

(19)



**Евразийское
патентное
ведомство**

(21) **202293045** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2023.02.15

(51) Int. Cl. *C12N 9/22* (2006.01)
C12N 15/113 (2010.01)

(22) Дата подачи заявки
2021.04.23

(54) **РНК-НАПРАВЛЯЕМЫЕ НУКЛЕАЗЫ, ИХ АКТИВНЫЕ ФРАГМЕНТЫ, ВАРИАНТЫ И СПОСОБЫ ПРИМЕНЕНИЯ**

(31) 63/014,970; 63/077,211

(32) 2020.04.24; 2020.09.11

(33) US

(86) PCT/US2021/028843

(87) WO 2021/217002 2021.10.28

(71) Заявитель:
ЛАЙФЭДИТ ТЕРАПЬЮТИКС, ИНК.
(US)

(72) Изобретатель:

**Боуэн Тайсон Д., Койл Майкл, Кроули
Александра Брайнер, Элич Тэдд Д.**
(US)

(74) Представитель:

**Веселицкий М.Б., Кузенкова Н.В.,
Каксис Р.А., Белоусов Ю.В., Куликов
А.В., Кузнецова Е.В., Соколов Р.А.,
Кузнецова Т.В.** (RU)

(57) В изобретении описывают составы и способы связывания с представляющей интерес последовательностью-мишенью. Составы находят применение для расщепления или модификации представляющей интерес последовательности-мишени, визуализации представляющей интерес последовательности-мишени и модификации экспрессии представляющей интерес последовательности. Составы содержат полипептиды РНК-направляемой нуклеазы (RGN), РНК CRISPR, трансактивирующие РНК CRISPR, направляющие РНК и молекулы нуклеиновой кислоты, кодирующие их. Также предусмотрены векторы и клетки-хозяева, содержащие молекулы нуклеиновой кислоты. Кроме того, предложены системы RGN для связывания представляющей интерес последовательности-мишени, где система RGN содержит полипептид РНК-направляемой нуклеазы и одну или несколько направляющих РНК.

A1

202293045

202293045

A1

РНК-НАПРАВЛЯЕМЫЕ НУКЛЕАЗЫ, ИХ АКТИВНЫЕ ФРАГМЕНТЫ, ВАРИАНТЫ И СПОСОБЫ ПРИМЕНЕНИЯ

5

Перекрестные сведения на родственные заявки

По настоящей заявке испрашивается приоритет на основе предварительной заявки на патент США US 63/014970, поданной 24 апреля 2020 г., и предварительной заявки на патент США US 63/077211, поданной 15 сентября 10 2021 г., каждая из которых полностью включена в настоящий документ посредством ссылки.

Заявление, касающееся перечня последовательностей

Перечень последовательностей, связанный с этой заявкой, предоставлен в формате ASCII вместо бумажной копии и настоящим включен посредством 15 ссылки в описание. Копия ASCII, обозначенная L103438_1200WO_0084_1_SL и имеющая размер 878362 байта, была создана 21 апреля 2021 г. и направляется через электронную систему подачи документов EFS-Web.

Область техники, к которой относится изобретение

Настоящее изобретение относится к молекулярной биологии и геному 20 редактированию.

Предшествующий уровень техники

В последние годы направленное редактирование или модификация генома становится важным инструментом фундаментальных и прикладных исследований. Первоначальные методы включали сконструированные нуклеазы, 25 такие как мегануклеазы, слитые белки «цинковые пальцы» или TALEN, что требовало создания химерных нуклеаз со сконструированными, программируемыми, специфичными для последовательности ДНК-связывающими доменами, специфичными для каждой конкретной целевой последовательности (последовательности-мишени). РНК-направляемые 30 нуклеазы, такие как ассоциированные с белком кластеризованные регулярно чередующиеся короткие палиндромные повторы бактериальной системы (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR-Cas), позволяют нацеливаться на конкретные последовательности путем создания комплекса нуклеазы с направляющей РНК, которая специфически гибридизуется

с определенной нуклеиновой последовательностью-мишенью. Получение специфичных к мишени направляющих РНК дешевле и эффективнее, чем создание химерных нуклеаз под каждую последовательность-мишень. Такие РНК-направляемые нуклеазы (RNA-guided nucleases – RGN) могут быть
5 использованы для редактирования геномов по выбору путем введения путем введения специфического для последовательности двухцепочечного разрыва, который репарируется с помощью допускающего ошибки метода негомологичного соединения концов (Non-Homologous End-Joining NHEJ) для введения мутации в определенное место генома. В другом варианте
10 гетерологичная ДНК может быть введена в геномный сайт с помощью направленной по гомологии репарации. РНК-направляемые нуклеазы (RGN) также можно использовать для редактирования оснований при слиянии с дезаминазой.

15 Краткое описание изобретения

Представлены составы и способы связывания представляющей интерес целевой последовательности. Составы находят применение в очистке или модификации представляющей интерес целевой последовательности, обнаружении представляющей интерес целевой последовательности и
20 модификации экспрессии представляющей интерес последовательности. Составы включают полипептиды РНК-направляемой нуклеазы (RGN), РНК CRISPR (crRNA), транскриптивирующие РНК CRISPR (tracrRNA), направляющие РНК (gRNA), кодирующие их молекулы нуклеиновой кислоты, а так же векторы и клетки-хозяева, содержащие молекулы нуклеиновой кислоты. Кроме того
25 предусмотрены системы RGN для связывания представляющей интерес последовательности-мишени, причем система RGN содержит полипептид РНК-направляемой нуклеазы и одну или несколько направляющих РНК. Таким образом, способы, представленные в настоящем описании, направлены на связывание представляющей интерес целевой последовательности и, в
30 некоторых вариантах осуществления настоящего изобретения, расщепление или модификацию представляющей интерес целевой последовательности. Представляющая интерес последовательность-мишень может быть модифицирована, например, в результате негомологичного соединения концов,

направленной по гомологии репарации с введенной донорной последовательностью или редактирования оснований.

Подробное описание изобретения

5 Специалист в области охвата настоящего изобретения может представить различные модификации и варианты осуществления настоящего изобретения, исходя из пояснений, представленных в настоящем описании и связанных с ними фигурах. Следует учитывать, что настоящее изобретение не ограничивается конкретными представленными в настоящем описании
10 вариантами его осуществления и что модификации и другие варианты осуществления настоящего изобретения также относятся к области охвата настоящего изобретения. Хотя в описании настоящего изобретения употребляются конкретные термины, они используются только в общем, описательном смысле, а не в целях ограничения.

15

I. Обзор

РНК-направляемые нуклеазы (RNA-guided nucleases – RGN) позволяют целенаправленно манипулировать определенным сайтом (сайтами) в геноме и полезны в контексте нацеливания на гены для терапевтических и
20 исследовательских применений. РНК-направляемые нуклеазы использовались для геномной инженерии в различных организмах, включая млекопитающих, например, стимулируя негомологичное соединение концов и гомологичную рекомбинацию. Составы и способы, представленные в настоящем изобретении, полезны для создания одно- или двухцепочечных разрывов в полинуклеотидах, модификации полинуклеотидов, обнаружения конкретного сайта в
25 полинуклеотиде или модификации экспрессии конкретного гена.

Описанные в настоящем изобретении РНК-направляемые нуклеазы, могут изменять экспрессию генов путем модификации последовательности-мишени. В конкретном варианте осуществления настоящего изобретения РНК-
30 направляемые нуклеазы направляются к последовательности-мишени направляющей РНК (guide RNA - gRNA) как часть системы кластеризованных регулярно чередующихся коротких палиндромных повторов (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR) РНК-направляемых нуклеаз. Нуклеазы RGN считаются «РНК-направляемыми», поскольку направляющие

РНК образуют комплекс с РНК-направляемыми нуклеазами для нацеливания РНК-направляемых нуклеаз для связывания с последовательностью-мишенью и, в некоторых вариантах осуществления настоящего изобретения, внесения одно- или двухцепочечного разрыва в последовательность-мишень. После расщепления последовательности-мишени разрыв может быть репарирован таким образом, что последовательность ДНК целевой последовательности модифицируется в процессе репарации. Таким образом, в настоящем изобретении предусматривают способы использования РНК-направляемых нуклеаз для модификации целевой последовательности в ДНК клеток-хозяев. Например, РНК-направляемые нуклеазы могут быть использованы для модификации последовательности-мишени в геномном локусе эукариотических или прокариотических клеток.

II. РНК-направляемые нуклеазы

В настоящем изобретении описаны РНК-направляемые нуклеазы. Используемый в настоящем изобретении термин «РНК-направляемая нуклеаза» (RGN) относится к полипептиду, который связывается с конкретной нуклеотидной последовательностью-мишенью специфичным для последовательности образом и направляется к целевой нуклеотидной последовательности молекулой направляющей РНК, которая образует комплекс с полипептидом и гибридизуется с последовательностью-мишенью. Хотя РНК-направляемая нуклеаза способна расщеплять последовательность-мишень при связывании, термин «РНК-направляемая нуклеаза» также охватывает мертвые РНК-направляемые нуклеазы, которые способны связываться с последовательностью-мишенью, но не расщеплять ее. Расщепление последовательности-мишени РНК-направляемой нуклеазой может привести к одно- или двухцепочечному разрыву. РНК-направляемые нуклеазы, способные расщеплять только одну нить двухцепочечной молекулы нуклеиновой кислоты, обозначаются в настоящем изобретении термином «никазы».

