности CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2948).

[0503] Вариант осуществления 29. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК содержит последовательность AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0504] Вариант осуществления 30. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК состоит из последовательности AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747).

[0505] Вариант осуществления 31. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК содержит последовательность AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

[0506] Вариант осуществления 32. Система по варианту осуществления 19 или варианту осуществления 20, в которой нацеливающая последовательность гРНК состоит из последовательности AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748).

[0507] Вариант осуществления 33. Система по любому из вариантов осуществления 1-32, дополнительно содержащая вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты BCL11A по сравнению с нацеливающей последовательностью гРНК первой гРНК.

[0508] Вариант осуществления 34. Система по варианту осуществления 33, в которой нацеливающая последовательность второй гРНК комплементарна последовательности целевой нуклеиновой кислоты, которая расположена с 5'- или 3'-конца от последовательности сайта связывания GATA1.

[0509] Вариант осуществления 35. Система по варианту осуществления 33, в которой каждая из первой и второй гРНК имеет нацеливающую последовательность, комплементарную последовательности в пределах промотора гена BCL11A.

[0510] Вариант осуществления 36. Система по любому из вариантов осуществления 1-35, в которой первая или вторая гРНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из последовательностей SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265 или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99%.

[0511] Вариант осуществления 37. Система по любому из вариантов осуществления 1-36, в которой первая или вторая гРНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265.

[0512] Вариант осуществления 38. Система по любому из вариантов осуществления 1-36, в которой первая или вторая гРНК имеет каркас, состоящий из последовательности, выбранной из группы, состоящей из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265.

[0513] Вариант осуществления 39. Система по варианту осуществления 38, в которой первая или вторая гРНК имеет каркас, состоящий из последовательности SEQ ID NO: 2238 или SEQ ID NO: 26800.

[0514] Вариант осуществления 40. Система по любому из вариантов осуществления

36-39, в которой нацеливающая последовательность связана с 3'-концом каркаса гРНК.

[0515] Вариант осуществления 41. Система по любому из вариантов осуществления 1-40, в которой белок CRISPR класса 2, типа V представляет собой вариант белка CasX, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, или по меньшей мере примерно 95%, или по меньшей мере примерно 95%, или по меньшей мере примерно 96%, или по меньшей мере примерно 97%, или по меньшей мере примерно 98%, или по меньшей мере примерно 99%.

[0516] Вариант осуществления 42. Система по варианту осуществления 41, в которой белок CRISPR класса 2, типа V представляет собой вариант белка CasX, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154.

[0517] Вариант осуществления 43. Система по варианту осуществления 41, в которой вариант белка CasX состоит из последовательности, выбранной из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154.

[0518] Вариант осуществления 44. Система по варианту осуществления 42, в которой вариант белка CasX состоит из последовательности, выбранной из группы, состоящей из SEQ ID NO: 126, 27043, 27046, 27050.

[0519] Вариант осуществления 45. Система по варианту осуществления 41, в которой вариант белка CasX содержит по меньшей мере одну модификацию относительно референсного белка CasX, имеющего последовательность, выбранную из SEQ ID NO: 1-3.

[0520] Вариант осуществления 46. Система по варианту осуществления 45, в которой по меньшей мере одна модификация содержит по меньшей мере одну аминокислотную замену, делецию или замену в домене варианта белка CasX относительно референсного белка CasX.

[0521] Вариант осуществления 47. Система по варианту осуществления 46, в которой домен выбран из группы, состоящей из домена связывания с нецелевой цепью (NTSB), домена загрузки целевой цепи (TSL), спирального домена I, спирального домена II, олигонуклеотид-связывающего домена (OBD) и домена расщепления ДНК RuvC.

[0522] Вариант осуществления 48. Система по любому из вариантов осуществления 41-47, в которой вариант белка CasX не содержит домена HNH.

[0523] Вариант осуществления 49. Система по любому из вариантов осуществления 41-48, в которой вариант белка CasX дополнительно содержит один или более сигналов ядерной локализации (NLS).

[0524] Вариант осуществления 50. Система по варианту осуществления 49, в которой один или более NLS выбраны из группы последовательностей, состоящей из PKKKRKV (SEQ ID NO: 168), KRPAATKKAGQAKKKK (SEQ ID NO: 169), PAAKRVKLD (SEQ ID NO: 170), RQRRNELKRSP (SEQ ID NO: 171),

NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 172), RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 173), VSRKRPRP (SEQ ID NO: 174), PPKKARED (SEQ ID NO: 175), PPKKKPL (SEQ ID NO: 176), SALIKKKKKMAP (SEQ ID NO: 177), DRLRR (SEQ ID NO: 178), PKQKKRK (SEQ ID NO: 179), RKLKKKIKKL (SEQ ID NO: 180), REKKKFLKRR (SEQ ID NO: 181), KRKGDEVGDGVDEVAKKKSKK (SEQ ID NO: 182), RKCLQAGMNLEARKTKK (SEQ ID NO: 183), PRPRKIPR (SEQ ID NO: 184), PPRKKRTVV (SEQ ID NO: 185), NLSKKKKRKREK (SEQ ID NO: 186), RRPSRPFRKP (SEQ ID NO: 187), KRPRSPSS (SEQ ID NO: 188), KRGINDRNFWRGENERKTR (SEQ ID NO: 189), PRPPKMARYDN (SEQ ID NO: 190), KRSFSKAF (SEQ ID NO: 191), KLKIKRPVK (SEQ ID NO: 192), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 193), PKTRRRPRRSQRKRPT (SEQ ID NO: 26792), SRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 194), KTRRRPRRSQRKRPT (SEQ ID NO: 195), RRKKRRPRRKKRR (SEQ ID NO: 196), PKKKSRRPKKKSRK (SEQ ID NO: 197), HKKKHPDASVNFSEFSK (SEQ ID NO: 198), QRPGPYDRPQRPGPYDRP (SEQ ID NO: 199), LSPSLSPLLSPSLSP (SEQ ID NO: 200), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 201), PKRGRGRPKRGRGR (SEQ ID NO: 202), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 203), PKKKRKVPPPPKKRKV (SEQ ID NO: 204), PAKRARRGYKC (SEQ ID NO: 27199), KLGPRKATGRW (SEQ ID NO: 27200), PRRKREE (SEQ ID NO: 27201), PYRGRKE (SEQ ID NO: 27202), PLRKRPRR (SEQ ID NO: 27203), PLRKRPRRGSPLRKRPRR (SEQ ID NO: 27204), PAAKRVKLDGGKRTADGSEFESPKKKRKV (SEQ ID NO: 27205), PAAKRVKLDGGKRTADGSEFESPKKKRKVGIIHGVPAA (SEQ ID NO: 27206), PAAKRVKLDGGKRTADGSEFESPKKKRKVAEAAAKEAAAKEAAKA (SEQ ID NO: 207), PAAKRVKLDGGKRTADGSEFESPKKKRKVPG (SEQ ID NO: 27208), KRKGSPERGERKRHW (SEQ ID NO: 27209), KRTADSQHSTPPKTKRKVEFEPKKKKRKV (SEQ ID NO: 27210) и PKKKRKVGGSKRTADSQHSTPPKTKRKVEFEPKKKKRKV (SEQ ID NO: 27211), причем один или более NLS связаны с белком CRISPR или со смежным NLS через линкерный пептид, при этом линкерный пептид выбран из группы, состоящей из RS, (G)_n (SEQ ID NO: 27212), (GS)_n (SEQ ID NO: 27213), (GSGGS)_n (SEQ ID NO: 214), (GGSGGS)_n (SEQ ID NO: 215), (GGGS)_n (SEQ ID NO: 216), GSGG (SEQ ID NO: 217), GGSGG (SEQ ID NO: 218), GSGSG (SEQ ID NO: 219), GSGGG (SEQ ID NO: 220), GGGSG (SEQ ID NO: 221), GSSSG (SEQ ID NO: 222), GPGP (SEQ ID NO: 223), GGP, PPP, PPAPPA (SEQ ID NO: 224), PPPG (SEQ ID NO: 27214), PPPGPPP (SEQ ID NO: 225), PPP(GGGS)_n (SEQ ID NO: 27215), (GGGS)_nPPP (SEQ ID NO: 27216), AEAAAKEAAAKEAAKA (SEQ ID NO: 27217) и TPPKTKRKVEFE (SEQ ID NO: 27218), где n равно от 1 до 5.

[0525] Вариант осуществления 51. Система по варианту осуществления 49 или варианту осуществления 50, в которой один или более NLS расположены на С-конце варианта белка CasX или рядом с ним.

[0526] Вариант осуществления 52. Система по варианту осуществления 49 или варианту осуществления 50, в которой один или более NLS расположены на N-конце варианта белка CasX или рядом с ним.

[0527] Вариант осуществления 53. Система по варианту осуществления 49 или варианту осуществления 50, содержащая один или более NLS, расположенных на N-конце и на C-конце варианта белка CasX или рядом с ними.

[0528] Вариант осуществления 54. Система по любому из вариантов осуществления 41-53, в которой вариант CasX способен образовывать рибонуклеопротеиновый комплекс (RNP) с гидовой нуклеиновой кислотой (гРНК).

[0529] Вариант осуществления 55. Система по варианту осуществления 54, в которой RNP варианта белка CasX и варианта гРНК демонстрирует по меньшей мере одну или более улучшенных характеристик по сравнению с RNP референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

[0530] Вариант осуществления 56. Система по варианту осуществления 55, в которой улучшенная характеристика выбрана из одного или более из группы, состоящей из улучшенного фолдинга варианта CasX; улучшенной аффинности связывания с гидовой нуклеиновой кислотой (гРНК); улучшенной аффинности связывания с целевой ДНК; улучшенной способности использовать более широкий спектр из одной или более последовательностей смежного с протоспейсером мотива (PAM), включая ATC, CTC, GTC или TTC, при редактировании целевой ДНК; улучшенного раскручивания целевой ДНК; повышенной активности редактирования; улучшенной эффективности редактирования; улучшенной специфичности редактирования; повышенной нуклеазной активности; улучшенной скорости расщепления последовательности целевой нуклеиновой кислоты; повышенной загрузки целевой цепи для двухцепочечного расщепления; сниженной загрузки целевой цепи для одноцепочечного никирования; сниженного нецелевого расщепления; улучшенного связывания с нецелевой цепью ДНК; улучшенной стабильности белка; улучшенной растворимости белка; улучшенного образования рибонуклеопротеинового комплекса (RNP); более высокой процентной доли компетентных в отношении расщепления RNP; улучшенной стабильности комплекса белок:гРНК (RNP); улучшенной растворимости комплекса белок:гРНК; улучшенного выхода белка; улучшенной экспрессии белка; и улучшенных характеристик слияния.

[0531] Вариант осуществления 57. Система по варианту осуществления 55 или варианту осуществления 56, в которой кратность улучшения характеристики RNP варианта белка CasX и варианта гРНК составляет от по меньшей мере примерно 1,1 до примерно 100 раз или более относительно RNP референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

[0532] Вариант осуществления 58. Система по варианту осуществления 55 или варианту осуществления 56, в которой кратность улучшения характеристики варианта белка CasX составляет по меньшей мере примерно 1,1, по меньшей мере примерно 2, по меньшей мере примерно 10, по меньшей мере примерно 100 раз или более относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

[0533] Вариант осуществления 59. Система по варианту осуществления 55 или варианту осуществления 56, в которой кратность улучшения характеристики варианта белка CasX составляет по меньшей мере примерно 1,1, по меньшей мере примерно 2, по меньшей мере примерно 10, по меньшей мере примерно 100 раз или более относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

[0534] Вариант осуществления 60. Система по любому из вариантов осуществления 55-59, в которой улучшенная характеристика включает эффективность редактирования, и RNP варианта белка CasX и варианта гНК характеризуется кратностью улучшения эффективности редактирования от 1,1 до 100 раз по сравнению с RNP референсного белка CasX с SEQ ID NO: 2 и гНК с любым из SEQ ID NO: 4-16.

[0535] Вариант осуществления 61. Система по любому из вариантов осуществления 54-59, в которой RNP, содержащий вариант CasX и вариант гНК, демонстрирует более высокую эффективность редактирования и/или связывание последовательности целевой нуклеиновой кислоты в клеточной системе анализа, когда любая из РАМ-последовательностей - TTC, ATC, GTC или CTC - расположена на расстоянии 1 нуклеотида от 5'-конца нецелевой цепи протоспейсера, имеющего идентичность с нацеливающей последовательностью гНК, по сравнению с эффективностью редактирования и/или связыванием в сопоставимой системе анализа для RNP, содержащего референсный белок CasX и референсную гНК.

[0536] Вариант осуществления 62. Система по варианту осуществления 61, в которой РАМ-последовательность представляет собой TTC.

[0537] Вариант осуществления 63. Система по варианту осуществления 62, в которой нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 17904-26789.

[0538] Вариант осуществления 64. Система по варианту осуществления 61, в которой РАМ-последовательность представляет собой ATC.

[0539] Вариант осуществления 65. Система по варианту осуществления 64, в которой нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-5625.

[0540] Вариант осуществления 66. Система по варианту осуществления 61, в которой РАМ-последовательность представляет собой CTC.

[0541] Вариант осуществления 67. Система по варианту осуществления 66, в которой нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 5626-13616.

[0542] Вариант осуществления 68. Система по варианту осуществления 61, в которой РАМ-последовательность представляет собой GTC.

[0543] Вариант осуществления 69. Система по варианту осуществления 66, в которой нацеливающая последовательность гНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 13617-17903.

[0544] Вариант осуществления 70. Система по любому из вариантов осуществления 61-68, в которой повышенная аффинность связывания с одной или более РАМ-последовательностями в по меньшей мере 1,5 раза выше по сравнению с аффинностью связывания любого из референсных белков CasX с SEQ ID NO: 1-3 для РАМ-последовательностей.

[0545] Вариант осуществления 71. Система по любому из вариантов осуществления 54-70, в которой RNP имеет на по меньшей мере 5%, на по меньшей мере 10%, на по меньшей мере 15% или на по меньшей мере 20% более высокую процентную долю компетентных в отношении расщепления RNP по сравнению с RNP референсного белка CasX и гРНК с SEQ ID NO: 4-16.

[0546] Вариант осуществления 72. Система по любому из вариантов осуществления 41-71, в которой вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий никазной активностью.

[0547] Вариант осуществления 73. Система по любому из вариантов осуществления 41-71, в которой вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий активностью расщепления двухцепочечных молекул.

[0548] Вариант осуществления 74. Система по любому из вариантов осуществления 1-54, в которой белок CasX представляет собой каталитически неактивный белок CasX (dCasX), и причем dCasX и гРНК сохраняют способность связываться с целевой нуклеиновой кислотой BCL11A.

Вариант осуществления 75. Система по варианту осуществления 74, в которой dCasX содержит мутацию по остаткам:

а. D672, E769 и/или D935, соответствующим белку CasX с SEQ ID NO:1; или

[0549] б. D659, E756 и/или D922, соответствующим белку CasX с SEQ ID NO: 2.

[0550] Вариант осуществления 76. Система по варианту осуществления 75, в которой мутация представляет собой замену остатка на аланин.

[0551] Вариант осуществления 77. Система по любому из вариантов осуществления 1-73, дополнительно содержащая нуклеиновую кислоту донорной матрицы.

[0552] Вариант осуществления 78. Система по варианту осуществления 77, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок гена BCL11A, выбранный из группы, состоящей из экзона BCL11A, интрона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A и последовательности сайта связывания GATA1.

[0553] Вариант осуществления 79. Система по варианту осуществления 78, в которой последовательность донорной матрицы содержит одну или более мутаций относительно соответствующего участка гена BCL11A дикого типа.

[0554] Вариант осуществления 80. Система по варианту осуществления 78 или варианту осуществления 79, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A, экзона 3 BCL11A, экзона 4 BCL11A, экзона 5

BCL11A, экзона 6 BCL11A, экзона 7 BCL11A, экзона 8 BCL11A и экзона 9 BCL11A.

[0555] Вариант осуществления 81. Система по варианту осуществления 80, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A и экзона 3 BCL11A.

[0556] Вариант осуществления 82. Система по любому из вариантов осуществления 77-81, в которой размер донорной матрицы находится в диапазоне от 10 до 15 000 нуклеотидов.

[0557] Вариант осуществления 83. Система по любому из вариантов осуществления 77-82, в которой донорная матрица представляет собой одноцепочечную ДНК-матрицу или одноцепочечную РНК-матрицу.

[0558] Вариант осуществления 84. Система по любому из вариантов осуществления 77-82, в которой донорная матрица представляет собой двухцепочечную ДНК-матрицу.

[0559] Вариант осуществления 85. Система по любому из вариантов осуществления 77-84, в которой донорная матрица содержит гомологичные плечи на 5'- и 3'-концах донорной матрицы, которые комплементарны последовательностям, фланкирующим сайты расщепления, в целевой нуклеиновой кислоте BCL11A, вводимой белком CRISPR класса 2, типа V, или рядом с ними.

[0560] Вариант осуществления 86. Нуклеиновая кислота, содержащая донорную матрицу по любому из вариантов осуществления 77-85.

[0561] Вариант осуществления 87. Нуклеиновая кислота, содержащая последовательность, которая кодирует CasX по любому из вариантов осуществления 41-76.

[0562] Вариант осуществления 88. Нуклеиновая кислота, содержащая последовательность, которая кодирует гРНК по любому из вариантов осуществления 1-39.

[0563] Вариант осуществления 89. Нуклеиновая кислота по варианту осуществления 87, в которой последовательность, которая кодирует белок CasX, является кодон-оптимизированной для экспрессии в эукариотической клетке.

[0564] Вариант осуществления Вектор, содержащий гРНК по любому из вариантов осуществления 1-39, белок CasX по любому из вариантов осуществления 41-76 или нуклеиновую кислоту по любому из вариантов осуществления 86-89.

[0565] Вариант осуществления 91. Вектор по варианту осуществления 90, дополнительно содержащий один или более промоторов.

[0566] Вариант осуществления 92. Вектор по варианту осуществления 90 или варианту осуществления 91, который выбран из группы, состоящей из ретровирусного вектора, лентивирусного вектора, аденовирусного вектора, аденоассоциированного вирусного (AAV) вектора, вектора на основе вируса простого герпеса (HSV), вирусоподобной частицы (VLP), частицы для доставки CasX (XDP), плазмиды, миникольца, наноплазмиды, ДНК-вектора и РНК-вектора.

[0567] Вариант осуществления 93. Вектор по варианту осуществления 92, который представляет собой вектор AAV.

[0568] Вариант осуществления 94. Вектор по варианту осуществления 93, причем вектор AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 или AAVRh10.

[0569] Вариант осуществления 95. Вектор по варианту осуществления 94, причем вектор AAV выбран из AAV1, AAV2, AAV5, AAV8 или AAV9.

[0570] Вариант осуществления 96. Вектор по варианту осуществления 94 или варианту осуществления 95, причем вектор AAV содержит нуклеиновую кислоту, содержащую следующие компоненты:

a. 5'-ITR;

b. 3'-ITR; и

c. транскрипционный, содержащий нуклеиновую кислоту по варианту осуществления 87, функционально связанную с первым промотором, и нуклеиновую кислоту по варианту осуществления 88, функционально связанную со вторым промотором.

[0571] Вариант осуществления 97. Вектор по варианту осуществления 96, в котором нуклеиновая кислота дополнительно содержит поли(A)-последовательность с 3'-конца от последовательности, кодирующей белок CasX.

[0572] Вариант осуществления 98. Вектор по варианту осуществления 96 или варианту осуществления 97, в котором нуклеиновая кислота дополнительно содержит один или более энхансерных элементов.

[0573] Вариант осуществления 99. Вектор по любому из вариантов осуществления 96-98, в котором одна частица AAV способна содержать нуклеиновую кислоту, причем частица AAV имеет все компоненты, необходимые для трансдукции и эффективной модификации целевой нуклеиновой в клетке-мишени.

[0574] Вариант осуществления 100. Вектор по варианту осуществления 92, который представляет собой ретровирусный вектор.

[0575] Вариант осуществления 101. Вектор по варианту осуществления 92, который представляет собой XDP, содержащую один или более компонентов полибелка gag.

[0576] Вариант осуществления 102. Вектор по варианту осуществления 101, в котором один или более компонентов полибелка gag выбраны из группы, состоящей из матричного белка (MA), нуклеокапсидного белка (NC), капсидного белка (CA), пептида p1, пептида p6, пептида P2A, пептида P2B, пептида P10, пептида p12, пептида PP21/24, пептида P12/P3/P8 и пептида P20.

[0577] Вариант осуществления 103. Вектор по варианту осуществления 101 или варианту осуществления 102, в котором XDP содержит один или более компонентов полибелка gag, белок CasX и гРНК.

[0578] Вариант осуществления 104. Вектор по варианту осуществления 103, в котором белок CasX и гРНК связаны друг с другом в RNP.

[0579] Вариант осуществления 105. Вектор по любому из вариантов осуществления 101-104, дополнительно содержащий донорную матрицу.

[0580] Вариант осуществления 106. Вектор по любому из вариантов осуществления

101-104, дополнительно содержащий псевдотипирующий гликопротеин вирусной оболочки или фрагмент антитела, который обеспечивает связывание и слияние XDP с клеткой-мишенью.

[0581] Вариант осуществления 107. Вектор по варианту осуществления по варианту осуществления, причем клетка-мишень выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), эмбриональной стволовой (ES) клетки, индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

[0582] Вариант осуществления 108. Клетка-хозяин, содержащая вектор по любому из вариантов осуществления 90-107.

[0583] Вариант осуществления 109. Клетка-хозяин по варианту осуществления 108, которая выбрана из группы, состоящей из ВНК, HEK293, HEK293T, NS0, SP2/0, клеток миеломы YO, мышинных клеток миеломы P3X63, PER, PER.C6, NIH3T3, COS, HeLa, CHO и дрожжевых клеток.

[0584] Вариант осуществления 110. Способ модификации последовательности целевой нуклеиновой кислоты BCL11A в популяции клеток, включающий введение в клетки популяции:

- a. системы по любому из вариантов осуществления 1-85;
- b. нуклеиновой кислоты по любому из вариантов осуществления 86-89;
- c. вектора по любому из вариантов осуществления 90-95;
- d. XDP по любому из вариантов осуществления 101-107; или
- e. комбинаций двух или более из пп. (a)-(d),

причем последовательность целевой нуклеиновой кислоты гена BCL11A клеток, на которую нацелена первая гРНК, модифицирована вариантом белка CasX.

[0585] Вариант осуществления 111. Способ по варианту осуществления 110, в котором модификация включает введение одноцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0586] Вариант осуществления 112. Способ по варианту осуществления 110, в котором модификация включает введение двухцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0587] Вариант осуществления 113. Способ по любому из вариантов осуществления 110-112, дополнительно включающий введение в клетки популяции второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты гена BCL11A по сравнению с первой гРНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток популяции.

[0588] Вариант осуществления 114. Способ по любому из вариантов осуществления 110-113, в котором модификация включает введение вставки, делеции, замены,

дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток популяции.

[0589] Вариант осуществления 115. Способ по варианту осуществления 110-114, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

[0590] Вариант осуществления 116. Способ по любому из вариантов осуществления 110-113, который включает вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток популяции.

[0591] Вариант осуществления 117. Способ по варианту осуществления 114, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (HITI).

[0592] Вариант осуществления 118. Способ по варианту осуществления 116 или варианту осуществления 117, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

[0593] Вариант осуществления 119. Способ по любому из вариантов осуществления 116-118, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в клетках популяции.

[0594] Вариант осуществления 120. Способ по любому из вариантов осуществления 110-119, в котором ген BCL11A клеток популяции модифицирован так, что экспрессия белка BCL11A понижается на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, в которых ген BCL11A не модифицирован.

[0595] Вариант осуществления 121. Способ по любому из вариантов осуществления 110-119, в котором ген BCL11A клеток популяции модифицирован так, что по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0596] Вариант осуществления 122. Способ по любому из вариантов осуществления 110-121, в котором клетки являются эукариотическими.

[0597] Вариант осуществления 123. Способ по варианту осуществления 122, в котором эукариотические клетки выбраны из группы, состоящей из клеток грызунов, клеток мыши, клеток крысы и клеток примата, не являющегося человеком.

[0598] Вариант осуществления 124. Способ по варианту осуществления 122, в котором эукариотические клетки представляют собой клетки человека.

[0599] Вариант осуществления 125. Способ по любому из вариантов осуществления 122-124, в котором эукариотическая клетка выбрана из группы, состоящей из

гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

[0600] Вариант осуществления 126. Способ по любому из вариантов осуществления 110-125, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vitro* или *ex vivo*.

[0601] Вариант осуществления 127. Способ по любому из вариантов осуществления 110-125, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vivo* у субъекта.

[0602] Вариант осуществления 128. Способ по варианту осуществления 127, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0603] Вариант осуществления 129. Способ по варианту осуществления 127, в котором субъект является человеком.

[0604] Вариант осуществления 130. Способ по любому из вариантов осуществления 127-129, который включает введение терапевтически эффективной дозы вектора AAV субъекту.

[0605] Вариант осуществления 131. Способ по варианту осуществления 130, в котором вектор AAV вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0606] Вариант осуществления 132. Способ по варианту осуществления 130, в котором вектор AAV вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

[0607] Вариант осуществления 133. Способ по любому из вариантов осуществления 127-129, включающий введение субъекту терапевтически эффективной дозы XDP.

[0608] Вариант осуществления 134. Способ по варианту осуществления 133, в котором XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

[0609] Вариант осуществления 135. Способ по варианту осуществления 133, в котором XDP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг

[0610] Вариант осуществления 136. Способ по любому из вариантов осуществления 128-135, в котором вектор или XDP вводят субъекту способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0611] Вариант осуществления 137. Способ по любому из вариантов осуществления 128-136, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0612] Вариант осуществления 138. Способ по любому из вариантов осуществления 128-137, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0613] Вариант осуществления 139. Способ по любому из вариантов осуществления 128-138, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0614] Вариант осуществления 140. Способ по любому из вариантов осуществления 110-139, дополнительно включающий приведение последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток в контакт с:

a. дополнительной нуклеазой CRISPR и гРНК, нацеленной на другой или перекрывающийся участок целевой нуклеиновой кислоты BCL11A по сравнению с первой гРНК;

b. полинуклеотидом, кодирующим дополнительную нуклеазу CRISPR и гРНК по п. (a);

c. вектором, содержащим полинуклеотид по п. (b); или

d. XDP, содержащей дополнительную нуклеазу CRISPR и гРНК по п. (a),

причем приведение в контакт приводит к модификации гена BCL11A в другом положении в последовательности по сравнению с последовательностью, на которую нацелена первая гРНК.

[0615] Вариант осуществления 141. Способ по варианту осуществления 140, причем дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий

последовательность, отличную от белка CasX по любому из предшествующих вариантов осуществления.

[0616] Вариант осуществления 142. Способ по варианту осуществления 140, причем дополнительная нуклеаза CRISPR не представляет собой белок CasX.

[0617] Вариант осуществления 143. Способ по варианту осуществления 142, в котором дополнительная нуклеаза CRISPR выбрана из группы, состоящей из Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, Cas13a, Cas13b, Cas13c, Cas13d, Cas14, Cpf1, C2cl, Csn2 и вариантов их последовательностей.

[0618] Вариант осуществления 144. Популяция клеток, модифицированная способом по любому из вариантов осуществления 110-143, причем клетки были модифицированы так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаружимом уровне.

[0619] Вариант осуществления 145. Популяция клеток, модифицированная способом по любому из вариантов осуществления 110-143, причем клетки были модифицированы так, что экспрессия белка BCL11A снижена на по меньшей мере 70%, на по меньшей мере 75%, на по меньшей мере 80%, на по меньшей мере 85%, на по меньшей мере 90% или на по меньшей мере 95% по сравнению с клетками, в которых ген BCL11A не был модифицирован.

[0620] Вариант осуществления 146. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий введение субъекту терапевтически эффективного количества клеток по варианту осуществления 144 или варианту осуществления 145.

[0621] Вариант осуществления 147. Способ по варианту осуществления 146, причем гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0622] Вариант осуществления 148. Способ по варианту осуществления 146 или варианту осуществления 147, в котором клетки являются аутологичными по отношению к субъекту, которому должны быть введены клетки.