Описываемые в настоящем изобретении РНК-направляемые нуклеазы включают APG06622, APG02787, APG06248, APG06007, APG02874, APG03850, APG07553, APG03031, APG09208, APG05586, APG08770, APG08167, APG01604, APG03021, APG06015, APG09344, APG07991, APG01868, APG02998, APG09298, APG06251, APG03066, APG01560, APG02777, APG05761, APG02479, APG08385, APG09217 и APG06657 РНК-направляемые нуклеазы, аминокислотные

последовательности которых указаны, соответственно, как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579, и активные фрагменты или их варианты, которые сохраняют способность связываться с целевой нуклеотидной последовательностью специфичным для РНК-направляемой нуклеазы образом. В некоторых из этих вариантов осуществления настоящего изобретения активный фрагмент или вариант РНК-направляемых нуклеаз APG06622, APG02787, APG06248, APG06007, APG02874, APG03850, APG07553, APG03031, APG09208, APG05586, APG08770, APG08167, APG01604, APG03021, APG06015, APG09344, APG07991, APG01868, APG02998, APG09298, APG06251, APG03066, APG01560, APG02777, APG05761, APG02479, APG08385, APG09217 или APG06657 способен расщеплять одно- или двухцепочечную последовательность-мишень. В некоторых вариантах осуществления настоящего изобретения активный вариант РНК-направляемых нуклеаз APG06622, APG02787, APG06248, APG06007, APG02874, APG03850, APG07553, APG03031, APG09208, APG05586, APG08770, APG08167, APG01604, APG03021, APG06015, APG09344, APG07991, APG01868, APG02998, APG09298, APG06251, APG03066, APG01560, APG02777, APG05761, APG02479, APG08385, APG09217 или APG06657 содержит аминокислотную последовательность, которая по меньшей мере на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентична аминокислотной последовательности, указанной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579. В определенных вариантах осуществления настоящего изобретения активный фрагмент РНК-направляемой нуклеазы APG06622, APG02787, APG06248, APG06007, APG02874, APG03850, APG07553, APG03031, APG09208, APG05586, APG08770, APG08167, APG01604, APG03021, APG06015, APG09344, APG07991, APG01868, APG02998, APG09298, APG06251, APG03066, APG01560, APG02777, APG05761, APG02479, APG08385, APG09217 или APG06657 содержит, по меньшей мере, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050 или более смежных аминокислотных остатков аминокислотной последовательности, указанной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579. Представленные в настоящем изобретении РНК-направляемые нуклеазы, могут содержать по меньшей мере один нуклеазный домен (например, ДНКазы,

РНКазы) и по меньшей мере один домен распознавания и/или связывания РНК для взаимодействия с направляющими РНК. Дополнительные домены, которые могут быть обнаружены в РНК-направляемых нуклеазах, представленных в настоящем изобретении, включают, но не ограничиваются ими: ДНК-связывающие домены, геликазные домены, домены, определяющие межбелковые взаимодействия, и домены димеризации. В конкретных вариантах осуществления настоящего изобретения предусматривают РНК-направляемые нуклеазы, последовательности которых могут быть по меньшей мере на 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентичны последовательности одного или нескольких доменов связывания ДНК, геликазного домена, домена, определяющего межбелковые взаимодействия, и домена димеризации.

Целевая нуклеотидная последовательность связывается с предусмотренной в настоящем описании РНК-направляемой нуклеазой, и гибридизуется с направляющей РНК, ассоциированной с РНК-направляемой нуклеазой. Целевая последовательность затем может быть расщеплена РНК-направляемой нуклеазой, если полипептид обладает нуклеазной активностью. Используемые в настоящем изобретении термины «отщеплять» или «расщепление» относятся к гидролизу по меньшей мере одной фосфодиэфирной связи в основной цепи целевой нуклеотидной последовательности, приводящему либо к одноцепочечным, либо к двухцепочечным разрывам в целевой последовательности. Описываемые в настоящем изобретении RGN могут расщеплять нуклеотиды внутри полинуклеотида, функционируя как эндонуклеаза, или могут функционировать как экзонуклеаза, удаляя нуклеотиды с 5' и/или 3'-конца полинуклеотида. В других вариантах осуществления настоящего изобретения описываемые RGN могут отщеплять нуклеотиды последовательности-мишени в любом положении полинуклеотида и, таким образом, функционировать и как эндонуклеаза, так и экзонуклеаза. Расщепление целевого полинуклеотида описываемыми RGN может приводить к ступенчатым разрывам или тупым концам.

Описываемые РНК-направляемые нуклеазы могут представлять собой полипептиды дикого типа, полученные из бактерией или архей. В другом варианте РНК-направляемые нуклеазы могут быть фрагментами полипептидов дикого типа. RGN дикого типа может быть модифицирована, например, для

изменения активности нуклеазы или изменения специфичности РАМ. В некоторых вариантах осуществления настоящего изобретения описываемая РНК-направляемая нуклеаза не встречается в природе.

В определенных вариантах осуществления настоящего изобретения РНК-направляемая нуклеаза действует как нуклеаза, расщепляя только одну нить целевой нуклеотидной последовательности. Такие РНК-направляемые нуклеазы имеют один функциональный нуклеазный домен. В конкретных вариантах осуществления настоящего изобретения нуклеаза способна расщеплять положительную цепь или отрицательную цепь. В некоторых из этих вариантов дополнительные домены нуклеазы были мутированы таким образом, что активность нуклеазы снижена или ликвидирована.

В других вариантах осуществления настоящего изобретения РНК-направляемая нуклеаза полностью лишена нуклеазной активности и упоминается в настоящем описании как мертвая или неактивная нуклеаза. Любой известный в данной области способ введения мутаций в аминокислотную последовательность, такой как ПЦР-опосредованный мутагенез и сайт-направленный мутагенез, может быть использован для генерации нуклеаз или нуклеазо-мертвых RGN. См., например, US 2014/0068797 и US 9790490, сущность каждого патента полностью включена в настоящее описание в виде ссылки.

РНК-направляемые нуклеазы, которые не обладают нуклеазной активностью, могут быть использованы для доставки химерного полипептида, полинуклеотида или малой молекулы с полезной нагрузкой в определенное местоположение генома. В некоторых из этих вариантов осуществления настоящего изобретения полипептид RGN или направляющая РНК могут быть объединены с детектируемой меткой, чтобы обеспечить обнаружение конкретной последовательности. В качестве примера, который не ограничивает рамок охвата настоящего изобретения, мертвая нуклеаза RGN может быть объединена с флуоресцентным белком и нацелена на конкретную последовательность, связанную с заболеванием для обнаружения этой последовательности.

В другом варианте мертвые нуклеазы RGN могут быть нацелены на определенные участки генома для изменения экспрессии желаемой последовательности. В некоторых вариантах осуществления настоящего

изобретения связывание мертвой РНК-направляемой нуклеазы с последовательностью-мишенью приводит к снижению экспрессии последовательности-мишени или гена, у которого транскрипция находится под контролем последовательности-мишени, путем интерференции в отношении связывания РНК-полимеразы или факторов транскрипции в целевой области генома. В других вариантах осуществления настоящего изобретения RGN (например, RGN без нуклеазной активности) или объединенная с нею направляющая РНК дополнительно содержит модулятор экспрессии, который при связывании с последовательностью-мишенью служит для подавления или активации экспрессии последовательности-мишени или гена, у которого транскрипция находится под контролем последовательности-мишени. В некоторых из этих вариантов осуществления настоящего изобретения модулятор экспрессии модулирует экспрессию последовательности-мишени или регулируемого гена посредством эпигенетических механизмов.

В других вариантах осуществления настоящего изобретения мертвые нуклеазы RGNs или RGN, обладающие только нуклеазной активностью, могут быть нацелены на конкретные местоположения в геноме для модификации последовательности целевого полинуклеотида путем слияния с полипептидом, редактирующим основание, например, дезаминазным полипептидом или его активным вариантом или фрагментом, который непосредственно химически модифицирует (например, дезаминирует) нуклеотидную основу, что приводит к превращению одного нуклеотидного основания в другое. Редактирующее основание полипептид может быть слит с RGN на его N-конце или C-конце. Дополнительно, полипептид, редактирующий основание, может быть слит к RGN с помощью пептидного линкера. Примером полипептида дезаминазы, не ограничивающим рамок охвата настоящего изобретения и применимым для таких составов и способов, является цитидиндезаминаза или аденозиндезаминаза (такая как редактор оснований адениндезаминазы, описанный Gaudelli и соавт., *Nature*, 2017, 551:464-471, US 2017/0121693 и US 2018/007301, WO 2018/027078, или любая из дезаминаз, описанных в WO 2020/139873 и US 63/07708, поданы 11 сентября 2020 года, US 63/146840, подан 8 февраля 2021 года, и US 63/164273, подан 22 марта 2021 года; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки). Кроме того, в области техники, к которой относится настоящее изобретение, известно, что некоторые

белки, полученные слиянием RGN и фермента, редактирующего основания, могут также содержать, по меньшей мере, один урацил-стабилизирующий полипептид, который увеличивает степень мутации цитидина, дезоксицитидина или цитозина в тимидин, дезокситимидин или тимин, соответственно, в молекуле нуклеиновой кислоты под действием дезаминазы. Неограничивающие примеры полипептидов, которые стабилизируют урацил, описаны в заявке на патент U.S. № 63/052175, подана 15 июля 2020 года, включая USP2 (SEQ ID NO: 1089), а так же домен ингибитора урацил-гликозилазы (UGI) (SEQ ID NO: 212), который может повысить эффективность редактирования основания.