[0623] Вариант осуществления 149. Способ по варианту осуществления 146 или варианту осуществления 147, в котором клетки являются аллогенными по отношению к субъекту, которому должны быть введены клетки.

[0624] Вариант осуществления 150. Способ по любому из вариантов осуществления 146-149, в котором клетки или их потомство сохраняются у субъекта в течение по меньшей мере одного месяца, двух месяцев, трех месяцев, четырех месяцев, пяти месяцев, шести месяцев, семи месяцев, восьми месяцев, девяти месяцев, десяти месяцев, одиннадцать месяцев, двенадцати месяцев, тринадцати месяцев, четырнадцати месяцев, пятнадцати месяцев, шестнадцати месяцев, семнадцати месяцев, восемнадцати месяцев, девятнадцати месяцев, двадцати месяцев, двадцати месяцев, двадцати двух месяцев, двадцати трех месяцев, двух лет, трех лет, четырех лет или пяти лет после введения модифицированных

клеток субъекту.

[0625] Вариант осуществления 151. Способ по любому из вариантов осуществления 146-150, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0626] Вариант осуществления 152. Способ по любому из вариантов осуществления 146-150, который приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0627] Вариант осуществления 153. Способ по любому из вариантов осуществления 146-150, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0628] Вариант осуществления 154. Способ по любому из вариантов осуществления 146-153, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0629] Вариант осуществления 155. Способ по любому из вариантов осуществления 146-153, в котором субъект является человеком.

[0630] Вариант осуществления 156. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий модификацию гена BCL11A в клетках субъекта, причем модификация включает приведение упомянутых клеток в контакт с терапевтически эффективной дозой:

- a. системы по любому из вариантов осуществления 1-85;
- b. нуклеиновой кислоты по любому из вариантов осуществления 86-89;
- c. вектора по любому из вариантов осуществления 90-95;
- d. XDP по любому из вариантов осуществления 101-104; или
- e. комбинаций двух или более из пп. (a)-(d),

причем ген BCL11A клеток, на который нацелена первая гРНК, модифицирован белком CasX.

[0631] Вариант осуществления 157. Способ по варианту осуществления 156, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0632] Вариант осуществления 158. Способ по любому из вариантов осуществления 156 или варианта осуществления 157, в котором клетка выбрана из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных

плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

[0633] Вариант осуществления 159. Способ по любому из вариантов осуществления 156-158, в котором модификация включает введение одноцепочечного разрыва в ген BCL11A клеток.

[0634] Вариант осуществления 160. Способ по любому из вариантов осуществления 156-158, в котором модификация включает введение двухцепочечного разрыва в ген BCL11A клеток.

[0635] Вариант осуществления 161. Способ по любому из вариантов осуществления 156-160, дополнительно включающий введение в клетки субъекта второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты по сравнению с первой гРНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток субъекта.

[0636] Вариант осуществления 162. Способ по любому из вариантов осуществления 156-161, в котором модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток.

[0637] Вариант осуществления 163. Способ по варианту осуществления 162, в котором модификация приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

[0638] Вариант осуществления 164. Способ по любому из вариантов осуществления 156-163, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, которые не были модифицированы.

[0639] Вариант осуществления 165. Способ по любому из вариантов осуществления 156-163, в котором ген BCL11A клеток субъекта модифицирован так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0640] Вариант осуществления 166. Способ по любому из вариантов осуществления 156-165, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0641] Вариант осуществления 167. Способ по любому из вариантов осуществления 156-166, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке

субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере, 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0642] Вариант осуществления 168. Способ по любому из вариантов осуществления 156-165, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0643] Вариант осуществления 169. Способ по любому из вариантов осуществления 156-161, который включает вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток.

[0644] Вариант осуществления 170. Способ по варианту осуществления 168, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (НТИ).

[0645] Вариант осуществления 171. Способ по варианту осуществления 168 или варианту осуществления 170, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

[0646] Вариант осуществления 172. Способ по любому из вариантов осуществления 166-171, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, которые не были модифицированы.

[0647] Вариант осуществления 173. Способ по любому из вариантов осуществления 166-171, в котором ген BCL11A клеток субъекта модифицирован так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

[0648] Вариант осуществления 174. Способ по любому из вариантов осуществления 166-173, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0649] Вариант осуществления 175. Способ по любому из вариантов осуществления 166-173, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0 по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере

мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере, 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0650] Вариант осуществления 176. Способ по любому из вариантов осуществления 166-173, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0651] Вариант осуществления 177. Способ по любому из вариантов осуществления 156-175, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

[0652] Вариант осуществления 178. Способ по любому из вариантов осуществления 156-175, в котором субъект является человеком.

[0653] Вариант осуществления 179. Способ по любому из вариантов осуществления 156-178, в котором вектор представляет собой вектор AAV, который вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0654] Вариант осуществления 180. Способ по любому из вариантов осуществления 156-178, в котором вектор представляет собой вектор AAV, который вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

[0655] Вариант осуществления 181. Способ по любому из вариантов осуществления 156-178, в котором XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

[0656] Вариант осуществления 182. Способ по любому из вариантов осуществления 156-178, в котором XDP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг.

[0657] Вариант осуществления 183. Способ по любому из вариантов осуществления 156-182, в котором вектор или XDP вводят субъекту способом введения, выбранным из

интрапаренохиматозного, внутривенного, внутриаартериального, внутрибрюшинного, внутрикапсульного, подкожного, внутримышечного, интраабдоминального или их комбинаций, причем способ введения представляет собой инъекцию, трансфузию или имплантацию.

[0658] Вариант осуществления 184. Способ по любому из вариантов осуществления 156-183, который приводит к улучшению по меньшей мере одной клинически значимой конечной точки у субъекта.

[0659] Вариант осуществления 185. Способ по варианту осуществления 184, который приводит к улучшению по меньшей мере одного клинически значимого параметра, выбранного из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

[0660] Вариант осуществления 186. Способ по варианту осуществления 184, который приводит к улучшению по меньшей мере двух клинически значимых параметров, выбранных из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

[0661] Вариант осуществления 187. Способ лечения субъекта с гемоглобинопатией, включающий:

- a. выделение индуцированных плюрипотентных стволовых клеток (iPSC) или гематопозитических стволовых клеток (HSC) у субъекта;
- b. модификацию целевой нуклеиновой кислоты BCL11A в iPSC или HSC способом по любому из вариантов осуществления 110-126;
- c. дифференцировку модифицированных iPSC или HSC в гематопозитическую клетку-предшественника; и
- d. имплантацию гематопозитической клетки-предшественника субъекту с

гемоглобинопатией,

причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0662] Вариант осуществления 188. Способ по варианту осуществления 187, в котором iPSC или HSC является аутологичной и выделена из костного мозга или из периферической крови субъекта.

[0663] Вариант осуществления 189. Способ по варианту осуществления 187, в котором iPSC или HSC является аллогенной и выделена из костного мозга или из периферической крови другого субъекта.

[0664] Вариант осуществления 190. Способ по любому из вариантов осуществления 187-189, в котором имплантация включает введение гематопозитической клетки-предшественника субъекту посредством трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0665] Вариант осуществления 191. Способ по любому из вариантов осуществления 187-190, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

[0666] Вариант осуществления 192. Способ повышения уровня эмбрионального гемоглобина (HbF) у субъекта посредством редактирования генома, включающий:

а. введение субъекту эффективной дозы вектора по любому из вариантов осуществления 90-95 или XDP по любому из вариантов осуществления 101-107, причем вектор или XDP доставляет систему CasX:гРНК в клетки субъекта;

б. редактирование целевой нуклеиновой кислоты BCL11A клеток субъекта с помощью CasX, на который нацелена первая гРНК;

с. редактирование включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в последовательности целевой нуклеиновой кислоты так, что экспрессия белка BCL11A снижается или устраняется,

причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

[0667] Вариант осуществления 193. Способ по варианту осуществления 192, который приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере, 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

[0668] Вариант осуществления 194. Способ по варианту осуществления 192 или варианту осуществления 193, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

[0669] Вариант осуществления 195. Способ по любому из вариантов осуществления 192-194, в котором клетка выбрана из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

[0670] Вариант осуществления 196. Способ по любому из вариантов осуществления 192-195, в котором целевая нуклеиновая кислота клеток отредактирована таким образом, что экспрессия белка BCL11A снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с целевой нуклеиновой кислотой клеток, которые не были отредактированы.

[0671] Вариант осуществления 197. Способ по любому из вариантов осуществления 192-196, в котором субъект выбран из группы, состоящей из мыши, крысы, свиньи и примата, не являющегося человеком.

[0672] Вариант осуществления 198. Способ по любому из вариантов осуществления 192-196, в котором субъект является человеком.

[0673] Вариант осуществления 199. Способ по любому из вариантов осуществления 192-198, в котором вектор вводят в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

[0674] Вариант осуществления 200. Способ по любому из вариантов осуществления 192-198, в котором XDP вводят в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг или по меньшей мере примерно 1×10^{16} частиц/кг.

[0675] Вариант осуществления 201. Способ по любому из вариантов осуществления

192-200, в котором вектор или XDP вводят способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

[0676] Вариант осуществления 202. Система по любому из вариантов осуществления 1-85, нуклеиновая кислота по любому из вариантов осуществления 86-89, вектор по любому из вариантов осуществления 90-95, XDP по любому из вариантов осуществления 101-104, клетка-хозяин по варианту осуществления 108 или варианту осуществления 109 или популяция клеток по варианту осуществления 144 или варианту осуществления 145 для применения в качестве лекарственного средства для лечения гемоглобинопатии.

[0677] Вариант осуществления 203. Система по варианту осуществления 1, в которой последовательность целевой нуклеиновой кислоты комплементарна последовательности нецелевой цепи, расположенной на расстоянии 1 нуклеотида от 3'-конца последовательности смежного с протоспейсером мотива (PAM).

[0678] Вариант осуществления 204. Система по варианту осуществления 203, в которой PAM-последовательность содержит мотив TC.

[0679] Вариант осуществления 205. Система по варианту осуществления 204, в которой PAM-последовательность содержит ATC, GTC, CTC или TTC.

[0680] Вариант осуществления 206. Система по любому из вариантов осуществления 203-205, в которой белок CRISPR класса 2, типа V содержит домен RuvC.

[0681] Вариант осуществления 207. Система по варианту осуществления 206, в которой домен RuvC создает смещенный двухцепочечный разрыв в последовательности целевой нуклеиновой кислоты.

[0682] Вариант осуществления 208. Система по любому из вариантов осуществления 203-207, в которой белок CRISPR класса 2, типа V не содержит нуклеазного домена HNH.

ПРИМЕРЫ

Пример 1. Создание вариантов конструкторов CasX

[0683] Чтобы создать конструктор CasX 488 (последовательности в таблице 6), кодон-оптимизированный конструктор CasX 119 (на основе конструктора CasX Stx2, кодирующего CasX *Planctomycetes* с SEQ ID NO: 2, с аминокислотными заменами и делециями) клонировали в плазмиду назначения (pStX) с применением стандартных способов клонирования. Чтобы создать конструктор CasX 491 (последовательности в таблице 6), кодон-оптимизированный конструктор CasX 484 (на основе конструктора CasX Stx2, кодирующего CasX *Planctomycetes* с SEQ ID NO: 2, с заменами и делециями определенных аминокислот, со слитым NLS и связанными гидовыми и ненацеливающими последовательностями) клонировали в плазмиду назначения (pStX) с применением стандартных способов клонирования. Конструктор CasX 1 (CasX с SEQ ID NO: 1) клонировали в вектор назначения с применением стандартных способов клонирования. Для построения CasX 488 ДНК конструктора CasX 119 амплифицировали методом ПЦР в двух реакциях с применением ДНК-полимеразы Q5 в соответствии с протоколом производителя и подходящих универсальных праймеров. Для построения кодон-оптимизированного CasX 491 ДНК конструктора CasX 484 амплифицировали методом ПЦР в двух реакциях с применением

ДНК-полимеразы Q5 в соответствии с протоколом производителя и подходящих праймеров. Конструкт CasX 1 амплифицировали методом ПЦР в двух реакциях с применением ДНК-полимеразы Q5 в соответствии с протоколом производителя и подходящих универсальных праймеров. Каждый из продуктов ПЦР очищали посредством гель-экстракции из 1% агарозного геля (Gold Bio, кат. № A-201-500) с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. Затем соответствующие фрагменты соединяли вместе с применением сборки Гибсона (New England BioLabs, кат. № E2621S) в соответствии с протоколом производителя. Сборными продуктами в pStx1 трансформировали химически компетентные бактериальные клетки Turbo Competent *E. coli*, высевали клетки на чашки с LB-агаром, содержащим канамицин. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiagen spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности сборки. Затем правильные клоны субклонировали в экспрессионный вектор млекопитающих pStx34 с применением клонирования с помощью ферментов рестрикции. Каркас pStx34 и клоны CasX 488 и 491 в pStx1 расщепляли с помощью XbaI и BamHI соответственно. Расщепленный каркас и соответствующие фрагменты вставки очищали посредством гель-экстракции из 1% агарозного геля (Gold Bio, кат. № A-201-500) с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. После этого очищенные каркас и вставку лигировали друг с другом с помощью T4 лигазы (New England Biolabs, кат. № M0202L) в соответствии с протоколом производителя. Лигированными продуктами трансформировали химически компетентные бактериальные клетки Turbo Competent *E. Coli*, и клетки высевали на чашки с LB-агаром, содержащим карбенициллин. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiagen spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности сборки.

[0684] Для построения CasX 515 (последовательности в таблице 6) ДНК конструкта CasX 491 амплифицировали методом ПЦР в двух реакциях с применением ДНК-полимеразы Q5 в соответствии с протоколом производителя и подходящих праймеров. Для построения CasX 527 (последовательности в таблице 6) ДНК конструкта CasX 491 амплифицировали методом ПЦР в двух реакциях с применением ДНК-полимеразы Q5 в соответствии с протоколом производителя и подходящих праймеров. Продукты ПЦР очищали посредством гель-экстракции из 1% агарозного геля с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. Каркас pStX расщепляли с помощью XbaI и SpeI для удаления фрагмента ДНК размером 2931 пара нуклеотидов между двумя сайтами в плазмиде pStx56. Фрагмент расщепленного каркаса очищали посредством гель-экстракции из 1% агарозного геля с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. Фрагменты вставки и каркаса затем соединяли вместе с использованием сборки Гибсона (New England BioLabs, кат. № E2621S) в соответствии с протоколом производителя. Сборными продуктами в

pStx56 трансформировали химически компетентные бактериальные клетки Turbo Competent *E. Coli*, и клетки высевали на чашки с LB-агаром, содержащим канамицин. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiagen spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности сборки. Плазмида pStx34 содержит промотор EF-1 α для белка, а также селективный маркер, обеспечивающий устойчивость как к пурамицину, так и к карбенициллину. Плазмида pStx56 содержит промотор EF-1 α для белка, а также селективный маркер, обеспечивающий устойчивость как к пурамицину, так и к канамицину. Последовательности, кодирующие нацеливающие последовательности, которые нацелены на представляющий интерес ген, были разработаны на основе местоположений РАМ CasX. Нацеливающую ДНК-последовательность заказывали в виде одноцепочечных ДНК-олигонуклеотидов (оцДНК) (Integrated DNA Technologies), состоящих из нацеливающей последовательности и обратного комплемента этой последовательности. Был проведен отжиг этих двух олигонуклеотидов, и их клонировали в pStx по отдельности или вместе методом сборки Golden Gate с применением ДНК-лигазы T4 и соответствующего фермента рестрикции для плазмиды. Продуктами сборки Golden Gate трансформировали химически компетентные или электрокомпетентные клетки, такие как NEB Turbo Competent *E. coli* (NEB, кат. № C2984I), и клетки высевали на чашки с LB-агаром, содержащим соответствующий антибиотик. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiaprep spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности лигирования.

[0685] Для построения CasX 535-537 (последовательности в таблице 6) ДНК конструктора CasX 515 амплифицировали методом ПЦР в двух реакциях для каждого конструктора с применением ДНК-полимеразы Q5 в соответствии с протоколом производителя. Для CasX 535 для амплификации применяли соответствующие праймеры. Для CasX 536 применяли соответствующие праймеры. Для CasX 537 применяли соответствующие праймеры. Продукты ПЦР очищали посредством гель-экстракции из 1% агарозного геля с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. Каркас pStx расщепляли с помощью XbaI и SpeI для удаления фрагмента ДНК размером 2931 пара нуклеотидов между двумя сайтами в плазмиде pStx56. Фрагмент расщепленного каркаса очищали посредством гель-экстракции из 1% агарозного геля с помощью набора Zymoclean Gel DNA Recovery Kit в соответствии с протоколом производителя. Затем фрагменты вставки и каркаса соединяли вместе с использованием сборки Гибсона в соответствии с протоколом производителя. Сборными продуктами в pStx56 трансформировали химически компетентные бактериальные клетки Turbo Competent *E. coli*, высевали клетки на чашки с LB-агаром, содержащим канамицин. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiagen spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности сборки. Плазмида pStx34

содержит промотор EF-1 α для белка, а также селективный маркер, обеспечивающий устойчивость как к пуромицину, так и к карбенициллину. Плазмида pStX56 содержит промотор EF-1 α для белка, а также селективный маркер, обеспечивающий устойчивость как к пуромицину, так и к канамицину. Последовательности, кодирующие нацеливающие последовательности, которые нацелены на представляющий интерес ген, были разработаны на основе местоположений РАМ CasX. Нацеливающую ДНК-последовательность заказывали в виде одноцепочечных ДНК-олигонуклеотидов (оцДНК) (Integrated DNA Technologies), состоящих из нацеливающей последовательности и обратного комплемента этой последовательности. Был проведен отжиг этих двух олигонуклеотидов, и их клонировали в pStX по отдельности или вместе методом сборки Golden Gate с применением ДНК-лигазы T4 и соответствующего фермента рестрикции для плазмиды. Продуктами сборки Golden Gate трансформировали химически или электрокомпетентные клетки, такие как NEB Turbo Competent *E. coli*, и клетки высевали на чашки с LB-агаром, содержащим соответствующий антибиотик. Отдельные колонии собирали и проводили мини-выделение с помощью набора Qiaprep spin Miniprep Kit в соответствии с протоколом производителя. Полученные плазмиды секвенировали по методу Сэнгера для обеспечения правильности лигирования.

[0686] Все последующие варианты CasX, такие как CasX 544 и CasX 660-664, 668, 670, 672, 676 и 677, клонировали с применением той же методологии, которая описана выше, с использованием сборки Гибсона с внутренними праймерами, несущими специфические мутации, и универсальными прямыми и обратными праймерами (различия заключались в сконструированных праймерах, несущих специфические мутации, а также в том, какой базовый конструктор CasX применялся). Контрольные плазмиды SaCas9 и SpyCas9 получали аналогично описанным выше плазмидам pStX, причем белок-кодирующие и гидовые области pStX заменяли на соответствующие белок-кодирующую область и гид. Нацеливающие последовательности для SaCas9 и SpyCas9 были либо взяты из литературы, либо рационально разработаны в соответствии с известными способами.

[0687] Экспрессию и извлечение конструкторов CasX проводили с применением стандартных методик, как в обобщенном виде описано ниже.

Очистка:

[0688] Замороженные пробы размораживали в течение ночи при 4 °C с перемешиванием на магнитной мешалке. Вязкость полученного лизата снижали посредством обработки и завершали лизис гомогенизацией, дважды обработав в NanoDeBEE (BEE International) под давлением 20 000 фунтов на квадратный дюйм. Лизат осветляли центрифугированием при 50 000 $\times$ g, 4 °C в течение 30 минут и собирали надосадочную жидкость. Очищенную надосадочную жидкость наносили на колонку Heparin 6 Fast Flow (Cytiva) с помощью АКТА Pure FPLC (Cytiva). Колонку промывали 5 объемами колонки буфера Heparin A (50 mM HEPES-NaOH, 250 mM NaCl, 5 mM MgCl₂, 0,5 mM TCEP, 10% глицерин, pH 8), затем 3 объемами колонки буфера Heparin B (буфер A с концентрацией NaCl, доведенной до 500 mM). Белок элюировали 1,75 объема колонки

буфера Heparin C (буфер А с концентрацией NaCl, доведенной до 1 М). Элюат наносили на колонку с StrepTactis HP (Cytiva) с помощью системы для ЖХБР. Колонку промывали 10 объемами колонки буфера Strep (50 mM HEPES-NaOH, 500 mM NaCl, 5 mM MgCl₂, 0,5 mM ТСЕР, 10% глицерин, pH 8). Белок элюировали из колонки с применением 1,65 объема колонки буфера Strep с добавлением 2,5 mM дестиобиотина. Фракции, содержащие CasX, объединяли, концентрировали при 4 °С с помощью центрифужного концентратора с пределом отсечения по молекулярной массе 50 кДа (Amicon) и очищали посредством эксклюзионной хроматографии на колонке Superdex 200 pg (Cytiva). Колонку уравнивали буфером SEC (25 mM фосфата натрия, 300 mM NaCl, 1 mM ТСЕР, 10% глицерин, pH 7,25) и проводили ЖХБР. Фракции, содержащие CasX, которые элюировались при соответствующей молекулярной массе, объединяли, концентрировали при 4 °С с помощью центрифужного концентратора с пределом отсечения по молекулярной массе 50 кД, делили на аликвоты и мгновенно замораживали в жидком азоте перед хранением при - 80 °С.

[0689] Вариант CasX 488: средний выход составил 2,7 мг очищенного белка CasX на литр культуры при чистоте 98,8% по оценке с помощью окрашивания коллоидным Кумасси.

[0690] Вариант CasX 491: средний выход составил 12,4 мг очищенного белка CasX на литр культуры при чистоте 99,4% по оценке с помощью окрашивания коллоидным Кумасси.

[0691] Вариант CasX 515: средний выход составил 7,8 мг очищенного белка CasX на литр культуры при чистоте 90% по оценке с помощью окрашивания коллоидным Кумасси.

[0692] Вариант CasX 526: средний выход составил 13,79 мг на литр культуры при чистоте 93%. Чистоту оценивали с помощью окрашивания коллоидным Кумасси.

Таблица 6. Последовательности ДНК и аминокислотные последовательности вариантов CasX

Конструкт	SEQ ID NO последовательности ДНК	SEQ ID NO аминокислотной последовательности
CasX 488	27155	123
CasX 491	27156	126
CasX 515	27157	133
CasX 527	27158	144
CasX 535	27159	26911
CasX 536	27160	26912
CasX 537	27161	26913
CasX 583	27162	26958
CasX 660	27163	27035
CasX 661	27164	27036
CasX 662	27165	27037
CasX 663	27166	27038
CasX 664	27167	27039
CasX 668	27168	27043
CasX 670	27169	27154

CasX 672	27170	27046
CasX 676	27171	27050
CasX 677	27172	27051

Пример 2. Создание гидовых РНК

[0693] Для создания одномолекулярных гидовых РНК и нацеливающих последовательностей создавали матрицы для транскрипции *in vitro* посредством проведения ПЦР с ДНК-полимеразой Q5, матричными праймерами для каждой основной цепи и праймерами для амплификации с промотором T7 и нацеливающей последовательностью. Последовательности ДНК-праймеров для промотора T7, гидовой и нацеливающей последовательности для гидов и нацеливающих последовательностей представлены в таблице 7 ниже. Гиды sg1, sg2, sg32, sg64, sg174 и sg235 соответствуют SEQ ID NO: 4, 5, 2104, 2106, 2238 и 26800 соответственно, за исключением того, что sg2, sg32 и sg64 были модифицированы дополнительным G на 5'-конце для повышения эффективности транскрипции (сравните последовательности в таблице 7 с таблицей 3). Нацеливающая последовательность 7.37 нацелена на бета 2-микроглобулин (B2M). После ПЦР-амплификации матрицы очищали и выделяли посредством экстракции фенолом - хлороформом - изоамиловым спиртом с последующим осаждением этанолом.

[0694] Транскрипцию *in vitro* проводили в буфере, содержащем 50 mM Tris, pH 8,0, 30 mM MgCl₂, 0,01% Triton X-100, 2 mM спермидина, 20 mM DTT, 5 mM нуклеозидтрифосфатов (NTP), 0,5 мкМ матрицы и 100 мкг/мл РНК-полимеразы T7. Реакционные смеси инкубировали при 37 °C в течение ночи. Добавляли 20 единиц ДНКазы I (Promega, № M6101) на 1 мл объема транскрипционной смеси и инкубировали в течение одного часа. РНК-продукты очищали электрофорезом в ПААГ в денатурирующих условиях, осаждали этанолом и ресуспендировали в 1× фосфатно-солевом буферном растворе. Для осуществления фолдинга оРНК пробы нагревали до 70 °C в течение 5 мин, а затем охлаждали до комнатной температуры. В реакционные смеси добавляли MgCl₂ до конечной концентрации 1 mM, нагревали до 50 °C в течение 5 мин, а затем охлаждали до комнатной температуры. Конечные продукты - гидовые РНК - хранили при -80 °C.

Таблица 7. Последовательности ДНК-праймеров для промотора T7, гидовой и нацеливающей последовательности для гидов

Праймер	SEQ ID NO	РНК-продукт	SEQ ID NO
Праймер с промотором T7	234	Используется для всех	
Прямой праймер основной цепи sg2	238	GGUACUGGCGCUUUUAUCUCAUUAC UUUGAGAGCCAUCACCAGCGACUAU	251
Обратный праймер основной цепи sg2	239	GUCGUAUGGGUAAAGCGCUUAUUUA UCGGAGAGAAAUCCGAUAAUAAGA	
Праймер со	240	AGCAUCAAAAGGGCCGAGAUGUCUCG	

спейсером sg2.7.37		CUCCG	
Прямой праймер основной цепи sg32	241	GGUACUGGCGCUUUUAUCUCAUUAC UUUGAGAGCCAUCACCAGCGACUAU GUCGUAUGGGUAAAGCGCCCUUCUUC GGAGGGAAGCAUCAAGGGCCGAGA UGUCUCG	252
Обратный праймер основной цепи sg32	242		
Праймер со спейсером sg32.7.37	243		
Прямой праймер основной цепи sg64	244	GGUACUGGCGCCUUUAUCUCAUUAC UUUGAGAGCCAUCACCAGCGACUAU GUCGUAUGGGUAAAGCGCUUACGGA CUUCGGUCCGUAAGAAGCAUCAAG GGCCGAGAUUGUCUCGCUCCG	253
Обратный праймер основной цепи sg64	245		
Праймер со спейсером sg64.7.37	246		
Прямой праймер основной цепи sg174	247	ACUGGCGCUUUUAUCUgAUUACUUU GAGAGCCAUCACCAGCGACUAUGUC GUAgUGGGUAAAGCUCCCUUCUCGG AGGGAGCAUCAAGGGCCGAGAU CUCGCUCCG	254
Обратный праймер основной цепи sg174	248		
Праймер со спейсером sg174.7.37	249		
Прямой праймер основной цепи sg235	НД	ACUGGCGCUUCUAUCUGAUUACUCU GAGCGCCAUCACCAGCGACUAUGUC GUAGUGGGUAAAGCCGCUUACGGAC UUCGGUCCGUAAGAGGCAUCAGAG	26800
Обратный праймер основной цепи sg235	НД		
Праймер со спейсером sg235.7.37	НД		

Пример 3. Оценка аффинности связывания с гидовой РНК

[0695] Очищенный CasX дикого типа и улучшенный CasX будут инкубировать с синтетической одномолекулярной гидовой РНК, содержащей Cy7.5-фрагмент на 3'-конце, в низкосолевого буферном растворе, содержащем хлорид магния и гепарин для предотвращения неспецифического связывания и агрегации. Концентрация оРНК будет постоянной - 10 пМ, в то время как белок будут титровать от 1 пМ до 100 мкМ в отдельных реакциях связывания. После того как реакция придет в равновесное состояние, пробы пропускают через вакуумный коллектор для анализа связывания с фильтром, используя

нитроцеллюлозную мембрану и положительно заряженную нейлоновую мембрану, которые связывают белок и нуклеиновую кислоту соответственно. На мембранах проведут визуализацию для идентификации гидовой РНК, и будут определены фракции связанной и несвязанной РНК по количеству флуоресценции на нитроцеллюлозной и нейлоновой мембране при каждой концентрации белка для расчета константы диссоциации комплекса белок-огРНК. Также будет проведен эксперимент с улучшенными вариантами огРНК для определения того, влияют ли эти мутации на аффинность к гиду белков дикого типа и мутантных белков. Авторы изобретения также выполняют анализ методом сдвига электрофоретической подвижности для качественного сравнения с результатами анализа связывания с фильтром и подтверждения того, что основным фактором ассоциации белков и РНК является связывание в растворе, а не агрегация.

Пример 4. Оценка аффинности связывания с целевой ДНК

[0696] Очищенный CasX дикого типа и улучшенный CasX будут объединены в комплекс с одномолекулярной гидовой РНК, несущей нацеливающую последовательность, комплементарную целевой нуклеиновой кислоте. RNP-комплекс будут инкубировать с двухцепочечной целевой ДНК, содержащей PAM и соответствующую последовательность целевой нуклеиновой кислоты с Су7.5-меткой на 5'-конце целевой цепи, в низкосолевого буферном растворе, содержащем хлорид магния и гепарин для предотвращения неспецифического связывания и агрегации. Концентрация целевой ДНК будет постоянной - 1 нМ, в то время как RNP будут титровать от 1 пМ до 100 мкМ в отдельных реакциях связывания. После того как реакция придет в равновесное состояние, пробы пропустят через нативный 5% полиакриламидный гель для отделения связанной и несвязанной целевой ДНК. Будет проведена визуализация геля для определения сдвигов подвижности целевой ДНК и будут рассчитаны фракции связанной и несвязанной ДНК при каждой концентрации белка для определения константы диссоциации тройного комплекса RNP-целевая ДНК.

Пример 5. Оценка различий в распознавании PAM *in vitro*

1. Сравнение референсных CasX и вариантов CasX

[0697] Анализ расщепления *in vitro* проводили с помощью CasX 2, CasX 119 и CasX 438 в комплексе с sg174.7.37, по существу как описано в примере 8. Применяли флуоресцентно-меченые целевые дцДНК со спейсером 7.37 и PAM TTC, CTC, GTC или ATC (последовательности приведены в таблице 8). Пробы отбирали во временных точках 0,25, 0,5, 1, 2, 5, 10, 30 и 60 минут. Для визуализации гелей применяли Cytiva Typhoon и проводили количественную оценку с помощью программного обеспечения IQTL 8.2. Кажущиеся константы скорости реакции первого порядка для расщепления нецелевой цепи ($k_{\text{расщепл.}}$) определяли для каждого комплекса CasX:огРНК на каждой мишени. Константы скорости для мишеней с PAM без TTC сравнивали с мишенями с PAM TTC, чтобы определить, изменялось ли относительное предпочтение к каждой PAM в данном варианте белка.

[0698] Для всех вариантов мишень с TTC поддерживала самую высокую скоростью

расщепления, затем следовала мишень с АТС, затем с СТС и, наконец, с ГТС (ФИГ. 10А-Д, таблица 9). Для каждой комбинации варианта CasX и РАМ NTC приведена скорость расщепления $k_{\text{расщепл.}}$. Для всех РАМ без NTC в скобках показано отношение скорости расщепления к скорости при наличии ТТС для этого варианта. Все РАМ без ТТС демонстрировали существенно сниженные скорости расщепления (в > 10 раз во всех случаях). Соотношение скорости расщепления данной РАМ без ТТС и РАМ ТТС для конкретного варианта оставалось по существу постоянным у всех вариантов. В случае мишени с СТС обеспечивалось расщепление со скоростью 3,5-4,3% от скорости расщепления мишени с ТТС; в случае мишени с ГТС обеспечивалось расщепление со скоростью 1,0-1,4%; а в случае мишени с АТС обеспечивалось расщепление со скоростью 6,5-8,3%. Исключение составляет вариант 491, для которого кинетика расщепления при наличии РАМ ТТС слишком быстрая, чтобы ее можно было точно измерить, что искусственно уменьшает кажущуюся разницу между РАМ ТТС и РАМ без ТТС. Сравнение относительных скоростей для 491 при наличии РАМ ГТС, СТС и АТС, которые попадают в поддающийся измерению диапазон, дает соотношения, сопоставимые с полученными для других вариантов при сравнении по РАМ без ТТС, что согласуется с повышением скоростей в этом ряду. В целом, различия между вариантами не являются достаточно существенными, чтобы предположить, что у них будут изменены относительные предпочтения к различным РАМ NTC. Однако более высокие базовые скорости расщепления у вариантов позволяют расщеплять мишени с РАМ АТС или СТС практически полностью в течение 10 минут, а кажущиеся $k_{\text{расщепл.}}$ сопоставимы или больше, чем $k_{\text{расщепл.}}$ у CasX 2 на РАМ ТТС (таблица 9). Такая повышенная скорость расщепления может пересекать пороговое значение, необходимое для эффективного редактирования генома в клетке человека, что объясняет очевидное увеличение гибкости требований к РАМ для этих вариантов.

Таблица 8. Последовательности ДНК-субстратов, применяемых в анализе *in vitro* расщепления РАМ

Гид*	Последовательность ДНК	SEQ ID NO
7.37 TTC РАМ TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATCTCGGC CCGAATGCTGTCAGCTTCA	27176
7.37 TTC РАМ NTS	TGAAGCTGACAGCA TT CGGGCCGAGATGTCTCGCTCCGTGG CCTTAGCTGTGCTCGCGCT	27177
7.37 CTC РАМ TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATCTCGGC CCGAGTGCTGTCAGCTTCA	27178
7.37 CTC РАМ NTS	TGAAGCTGACAGCA CT CGGGCCGAGATGTCTCGCTCCGTGG CCTTAGCTGTGCTCGCGCT	27179
7.37 GTC РАМ TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATCTCGGC CCGACTGCTGTCAGCTTCA	27180

Гид*	Последовательность ДНК	SEQ ID NO
7.37 GTC PAM NTS	TGAAGCTGACAGCAGTCGGGCCGAGATGTCTCGCTCCGTGG CCTTAGCTGTGCTCGCGCT	27181
7.37 ATC PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATCTCGGC CCGATTGCTGTCAGCTTCA	27182
7.37 ATC PAM NTS	TGAAGCTGACAGCAATCGGGCCGAGATGTCTCGCTCCGTGG CCTTAGCTGTGCTCGCGCT	27183

* PAM-последовательности для каждого выделены жирным шрифтом. TS - целевая цепь. NTS - нецелевая цепь.

Таблица 9. Кажущиеся скорости расщепления PAM NTC для вариантов CasX

Вариант	TTC	CTC	GTC	ATC
2	0,267 мин ⁻¹	9,29E-3 мин ⁻¹ (0,035)	3,75E-3 мин ⁻¹ (0,014)	1,87E-2 мин ⁻¹ (0,070)
119	8,33 мин ⁻¹	0,303 мин ⁻¹ (0,036)	8,64E-2 мин ⁻¹ (0,010)	0,540 мин ⁻¹ (0,065)
438	4,94 мин ⁻¹	0,212 мин ⁻¹ (0,043)	1,31E-2 мин ⁻¹ (0,013)	0,408 мин ⁻¹ (0,083)
491	16,42 мин ⁻¹	8,605 мин ⁻¹ (0,524)	2,447 мин ⁻¹ (0,149)	11,33 мин ⁻¹ (0,690)

2. Сравнение распознавания PAM с применением одного варианта CasX

Материалы и методы

[0699] Применяли флуоресцентно-меченые целевые дцДНК со спейсером 7.37 и PAM TTC, CTC, GTC, ATC, TTT, CTT, GTT или ATT (последовательности приведены в таблице 10). Были заказаны олигонуклеотиды с аминокислотной модификацией 5'-конца, которые метили с помощью N-гидроксисукцинимидного (NHS) сложного эфира Cy7.5 в случае олигонуклеотидов целевой цепи и NHS сложного эфира Cy5.5 в случае олигонуклеотидов нецелевой цепи. Целевые дцДНК получали посредством смешивания олигонуклеотидов в соотношении 1:1 в 1× буфере для расщепления (20 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1 mM TCEP, 5% глицерина, 10 mM MgCl₂), нагревания до 95 °C в течение 10 минут и охлаждения раствора до комнатной температуры.

[0700] Вариант 491 CasX объединяли в комплекс с sg174.7.37. Гид разбавляли в 1× буфере для расщепления до конечной концентрации 1,5 мкМ, а затем добавляли белок до конечной концентрации 1 мкМ. RNP инкубировали при 37 °C в течение 10 минут и затем помещали на лед.

[0701] Анализ расщепления проводили посредством разбавления RNP в буфере для расщепления до конечной концентрации 200 нМ и добавления целевой дцДНК до конечной концентрации 10 нМ. Пробы отбирали во временных точках 0,25, 0,5, 1, 2, 5 и 10 минут и

гасили реакцию посредством добавления к равному объему 95% формамида и 20 мМ ЭДТА. Продукты расщепления разделяли в 10% ПААГ с мочевиной. Для визуализации гелей использовали Amersham Typhoon и проводили количественную оценку с помощью программного обеспечения IQTL 8.2. Кажущиеся константы скорости реакции первого порядка для расщепления нецелевой цепи (красщепл.-) определяли для каждой целевой ДНК с применением GraphPad Prism.

Результаты

[0702] Исследовали относительную скорость расщепления 491.174 RNP на различных РАМ. В дополнение к тому, что эти данные также помогут в прогнозировании эффективности расщепления целевых и потенциальных нецелевых последовательностей в клетках, они также позволяют корректировать скорость расщепления синтетических целевых последовательностей. Авторы изобретения предположили, что в случае самоограничивающихся векторов AAV, когда в вектор могут быть добавлены новые протоспейсеры для обеспечения самонацеливания, скорость расщепления эписом может быть увеличена или уменьшена путем изменения РАМ.

[0703] Мы тестировали скорость расщепления RNP в отношении различных субстратов дцДНК, которые были идентичны по последовательности, за исключением РАМ. Этот экспериментальный подход должен позволить изолированно рассмотреть влияние собственно РАМ, не смешивая распознавание РАМ с эффектами, обусловленными спейсерной последовательностью и геномным контекстом. Были протестированы все РАМ NTC и NTT. Как и ожидалось, RNP расщепляли целевую ДНК с РАМ TTC быстрее, превращая по существу ее всю в продукт к первой временной точке (ФИГ. 11А). CTC расщеплялась примерно в два раза медленнее, хотя быстрое расщепление TTC усложняет определение точного значения красщепл. в таких условиях анализа, которые оптимизированы для регистрации более широкого диапазона скоростей расщепления (ФИГ. 11А, таблица 11). Целевая GTC расщеплялась медленнее всего среди РАМ NTC, при этом скорость расщепления была примерно в шесть раз ниже, чем в случае целевой TTC. Все РАМ NTT расщеплялись медленнее, чем все РАМ NTC, при этом TTT расщеплялась наиболее эффективно, а после нее следовала GTT (ФИГ. 11В, таблица 11). Относительная эффективность расщепления GTT среди всех РАМ NTT по сравнению с низкой скоростью расщепления GTC в сравнении со всеми РАМ NTC демонстрирует, что распознавание отдельных нуклеотидов РАМ зависит от контекста, причем конкретный нуклеотид в одном положении в РАМ влияет на зависимость предпочтения последовательности от других положений.

[0704] Тестируемые в настоящем изобретении РАМ-последовательности дают скорости расщепления, охватывающие три порядка величины, с сохранением активности расщепления в одной и той же спейсерной последовательности. Эти данные демонстрируют, что скорости расщепления в данной синтетической целевой последовательности могут быть легко модифицированы посредством изменения ассоциированной РАМ, что позволяет корректировать активность саморасщепления для

обеспечения эффективного нацеливания на геномную целевую последовательность перед расщеплением и элиминацией AAV эписомы.

Таблица 10. Последовательности ДНК-субстратов, применяемых в анализе расщепления PAM *in vitro***

РАМ и цепь	Спейсерная и РАМ-последовательность	SEQ ID NO
7.37 TTC PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC GA TGCTGTCAGCTTCA	27176
7.37 TTC PAM NTS	TGAAGCTGACAGCA TT CGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27177
7.37 CTC PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC GA TGCTGTCAGCTTCA	27178
7.37 CTC PAM NTS	TGAAGCTGACAGCA CT CGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27179
7.37 GTC PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC GA CTGCTGTCAGCTTCA	27180
7.37 GTC PAM NTS	TGAAGCTGACAGCA GT CGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27181
7.37 ATC PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC GA TTGCTGTCAGCTTCA	27182
7.37 ATC PAM NTS	TGAAGCTGACAGCA AT CGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27183
7.37 TTT PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC AA TGCTGTCAGCTTCA	27184
7.37 TTT PAM NTS	TGAAGCTGACAGCA TTT GGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27185
7.37 CTT PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC TA GTGCTGTCAGCTTCA	27186
7.37 CTT PAM NTS	TGAAGCTGACAGCA CTT GGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27187
7.37 GTT PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC TA CTGCTGTCAGCTTCA	27188
7.37 GTT PAM NTS	TGAAGCTGACAGCA TTT GGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27189
7.37 ATT PAM TS	AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATC TCGGCCCC TA TTGCTGTCAGCTTCA	27190
7.37 ATT PAM NTS	TGAAGCTGACAGCA ATT GGGGCCGAGATGTCTCGCTCC GTGGCCTTAGCTGTGCTCGCGCT	27191

* Показаны последовательности ДНК, применяемые для создания каждого субстрата дцДНК. РАМ-последовательности для каждого из них выделены жирным шрифтом. TS - целевая цепь. NTS - нецелевая цепь.

Таблица 11. Кажущиеся скорости расщепления PAM NTC и NTT для CasX 491.174

РАМ	TTC	ATC	CTC	GTC	TTT	ATT	CTT	GTT
k _{расщепл.} (мин ⁻¹)	15,6*	6,66	9,45	2,52	1,33	0,0675	0,0204	0,330

* Скорость расщепления TTC превышает разрешение этого анализа, поэтому полученное значение k_{расщепл.} следует рассматривать в качестве нижней границы.

Пример 6. Оценка нуклеазной активности для двухцепочечного расщепления

[0705] Очищенные CasX дикого типа и сконструированные варианты CasX будут объединены в комплекс с одномолекулярной гидовой РНК, несущей фиксированную нацеливающую последовательность для HRS. RNP-комплексы добавляют к буферу, содержащему $MgCl_2$ в конечной концентрации 100 нМ, и будут инкубировать с двухцепочечной целевой ДНК, несущей Cy7.5-метку на 5'-конце целевой или нецелевой цепи, в концентрации 10 нМ. Аликвоты реакционных смесей будут отбирать в фиксированных временных точках и гасить реакцию посредством добавления равного объема 50 мМ ЭДТА и 95% формамида. Будет проведено разделение проб в полиакриламидном геле в денатурирующих условиях для разделения расщепленных и нерасщепленных ДНК-субстратов. Результаты визуализируют и определяют скорости расщепления целевых и нецелевых цепей белком дикого типа и сконструированными вариантами. Для более четкого различия между изменениями в связывании с целевой последовательностью и скоростью катализа самой нуклеолитической реакции концентрацию белка будут титровать в диапазоне от 10 нМ до 1 мкМ, и при каждой концентрации будут определять скорости расщепления для получения псевдокривой Михаэлиса - Ментен и определения k_{cat} и K_M . Изменения K_M указывают на изменение связывания, а изменения k_{cat} указывают на изменение катализа.

Пример 7. Анализ PASS позволяет идентифицировать варианты белка CasX с разной специфичностью в отношении PAM-последовательности

[0706] Были проведены эксперименты для идентификации специфичности к PAM-последовательности белков CasX 2 (SEQ ID NO: 2), 491 (SEQ ID NO: 126), 515 (SEQ ID NO: 133), 533 (SEQ ID NO: 26909), 535 (SEQ ID NO: 26911), 668 (SEQ ID NO: 27043) и 672 (SEQ ID NO: 27046). Для обеспечения этого линию клеток HEK293 PASS_V1.01 или PASS_V1.02 обрабатывали вышеуказанными белками CasX в по меньшей мере двух повторных экспериментах и проводили секвенирование нового поколения (NGS) для расчета процентной доли редактирования с применением разнообразных спейсеров в их заданных целевых сайтах.

[0707] Материалы и методы. Для анализа вариантов клонированного белка применяли мультиплексный объединенный подход с помощью системы PASS. Вкратце, были созданы две объединенные линии клеток HEK293, названные PASS_V1.01 и PASS_V1.02. Каждая клетка в пределах пула содержала интегрированную в геном одномолекулярную гидовую РНК (огРНК), спаренную со специфическим целевым сайтом. После трансфекции конструктами, экспрессирующими белок, редактирование по специфической мишени специфическим спейсером можно количественно оценивать с помощью NGS. Каждая пара гида и мишени была разработана таким образом, чтобы можно было получать данные, относящиеся к активности, специфичности и нацеливаемости RNP-комплекса CasX-гид.

[0708] Спаренные последовательности спейсер-мишень были синтезированы компанией Twist Biosciences и получены в виде эквимольного пула олигонуклеотидов.

Этот пул амплифицировали методом ПЦР и клонировали методом Golden Gate для создания конечной библиотеки плазмид, названной p77. Каждая плазида содержала элемент, экспрессирующий ogPНК и целевой сайт, вместе с элементом, экспрессирующим GFP. Элемент, экспрессирующий ogPНК, состоял из промотора U6, направляющего транскрипцию каркаса гPНК 174 (SEQ ID NO: 2238), за которым следовала спейсерная последовательность, которая должна нацеливать RNP из гида и варианта CasX на заданный целевой сайт. Были сконструированы и синтезированы 250 возможных уникальных спаренных синтетических последовательностей спейсер-мишень. Затем из этой библиотеки плазмид получали пул лентивируса с помощью системы LentiX (Takara Bio USA, Inc) в соответствии с инструкциями производителя. После этого полученный вирусный препарат количественно оценивали с помощью количественной полимеразной цепной реакции (qPCR) и трансдуцировали в стандартную линию клеток HEK293 при низкой множественности инфекции, чтобы создать однокопийные интегрирования. Затем полученную линию клеток очищали с помощью цитометрии посредством сортировки клеток с активацией флуоресценции (FACS) для завершения получения PASS_V1.01 или PASS_V1.02. После этого линию клеток высевали в шестиугольный планшет и в двух повторностях либо обрабатывали водой, либо трансфицировали 2 мкг плазмиды p67, доставляемой с помощью реагента для трансфекции Lipofectamine (ThermoFisher) в соответствии с инструкциями производителя. Плазида p67 содержит промотор EF-1альфа, направляющий экспрессию белка CasX, меченного последовательностью сигнала ядерной локализации SV40. Через два дня обработанные клетки собирали, лизировали и экстрагировали геномную ДНК с помощью набора для выделения геномной ДНК (Zymo Research). Затем геномную ДНК амплифицировали методом ПЦР с помощью специально разработанных праймеров для создания ампликонов, совместимых с NGS компании Illumina, и проводили секвенирование на приборе NextSeq. Прочтения проб демультимплексировали и фильтровали в зависимости от их качества. После этого количественно определяли показатели результатов редактирования (доля прочтений с инделами) для каждой синтетической последовательности спейсер-мишень в обработанных пробах.

[0709] Для оценки специфичности РАМ-последовательности для белка CasX были категоризированы показатели результатов редактирования для четырех разных РАМ-последовательностей. Для целевых сайтов с РАМ TTC количественно оценивали 48 разных пар спейсер-мишень; для целевых сайтов с РАМ ATC, CTC и GTC количественно оценивали 14, 22 и 11 отдельных целевых сайтов соответственно. Для некоторых белков CasX повторные эксперименты проводили несколько десятков раз в течение нескольких месяцев. Для каждого из этих экспериментов рассчитывали среднюю эффективность редактирования для каждого из вышеописанных спейсеров. Затем рассчитали среднюю эффективность редактирования по четырем категориям РАМ-последовательности по данным всех таких экспериментов, а также стандартное отклонение этих измерений.

Результаты

[0710] В таблице 12 указана средняя эффективность редактирования для разных категорий РАМ и вариантов белка CasX, а также стандартное отклонение этих измерений. Также указано число измерений для каждой категории. Эти данные свидетельствуют о том, что сконструированные варианты CasX 491 и 515 являются специфичными в отношении канонической РАМ-последовательности TTC, тогда как другие сконструированные варианты CasX с большей или меньшей эффективностью действовали на протестированные РАМ-последовательности. В частности, средний ранговый порядок предпочтений РАМ для CasX 491 представляет собой $TTC \gg ATC > CTC > GTC$ или $TTC \gg ATC > GTC > CTC$ для CasX 515, в то время как CasX 2 дикого типа демонстрирует средний ранговый порядок $TTC \gg GTC > CTC > ATC$. Следует отметить, что для РАМ-последовательностей с более низкой эффективностью редактирования ошибка этих средних измерений является высокой. Напротив, варианты CasX 535, 668 и 672 обладают значительно более широким спектром распознавания РАМ с ранговым порядком $TTC > CTC > ATC > GTC$. Наконец, CasX 533 демонстрирует полностью измененный порядок ранжирования относительно характерного для CasX дикого типа (WT): $ATC > CTC \gg GTC > TTC$. Эти данные можно применять для конструирования максимально активных терапевтических молекул CasX для интересующей последовательности целевой ДНК.

[0711] В условиях экспериментов был идентифицирован набор белков CasX, которые более эффективно расщепляют двухцепочечную ДНК в клетках человека в целевых последовательностях ДНК, ассоциированных с РАМ-последовательностью TTC, ATC, CTC или GTC, что свидетельствует в пользу вариантов CasX с измененным спектром специфичности к РАМ относительно CasX 491 для неканонических РАМ (т. е. ATC, CTC и GTC).

Таблица 12. Средний показатель редактирования, осуществляемого избранными белками CasX на спейсерах, ассоциированных с РАМ-последовательностями TTC, ATC, CTC или GTC

Название CasX	РАМ-последовательность	Средняя процентная доля редактирования	Стандартное отклонение	Число измерений
2	ATC	0,40	1,35	336
2	CTC	0,46	2,29	528
2	GTC	0,69	6,27	264
2	TTC	5,28	7,34	1152
491	ATC	6,86	8,29	364
491	CTC	4,54	6,40	572
491	GTC	3,40	6,68	286
491	TTC	40,41	23,13	1248

515	ATC	4,47	5,49	252
515	CTC	3,36	4,80	396
515	GTC	3,65	10,75	198
515	TTC	36,75	24,89	864
533	ATC	47,50	15,86	96
533	CTC	25,90	14,74	28
533	GTC	6,34	8,36	44
533	TTC	0,87	3,05	22
535	ATC	9,70	10,20	56
535	CTC	11,77	13,59	88
535	GTC	7,62	15,04	44
535	TTC	29,29	18,78	192
668	ATC	44,69	24,40	56
668	CTC	46,14	26,57	88
668	GTC	30,48	24,06	44
668	TTC	55,34	28,59	192
672	ATC	25,51	20,85	56
672	CTC	30,05	22,95	88
672	GTC	14,21	13,38	44
672	TTC	52,36	27,64	192

Пример 8. Анализ расщепления *in vitro* с использованием CasX:гРНК

1. Сборка RNP

[0712] Очищенный CasX дикого типа и RNP CasX и одномолекулярной гидовой РНК (огРНК) либо получали непосредственно перед экспериментами, либо получали, мгновенно замораживали в жидком азоте и хранили при -80 °С для последующего применения. Для получения RNP-комплексов белок CasX инкубировали с огРНК при молярном соотношении 1:1,2. Вкратце, огРНК добавляли в буфер № 1 (25 мМ NaPi, 150 мМ NaCl, 200 мМ трегалозы, 1 мМ MgCl₂), затем к раствору огРНК медленно при перемешивании на вихревой мешалке добавляли CasX и инкубировали при 37 °С в течение 10 мин для образования RNP-комплексов. RNP-комплексы перед применением фильтровали через фильтры Costar 8160 с размером пор 0,22 мкм, предварительно смоченные 200 мкл буфера № 1. При необходимости пробу RNP концентрировали с помощью фильтра 0,5 мл Ultra с

пределом отсечения 100 кДа (Millipore, кат. № UFC510096) до получения желаемого объема. Образование компетентных RNP оценивали так, как описано ниже.

2. Определение компетентных в отношении расщепления фракций вариантов белка по сравнению с референсным CasX дикого типа

[0713] Способность вариантов CasX образовывать активные RNP в сравнении с референсным CasX определяли с помощью анализа расщепления *in vitro*. Мишень ДНК гена бета-2 микроглобулина (B2M) 7.37 для анализа расщепления была создана следующим образом. Олигонуклеотиды ДНК с последовательностью TGAAGCTGACAGCATTCGGGCCGAGATGTCTCGCTCCGTGGCCTTAGCTGTGCTCGC GCT (нецелевая цепь, NTS (SEQ ID NO: 27177)) и AGCGCGAGCACAGCTAAGGCCACGGAGCGAGACATCTCGGCCCGAATGCTGTCAGC TTCA (целевая цепь, TS (SEQ ID NO: 27176)) приобретали с 5'-флуоресцентными метками (LI-COR IRDye 700 и 800 соответственно). Целевые дцДНК образовывали посредством смешивания олигонуклеотидов в соотношении 1:1 в 1× буфере для расщепления (20 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1 mM TCEP, 5% глицерина, 10 mM MgCl₂), нагревания до 95 °C в течение 10 минут и охлаждения раствора до комнатной температуры.

[0714] RNP CasX восстанавливали с помощью указанных CasX и гидов (см. графики) в конечной концентрации 1 мкМ при 1,5-кратном избытке указанного гида, если не указано иное, в 1× буфере для расщепления (20 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1 mM TCEP, 5% глицерина, 10 mM MgCl₂) при 37 °C в течение 10 мин перед перемещением в лед до готовности к применению. Применяли целевую последовательность 7.37 вместе с огРНК, имеющей спейсеры, комплементарные целевой последовательности 7.37.

[0715] Готовили реакционные смеси с конечной концентрацией RNP 100 нМ и конечной концентрацией целевой последовательности 100 нМ. Реакции проводили при 37 °C и инициировали добавлением целевой ДНК 7.37. Аликвоты отбирали через 5, 10, 30, 60 и 120 минут и гасили реакцию добавлением 95% формамида с 20 mM ЭДТА. Пробы денатурировали нагреванием при 95 °C в течение 10 минут и разделяли в 10% ПААГ с мочевиной. Гели либо фотографировали с помощью LI-COR Odyssey CLx и количественно оценивали с помощью программного обеспечения LI-COR Image Studio, либо фотографировали с помощью Cytiva Typhoon и количественно оценивали с помощью программного обеспечения Cytiva IQTL. Полученные данные откладывали на графике и анализировали с помощью Prism. Авторы изобретения предположили, что в условиях анализа CasX работает как фермент однократного действия, на что указывает такой отмеченный факт, что субстехиометрические количества фермента неспособны расщеплять более чем стехиометрическое количество мишени даже при продолжительных временных интервалах, а вместо этого достигают плато, которое по масштабу соответствует количеству присутствующего фермента. Таким образом, фракция целевой последовательности, расщепленная в течение длительного времени с помощью эквимольного количества RNP, указывает на то, какая фракция RNP была образована надлежащим образом и обладает расщепляющей активностью. Результаты для дорожек с

продуктами расщепления аппроксимировали двухфазной моделью скорости, так как реакция расщепления явно отклоняется от монофазной модели в данном диапазоне концентраций, и для каждой из трех независимых повторов определяли плато. Рассчитали среднее значение и стандартное отклонение для определения активной фракции (таблица 13).

[0716] Кажущиеся активные (компетентные) фракции определяли для RNP, образованных из CasX 2+гид 174+спейсер 7.37, CasX 119+гид 174+спейсер 7.37, CasX 457+гид 174+спейсер 7.37, CasX 488+гид 174+спейсер 7.37 и CasX 491+гид 174+спейсер 7.37, как показано на ФИГ. 1. Определенные активные фракции показаны в таблице 13. Все варианты CasX имели фракции с более высокой активностью, чем CasX 2 дикого типа, что указывает на то, что в тестовых условиях сконструированные варианты CasX образуют значительно более активные и стабильные RNP с идентичным гидом по сравнению с CasX дикого типа. Это может быть связано с повышенной аффинностью к оgРНК, повышенной стабильностью или растворимостью в присутствии оgРНК или большей стабильностью компетентной в отношении расщепления конформации сконструированного комплекса CasX:оgРНК. На увеличение растворимости RNP указывало заметное уменьшение наблюдаемого осадка, образованного при добавлении CasX 457, CasX 488 или CasX 491 к оgРНК, по сравнению с CasX 2.