5

10 Следовательно, слитый белок может содержать описанную в настоящем изобретении RGN или её вариант, дезаминазу и необязательно по меньшей мере один урацил-стабилизирующий полипептид, такой как UGI или USP2. В определенных вариантах осуществления настоящего изобретения RGN, объединенная с полипептидом, редактирующим основания, представляет собой

15 никазу, расщепляющую цепь ДНК, на которую не воздействует полипептид, редактирующий основание (например, дезаминаза).

РНК-направляемые нуклеазы, слитые с полипептидом или доменом, могут быть разделены или соединены при помощи линкера. Используемый в настоящем изобретении термин «линкер», относится к химической группе или молекуле, связывающей две молекулы или фрагмента, например, домен связывания и домен расщепления нуклеазы. В некоторых вариантах осуществления настоящего изобретения линкер соединяет gRNA-связывающий домен РНК-направляемой нуклеазы и полипептид, редактирующий основание, такой как дезаминаза. В некоторых вариантах осуществления настоящего изобретения линкер соединяет мертвую нуклеазу RGN и дезаминазу. Как правило, линкер расположен между двумя соединяемыми группами, молекулами или иными фрагментами, и соединен с каждой из соединяемых двух частей посредством ковалентной связи. В некоторых вариантах осуществления настоящего изобретения линкер представляет собой аминокислоту или несколько аминокислот (например, пептид или белок). В некоторых вариантах осуществления настоящего изобретения линкер представляет собой органическую молекулу, группу, полимер или химический фрагмент. В некоторых вариантах осуществления настоящего изобретения длина линкера составляет 5-100 аминокислот, например, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

20

25

30

17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150 или 150-200 аминокислот в длину. Также предусматривают как более длинные, так и более короткие линкеры.

Рассматриваемые РНК-направляемые нуклеазы могут содержать по
5 меньшей мере один сигнал ядерной локализации (nuclear localization signal - NLS)
) для усиления транспортировки RGN в ядро клетки. Применение сигналов
ядерной локализации известно в области геномного редактирования. Обычно
NLS содержат последовательность основных аминокислот (см., например, Lange
и соавт., *J. Biol. Chem.*, 2007, 282:5101-5105). В конкретных вариантах
10 осуществления настоящего изобретения RGN содержит 2, 3, 4, 5, 6 или более
сигналов ядерной локализации. Сигнал (сигналы) ядерной локализации может
быть гетерологичным NLS. Примерами сигналов ядерной локализации, полезных
для описываемых в настоящем изобретении RGN, являются сигналы ядерной
локализации большого Т-антигена SV40, нуклеоплазмина и с-Мус (см.,
15 например, Ray и соавт., *Bioconjug Chem.*, 2015, 26(6):1004-7). В конкретных
вариантах осуществления настоящего изобретения RGN содержит
последовательность NLS, указанную как SEQ ID NO: 251 или 253. RGN может
содержать одну или более последовательностей NLS на N-конце, на С-конце или
как на N-, так и на С-конце. Например, RGN может содержать две
20 последовательности NLS в N-концевой области и четыре последовательности
NLS в С-концевой области.

Другие известные последовательности сигнала локализации, которые
локализируют полипептиды в определенном субклеточном местоположении
(положениях), также могут быть использованы для нацеливания на RGN,
25 включая, но не ограничиваясь ими, последовательности пластидной
локализации, последовательностями митохондриальной локализации и
сигнальные последовательности двойного нацеливания, которые нацелены как
на пластиды, так и на митохондрии (например, Nassoury, Morse, *Biochim. Biophys
Acta*, 2005, 1743:5-19; Kunze, Berger, *Front. Physiol.*, 2015,
30 [dx.doi.org/10.3389/fphys.2015.00259](https://doi.org/10.3389/fphys.2015.00259); Herrmann, Neupert, *IUBMB Life*, 2003,
55:219-225; Soll, *Curr. Opin. Plant Biol.*, 2002, 5:529-535; Carrie, Small, *Biochim.
Biophys. Acta*, 2013, 1833:253-259; Carrie и соавт., *FEBS J.*, 2009, 276:1187-1195;
Silva-Filho, *Curr. Opin. Plant Biol.*, 2003, 6:589-595; Peeters, Small, *Biochim.
Biophys. Acta*, 2001, 1541:54-63; Murcha и соавт., *J. Exp. Bot.*, 2014, 65:6301-6335;

Mackenzie, *Trends Cell Biol.*, 2005, 15:548-554; Glaser и соавт., *Plant Mol. Biol.*, 1998, 38:311-338).

В определенных вариантах осуществления настоящего изобретения предусмотренные РНК-направляемые нуклеазы содержат, по меньшей мере, один проникающий в клетку домен, который облегчает поглощение клеткой RGN. Проникающие в клетки домены известны в данной области техники, к которой относится настоящее изобретение. Они обычно содержат участки положительно заряженных аминокислотных остатков (т.е. поликатионные проникающие в клетки домены), чередующиеся полярные аминокислотные остатки и неполярные аминокислотные остатки (т.е. амфипатические проникающие в клетки домены) или гидрофобные аминокислотные остатки (т.е. гидрофобные проникающие в клетки домены) (см., например, Milletti F., *Drug Discov Today*, 2012, 17:850-860). Примером проникающего в клетку домена является трансактивирующий активатор транскрипции (ТАТ – trans-activating transcriptional activator) вируса иммунодефицита человека 1.

Сигнал ядерной локализации, сигнал пластидной локализации, сигнал митохондриальной локализации, сигнал локализации с двойным нацеливанием и/или проникающий в клетку домен могут быть расположены на амино-конце (N-конец), карбоксильном конце (С-конец) или ~~во~~ внутри РНК-направляемой нуклеазы.

Описываемые в настоящем изобретении RGN могут быть слиты с эффекторным доменом, таким как домен расщепления, домен дезаминазы или домен модулятора экспрессии, либо прямо, либо косвенно через линкерный пептид. Такой домен может быть расположен на N-конце, С-конце или внутри РНК-направляемой нуклеазы. В некоторых из этих вариантов осуществления настоящего изобретения RGN-компонент слитого белка представляет собой мертвую нуклеазу RGN.

В некоторых вариантах мет приводя к превращению одного нуклеинового основания в другое, и включает, но ими не ограничивается, цитидиндезаминазу или редактор оснований адениндезаминазы (см., например, Gaudelli и соавт., *Nature*, 2017, 551:464-471, US 2017/0121693, US 2018/0073012, US 9840699 и WO/2018/027078).

В некоторых вариантах осуществления настоящего изобретения эффекторный домен химерного белка RGN может быть доменом - модулятором

экспрессии, который представляет собой домен, который служит либо для усиления, либо для снижения регуляции транскрипции. Домен - модулятор экспрессии может быть доменом эпигенетической модификации, доменом-репрессором транскрипции или доменом активации транскрипции.

5 В некоторых из этих вариантов модулятор экспрессии химерного белка RGN содержит домен эпигенетической модификации, который ковалентно модифицирует ДНК или гистоновые белки для изменения структуры гистонов и/или хромосомной структуры без изменения последовательности ДНК, что приводит к изменениям в экспрессии генов (т.е. к повышению или понижению регуляции). Примеры эпигенетических модификаций, не ограничивающие рамок охвата настоящего изобретения, включают ацетилирование или метилирование остатков лизина, метилирование аргинина, фосфорилирование серина и треонина, убиквитинирование лизина и сумоилирование гистоновых белков, а также метилирование и гидроксиметилирование остатков цитозина в ДНК.

10 Примеры доменов эпигенетических модификаций, не ограничивающие рамок охвата настоящего изобретения, включают домены гистонацетилтрансферазы, домены деацетилазы гистонов, домены метилтрансферазы гистонов, домены деметилазы гистонов, домены метилтрансферазы ДНК и домены деметилазы ДНК.

20 В других вариантах осуществления настоящего изобретения модулятор экспрессии слитого белка содержит домен репрессора транскрипции, который взаимодействует с элементами контроля транскрипции и/или белками, регулирующими транскрипцию, такими как РНК-полимеразы и факторы транскрипции, для уменьшения или прекращения транскрипции по меньшей мере одного гена. Домены-репрессоры транскрипции известны в области техники, к которой относится настоящее изобретение, и включают, но ими не ограничиваются, Sp1-подобные репрессоры, IκB, и домены KRAB (Körperl Associated Box - KRAB).

25

30 В некоторых вариантах осуществления настоящего изобретения модулятор экспрессии слитого белка содержит домен активации транскрипции, взаимодействующий с элементами контроля транскрипции и/или белками, регулирующими транскрипцию, такими как РНК-полимеразы и транскрипционные факторы, для активации или увеличения транскрипции по меньшей мере одного гена. Известные в данной области техники домены

активации транскрипции включают, но ими не ограничиваются, домен активации вируса простого герпеса VP16 и домен активации NFAT.

Представленные в настоящей работе полипептиды RGN могут содержать выявляемую метку или метку, позволяющую проводить очистку. Детектируемая метка или метка, позволяющая проводить очистку, может быть расположена на N-конце, C-конце или внутри РНК-направляемой нуклеазы, как напрямую, так и через линкерный пептид. В некоторых из этих вариантов осуществления настоящего изобретения RGN компонент слитого белка представляет собой мертвую RGN нуклеазу. В некоторых вариантах осуществления настоящего изобретения RGN компонент слитого белка представляет собой RGN с активностью никазы.

Детектируемая метка – это молекула, которую можно обнаружить визуально или иным образом. Детектируемая метка может быть соединена с RGN как слитый белок (например, флуоресцентный белок) или может представлять собой небольшую конъюгированную с полипептидом RGN молекулу, которая может быть обнаружена визуально или другими способами. Детектируемые метки, которые соединяются с описанными в настоящем изобретении RGN, как присоединенные белки, представляют собой любой детектируемый белковый домен, включая, но ими не ограничиваясь, флуоресцентный белок или белковый домен, который может быть обнаружен с помощью специфического антитела. Примеры флуоресцентных белков включают, но ими не ограничиваются, зеленые флуоресцентные белки (например, GFP, EGFP, ZsGreen1) и желтые флуоресцентные белки (например, YFP, EYFP, ZsYellow1). Примеры малых молекул, используемых в качестве детектируемых меток, включают, но ими не ограничиваются, радиоактивные метки, такие как ^3H и ^{35}S .

Полипептиды RGN могут также содержать метку, позволяющую проводить очистку и представляющую собой любую молекулу, которая может быть использована для выделения белка из смеси (например, биологического образца или культуральной среды). Примеры меток, позволяющую проводить очистку, включают, но ими не ограничиваются, биотин, тус-метку, мальтозосвязывающий белок (MBP), глутатион-S-трансферазу (GST) и 3X FLAG метку.

II. Направляющая РНК

В настоящем изобретении описаны направляющие РНК и полинуклеотиды, кодирующие их. Используемый в настоящем изобретении термин «направляющая РНК» относится к нуклеотидной последовательности, имеющей достаточную степень комплементарности с целевой нуклеотидной последовательностью для гибридизации с ней и прямого специфического связывания с ней ассоциированной РНК-направляемой нуклеазы. Таким образом, соответствующая направляющая RGN РНК представляет собой одну или более молекул РНК (обычно, одну или две), которые могут связываться с RGN и нацеливать RGN на связывания с конкретной нуклеотидной последовательностью-мишенью, а в тех вариантах осуществления настоящего изобретения, в которых RGN обладает никазой или нуклеазной активностью, еще и расщеплять последовательность-мишень. В общем, направляющая РНК включает CRISPR РНК (crRNA) и транскрибируемую CRISPR РНК (tracrRNA). Нативные направляющие РНК, которые содержат как crRNA, так и tracrRNA, обычно содержат две отдельные молекулы РНК, которые гибридизуются друг с другом через повторяющуюся последовательность crRNA и анти-повторяющуюся последовательность tracrRNA.

Нативные прямые повторяющиеся последовательности в массиве CRISPR обычно имеют длину от 28 до 37 пар оснований, хотя длина может варьировать от примерно 23 п.н. до примерно 55 п.н. Спейсерные последовательности в массиве CRISPR обычно имеют длину от примерно 32 до примерно 38 п.н., хотя длина может составлять и от примерно 21 п.н. до примерно 72 п.н. Каждый массив CRISPR обычно содержит менее 50 единиц последовательности CRISPR-повтор-спейсер. CRISPR транскрибируются как часть длинного транскрипта, называемого первичным транскриптом CRISPR, который включает в себя большую часть массива CRISPR. Первичный транскрипт CRISPR расщепляется белками Cas с образованием crRNA или, в некоторых случаях, с образованием предшественников, пре-crRNA, которые расщепляются дополнительными белками Cas с образованием зрелых crRNA. Зрелые crRNA содержат спейсерную последовательность и повторы CRISPR. В тех вариантах осуществления настоящего изобретения, в которых пре-crRNA переводятся в зрелые (или обработанные) crRNA, созревание включает удаление примерно от одного до шести или более нуклеотидов с 5', 3' или с 5' и 3' конца. Для целей

редактирования генома или нацеливания на конкретную нуклеотидную представляющую интерес последовательность-мишень, нуклеотиды, удаляемые во время созревания молекулы пре-crRNA, не являются необходимыми для генерации или конструирования направляющей РНК.

5 CRISPR RNA (crRNA) содержит последовательность спейсера и последовательность повторов CRISPR. «Спейсерная последовательность» является нуклеотидной последовательностью, которая непосредственно гибридируется с представляющей интерес целевой нуклеотидной последовательностью-мишенью. Спейсерная последовательность
10 конструируется таким образом, чтобы быть полностью или частично комплементарной последовательности-мишени. В различных вариантах осуществления настоящего изобретения последовательность спейсера может содержать от 8 нуклеотидов до 30 нуклеотидов или более. Например, последовательность спейсера может составлять примерно 8, примерно 9,
15 примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29, примерно 30 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения,
20 последовательность спейсера может составлять 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения последовательность спейсера имеет от 10 до 26 нуклеотидов в длину или от 12 до 30 нуклеотидов в длину. В конкретных вариантах осуществления настоящего
25 изобретения длина последовательности спейсера составляет примерно 30 нуклеотидов. В некоторых вариантах осуществления настоящего изобретения степень комплементарности между последовательностью спейсера и соответствующей ей последовательностью-мишенью при оптимальном выравнивании с использованием подходящего алгоритма выравнивания
30 составляет от 50% до 99% или более, включая, но ими не ограничиваясь, примерно 50%, примерно 60%, примерно 70%, примерно 75%, примерно 80%, примерно 81%, примерно 82%, примерно 83%, примерно 84%, примерно 85%, примерно 86%, примерно 87%, примерно 88%, примерно 89%, примерно 90%, примерно 91%, примерно 92%, примерно 93%, примерно 94%, примерно 95%,

примерно 96%, примерно 97%, примерно 98%, примерно 99% или более. В некоторых вариантах осуществления настоящего изобретения степень комплементарности между последовательностью спейсера и соответствующей ей целевой последовательностью при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более. В некоторых вариантах осуществления настоящего изобретения последовательность спейсера не формирует вторичную структуру, которая может быть предсказана с использованием любого подходящего алгоритма укладки полинуклеотидов, известного в области техники, к которой относится настоящее изобретение, включая, но им не ограничиваясь, mFold (см., например, Zuker, Stiegler, *Nucleic Acids Res.*, 1981, 9:133-148) и RNAfold (см., например, Gruber и соавт., *Cell.*, 2008, 106(1):23-24).

Последовательность повторов CRISPR РНК содержит нуклеотидную последовательность, которая формирует структуру, либо сама по себе, либо совместно с гибридизированной tracrRNA, распознаваемой молекулой RGN. В некоторых вариантах осуществления настоящего изобретения последовательность повторов CRISPR РНК может содержать от примерно 8 нуклеотидов до примерно 30 нуклеотидов или более. Например, последовательность повторов CRISPR может составлять примерно 8, примерно 9, примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29, примерно 30 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения последовательность повторов CRISPR представляет собой 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения степень комплементарности между последовательностью повтора CRISPR и соответствующей ей последовательностью tracrRNA при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет примерно или более 50%, примерно 60%, примерно 70%, примерно 75%, примерно 80%, примерно 81%, примерно 82%, примерно 83%, примерно

84%, примерно 85%, примерно 86%, примерно 87%, примерно 88%, примерно 89%, примерно 90%, примерно 91%, примерно 92%, примерно 93%, примерно 94%, примерно 95%, примерно 96%, примерно 97%, примерно 98%, примерно 99% или более%. В некоторых вариантах осуществления настоящего

5 изобретения степень комплементарности между последовательностью повторов CRISPR и соответствующей ей последовательностью tracrRNA при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более.

10 В конкретных вариантах осуществления настоящего изобретения, последовательность повторов CRISPR содержит нуклеотидную последовательность, указанную как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124, или активный вариант, или его фрагмент, которые, будучи включены в состав направляющей РНК, способны
15 направить специфичное для последовательности связывание РНК-направляемой нуклеазы, представленной в настоящем изобретении, к последовательности-мишени. В некоторых вариантах осуществления настоящего изобретения активный вариант последовательности повтора CRISPR представляет собой последовательность дикого типа и содержит последовательность нуклеотидов,
20 которые, по меньшей мере, на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентичны нуклеотидной последовательности, указанной как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124. В некоторых вариантах осуществления настоящего изобретения фрагмент активной последовательности
25 повтора CRISPR представляет собой последовательность дикого типа и содержит, по меньшей мере, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 или 22 смежных нуклеотида нуклеотидной последовательности, представленной как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124.