3. Анализы расщепления *in vitro* - определение фракций, компетентных в отношении расщепления, для вариантов одномолекулярных гидов относительно референсных одномолекулярных гидов

[0717] С применением того же протокола для CasX 2.2.7.37, CasX 2.32.7.37, CasX 2.64.7.37 и CasX 2.174.7.37 также определяли, что фракции, компетентные в отношении расщепления, составляют $16\pm3\%$, $13\pm3\%$, $5\pm2\%$ и $22\pm5\%$, как показано на ФИГ. 2 и в таблице 10.

[0718] Второй набор гидов тестировали в других условиях для лучшего отделения вклада гида в образование RNP. Гиды 174, 175, 185, 186, 196, 214 и 215 со спейсером 7.37 смешивали с CasX 491 в конечных концентрациях 1 мкМ для гида и 1,5 мкМ для белка, а не с избытком гида, как ранее. Результаты представлены на ФИГ. 3 и в таблице 10. В таких условиях ограничения гида многие из этих гидов демонстрировали дополнительное улучшение по сравнению с гидом 174, при этом у гидов 185 и 196 компетентные фракции достигли $91\pm4\%$ и $91\pm1\%$ соответственно по сравнению с $80\pm9\%$ у гида 174.

[0719] Данные указывают на то, что оба из вариантов CasX и вариантов оgРНК способны образовывать более высокую долю активных RNP с гидовой РНК по сравнению с CasX дикого типа и оgРНК дикого типа. Кажущиеся скорости расщепления вариантов CasX 119, 457, 488 и 491 по сравнению с референсным CasX дикого типа определяли с помощью флуоресцентного анализа *in vitro* для расщепления целевой последовательности 7.37.

4. Анализ расщепления *in vitro* - определение $k_{\text{расщепл.}}$ для вариантов CasX по сравнению с референсным CasX дикого типа

[0720] RNP CasX восстанавливали с применением указанных CasX (см. ФИГ. 4) в конечной концентрации 1 мкМ при 1,5-кратном избытке указанного гида в 1× буфере для расщепления (20 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1 mM TCEP, 5% глицерина, 10 mM MgCl₂), при 37 °C в течение 10 мин перед перемещением в лед до готовности к применению. Готовили реакционные смеси с конечной концентрацией RNP 200 нМ и конечной концентрацией целевой последовательности 10 нМ. Реакции проводили при 37 °C, если не указано иное, и инициировали добавлением целевой ДНК. Аликвоты отбирали через 0,25, 0,5, 1, 2, 5 и 10 минут и гасили реакцию посредством добавления 95% формамида с 20 mM ЭДТА. Пробы денатурировали нагреванием при 95 °C в течение 10 минут и разделяли в 10% ПААГ с мочевиной. Гели фотографировали с помощью LI-COR Odyssey CLx и количественно оценивали с помощью программного обеспечения LI-COR Image Studio или фотографировали с помощью Cytiva Typhoon и количественно оценивали с помощью программного обеспечения Cytiva IQTL. Полученные данные откладывали на графике и анализировали с помощью Prism, и кажущуюся константу скорости реакции первого порядка для расщепления нецелевой цепи ($k_{\text{расщепл.}}$) определяли для каждой повторной комбинации CasX:огРНК по отдельности. Среднее значение и стандартное отклонение трех повторов с независимой аппроксимацией представлены в таблице 10, и результаты для дорожек с продуктами расщепления показаны на ФИГ. 5.

[0721] Кажущиеся константы скорости расщепления для CasX 2 дикого типа и вариантов CasX 119, 457, 488 и 491 CasX определяли с помощью гида 174 и спейсера 7.37 в каждом анализе (см. таблицу 10 и ФИГ. 4). Все варианты CasX имели улучшенные скорости расщепления относительно CasX 2 дикого типа. CasX 457 расщеплял медленнее, чем 119, несмотря на наличие более высокой компетентной фракции, как определено выше. CasX 488 и CasX 491 имели самые высокие скорости расщепления с большим отрывом от остальных вариантов; так как целевая ДНК была почти полностью расщеплена к первой временной точке, истинная скорость расщепления превышает разрешение этого анализа, и представленное значение $k_{\text{расщепл.}}$ следует принимать в качестве нижней границы.

[0722] Данные указывают на то, что варианты CasX имеют более высокий уровень активности, при этом скорости $k_{\text{расщепл.}}$ достигают превышения в по меньшей мере 30 раз по сравнению с CasX 2 дикого типа.

5. Анализ расщепления *in vitro*: сравнение вариантов гидов с гидами дикого типа

[0723] Также был проведен анализ расщепления с референсным CasX 2 дикого типа и референсным гидом 2 в сравнении с вариантами гидов 32, 64 и 174 для определения того, улучшают ли варианты расщепление. Эксперименты проводили так, как описано выше. Поскольку многие из полученных RNP не достигали полного расщепления целевой последовательности за время испытания, авторы изобретения определяли начальные скорости реакции (V_0), а не константы скорости реакции первого порядка. В первых двух временных точках (15 и 30 секунд) результаты вписывались в линию для каждой комбинации CasX:огРНК и повтора. Определяли среднее значение и стандартное отклонение угла наклона для трех повторов.

[0724] В условиях анализа значения V_0 для CasX 2 с гидами 2, 32, 64 и 174 составляли $20,4 \pm 1,4$ нМ/мин, $18,4 \pm 2,4$ нМ/мин, $7,8 \pm 1,8$ нМ/мин и $49,3 \pm 1,4$ нМ/мин (см. таблицу 13, а также ФИГ. 5 и ФИГ. 6). Гид 174 показал значительное улучшение скорости расщепления у полученного RNP (примерно в 2,5 раза относительно гида 2, см. ФИГ. 6), а активность гидов 32 и 64 была аналогична активности гида 2 или ниже. Следует отметить, что гид 64 обеспечивает скорость расщепления ниже, чем у гида 2, но его функциональность намного лучше *in vivo* (данные не показаны). Вероятно, некоторые изменения последовательности, внесенные для получения гида 64, улучшают транскрипцию *in vivo* за счет нуклеотида, участвующего в образовании триплекса. Улучшенная экспрессия гида 64, возможно, объясняет его улучшенную активность *in vivo*, хотя его сниженная стабильность может приводить к неправильному фолдингу *in vitro*.

[0725] Дополнительные эксперименты проводили с гидами 174, 175, 185, 186, 196, 214 и 215 со спейсером 7.37 и CasX 491 для определения относительных скоростей расщепления. Для снижения показателей кинетики расщепления до диапазона, поддающегося измерению с помощью нашего анализа, смеси реакций расщепления инкубировали при 10 °C. Результаты представлены на ФИГ. 7 и в таблице 13. В этих условиях гид 215 был единственным, который обеспечивал более высокую скорость расщепления, чем гид 174. Гид 196, который демонстрировал самую высокую активную фракцию RNP в условиях ограничения гида, имел кинетику, по существу аналогичную кинетике гида 174; это еще раз подчеркивает, что разные варианты приводят к улучшению разных характеристик.

[0726] Данные подтверждают, что в условиях анализа применение большинства вариантов гидов с CasX приводит к получению RNP с более высоким уровнем активности, чем в случае гида дикого типа, с улучшениями в начальной скорости расщепления в диапазоне от примерно 2 раз до > 6 раз. Числа в таблице 13 указывают, слева направо, вариант CasX, каркас оgPHK и спейсерную последовательность конструктора RNP. В названиях конструкторов RNP в таблице ниже вариант белка CasX, каркас гида и спейсер указаны в порядке слева направо.

6. Анализ расщепления *in vitro*: сравнение скорости расщепления и компетентной фракции для 515.174 и 526.174 относительно референса 2.2.

[0727] Авторы изобретения хотели сравнить сконструированные варианты белка CasX 515 и 526 в комплексе со сконструированным вариантом одномолекулярного гида 174 относительно референсного белка 2 дикого типа (SEQ ID NO: 2) и минимально модифицированного генно-инженерными методами варианта гида 2 (SEQ ID NO: 5). RNP-комплексы собирали так, как описано выше, с 1,5-кратным избытком гида. Анализы расщепления для определения $k_{\text{расщепл.}}$ и компетентной фракции выполняли так, как описано выше, в обоих случаях при 37 °C, и применяли разные временные точки для определения компетентной фракции для RNP дикого типа в сравнении со сконструированным RNP из-за значительно различающихся значений времени, необходимых для практически полного завершения реакций.

[0728] Полученные данные четко демонстрируют значительные улучшения активности RNP, когда используют как сконструированный белок, так и сконструированный гид. RNP 515.174 и 526.174 имели компетентные фракции 76% и 91% соответственно по сравнению с 16% для RNP 2.2 (ФИГ. 8, таблица 13). В кинетическом анализе как 515.174, так и 526.174 разрезали по существу всю целевую ДНК к первой временной точке, давая показатели, превышающие разрешение анализа, и приводя к расчетным скоростям расщепления 17,10 и 19,87 мин⁻¹ соответственно (ФИГ. 9, таблица 13). В отличие от этого, RNP 2.2 разрезал в среднем менее 60% целевой ДНК к конечной временной точке 10 минут и имеет расчетное значение $k_{\text{расщепл.}}$ почти на два порядка величины ниже, чем сконструированные RNP. Модификации, внесенные в белок и гид, приводят к получению RNP, которые более стабильны, с большей вероятностью образуют активные частицы, а также разрезают ДНК более эффективно при расчете показателя на одну частицу.

Таблица 13. Результаты анализа расщепления и анализа образования RNP

Конструкт RNP	$k_{\text{расщепл.}}^*$	Начальная скорость*	Компетентная фракция
2.2.7.37		20,4±1,4 нМ/мин	16±3%
2.32.7.37		18.4±2.4 нМ/мин	13±3%
2.64.7.37		7,8±1,8 нМ/мин	5±2%
2.174.7.37	0,51±0,01 мин ⁻¹	49,3±1,4 нМ/мин	22±5%
119.174.7.37	6,29±2,11 мин ⁻¹		35±6%
457.174.7.37	3,01±0,90 мин ⁻¹		53±7%
488.174.7.37	15,19 мин ⁻¹		67%
491.174.7.37	16,59 мин ⁻¹ /0,293 мин ⁻¹ (10°C)		83%/17% (с ограниченным количеством гида)
491.175.7.37	0,089 мин ⁻¹ (10°C)		5% (с ограниченным количеством гида)
491.185.7.37	0,227 мин ⁻¹ (10°C)		44% (с ограниченным количеством гида)
491.186.7.37	0,099 мин ⁻¹ (10°C)		11% (с ограниченным количеством гида)
491.196.7.37	0,292 мин ⁻¹		46% (с

	(10°C)		ограниченным количеством гида)
491.214.7.37	0,284 мин ⁻¹ (10°C)		30% (с ограниченным количеством гида)
491.215.7.37	0,398 мин ⁻¹ (10°C)		38% (с ограниченным количеством гида)
515.174.7.37	17,10 мин ^{-1**}		76%
526.174.7.37	19,87 мин ^{-1**}		91%

* Среднее ± стандартное отклонение

** Скорость превышает разрешение анализа

Пример 9. Исследование влияния длины спейсера на кинетику расщепления *in vitro*

[0729] Исследовали активность расщепления *in vitro* рибонуклеопротеиновыми комплексами (RNP) из двух вариантов CasX и гидовой РНК со спейсерами различной длины для определения того, какая длина спейсера обеспечивает наиболее эффективное расщепление целевой нуклеиновой кислоты, и изменяется ли предпочтительная длина спейсеров для белков.

Методы:

[0730] Исследовали активность расщепления *in vitro* рибонуклеопротеиновыми комплексами (RNP) из CasX и гидовой РНК со спейсерами различной длины для определения того, какая длина спейсера обеспечивает наиболее эффективное расщепление целевой нуклеиновой кислоты.

[0731] Вариант CasX 515 и 526 очищали так, как описано выше. Гиды с каркасом 174 (SEQ ID NO: 2238) получали посредством транскрипции *in vitro* (IVT). Матрицы для IVT образовывали с помощью ПЦР с использованием полимеразы Q5 (NEB M0491) в соответствии с рекомендованным протоколом, матричных олигонуклеотидов для каждой основной цепи каркаса и праймеров для амплификации с промотором T7 и спейсером 7.37 (GGCCGAGATGTCTCGTCCG; нацеливающий tdTomato (SEQ ID NO: 27192)) длиной 20 нуклеотидов или усеченным с 3'-конца до 18 или 19 нуклеотидов. Спейсерные последовательности, а также олигонуклеотиды, применяемые для создания каждой матрицы, показаны в таблице 14. Затем полученные матрицы применяли с РНК-полимеразой T7 для получения гидовых РНК в соответствии со стандартными протоколами. Гиды очищали с помощью электрофореза в полиакриламидном геле в денатурирующих условиях и заново проводили фолдинг перед применением.

[0732] RNP CasX восстанавливали посредством разбавления CasX до 1 мкМ в 1× буфере для расщепления (20 mM Tris-HCl, pH 7,5, 150 mM NaCl, 1 mM TCEP, 5% глицерина,

10 mM MgCl₂) и добавления огРНК до 1,2 мкМ и инкубации при 37 °С в течение 10 мин перед перемещением в лед до готовности к применению. Флуоресцентно-меченую целевую ДНК 7.37 приобретали в виде отдельных олигонуклеотидов в компании Integrated DNA Technologies (полные последовательности в таблице 14) и получали целевую дцДНК путем нагревания эквимольной смеси двух комплементарных цепей в 1× буфере для расщепления и медленного охлаждения до комнатной температуры.

[0733] RNP разводили в буфере для расщепления до конечной концентрации 200 нМ и инкубировали при 10 °С без встряхивания. Реакции расщепления инициировали добавлением целевой ДНК 7.37 до конечной концентрации 10 нМ. Пробы отбирали во временных точках 0,25, 0,5, 1, 2, 5, 10 и 30 минут. Отобранные во временных точках пробы гасили добавлением равного объема 95% формамида с 20 mM ЭДТА. Пробы денатурировали нагреванием при 95 °С в течение 10 минут и разделяли в 10% ПААГ с мочевиной. Для визуализации гелей использовали Amersham Typhoon и проводили анализ с помощью программного обеспечения IQTL. Полученные данные откладывали на графике и анализировали с помощью Prism. Расщепление нецелевой цепи аппроксимировали простой экспоненциальной функцией для определения кажущейся константы скорости реакции первого порядка ($k_{\text{расщепл.}}$).

Результаты

[0734] Проводили сравнение скоростей расщепления для вариантов CasX 515 и 526 в комплексе с огРНК с 18-, 19- или 20-нуклеотидными спейсерами, чтобы определить, какая длина спейсера приводит к наиболее эффективному расщеплению в случае каждого варианта белка. В соответствии с другими экспериментами, выполненными с огРНК, полученными транскрипцией *in vitro*, наилучшим для обоих вариантов белка с точки зрения эффективности был гид с 18-нуклеотидным спейсером (ФИГ. 12А и В, таблица 14). Спейсер длиной 18 нукл. обеспечивал в 1,4 раза более высокую скорость, чем спейсер длиной 20 нукл. в случае белка 515, и в 3 раза более высокую скорость, чем спейсер длиной 20 нукл. в случае белка 526. Активность спейсера длиной 19 нукл. была промежуточной в случае обоих белков, хотя, опять же, различие было более выраженным для варианта 526. В целом, наблюдалось, что спейсеры короче 20 нукл. обладают повышенной активностью при использовании целого диапазона белков, спейсеров и способов доставки, но степень улучшения и оптимальная длина спейсера варьировались. Эти данные показывают, что два сконструированных белка, которые являются достаточно похожими по последовательности (отличаются только двумя остатками), могут претерпевать изменения активности, вызванные длиной спейсера, которые сходны по направлению, но существенно различаются по степени.

Таблица 14. Значимые последовательности и олигонуклеотиды

Описание	Последовательность	SEQ ID NO

Нецелевая цепь целевой последовательности 7.37	IR700- TGAAGCTGACAGCATTCGGGCCGAGATG TCTCGCTCCGTGGCCTTAGCTGTGCTCGC GCT	27177
Целевая цепь целевой последовательности 7.37	IR800- AGCGCGAGCACAGCTAAGGCCACGGAGC GAGACATCTCGGCCCGAATGCTGTCAGCT TCA	27176
Спейсерная последовательность длиной 20 нукл.	GGCCGAGATGTCTCGCTCCG	27191
Спейсерная последовательность длиной 18 нукл.	GGCCGAGATGTCTCGCTC	27193
Спейсерная последовательность длиной 19 нукл.	GGCCGAGATGTCTCGCTCC	27194
Матричный прямой для каркаса 174	GAAATTAATACGACTCACTATAACTGGCG CTTTTATCTGATTACTTTGAGAGCCATCA CCAGCGACTATGTCTAGTGGGTAAAGCT	247
Матричный обратный для каркаса 174	CTTTGATGCTCCCTCCGAAGAGGGAGCTT TACCCACTACGACATAGTCGC	248
Праймер для амплификации T7	GAAATTAATACGACTCACTATA	234
Праймер со спейсером для каркаса 174 длиной 20 нукл.	CGGAGCGAGACATCTCGGCCCTTTGATGC TCCCTCC	249
Праймер со спейсером для каркаса 174 длиной 18 нукл.	GAGCGAGACATCTCGGCCCTTTGATGCTC CCTCC	27195
Праймер со спейсером для каркаса 174 длиной 19 нукл.	GGAGCGAGACATCTCGGCCCTTTGATGCT CCCTCC	27196

Таблица 15. Скорости расщепления RNP с укороченными спейсерами

Длина спейсера	k-расщепл. (мин ⁻¹) для 515	k-расщепл. (мин ⁻¹) для 526
18	0,215	0,427
19	0,182	0,282

20	0,150	0,143
----	-------	-------

Пример 10. Оценка аффинности связывания с гидовой РНК

[0735] Очищенный CasX дикого типа и улучшенный CasX будут инкубировать с синтетической одномолекулярной гидовой РНК, содержащей Cy7.5-фрагмент на 3'-конце, в низкосолеваемом буферном растворе, содержащем хлорид магния и гепарин для предотвращения неспецифического связывания и агрегации. Концентрация огРНК будет постоянной - 10 пМ, в то время как белок будут титровать от 1 пМ до 100 мкМ в отдельных реакциях связывания. После того как реакция придет в равновесное состояние, пробы пропустят через вакуумный коллектор для анализа связывания с фильтром, используя нитроцеллюлозную мембрану и положительно заряженную нейлоновую мембрану, которые связывают белок и нуклеиновую кислоту соответственно. На мембранах проведут визуализацию для идентификации гидовой РНК, и будут определены фракции связанной и несвязанной РНК по количеству флуоресценции на нитроцеллюлозной и нейлоновой мембране при каждой концентрации белка для расчета константы диссоциации комплекса белок-огРНК. Также будет проведен эксперимент с улучшенными вариантами огРНК для определения того, влияют ли эти мутации на аффинность к гиду белков дикого типа и мутантных белков. Авторы изобретения также выполняют анализ методом сдвига электрофоретической подвижности для качественного сравнения с результатами анализа связывания с фильтром и подтверждения того, что основным фактором ассоциации белков и РНК является связывание в растворе, а не агрегация.

Пример 11. Оценка аффинности связывания с целевой ДНК

[0736] Очищенный CasX дикого типа и улучшенный CasX будут объединены в комплекс с одномолекулярной гидовой РНК, несущей нацеливающую последовательность, комплементарную целевой нуклеиновой кислоте. RNP-комплекс будут инкубировать с двухцепочечной целевой ДНК, содержащей РАМ и соответствующую последовательность целевой нуклеиновой кислоты с Cy7.5-меткой на 5'-конце целевой цепи, в низкосолеваемом буферном растворе, содержащем хлорид магния и гепарин для предотвращения неспецифического связывания и агрегации. Концентрация целевой ДНК будет постоянной - 1 нМ, в то время как RNP будут титровать от 1 пМ до 100 мкМ в отдельных реакциях связывания. После того как реакция придет в равновесное состояние, пробы пропустят через нативный 5% полиакриламидный гель для отделения связанной и несвязанной целевой ДНК. Будет проведена визуализация геля для определения сдвигов подвижности целевой ДНК и будут рассчитаны фракции связанной и несвязанной ДНК при каждой концентрации белка для определения константы диссоциации тройного комплекса RNP-целевая ДНК. Ожидается, что эксперименты продемонстрируют улучшенную аффинность связывания RNP, содержащего вариант CasX и вариант гРНК, по сравнению с RNP, содержащим референсный CasX и референсную гРНК.

Пример 12. Оценка улучшенных характеристик экспрессии и растворимости вариантов CasX для получения RNP

[0737] CasX дикого типа и модифицированные варианты CasX будут

экспрессированы в BL21 (DE3) *E. coli* в идентичных условиях. Все белки будут находиться под контролем IPTG-индуцибельного промотора T7. Клетки будут выращивать до OD 0,6 в среде ТВ при 37 °C, после чего температура выращивания будет снижена до 16 °C и экспрессия будет индуцирована добавлением 0,5 mM IPTG. Клетки будут собирать после 18 часов экспрессии. Растворимые белковые фракции будут экстрагированы и проанализированы в геле для ДСН-ПААГ. Относительные уровни экспрессии растворимого CasX будут определять посредством окрашивания Кумасси. Белки будут очищены параллельно в соответствии с протоколом, приведенным выше, и будет проведено сравнение конечного выхода очищенного белка. Для определения растворимости очищенного белка конструкторы будут концентрировать в буфере для хранения вплоть до начала осаждения белка. Осажденный белок будет удален центрифугированием и будет измерена конечная концентрация растворимого белка для определения максимальной растворимости каждого варианта. Наконец, будут получены комплексы вариантов CasX с одномолекулярной гидовой РНК, которые будут сконцентрированы вплоть до начала образования осадка. Осажденные RNP будут удалены центрифугированием и будет измерена конечная концентрация растворимых RNP для определения максимальной растворимости каждого варианта, связанного с гидовой РНК.

Пример 13. Редактирование области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в клетках HEK293T

[0738] Для демонстрации способности CasX редактировать область связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* проводили эксперимент с применением варианта CasX 438, варианта гида 174 и спейсера, нацеленного на область связывания GATA1 в человеческом локусе эритроид-специфичного энхансера *BCL11A* в клетках HEK293T.

[0739] Клетки HEK293T поддерживали при 37 °C и 5% CO₂ в среде Fibroblast (FB), состоящей из среды Игла в модификации Дульбекко (DMEM; Corning Cellgro, № 10-013-CV) с добавлением 10% эмбриональной бычьей сыворотки (FBS; Seradigm, № 1500-500) и 100 ЕД/мл пенициллина и 100 мг/мл стрептомицина (100× Pen-Strep; GIBCO, № 15140-122), которая может дополнительно включать пируват натрия (100×, Thermofisher, № 11360070), неосновные аминокислоты (100×, Thermofisher, № 11140050), буфер HEPES (100×, Thermofisher, № 15630080) и 2-меркаптоэтанол (1000×, Thermofisher, № 21985023).

[0740] Для этого эксперимента клетки HEK293T высевали по 20-40 тыс. клеток/лунка в 96-луночном планшете в 100 мкл среды FB и культивировали в инкубаторе при 37 °C и 5% CO₂. На следующий день клетки трансфицировали при конфлюэнтности примерно 75%. Конструкт CasX и гида (см. последовательности в таблице 16) трансфицировали в клетки HEK293T по 100-500 нг на лунку с использованием Lipofectamine 3000 в соответствии с протоколом производителя, используя по 3 лунки на конструкт в качестве повторов. В качестве отрицательного контроля применяли ненацеливающую плазмиду. В качестве эталонного контроля применяли конструкт SpyCas9 и гида, нацеливающийся на ту же область. Успешно трансфицированные клетки

отбирали с помощью пурамицина при 0,3-3 мкг/мл в течение 24-48 часов с последующим восстановлением в среде FB. Впоследствии клетки каждой пробы из эксперимента лизировали и экстрагировали геном в соответствии с протоколом производителя и стандартными практиками. Редактирование в клетках каждой экспериментальной пробы анализировали с помощью анализа NGS. Вкратце, геномную ДНК амплифицировали методом ПЦР с праймерами, специфичными к интересующей целевой области генома с образованием целевого ампликона. Эти праймеры содержат дополнительную последовательность на 5'-концах для введения рида Illumina и 2 последовательностей. Кроме того, они содержат случайную последовательность длиной 16 нукл., которая функционирует как уникальный молекулярный идентификатор (UMI). Качественную и количественную оценку ампликона проводили с помощью набора Fragment Analyzer DNA analyzer kit (Agilent, дцДНК 35-1500 п. о.). Ампликоны секвенировали на Illumina Miseq в соответствии с инструкциями производителя. Необработанные файлы fastq, полученные в результате секвенирования, обрабатывали следующим образом. (1) Последовательности обрезали для обеспечения качества и для адаптерных последовательностей с помощью программы cutadapt (вер. 2.1). (2) Последовательности из рида 1 и рида 2 объединяли в одну последовательность вставки с помощью программы flash2 (вер. 2.2.00). (3) Консенсусные последовательности вставки проводили через программу CRISPResso2 (вер. 2.0.29) вместе с ожидаемой последовательностью ампликона и спейсерной последовательностью. В этой программе количественно определяется процентная доля ридов, которые были модифицированы в окне вокруг 3'-конца спейсера (окно размером 30 п. о., центрированное по -3 п. о. от 3'-конца спейсера). Активность молекулы CasX количественно определяли как общую процентную долю ридов, содержащих вставки и/или делеции в любом месте в пределах этого окна.

Таблица 16. Гидовые последовательности

Спейсер	Спейсерная последовательность	SEQ ID NO	Последовательность гида 174	SEQ ID NO	Последовательность в гид+спейсер	SEQ ID NO
21.1	UGGAGCCU GUGAUAAA AGCA	22	ACUGGCGCUU UUAUCUGAUU ACUUUGAGAG CCAUCACCAG CGACUAUGUC GUAGUGGGUA AAGCUCCCUC UUCGGAGGGA GCAUCAAG	2238	ACUGGCGCUUUU AUCUGAUUACUU UGAGAGCCAUCA CCAGCGACUAUG UCGUAGUGGGUA AAGCUCCCUCUU CGGAGGGAGCAU CAAAGUGGAGCC UGUGAUAAAAGC A	271

[0741] Результаты: на графике, представленном на ФИГ. 18, показаны результаты анализа методом NGS CasX-опосредованного редактирования области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в клетках НЕК293Т через 5 дней после трансфекции. Каждая точка данных представляет собой средний показатель ридов NGS для результатов редактирования, полученных в отдельных условиях обработки. Результаты указывают на то, что CasX и гид были способны редактировать локус эритроид-специфичного энхансера *BCL11A* со средним уровнем редактирования 90%, в то время как конструктор SpyCas9 демонстрировал средний уровень редактирования 80%. При использовании конструктора с ненацеливающим спейсером редактирование отсутствовало (данные не показаны). Этот пример демонстрирует, что CasX с соответствующим гидом были способны редактировать локус эритроид-специфичного энхансера *BCL11A* в клетках НЕК293Т. Эксперименты с вариантами CasX 668, 672, 676 и гРНК 235 будут проведены в аналогичных условиях и, как ожидается, будут приводить к аналогичной эффективности редактирования.

Пример 14. Редактирование области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в клетках K562

[0742] Для демонстрации способности CasX редактировать локус эритроид-специфичного энхансера *BCL11A* проводили эксперимент с использованием вариантов CasX 119 и 491, варианта каркаса 174 и спейсера, нацеленного на область связывания GATA1 в человеческом локусе эритроид-специфичного энхансера *BCL11A* в клетках K562.

[0743] Клетки K562 поддерживали при 37 °C и 5% CO₂ в среде, состоящей из RPMI (RPMI; Thermofisher, № 11875119) с добавлением 10% эмбриональной бычьей сыворотки (FBS; Seradigm, № 1500-500) и 100 ЕД/мл пенициллина и 100 мг/мл стрептомицина (100× Pen-Strep; GIBCO, № 15140-122), которая может дополнительно включать пируват натрия (100×, Thermofisher, № 11360070) и буфер HEPES (100×, Thermofisher, № 15630080).

[0744] В этом эксперименте CasX и гид, нацеливающий на область связывания GATA1 в локусе *BCL11A*, вводили в клетки K562 с помощью двух разных способов доставки, RNP и XDP (RNP, упакованные в XDP). В первой экспериментальной группе CasX RNP, нацеливающийся на область связывания GATA1 в локусе *BCL11A* (спейсерную последовательность см. в таблице), был получен стандартными способами. Вкратце, каждым CasX RNP (см. последовательности в таблице) в количестве 10-100 пмоль на одно условие трансдуцировали 100-500 тыс. клеток K562 с помощью набора Lonza nucleofector kit в соответствии с протоколом производителя, используя по 3 лунки на конструктор в качестве повторы. Клетки культивировали в среде RPMI с добавками при 37 °C и 5% CO₂.

[0745] Во второй экспериментальной группе XDP, в которых инкапсулирован CasX, нацеленный на область связывания GATA1 в локусе *BCL11A*, получали так, как описано ниже. Вкратце, XDP получали с помощью четырех структурных плазмид: pXDP17, pSG0010, pGP2 и pXDP3. Плазмида pXDP17 экспрессирует последовательность gag ВИЧ-1, за которой следует версия 491 CasX. pSG0010 представляет собой каркас 174 со спейсером 21.1 (см. последовательность ниже), нацеленный на BC11A, экспрессируемый

под промотором U6. pGP2 экспрессирует нацеливающий фрагмент VSV-G. pXDP3 экспрессирует полибелок gag ВИЧ-1 без присоединенной молекулы CasX. Для получения XDP клетки LentiX от компании Takara разделяли и высевали за 24 часа до трансфекции плазмидной ДНК. Смешивали 89 мкг плазмиды pSG0010, 366 мкг плазмиды pXDP0017, 30 мкг плазмиды pXDP0003 и 1,7 мкг плазмиды pGP2 с Opti-MEM и PEI, а затем добавляли к клеточной культуре. Среды меняли на Opti-MEM через 16 часов после трансфекции. Через 54 часа после трансфекции среды собирали и концентрировали посредством центрифугирования. XDP ресуспендировали в буфере 1 с 150 мМ NaCl и замораживали при -150 °C. В день проведения эксперимента XDP размораживали на льду и сразу же использовали на клетках.

[0746] Клетки K562 высевали по 30-50 тыс. клеток/лунка в 96-луночный планшет, трансдуцировали XDP в диапазоне разных значений множественности заражения (MOI) и культивировали в среде RPMI с добавками при 37 °C и 5% CO₂.

[0747] Через четыре дня с помощью анализа NGS проводили анализ редактирования в клетках из каждой экспериментальной пробы, трансдуцированной RNP или XDP. Вкратце, геномную ДНК амплифицировали методом ПЦР с праймерами, специфичными к интересующей целевой области генома с образованием целевого ампликона. Эти праймеры содержат дополнительную последовательность на 5'-концах для введения ридов Illumina и 2 последовательностей. Кроме того, они содержат случайную последовательность длиной 16 нукл., которая функционирует как уникальный молекулярный идентификатор (UMI). Качественную и количественную оценку ампликона проводили с помощью набора Fragment Analyzer DNA analyzer kit (Agilent, дцДНК 35-1500 п. о.). Ампликоны секвенировали на Illumina Miseq в соответствии с инструкциями производителя. Необработанные файлы fastq, полученные в результате секвенирования, обрабатывали следующим образом. (1) Последовательности обрезают для обеспечения качества и для адаптерных последовательностей с помощью программы cutadapt (вер. 2.1). (2) Последовательности из ридов 1 и ридов 2 объединяли в одну последовательность вставки с помощью программы flash2 (вер. 2.2.00). (3) Консенсусные последовательности вставки проводили через программу CRISPResso2 (вер. 2.0.29) вместе с ожидаемой последовательностью ампликона и спейсерной последовательностью. В этой программе количественно определяется процентная доля ридов, которые были модифицированы в окне вокруг 3'-конца спейсера (окно размером 30 п. о., центрированное по -3 п. о. от 3'-конца спейсера). Активность молекулы CasX количественно определяли как общую процентную долю ридов, содержащих вставки и/или делеции в любом месте в пределах этого окна.

[0748] Результаты: на графике, представленном на ФИГ. 19, показаны результаты анализа методом NGS CasX-опосредованного редактирования области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в клетках K562 через 4 дня после трансдукции RNP. Каждая точка данных представляет собой средний показатель ридов NGS для результатов редактирования, полученных в отдельных условиях обработки.

Результаты показывают, что CasX и гид были способны редактировать локус эритроид-специфичного энхансера *BCL11A* дозозависимым образом, при этом вариант 491 CasX неизменно демонстрирует более высокий уровень редактирования относительно варианта 119 CasX. Этот пример демонстрирует, что в условиях анализа CasX с соответствующим гидом были способны редактировать локус эритроид-специфичного энхансера *BCL11A* в клетках K562.

[0749] На графике, представленном на ФИГ. 20, показаны результаты анализа методом NGS CasX-опосредованного редактирования области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в клетках K562 через 4 дня после трансдукции XDP. Каждая точка данных представляет собой средний показатель ридов NGS для результатов редактирования, полученных в отдельных условиях обработки. Результаты указывают на то, что CasX и гид были способны редактировать локус эритроид-специфичного энхансера *BCL11A* дозозависимым образом. Этот пример демонстрирует, что CasX с соответствующим гидом были способны редактировать локус эритроид-специфичного энхансера *BCL11A* в клетках K562. Эксперименты с вариантами CasX 668, 672, 676 и гРНК 235 будут проведены в аналогичных условиях и, как ожидается, будут приводить к аналогичной эффективности редактирования.

Пример 15. Редактирование области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в гематopoэтических стволовых клетках

[0750] Для демонстрации способности CasX редактировать локус эритроид-специфичного энхансера *BCL11A* проводили эксперименты с помощью вариантов CasX 119 и 491, варианта каркаса 174 и спейсера, нацеленного на область связывания GATA1 в человеческом локусе эритроид-специфичного энхансера *BCL11A* в гематopoэтических стволовых клетках (HSC) CD34⁺.

[0751] HSC культивировали в среде StemSpan SFEM II (Stem Cell, № 9605) с добавлением CC100 (Stem Cell, № 2697) и поддерживали при 37 °C и 5% CO₂. В этом эксперименте CasX и гид, нацеливающий на область связывания GATA1 локуса *BCL11A*, вводили в HSC с помощью двух разных способов доставки - RNP и XDP. В первой экспериментальной группе CasX RNP, нацеливающий на область связывания GATA1 в локусе *BCL11A* (см. спейсерную последовательность в таблице), получали стандартными способами. Каждым CasX RNP (см. последовательности в таблице) в количестве 10-100 пмоль на одно условие трансдуцировали 100-500 тыс. HSC с помощью набора Lonza nucleofector kit в соответствии с протоколом производителя, используя по 3 лунки на конструкт в качестве повторов. Клетки культивировали в среде SFEM II с добавками при 37 °C и 5% CO₂.

[0752] Во второй экспериментальной группе XDP, в которых инкапсулирован CasX, нацеленный на область связывания GATA1 в локусе *BCL11A*, получали так, как описано ниже. Вкратце, XDP получали с помощью четырех структурных плазмид: pXDP17, pSG0010, pGP2 и pXDP3. Плазида pXDP17 экспрессирует последовательность gag ВИЧ-1, за которой следует версия 491 CasX. pSG0010 представляет собой каркас 174 со

спейсером 21.1 (см. последовательность ниже), нацеленный на BC11A, экспрессируемый под промотором U6. pGP2 экспрессирует нацеливающий фрагмент VSV-G. pXDP3 экспрессирует полибелок gag ВИЧ-1 без присоединенной молекулы CasX. Для получения XDP клетки LentiX от компании Takara разделяли и высевали за 24 часа до трансфекции плазмидной ДНК. Смешивали 89 мкг плазмиды pSG0010, 366 мкг плазмиды pXDP0017, 30 мкг плазмиды pXDP0003 и 1,7 мкг плазмиды pGP2 с Opti-MEM и PEI, а затем добавляли к клеточной культуре. Среды меняли на Opti-MEM через 16 часов после трансфекции. Через 54 часа после трансфекции среды собирали и концентрировали посредством центрифугирования. XDP ресуспендировали в буфере 1 с 150 мМ NaCl и замораживали при -150 °C. В день проведения эксперимента XDP размораживали на льду и сразу же использовали на клетках.

[0753] HSC высевали в количестве по 30-50 тыс. клеток/лунка в 96-луночный планшет, трансдуцировали XDP в диапазоне разных значений MOI и культивировали в среде SFEM II с добавками при 37 °C и 5% CO₂. Через четыре дня с помощью анализа NGS проводили анализ редактирования при каждом условии эксперимента в клетках из проб, трансдуцированных RNP или XDP. Вкратце, клетки для каждой пробы из эксперимента лизировали и геном экстрагировали в соответствии с протоколом производителя и стандартными практиками. Редактирование в клетках каждой экспериментальной пробы анализировали с помощью анализа NGS. Вкратце, геномную ДНК амплифицировали методом ПЦР с праймерами, специфичными к интересующей целевой области генома с образованием целевого ампликона. Эти праймеры содержат дополнительную последовательность на 5'-концах для введения рида Illumina и 2 последовательностей. Кроме того, они содержат случайную последовательность длиной 16 нукл., которая функционирует как уникальный молекулярный идентификатор (UMI). Качественную и количественную оценку ампликона проводили с помощью набора Fragment Analyzer DNA analyzer kit (Agilent, дцДНК 35-1500 п. о.). Ампликоны секвенировали на Illumina Miseq в соответствии с инструкциями производителя. Необработанные файлы fastq, полученные в результате секвенирования, обрабатывали следующим образом. (1) Последовательности обрезали для обеспечения качества и для адаптерных последовательностей с помощью программы cutadapt (вер. 2.1). (2) Последовательности из рида 1 и рида 2 объединяли в одну последовательность вставки с помощью программы flash2 (вер. 2.2.00). (3) Консенсусные последовательности вставки проводили через программу CRISPResso2 (вер. 2.0.29) вместе с ожидаемой последовательностью ампликона и спейсерной последовательностью. В этой программе количественно определяется процентная доля ридов, которые были модифицированы в окне вокруг 3'-конца спейсера (окно размером 30 п. о., центрированное по -3 п. о. от 3'-конца спейсера). Активность молекулы CasX количественно определяли как общую процентную долю ридов, содержащих вставки и/или делеции в любом месте в пределах этого окна.

[0754] Результаты: на графике, представленном на ФИГ. 21, показаны результаты анализа NGS CasX-опосредованного редактирования области связывания GATA1 в локусе

эритроид-специфичного энхансера *BCL11A* в HSC через 4 дня после трансдукции RNP. Каждая точка данных представляет собой средний показатель ридов NGS для результатов редактирования, полученных в отдельных условиях обработки. Результаты показывают, что CasX и гид были способны редактировать локус эритроид-специфичного энхансера *BCL11A* дозозависимым образом, при этом вариант 491 CasX неизменно демонстрирует более высокий уровень редактирования относительно варианта 119 CasX. Этот пример демонстрирует, что в условиях анализа CasX с соответствующим гидом были способны редактировать локус эритроид-специфичного энхансера *BCL11A* в HSC. На графике, представленном на ФИГ. 22, показаны результаты анализа NGS CasX-опосредованного редактирования области связывания GATA1 в локусе эритроид-специфичного энхансера *BCL11A* в HSC через 4 дня после трансдукции XDP. Каждая точка данных представляет собой средний показатель ридов NGS для результатов редактирования, полученных в отдельных условиях обработки. Результаты указывают на то, что CasX и гид были способны редактировать локус эритроид-специфичного энхансера *BCL11A* дозозависимым образом. Этот пример демонстрирует, что CasX с соответствующим гидом были способны редактировать локус эритроид-специфичного энхансера *BCL11A* в HSC. Эксперименты с вариантами CasX 668, 672, 676 и гРНК 235 будут проведены в аналогичных условиях и, как ожидается, будут приводить к аналогичной эффективности редактирования.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Система, содержащая белок CRISPR класса 2, типа V и первую гидовую рибонуклеиновую кислоту (гРНК), причем гРНК содержит нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, содержащей ген связывающегося с полипиримидиновым трактом белка 1 (BCL11A).

2. Система по п. 1, в которой гРНК содержит нацеливающую последовательность, комплементарную последовательности целевой нуклеиновой кислоты, выбранной из группы, состоящей из:

- a. интрона BCL11A;
- b. экзона BCL11A;
- c. соединения интрон-экзон BCL11A;
- d. регуляторного элемента BCL11A; и
- e. межгенной области.

3. Система по п. 1 или п. 2, в которой ген BCL11A содержит последовательность дикого типа.

4. Система по любому из пп. 1-3, в которой гРНК представляет собой одномолекулярную гРНК (огРНК).

5. Система по любому из пп. 1-4, в которой гРНК представляет собой двухмолекулярную гРНК (дгРНК).

6. Система по любому из пп. 1-5, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789, или последовательность, идентичную ей на по меньшей мере примерно 65%, по меньшей мере примерно 75%, по меньшей мере примерно 85% или по меньшей мере примерно 95%.

7. Система по любому из пп. 1-5, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-26789.

8. Система по п. 7, в которой у нацеливающей последовательности удален один нуклеотид с 3'-конца последовательности.

9. Система по п. 7, в которой у нацеливающей последовательности удалены два нуклеотида с 3'-конца последовательности.

10. Система по п. 7, в которой у нацеливающей последовательности удалены три нуклеотида с 3'-конца последовательности.

11. Система по п. 7, в которой у нацеливающей последовательности удалены четыре нуклеотида с 3'-конца последовательности.

12. Система по п. 7, в которой у нацеливающей последовательности удалены пять нуклеотидов с 3'-конца последовательности.

13. Система по любому из пп. 1-12, в которой нацеливающая последовательность гРНК комплементарна последовательности экзона BCL11A.

14. Система по п. 13, в которой нацеливающая последовательность гРНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A, последовательности экзона 3 BCL11A, последовательности экзона 4 BCL11A, последовательности экзона 5 BCL11A, последовательности экзона 6 BCL11A, последовательности экзона 7 BCL11A, последовательности экзона 8 BCL11A и последовательности экзона 9 BCL11A.

15. Система по п. 14, в которой нацеливающая последовательность гРНК комплементарна последовательности, выбранной из группы, состоящей из последовательности экзона 1 BCL11A, последовательности экзона 2 BCL11A и последовательности экзона 3 BCL11A.

16. Система по любому из пп. 1-12, в которой нацеливающая последовательность гРНК комплементарна последовательности регуляторного элемента BCL11A.

17. Система по п. 16, в которой нацеливающая последовательность гРНК комплементарна последовательности промотора гена BCL11A.

18. Система по п. 16, в которой нацеливающая последовательность гРНК комплементарна последовательности энхансерного регуляторного элемента.

19. Система по п. 18, в которой нацеливающая последовательность гРНК комплементарна последовательности, которая содержит эритроид-специфичный энхансерный сайт связывания GATA1 (GATA1) гена BCL11A.

20. Система по п. 16, в которой нацеливающая последовательность гРНК комплементарна последовательности, которая расположена с 5'-конца от сайта связывания GATA1 гена BCL11A.

21. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК содержит последовательность UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

22. Система по п. 19, в которой нацеливающая последовательность гРНК состоит из последовательности UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22).

23. Система по п. 18, в которой нацеливающая последовательность гРНК содержит последовательность UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

24. Система по п. 18, в которой нацеливающая последовательность гРНК состоит из последовательности UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23).

25. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК содержит последовательность CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

26. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК состоит из последовательности CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949).

27. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК содержит последовательность GAGGCCAAACCCUCCUGGA (SEQ ID NO: 2948) или

последовательность, идентичную ей на по меньшей мере 90% или 95%.

28. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК состоит из последовательности CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2948).

29. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК содержит последовательность AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

30. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК состоит из последовательности AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747).

31. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК содержит последовательность AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748) или последовательность, идентичную ей на по меньшей мере 90% или 95%.

32. Система по п. 19 или п. 20, в которой нацеливающая последовательность гРНК состоит из последовательности AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748).

33. Система по любому из пп. 1-32, дополнительно содержащая вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты BCL11A по сравнению с нацеливающей последовательностью гРНК первой гРНК.

34. Система по п. 33, в которой нацеливающая последовательность второй гРНК комплементарна последовательности целевой нуклеиновой кислоты, которая расположена с 5'- или 3'-конца от последовательности сайта связывания GATA1.

35. Система по п. 33, в которой каждая из первой и второй гРНК имеет нацеливающую последовательность, комплементарную последовательности в пределах промотора гена BCL11A.

36. Система по любому из пп. 1-35, в которой первая или вторая гРНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98%, по меньшей мере примерно 99%.

37. Система по любому из пп. 1-36, в которой первая или вторая гРНК имеет каркас, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265.

38. Система по любому из пп. 1-36, в которой первая или вторая гРНК имеет каркас, состоящий из последовательности, выбранной из группы, состоящей из SEQ ID NO: 2238-2285, 26794-26839 и 27219-27265.

39. Система по п. 38, в которой первая или вторая гРНК имеет каркас, состоящий из последовательности с SEQ ID NO: 2238 или SEQ ID NO: 26800.

40. Система по любому из пп. 36-39, в которой нацеливающая последовательность

связана с 3'-концом каркаса гРНК.

41. Система по любому из пп. 1-40, в которой белок CRISPR класса 2, типа V представляет собой вариант белка CasX, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154, или последовательность, идентичную последовательности на по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80%, по меньшей мере примерно 90%, или по меньшей мере примерно 95%, или по меньшей мере примерно 95%, или по меньшей мере примерно 96%, или по меньшей мере примерно 97%, или по меньшей мере примерно 98%, или по меньшей мере примерно 99%.

42. Система по п. 41, в которой белок CRISPR класса 2, типа V представляет собой вариант белка CasX, содержащий последовательность, выбранную из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154.

43. Система по п. 41, в которой вариант белка CasX состоит из последовательности, выбранной из группы, состоящей из SEQ ID NO: 59, 72-99, 101-148 и 26908-27154.

44. Система по п. 42, в которой вариант белка CasX состоит из последовательности, выбранной из группы, состоящей из SEQ ID NO: 126, 27043, 27046, 27050.

45. Система по п. 41, в которой вариант белка CasX содержит по меньшей мере одну модификацию относительно референсного белка CasX, имеющего последовательность, выбранную из SEQ ID NO: 1-3.

46. Система по п. 45, в которой по меньшей мере одна модификация содержит по меньшей мере одну аминокислотную замену, делецию или замену в домене варианта белка CasX относительно референсного белка CasX.

47. Система по п. 46, в которой домен выбран из группы, состоящей из домена связывания с нецелевой цепью (NTSB), домена загрузки целевой цепи (TSL), спирального домена I, спирального домена II, олигонуклеотид-связывающего домена (OBD) и домена расщепления ДНК RuvC.

48. Система по любому из пп. 41-47, в которой вариант белка CasX не содержит домена HNH.

49. Система по любому из пп. 41-48, в которой вариант белка CasX дополнительно содержит один или более сигналов ядерной локализации (NLS).

50. Система по п. 49, в которой один или более NLS выбраны из группы последовательностей, состоящей из PKKKRKV (SEQ ID NO: 168), KRPAATKKAGQAKKKK (SEQ ID NO: 169), PAAKRVKLD (SEQ ID NO: 170), RQRRNELKRSP (SEQ ID NO: 171), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 172), RMRIZFKNKGKDTAELRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 173), VSRKRPRP (SEQ ID NO: 174), PPKKARED (SEQ ID NO: 175), PQPKKKPL (SEQ ID NO: 176), SALIKKKKKMAP (SEQ ID NO: 177), DRLRR (SEQ ID NO: 178), PKQKKRK (SEQ ID NO: 179), RKLKKKIKKL (SEQ ID NO: 180), REKKKFLKRR (SEQ ID NO: 181),

KRKGDEV DGVDEVAKKKSKK (SEQ ID NO: 182), RKCLQAGMNLEARKTKK (SEQ ID NO: 183), PRPRKIPR (SEQ ID NO: 184), PPRKKRTVV (SEQ ID NO: 185), NLSKKKKRKREK (SEQ ID NO: 186), RRPSRPFRKP (SEQ ID NO: 187), KRPRSPSS (SEQ ID NO: 188), KRGINDRNFWRGENERKTR (SEQ ID NO: 189), PRPPKMARYDN (SEQ ID NO: 190), KRSFSKAF (SEQ ID NO: 191), KLKIKRPVK (SEQ ID NO: 192), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 193), PKTRRRPRRSQRKRPT (SEQ ID NO: 26792), SRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 194), KTRRRPRRSQRKRPT (SEQ ID NO: 195), RRKKRRPRRKKRR (SEQ ID NO: 196), PKKKS RKP KKKSRK (SEQ ID NO: 197), HKKKHPDASVNFSEFSK (SEQ ID NO: 198), QRP GPYDRP QRP GPYDRP (SEQ ID NO: 199), LSPSLSP LSPSLSP L (SEQ ID NO: 200), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 201), PKRGRGRPKRGRGR (SEQ ID NO: 202), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 203), PKKKRKVPPPPKKKRKV (SEQ ID NO: 204), PAKRARRGYKC (SEQ ID NO: 27199), KLGPRKATGRW (SEQ ID NO: 27200), PRRKREE (SEQ ID NO: 27201), PYRGRKE (SEQ ID NO: 27202), PLRKRPRR (SEQ ID NO: 27203), PLRKRPRRGSPLRKRPRR (SEQ ID NO: 27204), PAAKRVKLDGGKRTADGSEFESP KKKRKV (SEQ ID NO: 27205), PAAKRVKLDGGKRTADGSEFESP KKKRKVGIHGV PAA (SEQ ID NO: 27206), PAAKRVKLDGGKRTADGSEFESP KKKRKV A E A A A K E A A A K E A A A K A (SEQ ID NO: 207), PAAKRVKLDGGKRTADGSEFESP KKKRKV P G (SEQ ID NO: 27208), KRKGSPERGERKRHW (SEQ ID NO: 27209), KRTADSQHSTPPKTKRKVEFEP KKKRKV (SEQ ID NO: 27210) и PKKKRKVGGSKRTADSQHSTPPKTKRKVEFEP KKKRKV (SEQ ID NO: 27211), причем один или более NLS связаны с белком CRISPR или со смежным NLS через линкерный пептид, при этом линкерный пептид выбран из группы, состоящей из RS, (G)_n (SEQ ID NO: 27212), (GS)_n (SEQ ID NO: 27213), (GSGGS)_n (SEQ ID NO: 214), (GGSGGS)_n (SEQ ID NO: 215), (GGGS)_n (SEQ ID NO: 216), GGSG (SEQ ID NO: 217), GGSGG (SEQ ID NO: 218), GSGSG (SEQ ID NO: 219), GSGGG (SEQ ID NO: 220), GGGSG (SEQ ID NO: 221), GSSSG (SEQ ID NO: 222), GPGP (SEQ ID NO: 223), GGP, PPP, PPAPPA (SEQ ID NO: 224), PPPG (SEQ ID NO: 27214), PPPGPPP (SEQ ID NO: 225), PPP(GGGS)_n (SEQ ID NO: 27215), (GGGS)_nPPP (SEQ ID NO: 27216), A E A A A K E A A A K E A A A K A (SEQ ID NO: 27217) и TPPKTKRKVEFE (SEQ ID NO: 27218), где n равно от 1 до 5.

51. Система по п. 49 или п. 50, в которой один или более NLS расположены на С-конце варианта белка CasX или рядом с ним.

52. Система по п. 49 или п. 50, в которой один или более NLS расположены на N-конце варианта белка CasX или рядом с ним.

53. Система по п. 49 или п. 50, содержащая один или более NLS, расположенных на N-конце и на С-конце варианта белка CasX или рядом с ними.

54. Система по любому из пп. 41-53, в которой вариант CasX способен образовывать рибонуклеопротеиновый комплекс (RNP) с гидовой нуклеиновой кислотой (гРНК).

55. Система по п. 54, в которой RNP варианта белка CasX и варианта гРНК демонстрирует по меньшей мере одну или более улучшенных характеристик по сравнению с RNP референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК,

содержащей последовательность с любым из SEQ ID NO: 4-16.

56. Система по п. 55, в которой улучшенная характеристика выбрана из одного или более из группы, состоящей из улучшенного фолдинга варианта CasX; улучшенной аффинности связывания с гидовой нуклеиновой кислотой (гРНК); улучшенной аффинности связывания с целевой ДНК; улучшенной способности использовать более широкий спектр из одной или более последовательностей смежного с протоспейсером мотива (PAM), включая АТС, СТС, ГТС или ТТС, при редактировании целевой ДНК; улучшенного раскручивания целевой ДНК; повышенной активности редактирования; улучшенной эффективности редактирования; улучшенной специфичности редактирования; повышенной нуклеазной активности; улучшенной скорости расщепления последовательности целевой нуклеиновой кислоты; повышенной загрузки целевой цепи для двухцепочечного расщепления; сниженной загрузки целевой цепи для одноцепочечного никирования; сниженного нецелевого расщепления; улучшенного связывания с нецелевой цепью ДНК; улучшенной стабильности белка; улучшенной растворимости белка; улучшенного образования рибонуклеопротеинового комплекса (RNP); более высокой процентной доли компетентных в отношении расщепления RNP; улучшенной стабильности комплекса белок:гРНК (RNP); улучшенной растворимости комплекса белок:гРНК; улучшенного выхода белка; улучшенной экспрессии белка; и улучшенных характеристик слияния.

57. Система по п. 55 или п. 56, в которой кратность улучшения характеристики RNP варианта белка CasX и варианта гРНК составляет от по меньшей мере примерно 1,1 до примерно 100 раз или более относительно RNP референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

58. Система по п. 55 или п. 56, в которой кратность улучшения характеристики варианта белка CasX составляет по меньшей мере примерно 1,1, по меньшей мере примерно 2, по меньшей мере примерно 10, по меньшей мере примерно 100 раз или более относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

59. Система по п. 55 или п. 56, в которой кратность улучшения характеристики варианта белка CasX составляет по меньшей мере примерно 1,1, по меньшей мере примерно 2, по меньшей мере примерно 10, по меньшей мере примерно 100 раз или более относительно референсного белка CasX с SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 и гРНК, содержащей последовательность с любым из SEQ ID NO: 4-16.

60. Система по любому из пп. 55-59, в которой улучшенная характеристика включает эффективность редактирования, и RNP варианта белка CasX и варианта гРНК характеризуется кратностью улучшения эффективности редактирования от 1,1 до 100 раз по сравнению с RNP референсного белка CasX с SEQ ID NO: 2 и гРНК с любым из SEQ ID NO: 4-16.

61. Система по любому из пп. 54-59, в которой RNP, содержащий вариант CasX и

вариант гРНК, демонстрирует более высокую эффективность редактирования и/или связывание последовательности целевой нуклеиновой кислоты в клеточной системе анализа, когда любая из РАМ-последовательностей - TTC, ATC, GTC или CTC - расположена на расстоянии 1 нуклеотида от 5'-конца нецелевой цепи протоспейсера, имеющего идентичность с нацеливающей последовательностью гРНК, по сравнению с эффективностью редактирования и/или связыванием в сопоставимой системе анализа для RNP, содержащего референсный белок CasX и референсную гРНК.

62. Система по п. 61, в которой РАМ-последовательность представляет собой TTC.

63. Система по п. 62, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 17904-26789.

64. Система по п. 61, в которой РАМ-последовательность представляет собой ATC.

65. Система по п. 64, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 272-2100 и 2286-5625.

66. Система по п. 61, в которой РАМ-последовательность представляет собой CTC.

67. Система по п. 66, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 5626-13616.

68. Система по п. 61, в которой РАМ-последовательность представляет собой GTC.

69. Система по п. 66, в которой нацеливающая последовательность гРНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 13617-17903.

70. Система по любому из пп. 61-68, в которой повышенная аффинность связывания одной или более РАМ-последовательностей в по меньшей мере 1,5 раза выше по сравнению с аффинностью связывания любого из референсных белков CasX с SEQ ID NO: 1-3 для РАМ-последовательностей.