30 В некоторых вариантах осуществления настоящего изобретения crRNA не встречается в природе. В некоторых вариантах осуществления настоящего изобретения конкретная последовательность повторов CRISPR не соединена в природе со сконструированной последовательностью спейсеров, в таком случае последовательность повторов CRISPR считается гетерологичной в отношении

последовательности спейсеров. В некоторых вариантах осуществления настоящего изобретения последовательность спейсеров представляет собой сконструированную последовательность, не встречающуюся в природе.

Транс-активирующая CRISPR РНК (tracrRNA) содержит нуклеотидную последовательность, содержащую область, обладающую достаточной комплементарностью для гибридизации с последовательностью повтора CRISPR (crRNA), которая упоминается в настоящем изобретении как область анти-повтора. В некоторых вариантах осуществления настоящего изобретения молекула tracrRNA дополнительно содержит область с вторичной структурой (например, шпильку) или образует вторичную структуру при гибридизации с соответствующей ей crRNA. В некоторых вариантах осуществления настоящего изобретения область tracrRNA, которая полностью или частично комплементарна последовательности повторов CRISPR, находится на 5'-конце молекулы, а 3'-конец tracrRNA содержит вторичную структуру. Эта область вторичной структуры обычно включает в себя несколько структур типа шпильки, включая связующую шпильку, которая находится рядом с последовательностью анти-повтора. Связующая шпилька образует ядро взаимодействия между направляющей РНК и RGN и находится на пересечении направляющей РНК, RGN и ДНК-мишени. Связующая шпилька часто имеет консервативную нуклеотидную последовательность в основании стержня шпильки с мотивом UNANNC (SEQ ID NO: 132), который обнаружен во многих шпильках в tracrRNA. Интересно, что некоторые из RGN по настоящему изобретению используют tracrRNA, которые содержат неканонические последовательности в основании стержня связующих шпилек, включая UNANNA, UNANNU, UNANNG и CNANNC (указаны, как SEQ ID NO: 129, 130, 131 и 133, соответственно). На 3'-конце tracrRNA часто находятся терминаторные шпильки, различающиеся по структуре и количеству, однако обычно это GC-обогащенная Rho-независимая транскрипционная терминаторная шпильку, за которой следует U-цепочка на 3'-конце. См., например, Briner и соавт. *Molecular Cell.*, 2014, 56:333-339; Briner, Barrangou *Cold Spring. Harb. Protoc.*, 2016, doi: 10.1101/pdb.top090902, US 2017/02756488; сущность каждой приведенной публикации и патента полностью включена в настоящее изобретение в виде ссылки.

В различных вариантах осуществления настоящего изобретения область анти-повтора tracrRNA, которая полностью или частично комплементарна последовательности повтора CRISPR, содержит от примерно 8 нуклеотидов до примерно 30 нуклеотидов или более. Например, область образования пар между основаниями последовательности анти-повтора tracrRNA и последовательности повтора CRISPR может составлять примерно 8, примерно 9, примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29, примерно 30 или более нуклеотидов в длину. В конкретных вариантах осуществления настоящего изобретения область образования пар между основаниями последовательности анти-повтора tracrRNA и последовательности повтора CRISPR составляет 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения степень комплементарности между последовательностью повтора CRISPR и соответствующей ей последовательностью анти-повтора tracrRNA при оптимальном выравнивании с использованием подходящего алгоритма выравнивания составляет примерно или более чем примерно 50%, примерно 60%, примерно 70%, примерно 75%, примерно 80%, примерно 81%, примерно 82%, примерно 83%, примерно 84%, примерно 85%, примерно 86%, примерно 87%, примерно 88%, примерно 89%, примерно 90%, примерно 91%, примерно 92%, примерно 93%, примерно 94%, примерно 95%, примерно 96%, примерно 97%, примерно 98%, примерно 99% или более. В некоторых вариантах осуществления настоящего изобретения степень комплементарности между последовательностью повтора CRISPR и соответствующей ей последовательностью анти-повтора tracrRNA при оптимальном выравнивании с использованием соответствующего алгоритма выравнивания составляет 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более.

В некоторых вариантах осуществления настоящего изобретения tracrRNA полностью может содержать в длину от примерно 60 нуклеотидов до более чем примерно 210 нуклеотидов. Например, tracrRNA может составлять примерно 60, примерно 65, примерно 70, примерно 75, примерно 80, примерно 85, примерно

90, примерно 95, примерно 100, примерно 105, примерно 110, примерно 115, примерно 120, примерно 125, примерно 130, примерно 135, примерно 140, примерно 150, примерно 160, примерно 170, примерно 180, примерно 190, примерно 200, примерно 210 или более нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения tracrRNA состоит из 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 190, 200, 210 или более нуклеотидов в длину. В конкретных вариантах осуществления настоящего изобретения tracrRNA имеет длину от примерно 80 до примерно 90 нуклеотидов, включая примерно 80, примерно 81, примерно 82, примерно 83, примерно 84, примерно 85, примерно 86, примерно 87, примерно 88, примерно 89 и примерно 90 нуклеотидов в длину. В некоторых вариантах осуществления настоящего изобретения tracrRNA имеет длину от 80 до 90 нуклеотидов, включая 80, 81, 82, 83, 84, 85, 86, 87, 88, 89 и 90 нуклеотидов в длину.

В некоторых вариантах осуществления настоящего изобретения tracrRNA содержит нуклеотидную последовательность, указанную как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125, или же ее активный вариант или фрагмент, который, будучи включен в направляющую РНК, способен направлять специфичное для последовательности связывание РНК-направляемой нуклеазы, описываемой в настоящем изобретении, к последовательности-мишени. В некоторых вариантах осуществления настоящего изобретения активный вариант последовательности tracrRNA дикого типа содержит нуклеотидную последовательность, которая по меньшей мере на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентична нуклеотидной последовательности, указанной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125. В некоторых вариантах осуществления настоящего изобретения фрагмент активной последовательности tracrRNA дикого типа содержит, по меньшей мере 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 или более смежных нуклеотидов к нуклеотидной последовательности, указанной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125.

Две полинуклеотидные последовательности можно считать по существу комплементарными, когда они гибридизируются друг с другом в жестких

условиях. Аналогично, считается, что RGN связывается с конкретной последовательностью-мишенью специфичным для последовательности образом, если направляющая РНК, связанная с RGN, связывается с последовательностью-мишенью в жестких условиях. Под «жесткими условиями» или «жесткими условиями гибридизации» подразумеваются условия, при которых две полинуклеотидные последовательности будут гибридизироваться друг с другом в значительно большей степени, чем с другими последовательностями (например, по меньшей мере в 2 раза по сравнению с исходным состоянием). Жесткие условия проведения реакции зависят от последовательности и будут различными в разных условиях. Как правило, жесткими условиями проведения реакции будут те, при которых концентрация соли составляет примерно менее 1,5 М ионов Na, обычно концентрация ионов Na (или другой соли) составляет от 0,01 до 1,0 М при pH от 7,0 до 8,3, а температура составляет по меньшей мере примерно 30°C для коротких последовательностей (например, от 10 до 50 нуклеотидов) и по меньшей мере примерно 60°C для длинных последовательностей (например, более 50 нуклеотидов). Жесткие условия проведения реакции также могут быть созданы за счет добавления дестабилизирующих агентов, таких как формамид. Примеры менее жестких условий проведения реакции включают гибридизацию в буферном растворе 30-35% формамида, 1 М NaCl, 1% SDS (додецилсульфат натрия) при 37°C и промывку в 1X - 2X SSC (20X SSC = 3,0 М NaCl/0,3 М тринатриевая соль цитрата) при 50-55°C. Пример условий проведения реакции средней жесткости включают гибридизацию в 40-45% формамиде, 1,0 М NaCl, 1% SDS при 37°C и промывку в 0,5X - 1X SSC при 55-60°C. Примерные очень жесткие условия проведения реакции включают гибридизацию в 50% формамиде, 1 М NaCl, 1% SDS при 37°C и промывку в 0,1X SSC при 60-65°C. При необходимости промывочные буферы могут содержать от 0,1% до 1% SDS. Продолжительность гибридизации обычно составляет примерно менее 24 ч, обычно от 4 до 12 ч. Продолжительность промывки обычно занимает время, по меньшей мере, достаточное для достижения равновесия.

T_m – это температура (при определенной ионной силе и pH), при которой 50% комплементарной целевой последовательности гибридизируется с идеально спариваемой последовательностью. Для гибридизации ДНК-ДНК T_m может быть примерно определена из уравнения Meinkoth и Wahl, *Anal. Biochem.*, 2984,

138:267-284: $T_m = 81,5^{\circ}\text{C} + 16,6(\log M) + 0,41(\%GC) - 0,61(\% \text{ form}) - 500/L$; где M - молярность одновалентных катионов, % GC - процентное содержание гуанозиновых и цитозиновых нуклеотидов в ДНК, % form - процентное содержание формамида в растворе для гибридизации, а L - длина гибридного фрагмента ДНК-ДНК в парах оснований. Как правило, для жестких условий проведения реакции температуру выбирают примерно на 5°C ниже термической точки плавления (T_m) для конкретной последовательности и комплементарной ей цепи при установленных ионной силе и pH. Однако при очень жестких условиях можно проводить гибридизацию и/или промывку при температуре на 1, 2, 3 или 4°C ниже термической точки плавления (T_m); при средних жестких условиях можно проводить гибридизацию и/или промывку при температуре на 6, 7, 8, 9 или 10°C ниже, чем термическая точка плавления (T_m); не очень жесткие условия предполагают проведение реакции гибридизации и/или промывки при 11, 12, 13, 14, 15 или на 20°C ниже термической точки плавления (T_m). Применяя уравнение, условия проведения гибридизации и промывки, а также нужную T_m , специалисты понимают, что вариации жесткости растворов для условий проведения реакции гибридизации и/или промывки по своей сути достаточно описаны в настоящем изобретении. Подробное руководство по гибридизации нуклеиновых кислот содержится в кн.: Tijssen «Laboratory Techniques in Biochemistry, Molecular Biology—Hybridization with Nucleic Acid Probes», часть I, глава 2, изд-во Elsevier, Нью-Йорк, 1993; в кн.: «Current Protocols in Molecular Biology», 1995, под ред. Ausubel и соавт., глава 2, изд-во Greene Publishing, Wiley-Interscience, Нью-Йорк. См. кн.: Sambrook и соавт., «Molecular Cloning: A Laboratory Manual», 1989, 2ое изд., изд-во Cold Spring Harbor Laboratory Press, Plainview, Нью-Йорк.

Используемый в настоящем изобретении термин «специфичная последовательность» может также быть отнесен к связыванию с целевой последовательностью-мишенью с частотой большей, чем связывание с рандомизированной фоновой последовательностью.

Направляющая РНК может быть отдельной направляющей РНК или двунаправленной направляющей РНК. Отдельная направляющая РНК содержит crRNA и tracrRNA на одной молекуле РНК, тогда как двунаправленная направляющая РНК содержит crRNA и tracrRNA, на двух различных молекулах РНК, гибридизированных друг с другом, по меньшей мере через часть

последовательности повтора CRISPR, находящегося на crRNA, и по меньшей мере часть tracrRNA, которая может быть полностью или частично комплементарна последовательности повтора CRISPR, находящегося на crRNA. В некоторых вариантах осуществления настоящего изобретения, в которых направляющая РНК представляет собой отдельную направляющую РНК, crRNA и tracrRNA разделены линкерной нуклеотидной последовательностью. В целом, линкерная нуклеотидная последовательность представляет собой последовательность, которая не содержит комплементарные основания, чтобы избежать образования вторичной структуры внутри или включения нуклеотидов линкерной нуклеотидной последовательности. В некоторых вариантах осуществления настоящего изобретения линкерная нуклеотидная последовательность между crRNA и tracrRNA составляет 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 11, по меньшей мере 12 или более нуклеотидов в длину. В конкретных вариантах осуществления настоящего изобретения линкерная нуклеотидная последовательность отдельной направляющей РНК имеет длину по меньшей мере 4 нуклеотида. В определенных вариантах осуществления настоящего изобретения линкерная нуклеотидная последовательность представляет собой нуклеотидную последовательность, указанную как SEQ ID NO: 249.

Отдельная направляющая РНК, как и двунаправленная направляющая РНК, могут быть синтезированы химическим путем или методом транскрипции *in vitro*. Тесты для определения специфичного по последовательности связывания между RGN и направляющей РНК известны в области техники, к которой относится настоящее изобретение, и включает, но им не ограничиваются, метод связывания *in vitro* между экспрессируемой RGN и направляющей РНК, содержащими выявляемую метку (например, биотин), которая может быть использована для детекции в методе соосаждения, когда комплекс направляющая РНК:RGN захватывается при помощи детектируемой метки (например, гранулами стрептавидина). Контрольная направляющая РНК с неродственной направляющей РНК последовательностью или структурой может быть использована в качестве отрицательного контроля для неспецифического связывания RGN с РНК. В определенных вариантах осуществления настоящего изобретения направляющая РНК указана как SEQ ID NO: 4, 11, 18, 25, 32, 39, 46,

53, 59, 66, 73, 79, 86, 92, 99, 106, 113, 120 или 126, где разделительная последовательность может быть любой последовательностью и обозначается как poly-N последовательность.

5 В определенных вариантах осуществления настоящего изобретения направляющая РНК может быть введена в клетку-мишень, органеллу или эмбрион в виде молекулы РНК. Направляющая РНК может быть транскрибирована *in vitro* или химически синтезирована. В других вариантах осуществления настоящего изобретения нуклеотидная последовательность, кодирующая направляющую РНК, вводится в клетку, органеллу или эмбрион. В 10 некоторых из этих вариантов осуществления настоящего изобретения нуклеотидная последовательность, кодирующая направляющую РНК, функционально связана с промотором (например, промотором РНК-полимеразы III). Промотор может быть нативным или гетерологичным по отношению к нуклеотидной последовательности, кодирующей направляющую РНК.

15 В некоторых вариантах осуществления настоящего изобретения направляющая РНК может быть введена в клетку-мишень, органеллу или эмбрион в виде рибонуклеопротеинового комплекса, как описано в настоящем изобретении, в котором направляющая РНК связана с полипептидом РНК-направляемой нуклеазой.

20 Направляющая РНК направляет ассоциированную РНК-направляемую нуклеазу к конкретной нуклеотидной последовательности-мишени, представляющей интерес, посредством гибридизации направляющей РНК с нуклеотидной последовательностью-мишенью. Нуклеотидная последовательность-мишень может содержать ДНК, РНК или их комбинацию и 25 может быть одноцепочечной или двухцепочечной. Нуклеотидная последовательность-мишень может быть геномной ДНК (т.е. хромосомной ДНК), плазмидной ДНК или молекулой РНК (например, мессенджер-РНК, рибосомной РНК, транспортной РНК, микро-РНК, малой интерферирующей РНК). Целевая нуклеотидная последовательность может быть связана (и в 30 некоторых вариантах осуществления настоящего изобретения расщеплена) РНК-направляемой нуклеазой *in vitro* или в клетке. Хромосомная последовательность, на которую нацелена RGN, может быть ядерной, пластидной или митохондриальной хромосомной последовательностью. В некоторых вариантах

осуществления настоящего изобретения нуклеотидная последовательность-мишень является уникальной в геноме-мишени.

Нуклеотидная последовательность-мишень примыкает к мотиву, прилегающему к протоспейсеру (Protospacer Adjacent Motif - PAM).

5 Протоспейсер обычно представляет собой последовательность примерно от 1 до 10 нуклеотидов, включающую примерно 1, примерно 2, примерно 4, примерно 5, примерно 7, примерно 9 или примерно 10 нуклеотидов из нуклеотидной последовательности-мишени. В конкретных вариантах осуществления настоящего изобретения PAM находится в пределах от 1 до 10 нуклеотидов
10 целевой нуклеотидной последовательности и содержит 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 нуклеотидов в длину. PAM может быть на 5' или 3' конце последовательности-мишени. В конкретных вариантах осуществления настоящего изобретения PAM находится на 3' конце последовательности-мишени для представленных в настоящем изобретении RGN. Как правило, PAM представляет собой
15 консенсусную последовательность примерно из 3-4 нуклеотидов, но в конкретных вариантах длина может быть из 2, 3, 4, 5, 6, 7, 8, 9 или более нуклеотидов. В различных вариантах осуществления настоящего изобретения последовательность PAM, узнаваемая представленной в настоящем изобретении RGN, содержит консенсусную последовательность, указанную как SEQ ID NO:
20 7, 14, 21, 28, 35, 42, 49, 62, 69, 79, 82, 95, 102, 109 или 116.

В конкретных вариантах осуществления настоящего изобретения РНК-направляемая нуклеаза, содержащая последовательность, указанную как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, или ее активный вариант, или фрагмент связывает нуклеотидную
25 последовательность-мишень, примыкающую к последовательности PAM, указанной как SEQ ID NO: 7, 14, 21, 28, 35, 42, 49, 62, 69, 79, 82, 95, 102, 109 или 116. В некоторых из этих вариантов осуществления настоящего изобретения RGN связывается с направляющей последовательностью, содержащей повторяющуюся последовательность CRISPR, указанную как SEQ ID NO: 2, 9,
30 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124, соответственно, или же с ее активным вариантом или фрагментом и последовательностью tracrRNA, указанной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125, соответственно, или ее

активным вариантом или фрагментом. Представленные системы RGN описаны ниже в настоящем изобретении в примерах 1-3 и табл. 1 и 2.

Были получены варианты RGN APG05586 (SEQ ID NO: 63), которые имеют аминокислотные последовательности, указанные как SEQ ID NO: 570-579. RGN, содержащие любую из последовательностей, указанных как SEQ ID NO: 63 и 570-579, могут связывать последовательность нуклеотидов-мишеней, смежную с последовательностью PAM, указанной как SEQ ID NO: 79. В некоторых вариантах осуществления настоящего изобретения варианты RGN APG05586 связываются с направляющей последовательностью, содержащей последовательность повтора CRISPR, указанную как SEQ ID NO: 64, а также может содержать последовательность tracrRNA, указанную как SEQ ID NO: 65. Эти системы RGN описаны в настоящем изобретении ниже в примере 5.