71. Система по любому из пп. 54-70, в которой RNP имеет на по меньшей мере 5%, на по меньшей мере 10%, на по меньшей мере 15% или на по меньшей мере 20% более высокую процентную долю компетентных в отношении расщепления RNP в сравнении с RNP референсного белка CasX и гРНК с SEQ ID NO: 4-16.

72. Система по любому из пп. 41-71, в которой вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий уникальной активностью.

73. Система по любому из пп. 41-71, в которой вариант белка CasX содержит домен расщепления ДНК RuvC, обладающий активностью расщепления двухцепочечных молекул.

74. Система по любому из пп. 1-54, в которой белок CasX представляет собой каталитически неактивный белок CasX (dCasX), и при этом dCasX и гРНК сохраняют способность связываться с целевой нуклеиновой кислотой BCL11A.

75. Система по п. 74, в которой dCasX содержит мутацию по остаткам:

а. D672, E769 и/или D935, соответствующим белку CasX с SEQ ID NO: 1; или

б. D659, E756 и/или D922, соответствующим белку CasX с SEQ ID NO: 2.

76. Система по п. 75, в которой мутация представляет собой замену остатка на

аланин.

77. Система по любому из пп. 1-73, дополнительно содержащая нуклеиновую кислоту донорной матрицы.

78. Система по п. 77, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок гена BCL11A, выбранный из группы, состоящей из экзона BCL11A, интрона BCL11A, соединения интрон-экзон BCL11A, регуляторного элемента BCL11A и последовательности сайта связывания GATA1.

79. Система по п. 78, в которой последовательность донорной матрицы содержит одну или более мутаций относительно соответствующего участка гена BCL11A дикого типа.

80. Система по п. 78 или п. 79, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A, экзона 3 BCL11A, экзона 4 BCL11A, экзона 5 BCL11A, экзона 6 BCL11A, экзона 7 BCL11A, экзона 8 BCL11A и экзона 9 BCL11A.

81. Система по п. 80, в которой донорная матрица содержит нуклеиновую кислоту, содержащую по меньшей мере участок экзона BCL11A, выбранного из группы, состоящей из экзона 1 BCL11A, экзона 2 BCL11A и экзона 3 BCL11A.

82. Система по любому из пп. 77-81, в которой размер донорной матрицы находится в диапазоне от 10 до 15 000 нуклеотидов.

83. Система по любому из пп. 77-82, в которой донорная матрица представляет собой одноцепочечную ДНК-матрицу или одноцепочечную РНК-матрицу.

84. Система по любому из пп. 77-82, в которой донорная матрица представляет собой двухцепочечную ДНК-матрицу.

85. Система по любому из пп. 77-84, в которой донорная матрица содержит гомологичные плечи на 5'- и 3'-концах донорной матрицы, которые комплементарны последовательностям, фланкирующим сайты расщепления в целевой нуклеиновой кислоте BCL11A, вводимой белком CRISPR класса 2, типа V, или рядом с ними.

86. Нуклеиновая кислота, содержащая донорную матрицу по любому из пп. 77-85.

87. Нуклеиновая кислота, содержащая последовательность, которая кодирует CasX по любому из пп. 41-76.

88. Нуклеиновая кислота, содержащая последовательность, которая кодирует гРНК по любому из пп. 1-40.

89. Нуклеиновая кислота по п. 87, в которой последовательность, которая кодирует белок CasX, является кодон-оптимизированной для экспрессии в эукариотической клетке.

90. Вектор, содержащий гРНК по любому из пп. 1-40, белок CasX по любому из пп. 41-76 или нуклеиновую кислоту по любому из пп. 86-89.

91. Вектор по п. 90, дополнительно содержащий один или более промоторов.

92. Вектор по п. 90 или п. 91, выбранный из группы, состоящей из ретровирусного вектора, лентивирусного вектора, аденовирусного вектора, аденоассоциированного

вирусного (AAV) вектора, вектора на основе вируса простого герпеса (HSV), вирусоподобной частицы (VLP), частицы для доставки CasX (XDP), плазмиды, миникольца, наноплазмиды, ДНК-вектора и РНК-вектора.

93. Вектор по п. 92, представляющий собой вектор AAV.

94. Вектор по п. 93, причем вектор AAV выбран из AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74 или AAVRh10.

95. Вектор по п. 94, причем вектор AAV выбран из AAV1, AAV2, AAV5, AAV8 или AAV9.

96. Вектор по п. 94 или п. 95, причем вектор AAV содержит нуклеиновую кислоту, содержащую следующие компоненты:

a. 5'-ITR;

b. 3'-ITR; и

c. трансген, содержащий нуклеиновую кислоту по п. 87, функционально связанную с первым промотором, и нуклеиновую кислоту по п. 88, функционально связанную со вторым промотором.

97. Вектор по п. 96, в котором нуклеиновая кислота дополнительно содержит последовательность поли(А) с 3'-конца от последовательности, кодирующей белок CasX.

98. Вектор по п. 96 или п. 97, в котором нуклеиновая кислота дополнительно содержит один или более энхансерных элементов.

99. Вектор по любому из пп. 96-98, в котором одна частица AAV способна содержать нуклеиновую кислоту, причем частица AAV имеет все компоненты, необходимые для трансдукции и эффективной модификации целевой нуклеиновой в клетке-мишени.

100. Вектор по п. 92, представляющий собой ретровирусный вектор.

101. Вектор по п. 92, представляющий собой XDP, содержащую один или более компонентов полибелка gag.

102. Вектор по п. 101, в котором один или более компонентов полибелка gag выбраны из группы, состоящей из матриксного белка (МА), нуклеокапсидного белка (NC), капсидного белка (CA), пептида р1, пептида р6, пептида Р2А, пептида Р2В, пептида Р10, пептида р12, пептида РР21/24, пептида Р12/Р3/Р8 и пептида Р20.

103. Вектор по п. 101 или п. 102, в котором XDP содержит один или более компонентов полибелка gag, вариант белка CasX и гРНК.

104. Вектор по п. 103, в котором вариант белка CasX и гРНК связаны друг с другом в RNP.

105. Вектор по любому из пп. 101-104, дополнительно содержащий донорную матрицу.

106. Вектор по любому из пп. 101-104, дополнительно содержащий псевдотипирующий гликопротеин вирусной оболочки или фрагмент антитела, который обеспечивает связывание и слияние XDP с клеткой-мишенью.

107. Вектор по п. 106, причем клетка-мишень выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника

(NPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), эмбриональной стволовой (ES) клетки, индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

108. Клетка-хозяин, содержащая вектор по любому из пп. 90-107.

109. Клетка-хозяин по п. 108, выбранная из группы, состоящей из клеток-фибробластов почки новорожденного хомяка (BHK), клеток эмбриональной почки человека 293 (HEK293), клеток эмбриональной почки человека 293T (HEK293T), клеток NS0, клеток SP2/0, клеток миеломы YO, мышинных клеток миеломы P3X63, клеток PER, клеток PER.C6, клеток гибридомы, клеток NIH3T3, клеток, происходящих от CV-1 (обезьяньих), с генетическим материалом SV40 (COS), клеток HeLa, клеток яичника китайского хомячка (CHO) и дрожжевых клеток.

110. Способ модификации последовательности целевой нуклеиновой кислоты BCL11A в популяции клеток, включающий введение в клетки популяции:

- a. системы по любому из пп. 1-85;
- b. нуклеиновой кислоты по любому из пп. 86-89;
- c. вектора по любому из пп. 90-100;
- d. XDP по любому из пп. 101-107; или
- e. комбинаций двух или более из пп. (a)-(d),

причем последовательность целевой нуклеиновой кислоты гена BCL11A клеток, на которую нацелена первая гРНК, модифицирована вариантом белка CasX.

111. Способ по п. 110, в котором модификация включает введение одноцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

112. Способ по п. 110, в котором модификация включает введение двухцепочечного разрыва в последовательность целевой нуклеиновой кислоты гена BCL11A клеток популяции.

113. Способ по любому из пп. 110-112, дополнительно включающий введение в клетки популяции второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты гена BCL11A по сравнению с первой гРНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток популяции.

114. Способ по любому из пп. 110-113, в котором модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток популяции.

115. Способ по пп. 110-114, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

116. Способ по любому из пп. 110-113, включающий вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток популяции.

117. Способ по п. 114, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (НТИ).

118. Способ по п. 116 или п. 117, в котором последовательность сайта связывания GATA1 целевой нуклеиновой кислоты модифицирована.

119. Способ по любому из пп. 116-118, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в клетках популяции.

120. Способ по любому из пп. 110-119, в котором ген BCL11A клеток популяции модифицирован так, что экспрессия белка BCL11A снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с клетками, в которых ген BCL11A не модифицирован.

121. Способ по любому из пп. 110-119, в котором ген BCL11A клеток популяции модифицирован так, что по меньшей мере примерно 10%, по меньшей мере примерно 20%, по меньшей мере примерно 30%, по меньшей мере примерно 40%, по меньшей мере примерно 50%, по меньшей мере примерно 60%, по меньшей мере примерно 70%, по меньшей мере примерно 80% или по меньшей мере примерно 90% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

122. Способ по любому из пп. 110-121, в котором клетки являются эукариотическими.

123. Способ по п. 122, в котором эукариотические клетки выбраны из группы, состоящей из клеток грызунов, клеток мыши, клеток крысы и клеток приматов, не являющихся человеком.

124. Способ по п. 122, в котором эукариотические клетки представляют собой клетки человека.

125. Способ по любому из пп. 122-124, в котором эукариотическая клетка выбрана из группы, состоящей из гематопозитической стволовой клетки (HSC), гематопозитической клетки-предшественника (HPC), клетки CD34+, мезенхимальной стволовой клетки (MSC), индуцированной плюрипотентной стволовой клетки (iPSC), общей миелоидной клетки-предшественника, клетки-проэритробласта и клетки-эритробласта.

126. Способ по любому из пп. 110-125, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vitro* или *ex vivo*.

127. Способ по любому из пп. 110-125, в котором модификация последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток происходит *in vivo* у субъекта.

128. Способ по п. 127, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

129. Способ по п. 127, в котором субъект является человеком.

130. Способ по любому из пп. 127-129, включающий введение субъекту терапевтически эффективной дозы вектора AAV.

131. Способ по п. 130, в котором вектор AAV вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

132. Способ по п. 130, в котором вектор AAV вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

133. Способ по любому из пп. 127-129, включающий введение субъекту терапевтически эффективной дозы XDP.

134. Способ по п. 133, в котором XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

135. Способ по п. 133, в котором XDP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг.

136. Способ по любому из пп. 128-135, в котором вектор или XDP вводят субъекту способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

137. Способ по любому из пп. 128-136, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

138. Способ по любому из пп. 128-137, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0,

по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

139. Способ по любому из пп. 128-138, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

140. Способ по любому из пп. 110-139, дополнительно включающий приведение последовательности целевой нуклеиновой кислоты гена BCL11A популяции клеток в контакт с:

a. дополнительной нуклеазой CRISPR и гРНК, нацеленной на другой или перекрывающийся участок целевой нуклеиновой кислоты BCL11A по сравнению с первой гРНК;

b. полинуклеотидом, кодирующим дополнительную нуклеазу CRISPR и гРНК по п. (a);

c. вектором, содержащим полинуклеотид по п. (b); или

d. XDP, содержащей дополнительную нуклеазу CRISPR и гРНК по п. (a),

причем приведение в контакт приводит к модификации гена BCL11A в другом положении в последовательности по сравнению с последовательностью, на которую нацелена первая гРНК.

141. Способ по п. 140, в котором дополнительная нуклеаза CRISPR представляет собой белок CasX, имеющий последовательность, отличную от белка CasX по любому из предшествующих пунктов.

142. Способ по п. 140, в котором дополнительная нуклеаза CRISPR не является белком CasX.

143. Способ по п. 142, в котором дополнительная нуклеаза CRISPR выбрана из группы, состоящей из Cas9, Cas12a, Cas12b, Cas12c, Cas12d (CasY), Cas12j, Cas12k, Cas13a, Cas13b, Cas13c, Cas13d, Cas14, Cpf1, C2cl, Csn2 и вариантов их последовательностей.

144. Популяция клеток, модифицированная способом по любому из пп. 110-143, в которой клетки были модифицированы так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

145. Популяция клеток, модифицированная способом по любому из пп. 110-143, в которой клетки были модифицированы так, что экспрессия белка BCL11A снижена на по меньшей мере 70%, на по меньшей мере 75%, на по меньшей мере 80%, на по меньшей мере 85%, на по меньшей мере 90% или на по меньшей мере 95% по сравнению с клетками, в которых ген BCL11A не был модифицирован.

146. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий введение субъекту терапевтически эффективного количества клеток по п. 144 или п. 145.

147. Способ по п. 146, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

148. Способ по п. 146 или п. 147, в котором клетки являются аутологичными по отношению к субъекту, которому должны быть введены клетки.

149. Способ по п. 146 или п. 147, в котором клетки являются аллогенными по отношению к субъекту, которому должны быть введены клетки.

150. Способ по любому из пп. 146-149, в котором клетки или их потомство сохраняются у субъекта в течение по меньшей мере одного месяца, двух месяцев, трех месяцев, четырех месяцев, пяти месяцев, шести месяцев, семи месяцев, восьми месяцев, девяти месяцев, десяти месяцев, одиннадцати месяцев, двенадцати месяцев, тринадцати месяцев, четырнадцати месяцев, пятнадцати месяцев, шестнадцати месяцев, семнадцати месяцев, восемнадцати месяцев, девятнадцати месяцев, двадцати месяцев, двадцати одного месяца, двадцати двух месяцев, двадцати трех месяцев, двух лет, трех лет, четырех лет или пяти лет после введения модифицированных клеток субъекту.

151. Способ по любому из пп. 146-150, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

152. Способ по любому из пп. 146-150, который приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

153. Способ по любому из пп. 146-150, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

154. Способ по любому из пп. 146-153, в котором субъект выбран из группы, состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

155. Способ по любому из пп. 146-153, в котором субъект представляет собой человека.

156. Способ лечения гемоглобинопатии у нуждающегося в этом субъекта, включающий модификацию гена BCL11A в клетках субъекта, причем модификация включает приведение упомянутых клеток в контакт с терапевтически эффективной дозой:

- a. системы по любому из пп. 1-85;
- b. нуклеиновой кислоты по любому из пп. 86-89;
- c. вектора по любому из пп. 90-100;
- d. XDP по любому из пп. 101-104; или
- e. комбинаций двух или более из пп. (a)-(d),

при этом ген BCL11A клеток, на который нацелена первая гРНК, модифицирован белком CasX.

157. Способ по п. 156, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

158. Способ по любому из п. 156 или п. 157, в котором клетка выбрана из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

159. Способ по любому из пп. 156-158, в котором модификация включает введение одноцепочечного разрыва в ген BCL11A клеток.

160. Способ по любому из пп. 156-158, в котором модификация включает введение двухцепочечного разрыва в ген BCL11A клеток.

161. Способ по любому из пп. 156-160, дополнительно включающий введение в клетки субъекта второй гРНК или нуклеиновой кислоты, кодирующей вторую гРНК, причем вторая гРНК имеет нацеливающую последовательность, комплементарную другому или перекрывающемуся участку целевой нуклеиновой кислоты по сравнению с первой гРНК, что приводит к дополнительному разрыву в целевой нуклеиновой кислоте BCL11A клеток субъекта.

162. Способ по любому из пп. 156-161, в котором модификация включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в ген BCL11A клеток.

163. Способ по п. 162, в котором модификация приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

164. Способ по любому из пп. 156-163, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно на 90% по сравнению с клетками, которые не были модифицированы.

165. Способ по любому из пп. 156-163, в котором ген BCL11A клеток субъекта модифицированы так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

166. Способ по любому из пп. 156-165, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

167. Способ по любому из пп. 156-166, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере

0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

168. Способ по любому из пп. 156-165, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

169. Способ по любому из пп. 156-161, включающий вставку донорной матрицы в сайт (-ы) разрыва последовательности целевой нуклеиновой кислоты гена BCL11A клеток.

170. Способ по п. 168, в котором вставка донорной матрицы опосредована направляемой гомологией репарацией (HDR) или независимой от гомологии нацеленной интеграцией (НТИ).

171. Способ по п. 168 или п. 170, в котором вставка донорной матрицы приводит к нокдауну или нокауту гена BCL11A в модифицированных клетках субъекта.

172. Способ по любому из пп. 166-171, в котором ген BCL11A клеток модифицирован так, что экспрессия белка BCL11A модифицированными клетками снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно на 90% по сравнению с клетками, которые не были модифицированы.

173. Способ по любому из пп. 166-171, в котором ген BCL11A клеток субъекта модифицированы так, что по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% модифицированных клеток не экспрессируют белок BCL11A на обнаруживаемом уровне.

174. Способ по любому из пп. 166-173, который приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

175. Способ по любому из пп. 166-173, который приводит к соотношению HbF и гемоглобина S (HbS) в кровотоке субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

176. Способ по любому из пп. 166-173, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

177. Способ по любому из пп. 156-175, в котором субъект выбран из группы,

состоящей из грызуна, мыши, крысы и примата, не являющегося человеком.

178. Способ по любому из пп. 156-175, в котором субъект представляет собой человека.

179. Способ по любому из пп. 156-178, в котором вектор представляет собой AAV, и его вводят субъекту в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

180. Способ по любому из пп. 156-178, в котором вектор представляет собой AAV, и его вводят субъекту в дозе по меньшей мере от примерно 1×10^5 вг/кг до примерно 1×10^{16} вг/кг, по меньшей мере от примерно 1×10^6 вг/кг до примерно 1×10^{15} вг/кг или по меньшей мере от примерно 1×10^7 вг/кг до примерно 1×10^{14} вг/кг.

181. Способ по любому из пп. 156-178, в котором XDP вводят субъекту в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг, по меньшей мере примерно 1×10^{16} частиц/кг.

182. Способ по любому из пп. 156-178, в котором XDP вводят субъекту в дозе по меньшей мере от примерно 1×10^5 частиц/кг до примерно 1×10^{16} частиц/кг, или по меньшей мере от примерно 1×10^6 частиц/кг до примерно 1×10^{15} частиц/кг, или по меньшей мере от примерно 1×10^7 частиц/кг до примерно 1×10^{14} частиц/кг.

183. Способ по любому из пп. 156-182, в котором вектор или XDP вводят субъекту способом введения, выбранным из интрапаренохиматозного, внутривенного, внутриартериального, внутрибрюшинного, внутрикапсульного, подкожного, внутримышечного, интраабдоминального или их комбинаций, причем способ введения представляет собой инъекцию, трансфузию или имплантацию.

184. Способ по любому из пп. 156-183, который приводит к улучшению по меньшей мере одной клинически значимой конечной точки у субъекта.

185. Способ по п. 184, который приводит к улучшению по меньшей мере одного клинически значимого параметра, выбранного из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития

тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

186. Способ по п. 184, который приводит к улучшению по меньшей мере двух клинически значимых параметров, выбранных из группы, состоящей из развития заболевания органа-мишени, альбуминурии, гипертензии, гипостении, гипостенурии, диастолической дисфункции, функциональной способности переносить физическую нагрузку, острого коронарного синдрома, болевых явлений, степени выраженности боли, анемии, гемолиза, гипоксии тканей, нарушения функции органов, аномальных значений гематокрита, детской смертности, частоты возникновения инсультов, уровней гемоглобина по сравнению с исходным показателем, уровней HbF, сниженной частоты развития тромбоэмболии легочной артерии, частоты возникновения вазоокклюзионных кризов, концентрации гемоглобина S в эритроцитах, частоты госпитализаций, концентрации железа в печени, потребности в трансфузии крови и оценки качества жизни.

187. Способ лечения субъекта с гемоглобинопатией, включающий:

a. выделение индуцированных плюрипотентных стволовых клеток (iPSC) или гематопозитических стволовых клеток (HSC) у субъекта;

b. модификацию целевой нуклеиновой кислоты BCL11A в iPSC или HSC способом по любому из пп. 110-126;

c. дифференцировку модифицированных iPSC или HSC в гематопозитическую клетку-предшественника; и

d. имплантацию гематопозитической клетки-предшественника субъекту с гемоглобинопатией,

причем способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

188. Способ по п. 187, в котором iPSC или HSC является аутологичной и выделена из костного мозга или из периферической крови субъекта.

189. Способ по п. 187, в котором iPSC или HSC является аллогенной и выделена из костного мозга или из периферической крови другого субъекта.

190. Способ по любому из пп. 187-189, в котором имплантация включает введение гематопозитической клетки-предшественника субъекту посредством трансплантации, местной инъекции, системной инфузии или их комбинаций.

191. Способ по любому из пп. 187-190, в котором гемоглобинопатия представляет собой серповидноклеточную анемию или бета-талассемию.

192. Способ повышения уровня эмбрионального гемоглобина (HbF) у субъекта посредством редактирования генома, включающий:

a. введение субъекту эффективной дозы вектора по любому из пп. 90-100 или XDP

по любому из пп. 101-107, причем вектор или XDP доставляет систему CasX:гРНК в клетки субъекта;

б. редактирование целевой нуклеиновой кислоты BCL11A клеток субъекта с помощью CasX, на который нацелена первая гРНК;

с. редактирование включает введение вставки, делеции, замены, дублирования или инверсии одного или более нуклеотидов в последовательности целевой нуклеиновой кислоты так, что экспрессия белка BCL11A снижается или устраняется,

при этом способ приводит к повышению уровней гемоглобина F (HbF) в кровотоке субъекта на по меньшей мере примерно 5%, на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40% или на по меньшей мере примерно 50% по сравнению с уровнями HbF у субъекта до лечения.

193. Способ по п. 192, который приводит к соотношению HbF и гемоглобина S (HbS) у субъекта по меньшей мере 0,01:1,0, по меньшей мере 0,025:1,0, по меньшей мере 0,05:1,0, по меньшей мере 0,075:1,0, по меньшей мере 0,1:1,0, по меньшей мере 0,2:1,0, по меньшей мере 0,3:1,0, по меньшей мере 0,4:1,0, по меньшей мере 0,5:1,0, по меньшей мере 0,75:1,0, по меньшей мере 1,0:1,0, по меньшей мере 1,25:1,0, по меньшей мере 1,5:1,0 или по меньшей мере 1,75:1,0.

194. Способ по п. 192 или п. 193, который приводит к уровням HbF по меньшей мере примерно 5%, или по меньшей мере примерно 10%, или по меньшей мере примерно 20%, или по меньшей мере примерно 30% от общего гемоглобина в кровотоке субъекта.

195. Способ по любому из пп. 192-194, в котором клетки выбраны из группы, состоящей из гематопозитических стволовых клеток (HSC), гематопозитических клеток-предшественников (HPC), клеток CD34+, мезенхимальных стволовых клеток (MSC), индуцированных плюрипотентных стволовых клеток (iPSC), общих миелоидных клеток-предшественников, клеток-проэритробластов и клеток-эритробластов.

196. Способ по любому из пп. 192-195, в котором целевая нуклеиновая кислота клеток отредактирована так, что экспрессия белка BCL11A снижена на по меньшей мере примерно 10%, на по меньшей мере примерно 20%, на по меньшей мере примерно 30%, на по меньшей мере примерно 40%, на по меньшей мере примерно 50%, на по меньшей мере примерно 60%, на по меньшей мере примерно 70%, на по меньшей мере примерно 80% или на по меньшей мере примерно 90% по сравнению с целевой нуклеиновой кислотой клеток, которые не были отредактированы.

197. Способ по любому из пп. 192-196, в котором субъект выбран из группы, состоящей из мыши, крысы, свиньи и примата, не являющегося человеком.

198. Способ по любому из пп. 192-196, в котором субъект представляет собой человека.

199. Способ по любому из пп. 192-198, в котором вектор вводят в дозе по меньшей мере примерно 1×10^5 векторных геномов/кг (вг/кг), по меньшей мере примерно 1×10^6 вг/кг, по меньшей мере примерно 1×10^7 вг/кг, по меньшей мере примерно 1×10^8 вг/кг, по меньшей

мере примерно 1×10^9 вг/кг, по меньшей мере примерно 1×10^{10} вг/кг, по меньшей мере примерно 1×10^{11} вг/кг, по меньшей мере примерно 1×10^{12} вг/кг, по меньшей мере примерно 1×10^{13} вг/кг, по меньшей мере примерно 1×10^{14} вг/кг, по меньшей мере примерно 1×10^{15} вг/кг или по меньшей мере примерно 1×10^{16} вг/кг.

200. Способ по любому из пп. 192-198, в котором XDP вводят в дозе по меньшей мере примерно 1×10^5 частиц/кг, по меньшей мере примерно 1×10^6 частиц/кг, по меньшей мере примерно 1×10^7 частиц/кг, по меньшей мере примерно 1×10^8 частиц/кг, по меньшей мере примерно 1×10^9 частиц/кг, по меньшей мере примерно 1×10^{10} частиц/кг, по меньшей мере примерно 1×10^{11} частиц/кг, по меньшей мере примерно 1×10^{12} частиц/кг, по меньшей мере примерно 1×10^{13} частиц/кг, по меньшей мере примерно 1×10^{14} частиц/кг, по меньшей мере примерно 1×10^{15} частиц/кг или по меньшей мере примерно 1×10^{16} частиц/кг.

201. Способ по любому из пп. 192-200, в котором вектор или XDP вводят способом введения, выбранным из трансплантации, местной инъекции, системной инфузии или их комбинаций.

202. Система по любому из пп. 1-85, нуклеиновая кислота по любому из пп. 86-89, вектор по любому из пп. 90-95, XDP по любому из пп. 101-104, клетка-хозяин по п. 108 или п. 109 или популяция клеток по п. 144 или п. 145 для применения в качестве лекарственного средства для лечения гемоглобинопатии.

203. Система по п. 1, в которой последовательность целевой нуклеиновой кислоты комплементарна последовательности нецелевой цепи, расположенной на расстоянии 1 нуклеотида от 3'-конца последовательности смежного с протоспейсером мотива (PAM).

204. Система по п. 203, в которой PAM-последовательность содержит мотив TC.

205. Система по п. 204, в которой PAM-последовательность содержит ATC, GTC, CTC или TTC.

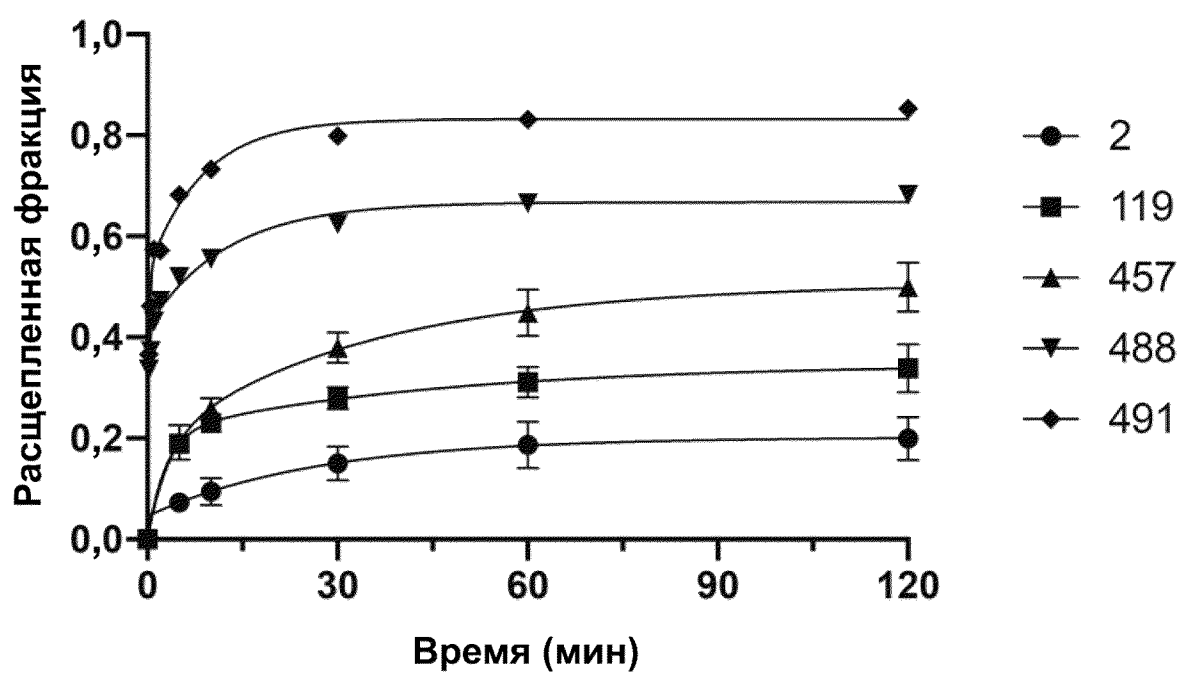
206. Система по любому из пп. 203-205, в которой белок CRISPR класса 2, типа V содержит домен RuvC.

207. Система по п. 206, в которой домен RuvC создает смещенный двухцепочечный разрыв в последовательности целевой нуклеиновой кислоты.

208. Система по любому из пп. 203-207, в которой белок CRISPR класса 2, типа V не содержит нуклеазного домена HNH.

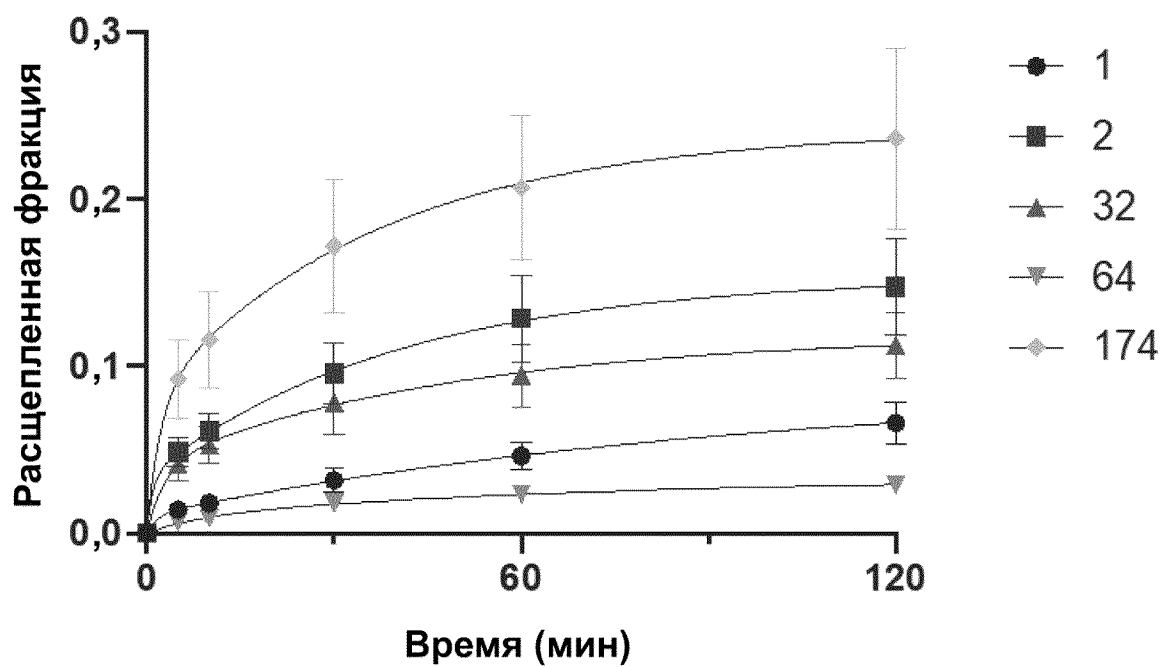
По доверенности

Активные фракции вариантов белка



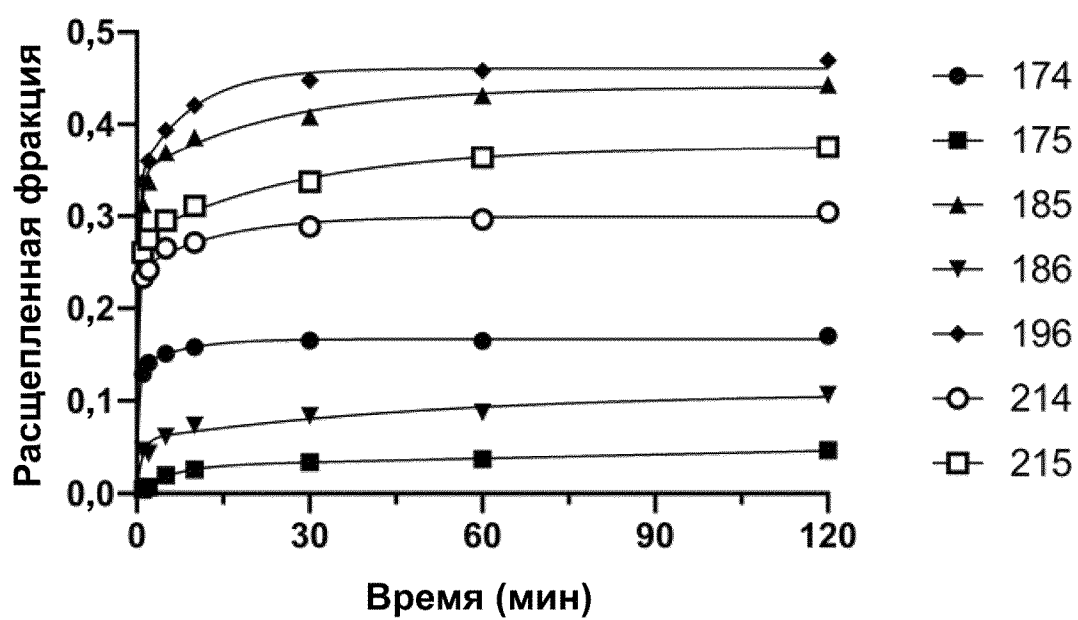
ФИГ. 1

Активные фракции вариантов огРНК



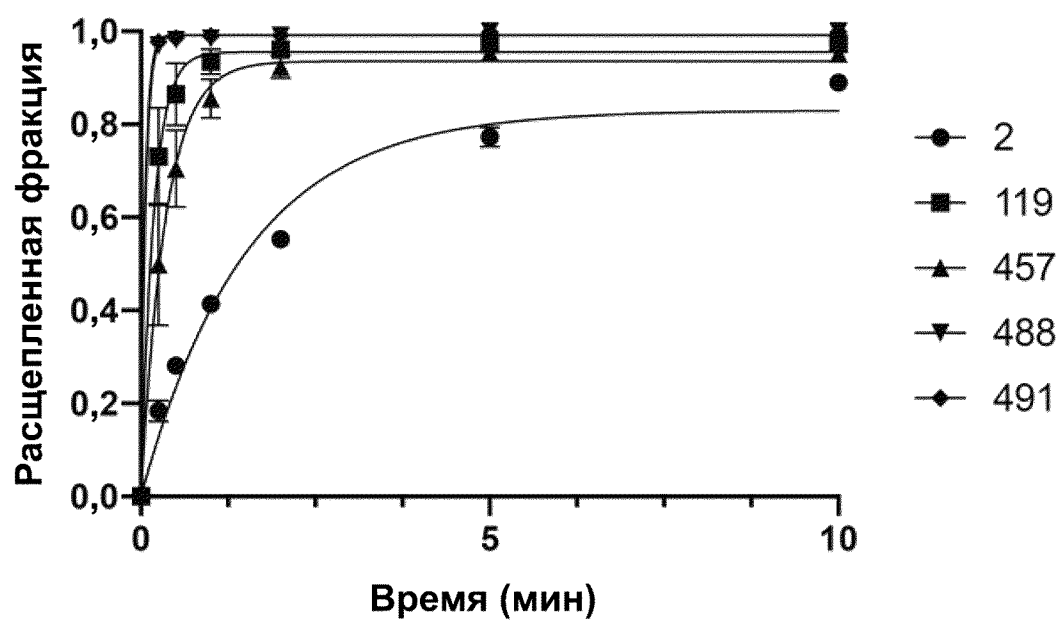
ФИГ. 2

Компетентные фракции в условиях ограничения гида



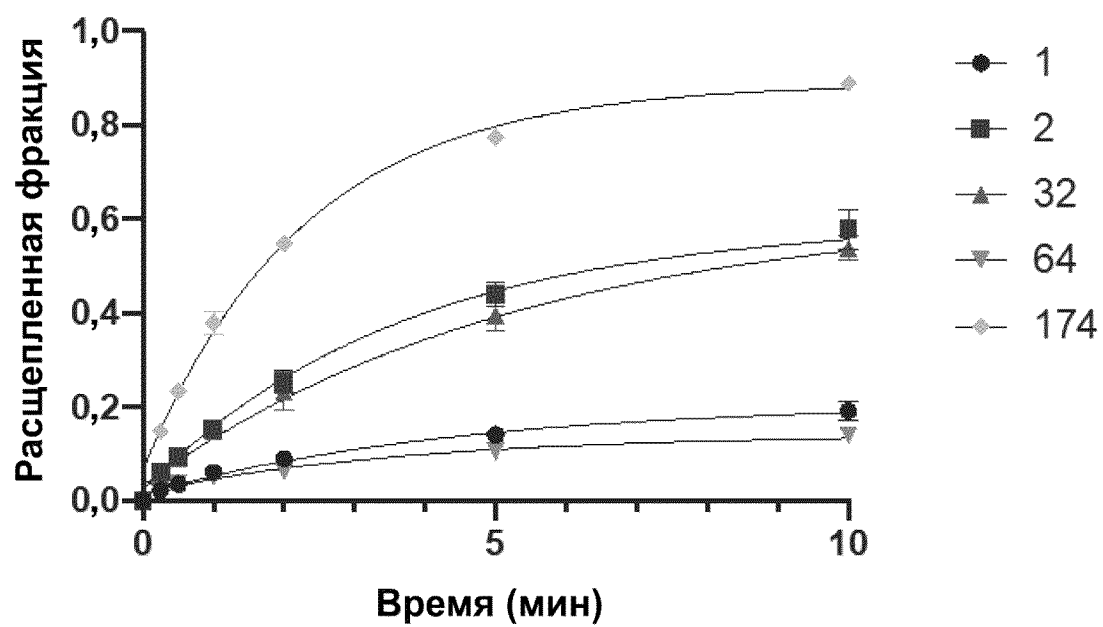
ФИГ. 3

Скорости расщепления вариантов белка



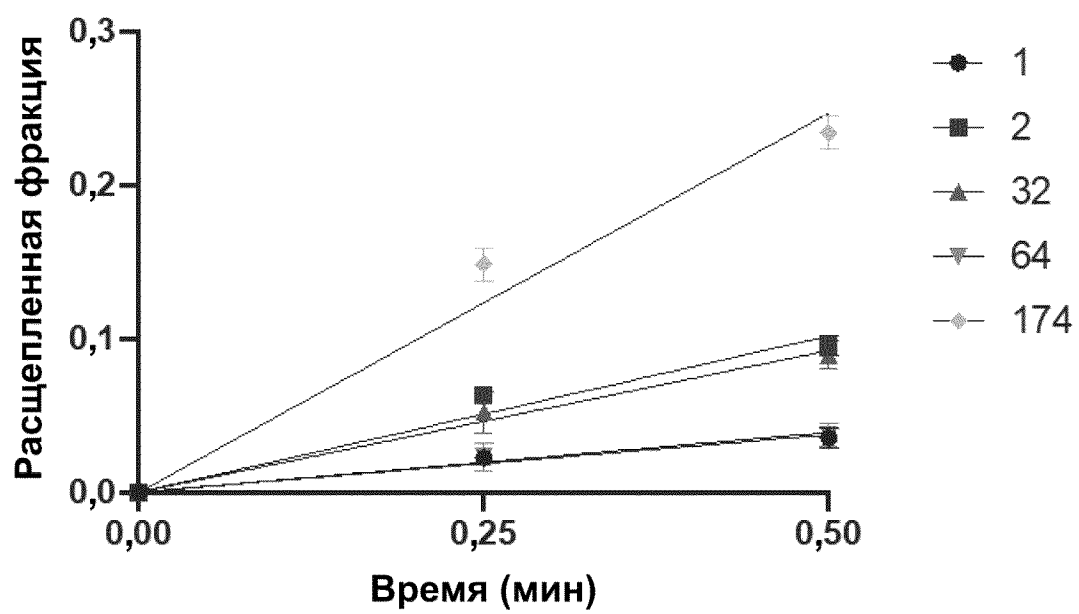
ФИГ. 4

Скорости расщепления вариантов ogРНК



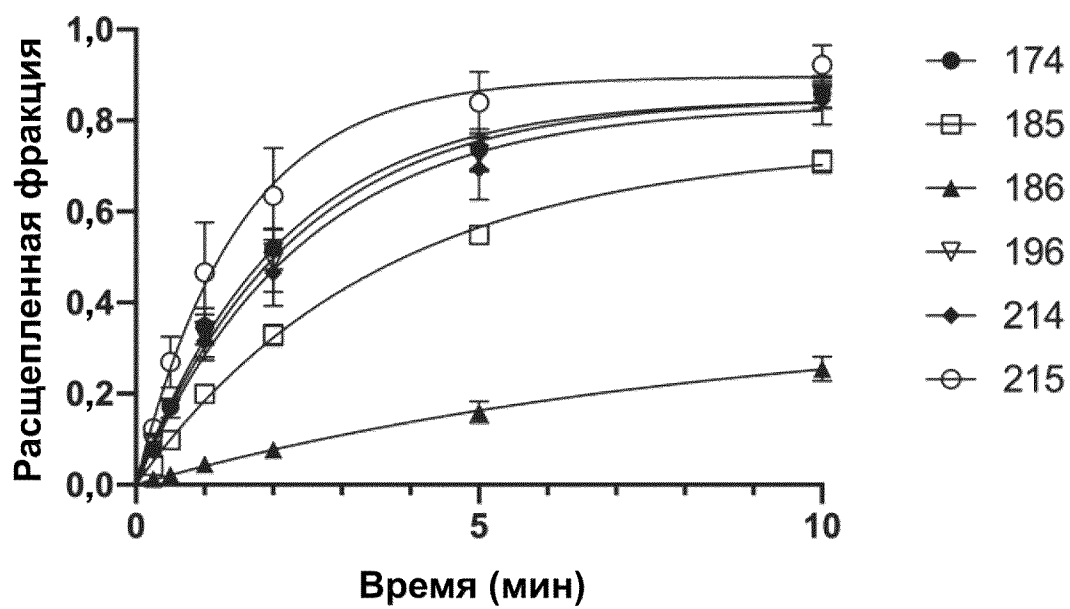
ФИГ. 5

Начальные скорости вариантов оgРНК

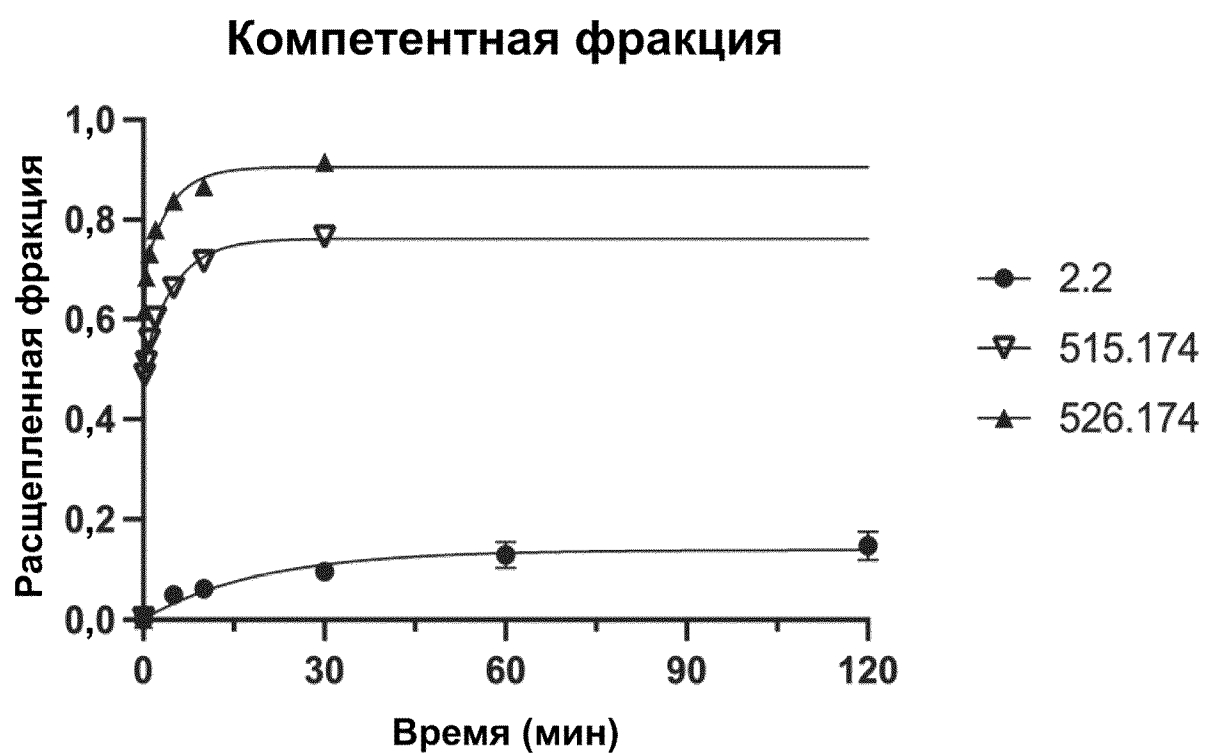


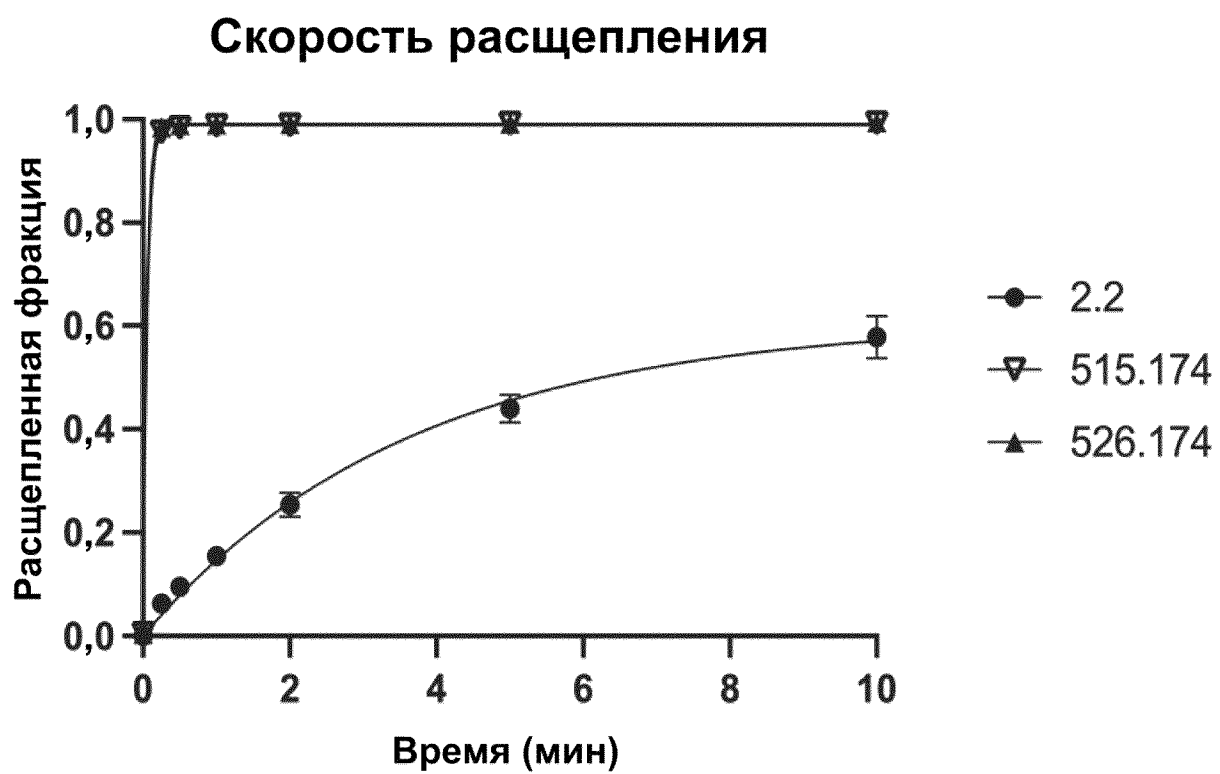
ФИГ. 6

Скорости расщепления вариантов огРНК при 10 °С



ФИГ. 7

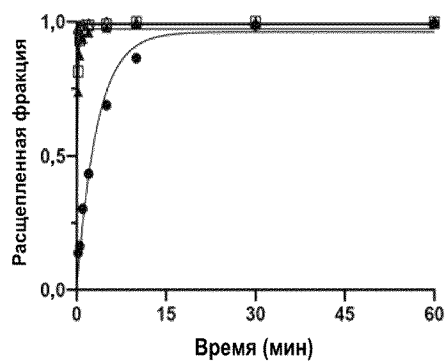
**ФИГ. 8**

**ФИГ. 9**

Скорость расщепления на РАМ NTC

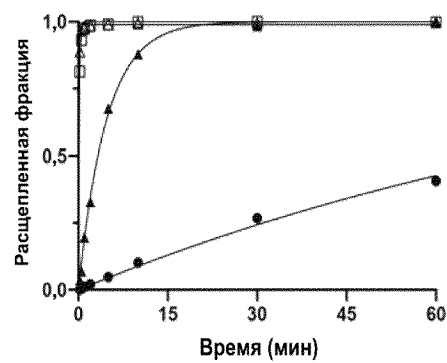
ФИГ. 10А

ТТС



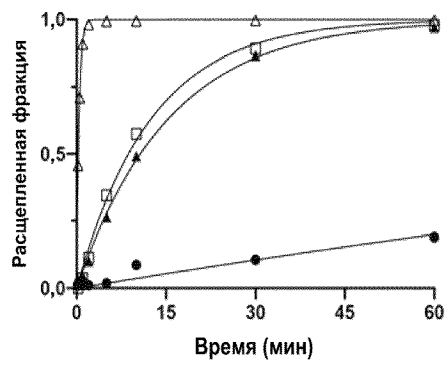