В области техники, к которой относится настоящее изобретение, хорошо известно, что специфичность последовательности PAM для конкретного фермента нуклеазы зависит от концентрации фермента (см., например, Karvelis и соавт., *Genome Biol.*, 2015, 16:253), концентрация может быть изменена путем изменения промотора, используемого для экспрессии RGN или количества рибонуклеопротеинового комплекса, доставляемого в клетку, органеллу или эмбрион.

После распознавания соответствующей последовательности PAM, RGN может расщепить целевую нуклеотидную последовательность в конкретном сайте расщепления. Используемый в настоящем изобретении термин «сайт расщепления» состоит из двух конкретных нуклеотидов нуклеотидной последовательности-мишени, между которыми нуклеотидная последовательность и расщепляется с помощью RGN. Сайт расщепления может содержать 1-й и 2-й, 2-й и 3-й, 3-й и 4-й, 4-й и 5-й, 5-й и 6-й, 7-й и 8-й или 8-й и 9-й нуклеотиды от PAM в любом направлении 5' или 3'. В некоторых вариантах осуществления настоящего изобретения сайт расщепления может находиться на расстоянии более 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 нуклеотидов от PAM либо в 5'-, либо в 3'-направлении. Поскольку RGN могут расщеплять целевую нуклеотидную последовательность, со смещением концов, в некоторых вариантах осуществления настоящего изобретения сайт расщепления определяется исходя из расстояния двух нуклеотидов от PAM на положительной

(+) цепи полинуклеотида и расстояния двух нуклеотидов от PAM на отрицательной (-) цепи полинуклеотида.

III. Нуклеотиды, кодирующие РНК-направляемые нуклеазы, CRISPR РНК и/или tracrRNA

Настоящее изобретение представляет полинуклеотиды, содержащие описанные в настоящем изобретении CRISPR РНК, tracrRNA и/или sgRNA, и полинуклеотиды, содержащие нуклеотидные последовательности, кодирующие описанные в настоящем изобретении РНК-направляемые нуклеазы, CRISPR РНК, tracrRNA и/или sgRNA. Описанные в настоящем изобретении полинуклеотиды включают полинуклеотиды, содержащие или кодирующие последовательность повтора CRISPR, содержащую нуклеотидную последовательность, указанную как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124 или ее активный вариант или фрагмент, который при включении в направляющую РНК способен направить специфичное для последовательности связывание ассоциированной РНК-направляемой нуклеазы с представляющей интерес последовательностью-мишенью. Также в настоящем изобретении представлены полинуклеотиды, содержащие или кодирующие tracrRNA, содержащие нуклеотидную последовательность указанную как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125 или ее активный вариант или фрагмент, который при включении в направляющую РНК способен направлять специфичное для последовательности связывание ассоциированной РНК-направляемой нуклеазы с интересующей последовательностью-мишенью. Также в настоящем изобретении предоставлены полинуклеотиды, кодирующие РНК-направляемую нуклеазу, содержащую аминокислотную последовательность, указанную как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, и ее активный вариант или фрагмент, обладающий способностью связываться с последовательностью-мишеней специфическим для последовательности образом.

Использование в настоящем изобретении терминов «полинуклеотид» или «молекула нуклеиновой кислоты» не ограничивает настоящее изобретение ДНК-полинуклеотидами. Специалистам в области техники, к которой относится настоящее изобретение, понятно, что полинуклеотиды предполагают также

рибонуклеотиды (РНК) и комбинации рибонуклеотидов и дезоксирибонуклеотидов. Дезоксирибонуклеотиды и рибонуклеотиды могут быть как встречающимися в природе молекулами, так и синтетическими аналогами. К ним относятся пептидные нуклеиновые кислоты (PNA), химеры ПНК-ДНК, замкнутые нуклеиновые кислоты (LNA) и фосфотионат-связанные последовательности. Полинуклеотиды, описанные в настоящем изобретении, охватывают все формы последовательностей, включая, но ими не ограничиваясь, одноцепочечные формы, двухцепочечные формы, гибриды ДНК-РНК, триплексные структуры, структуры типа «шпилька» и т.п.

10 Молекула нуклеиновой кислоты, кодирующая RGN, может быть кодон-оптимизированной для экспрессии в представляющем интерес организме. «Кодон-оптимизированная» кодирующая последовательность представляет собой полинуклеотидную кодирующую последовательность с частотой использования кодонов, имитирующей частоту предпочтительного

15 использования кодонов или условий транскрипции конкретной клеткой-хозяином. Экспрессия в конкретной клетке-хозяине или организме усиливается в результате изменения одного или более кодонов на уровне нуклеиновой кислоты таким образом, что транслируемая аминокислотная последовательность не изменяется. Молекулы нуклеиновой кислоты могут быть кодон-оптимизированы

20 полностью или частично. Таблицы кодонов и другие ссылки, содержащие информацию о предпочтениях для широкого круга организмов доступны в области техники, к которой относится настоящее изобретение (для обсуждения использования предпочтительного для растений кодона см., например, Campbell, Gowri, *Plant Physiol.*, 1990, 92:1-11). В области техники, к которой относится

25 настоящее изобретение, существуют методики синтеза кодон-оптимизированных кодирующих последовательностей генов растений или млекопитающих (например, человека. см., например, US 5380831, US 5436391, Muggay и соавт., *Nucleic Acids Res.*, 1989, 17:477-498; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки).

30 Полинуклеотиды, кодирующие RGN, crRNA, tracrRNA и/или sgRNA, описанные в настоящем изобретении, могут быть представлены в кассетах экспрессии для экспрессии *in vitro* или экспрессии в клетке, органелле, эмбрионе или организме, представляющем интерес. Кассета включает 5' и 3' регуляторные последовательности, функционально связанные с полинуклеотидом,

кодирующим RGN, crRNA, tracrRNA и/или sgRNA, описанные в настоящем изобретении, что позволяет экспрессировать полинуклеотиды. Кассета может дополнительно содержать по меньшей мере один дополнительный ген или генетический элемент для ко-трансформации в организме. Если в кассету

5 включены дополнительные гены или элементы, то компоненты кассеты будут функционально связаны. Используемый в настоящем изобретении термин «функционально связанный» означает функциональную связь между двумя или более элементами. Например, функциональная связь между промотором и интересующей кодирующей областью (например, областью, кодирующей RGN,

10 crRNA, tracrRNA и/или sgRNA) является функциональной связью, обеспечивающей экспрессию интересующей кодирующей области. Функционально связанные элементы могут быть смежными или несмежными. Когда речь идет о соединении двух областей кодирования белка, функциональная связь подразумевает, что области кодирования находятся в

15 одной рамке считывания. В качестве альтернативы, дополнительный ген (гены) или элемент (элементы) могут быть представлены на нескольких кассетах экспрессии. Например, нуклеотидная последовательность, кодирующая описанную в настоящем изобретении RGN, может присутствовать в одной кассете экспрессии, тогда как нуклеотидная последовательность, кодирующая

20 crRNA, tracrRNA или целую направляющую РНК, может находиться на отдельной кассете экспрессии. Такая кассета экспрессии содержит множеством сайтов рестрикции и/или сайтов рекомбинации для вставки полинуклеотидов, которые должны регулироваться на уровне транскрипции. Кассета экспрессии может дополнительно содержать селективируемый маркерный ген.

25 Кассета экспрессии может включать в направлении транскрипции 5'-3' транскрипционную (и, в некоторых вариантах осуществления настоящего изобретения, трансляционную) область инициации транскрипции (т.е. промотор), кодирующую RGN-, crRNA-, tracrRNA- и/или sgRNA-полинуклеотид по настоящему изобретению, а также транскрипционную (и, в некоторых

30 вариантах осуществления настоящего изобретения, трансляционную) функциональную область терминации (т.е. область терминации), действующую в исследуемом организме. Промоторы, описываемые в настоящем изобретении, способны направлять или стимулировать экспрессию кодирующей последовательности в клетке-хозяине. Регуляторные области (например,

промоторы, транскрипционные регуляторные области и области терминации трансляции) могут быть эндогенными или гетерологичными для клетки-хозяина или друг для друга. Используемый в настоящем изобретении термин «гетерологичная» в отношении последовательности означает

5 последовательность, которая происходит от чужеродного вида или, если от того же вида, то она существенно изменена от своей нативной формы по составу и/или геномному локусу вследствие преднамеренного вмешательства человека. Используемый в настоящем изобретении химерный ген включает кодирующую последовательность, функционально связанную с областью инициации
10 транскрипции, которая гетерологична кодирующей последовательности.

Оптимальные области терминации, такие как области терминации октопин-синтазы и нопалин-синтазы, доступны на Ti-плазмиде *A. tumefaciens*. См. также Guerineau с соавт., *Mol. Gen. Genet.* 1991, 262:141-144; Proudfoot, *Cell.*, 1991, 64:671-674; Sanfason и соавт., *Genes Dev.*, 1991, 5:141-149; Mogen и соавт., *Plant Cell.*, 1990, 2:1261-1272; Munroe и соавт., *Gene*, 1990, 91:151-158; Ballas и соавт.,
15 *Nucleic Acids Res.*, 1989, 17:7891-7903; Joshi и соавт., *Nucleic Acids Res.*, 1987, 15:9627-9639; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки.