ФИГ. 10В

СТС



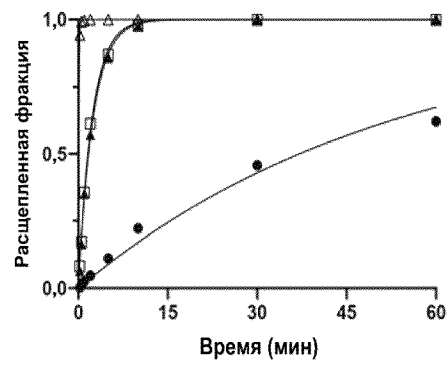
ФИГ. 10С

ГТС



ФИГ. 10Д

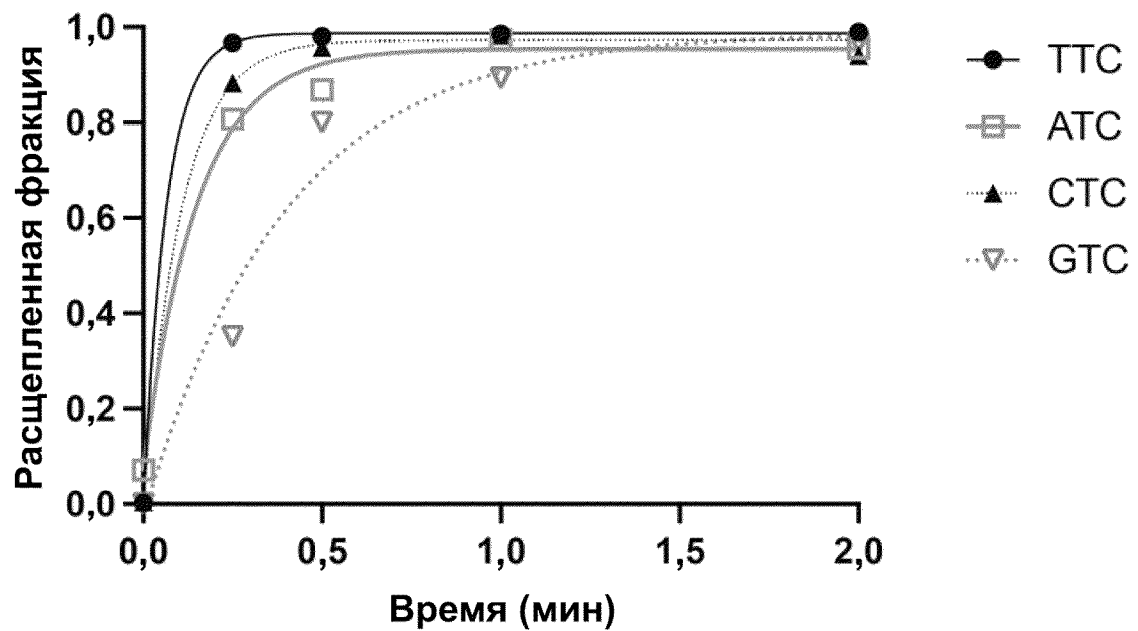
АТС



● 2 ▲ 438
 □ 119 △ 491

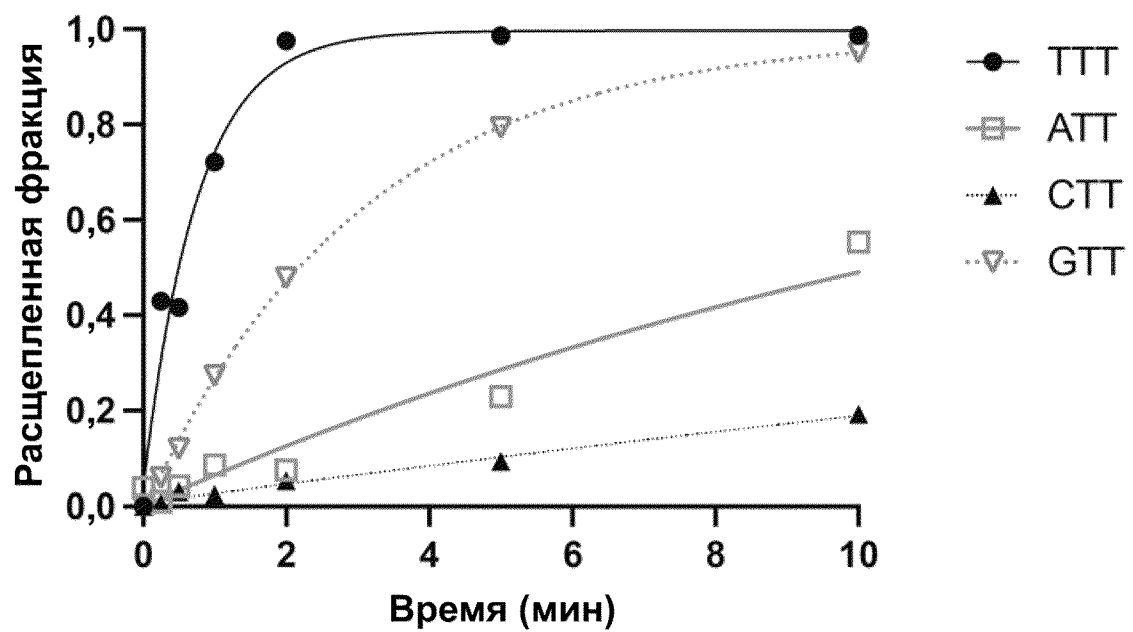
ФИГ. 11А

РАМ NTC



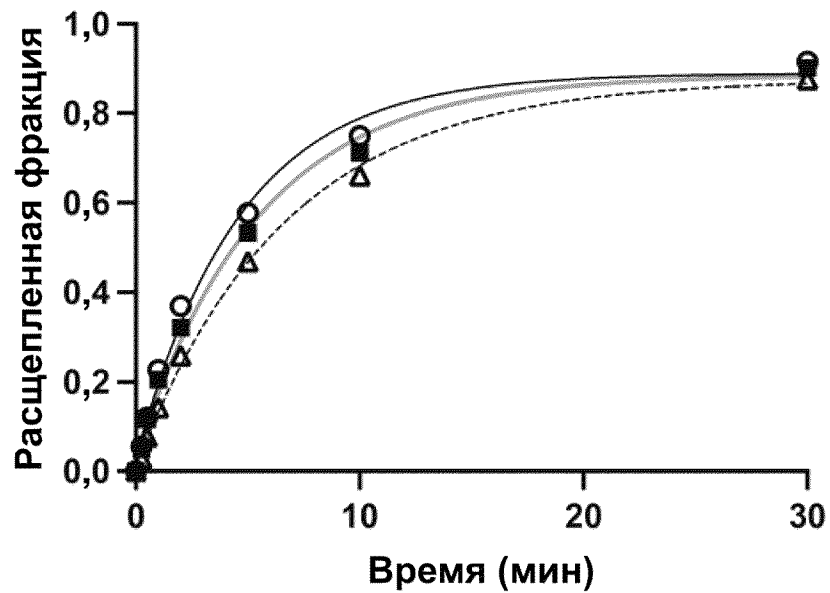
ФИГ. 11В

РАМ NTT



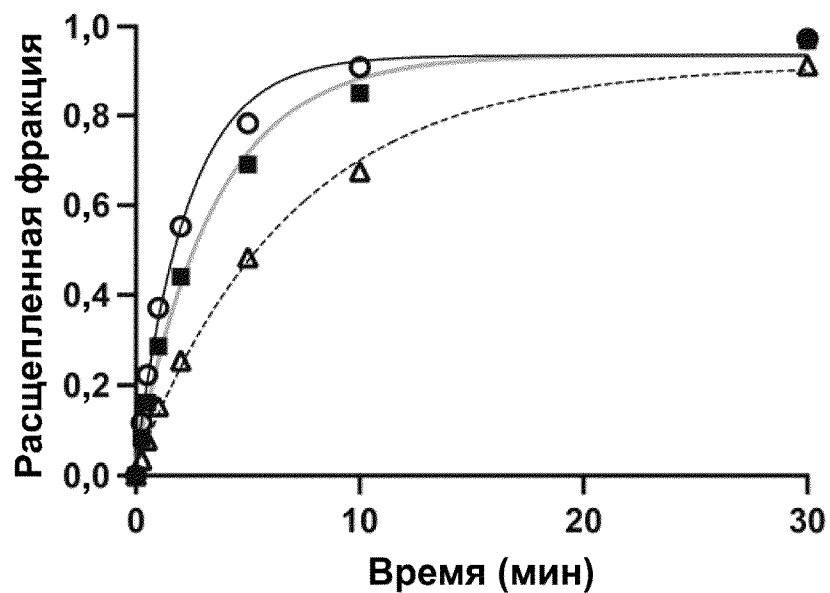
ФИГ. 12А

Вариант 515



ФИГ. 12В

Вариант 526

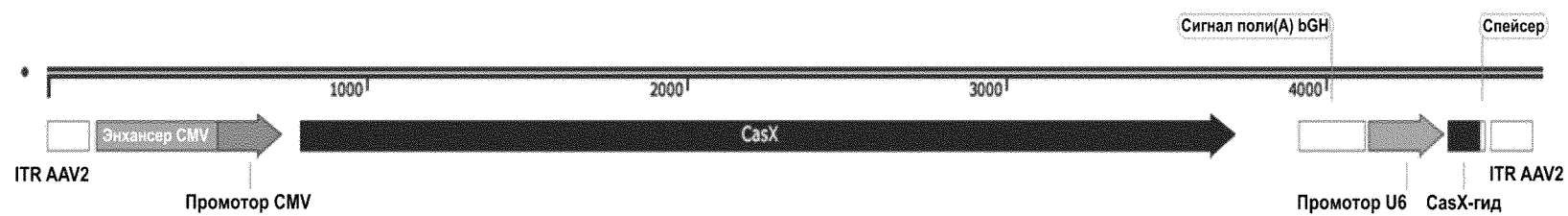


Длина спейсера

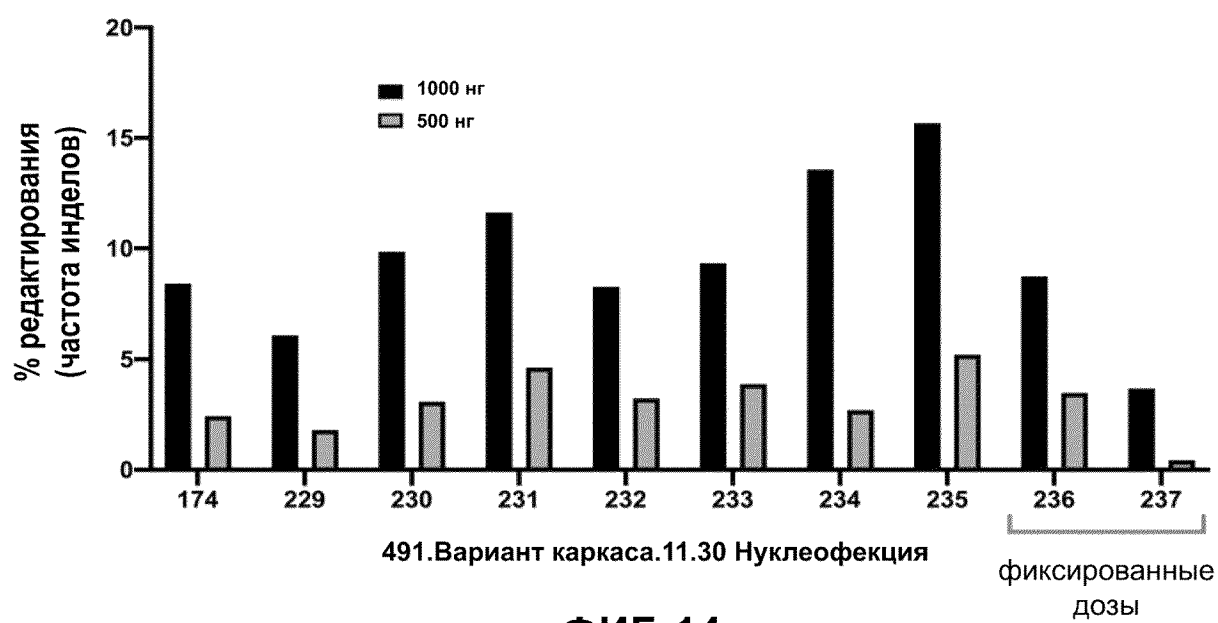
○ 18

■ 19

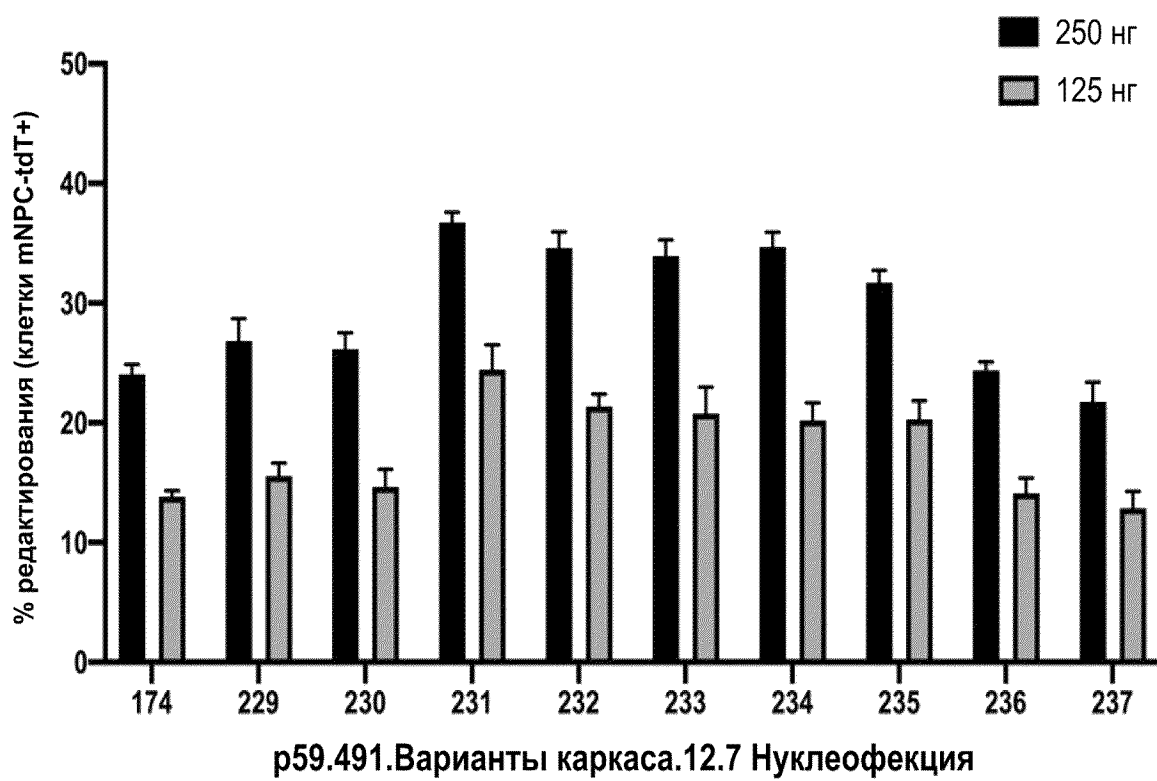
△ 20



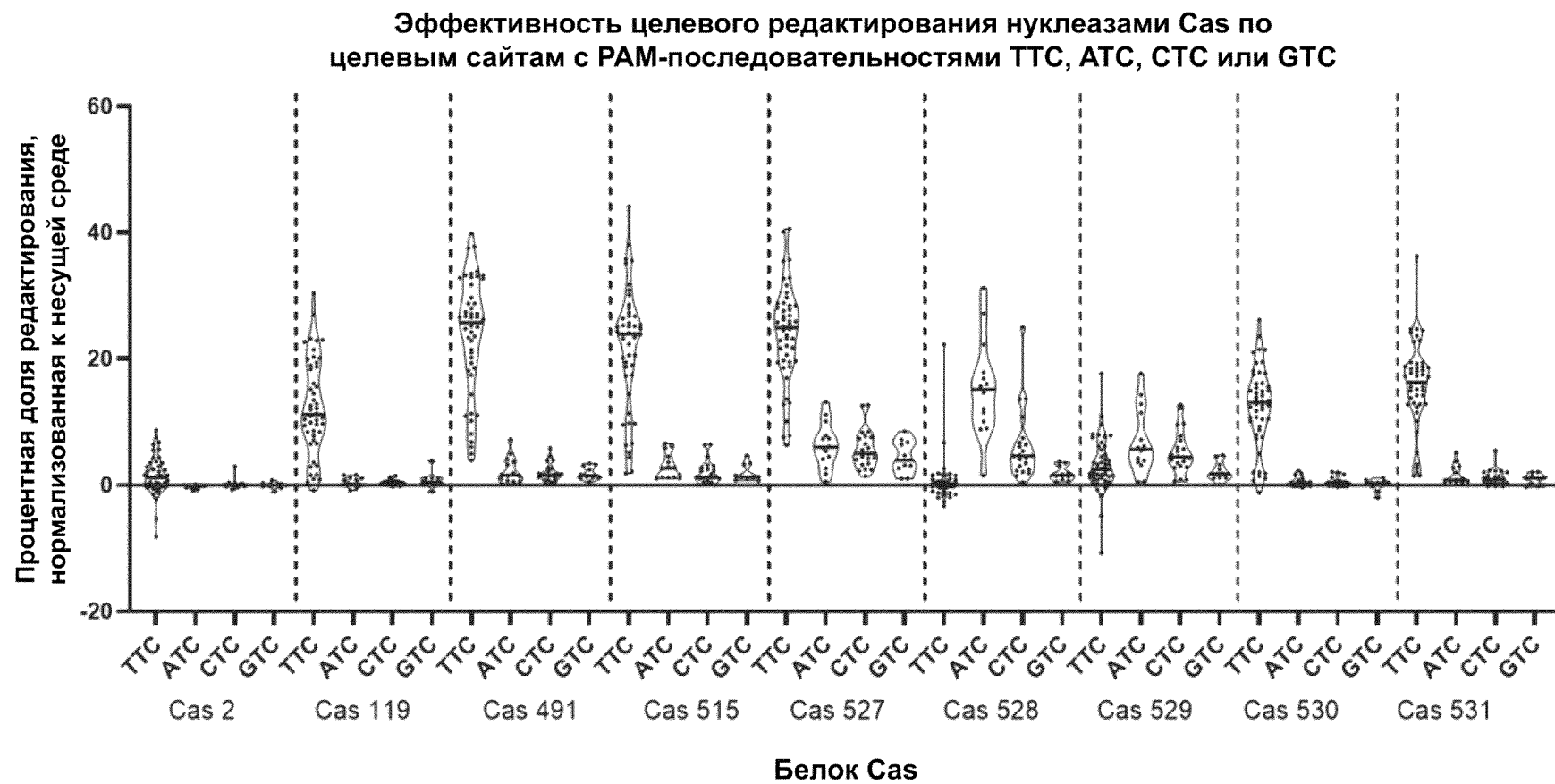
ФИГ. 13



ФИГ. 14

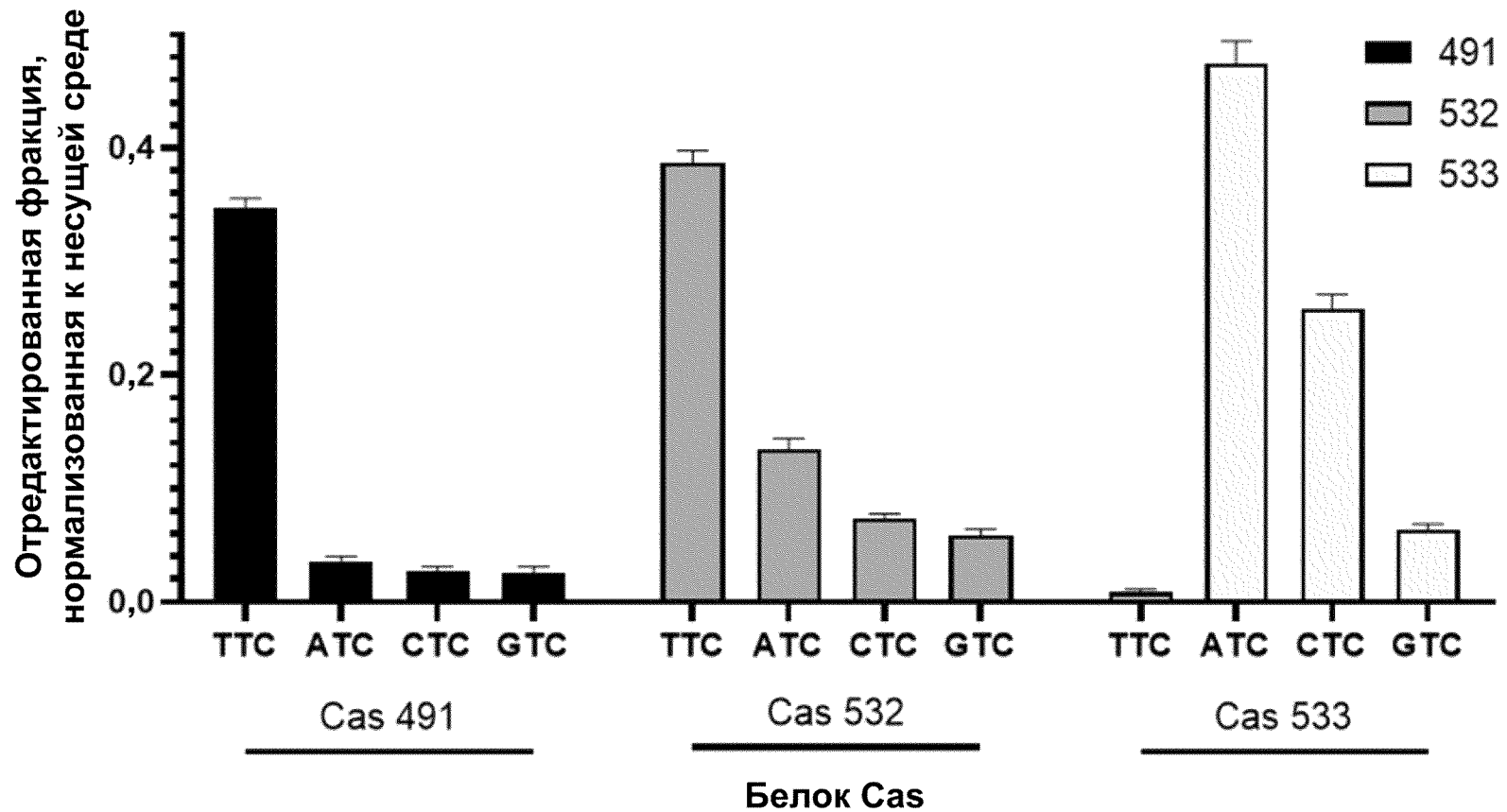


ФИГ. 15



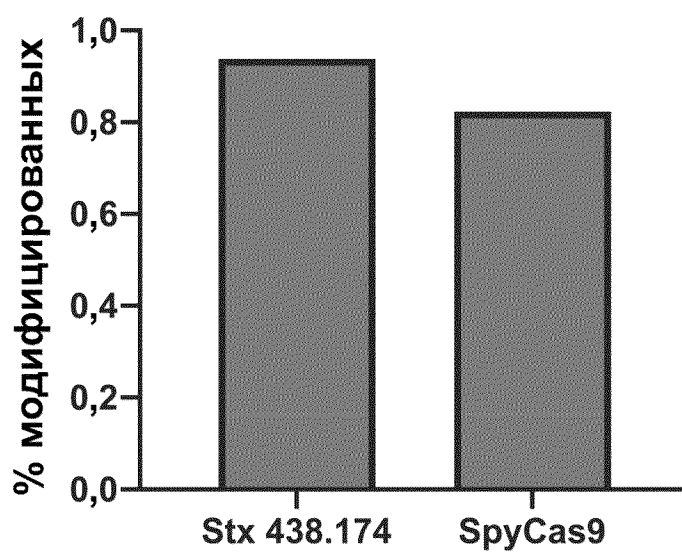
ФИГ. 16

Эффективность целевого редактирования нуклеазами Cas по
целевым сайтам с РАМ-последовательностями TTC, ATC, CTC или GTC



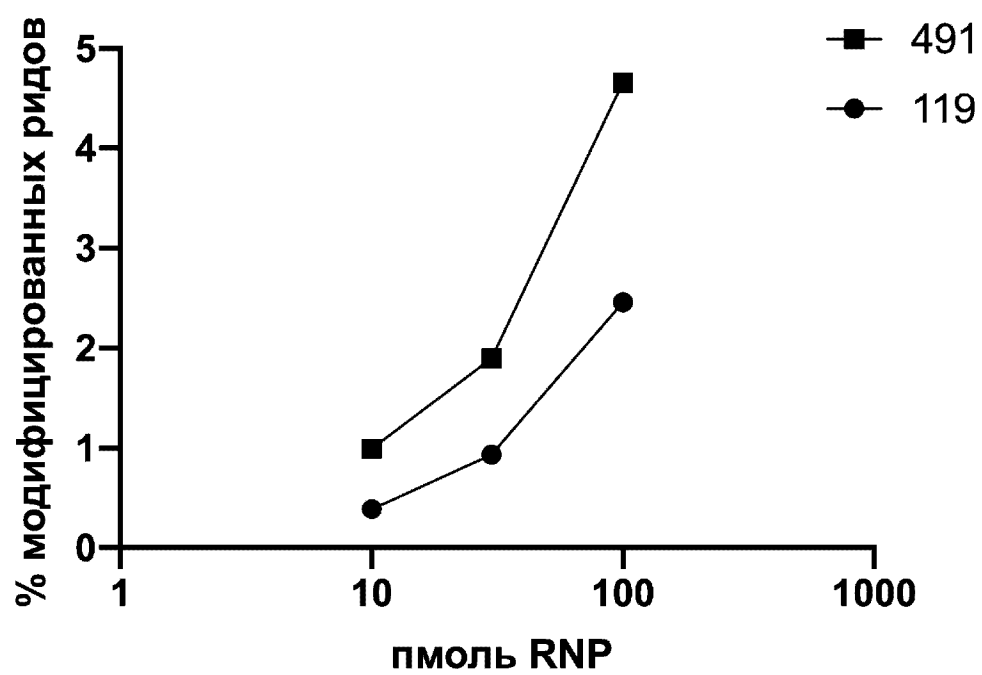
ФИГ. 17

**% модифицированного
BCL11A в HEK293T**

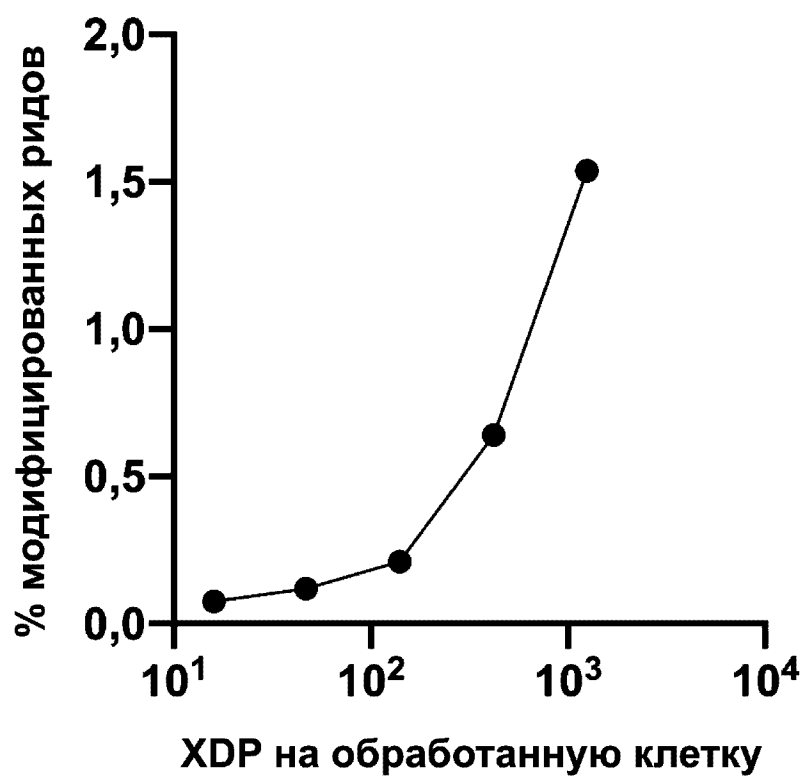


Лечение

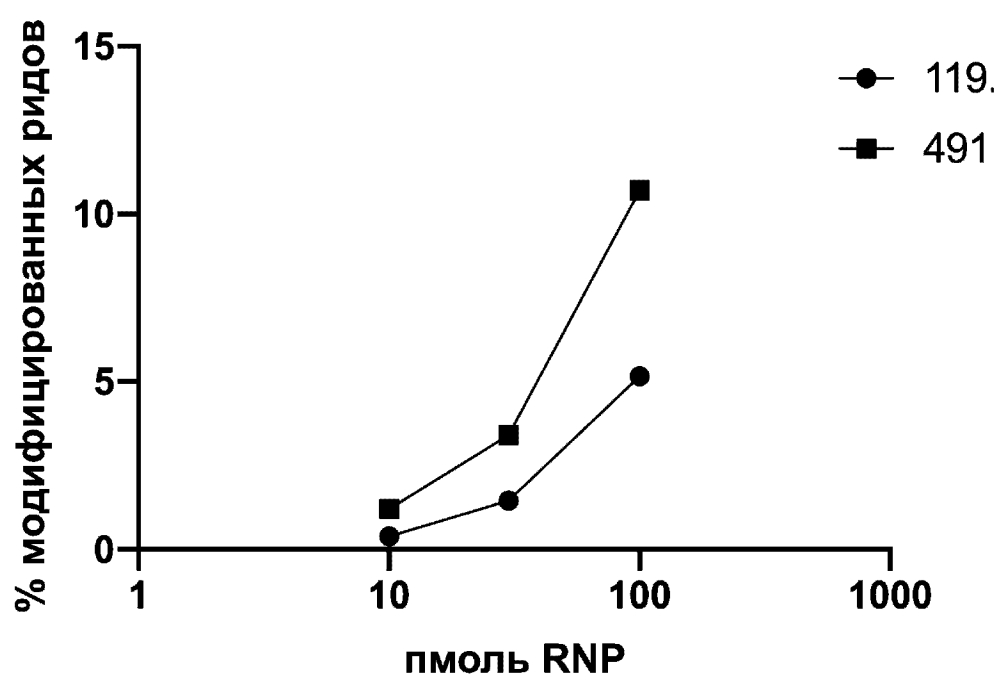
ФИГ. 18



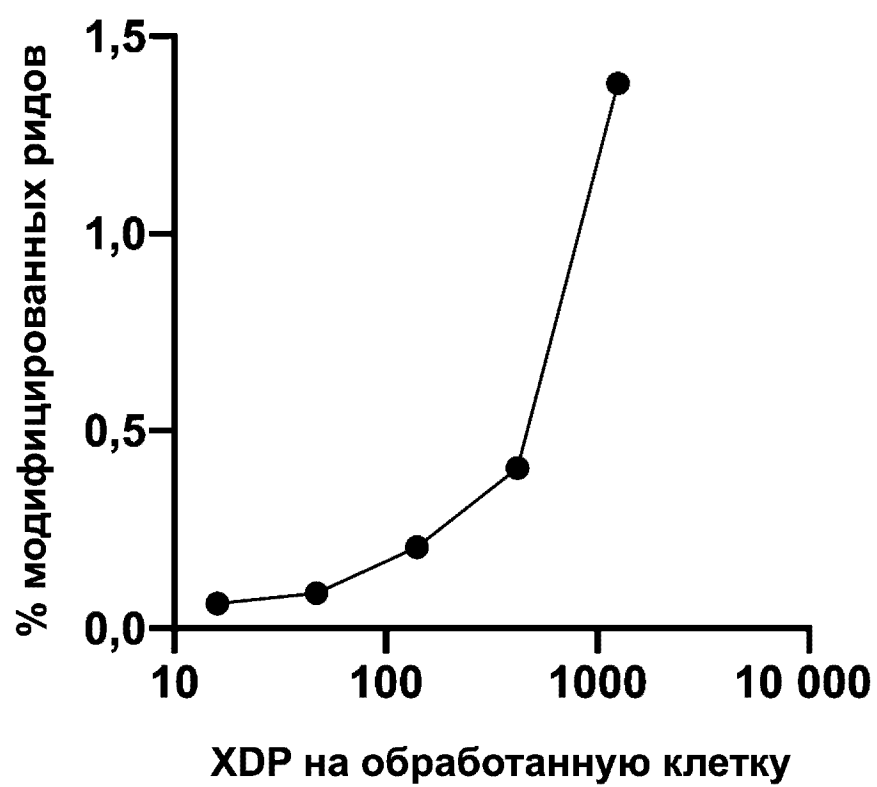
ФИГ. 19



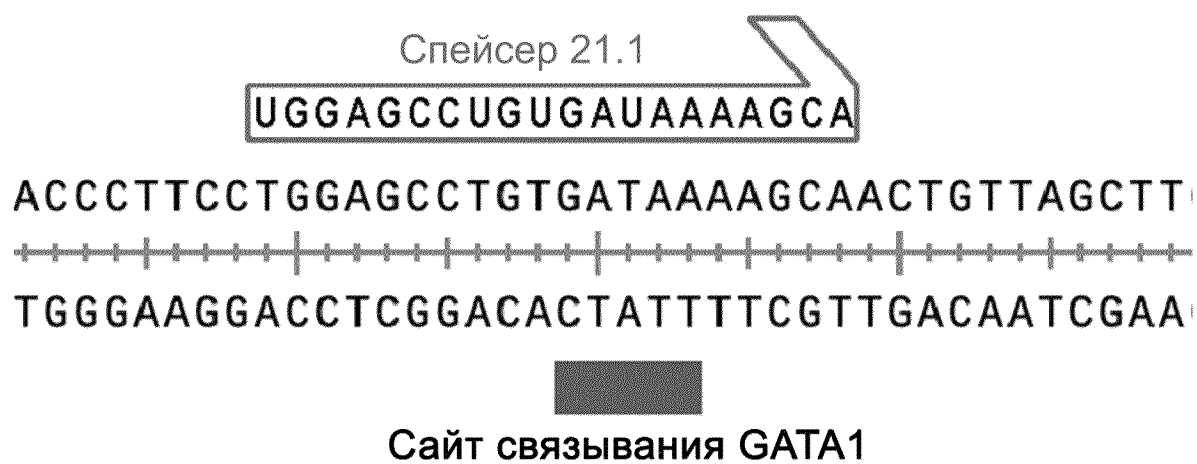
ФИГ. 20



ФИГ. 21



ФИГ. 22



ФИГ. 23