Дополнительные регуляторные сигналы включают, но ими не
20 ограничиваются, сайты начала инициации транскрипции, операторы, активаторы, энхансеры, другие регуляторные элементы, сайты связывания рибосом, инициации кодонов, сигналы терминации и т. п. См., например, US 5039523, US 4853331; EPO 0480762A2; кн.: Sambrook и соавт., «Molecular Cloning: A Laboratory Manual», под ред. Maniatis и соавт., (изд-во Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Нью-Йорк, 1992), в дальнейшем «Sambrook 11»; кн.: Davis и соавт., «Advanced Bacterial Genetics (изд-во Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Нью-Йорк, 1980); и ссылки, приведенные в данных публикациях.

При приготовлении кассеты для экспрессии можно манипулировать
30 различными фрагментами ДНК, чтобы обеспечить последовательности ДНК в надлежащей ориентации и, при необходимости, в надлежащей рамке считывания. Для этого можно использовать адаптеры или линкеры для соединения фрагментов ДНК или другие манипуляции для создания оптимальных сайтов рестрикции, удаления лишней ДНК, удаления сайтов

рестрикции и т.п. С этой целью могут быть использованы методики мутагенеза *in vitro*, репарации праймера, рестрикции, отжига, повторные замены, например, транзиции и трансверсии.

В практике применения настоящего изобретения может быть использован ряд промоторов. Промоторы могут быть выбраны в зависимости от желаемого результата. Нуклеиновые кислоты могут быть объединены с конститутивными, индуцибельными, специфическими для стадии роста, специфическими для типа клеток, тканепредпочтительными, тканеспецифичными или иными промоторами для экспрессии в интересующем организме. См., например, промоторы, описанные в WO 99/43838, в US 8575425; US 7790846; US 8147856; US 8586832; US 7772369; US 7534939; US 6072050; US 5659026; US 5608149; US 5608144; US 5604121; US 5569597; US 5466785; US 5399680; US 5268463; US 5608142; US 6177611; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки.

Для экспрессии в растениях к конститутивным промоторам также относят промотор CaMV 35S (Odell и соавт., *Nature.*, 1985, 313:810-812), актин риза (McElroy и соавт., *Plant Cell.*, 1990, 2:163-171), убиквитин (Christensen и соавт., *Plant Mol. Biol.*, 1989, 12:619-632, Christensen и соавт., *Plant Mol. Biol.*, 1992, 18:675-689), pEMU (Last и соавт., *Theor. Appl. Genet.*, 1991, 81:581-588), MAS (Velten и соавт., *EMBO J.*, 1984, 3:2723-2730).

Примерами индуцибельных промоторов являются промотор Adh1, который индуцируется гипоксией или холодовым стрессом, промотор Hsp70, который индуцируется тепловым стрессом, промотор PPDК и промотор пепкарбоксилазы, которые индуцируются светом. Также полезными являются промоторы, которые являются химически индуцируемыми, например, промотор In2-2, индуцируемый сафенером (US 5364780), промотор, индуцируемый ауксином и специфичный для тапетума, но также активный в каллусе (PCT US01/22169), промоторы, реагирующие на стероиды (см., например, промотор ERE, индуцируемый эстрогеном, и глюкокортикоид-индуцируемый промотор, описанные в публикации Schena и соавт., *Proc. Natl. Acad. Sci. USA.*, 1991, 88:10421-10425; McNellis и соавт., *Plant J.*, 1998, 14(2):247-257), а так же тетрациклин-индуцируемые и тетрациклин-репресслируемые промоторы (см., например, Gatz и соавт., *Mol. Gen. Genet.*, 1991, 227:229-237, US 5814618 и US 5789156); сущность

каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки.

Тканеспецифичные или тканепредпочтительные промоторы могут быть использованы для направленной экспрессии экспрессируемой конструкции в
5 определенные ткани. В некоторых вариантах осуществления настоящего изобретения тканеспецифические или тканепредпочтительные промоторы активны в тканях растений. Примеры промоторов под контролем развития у растений включают промоторы, которые предпочтительно иницируют транскрипцию в определенных тканях, таких как листья, корни, плоды, семена
10 или цветы. «Тканеспецифический» промотор – это промотор, который иницирует транскрипцию только в определенных тканях. В отличие от конститутивной экспрессии генов, тканеспецифическая экспрессия является результатом нескольких взаимодействующих уровней регуляции генов. Поэтому для достижения эффективной и надежной экспрессии трансгенов в
15 определенные ткани предпочтительно использовать промоторы из гомологичных или близкородственных видов растений. В некоторых вариантах осуществления настоящего изобретения экспрессия предполагает использование тканепредпочтительного промотора. «Тканепредпочтительный» промотор – это промотор, который иницирует транскрипцию преимущественно, не обязательно
20 полностью или исключительно, в определенных тканях.

В некоторых вариантах осуществления настоящего изобретения молекулы нуклеиновой кислоты, кодирующие RGN, crRNA и/или tracrRNA, содержат промотор, специфичный для клеточного типа. Промотор, специфичный для клеточного типа – это промотор, который в первую очередь обеспечивает
25 экспрессию в определенных типах клеток в одном или нескольких органах. Некоторые примеры растительных клеток, в которых могут быть преимущественно активны клеточно-специфичные промоторы растений, включают, например, клетки ВЕТЛ, клетки сосудов корней, листьев, клетки стебля и ствольные клетки. Молекулы нуклеиновой кислоты могут также
30 содержать промоторы, предпочтительного клеточного типа. Промотор «предпочтительного типа клеток» – это промотор, который стимулирует экспрессию преимущественно, но не обязательно полностью или исключительно, в определенных типах клеток в одном или нескольких органах. Некоторые примеры растительных клеток, в которых промоторы

предпочтительного клеточного типа, функционирующие в растениях, могут быть предпочтительно активны, включают, например, клетки BETL, клетки сосудов корней, листьев, клетки стебля и стволовые клетки.

5 Последовательности нуклеиновых кислот, кодирующие RGN, crRNA, tracrRNA и/или sgRNA, могут быть функционально связаны с промоторной последовательностью, которая распознается фаговой РНК-полимеразой, например, для синтеза мРНК *in vitro*. В таких вариантах осуществления настоящего изобретения транскрибированная *in vitro* РНК может быть очищена для использования в методиках, описанных в настоящем изобретении.

10 Например, последовательность промотора может быть последовательностью промотора T7, T3 или SP6 или вариацией последовательности промотора T7, T3 или SP6. В таких вариантах осуществления настоящего изобретения экспрессированные белки и/или РНК могут быть очищены для использования в методиках модификации генома, описанных в настоящем изобретении.

15 В некоторых вариантах осуществления настоящего изобретения полинуклеотид, кодирующий RGN, crRNA, tracrRNA и/или sgRNA, также может быть связан с сигналом полиаденилирования (например, сигналом SV40 polyA и другими функциональными сигналами в растениях) и/или, по меньшей мере, одной последовательностью терминации транскрипции. Дополнительно,

20 последовательность, кодирующая RGN, также может быть связана с последовательностью (последовательностями), кодирующей по меньшей мере один сигнал ядерной локализации, по меньшей мере один проникающий в клетку домен и/или по меньшей мере один сигнальный пептид, способный

25 транспортировать белки в определенные субклеточные местоположения, как описано в настоящем изобретении.

Полинуклеотид, кодирующий RGN, crRNA, tracrRNA и/или sgRNA, может присутствовать в векторе или множестве векторов. Используемый в настоящем изобретении термин «вектор» обозначает полинуклеотидную композицию, предназначенную для переноса, доставки или введения нуклеиновой кислоты в

30 клетку-хозяина. Оптимальные векторы включают плазмидные векторы, фагмиды, космиды, искусственные/мини-хромосомы, транспозоны и вирусные векторы (например, лентивирусный вектор, адено-ассоциированный вирусный вектор, бакуловирусный вектор). Вектор может содержать дополнительные последовательности контроля экспрессии (например, энхансерные

последовательности, последовательности Козака, последовательности полиаденилирования, последовательности терминации транскрипции), селективируемые маркерные последовательности (например, гены устойчивости к антибиотикам), источники репликации и т.п. Дополнительную информацию можно найти в кн.: Ausubel и соавт., «Current Protocols in Molecular Biology», изд-во John Wiley & Sons, Нью-Йорк, 2003 или Sambrook, Russell, «Molecular Cloning: A Laboratory Manual», изд-во Cold Spring Harbor Press, Cold Spring Harbor, Нью-Йорк, 2001, 3е издание.

Вектор может также включать селективируемый маркерный ген для отбора трансформированных клеток. Селективируемые маркерные гены используют для отбора трансформированных клеток или тканей. Маркерные гены включают гены, кодирующие устойчивость к антибиотикам, например, гены, кодирующие неомидин фосфотрансферазу II (NEO) и гигромицин фосфотрансферазу (HPT), а также гены, обеспечивающие устойчивость к гербицидным соединениям, таким как глюофосинат аммония, бромоксирил, имидазолиноны и 2,4-дихлорфеноксиацетат (2,4-D).

В некоторых вариантах осуществления настоящего изобретения кассета экспрессии или вектор, включающий последовательность, кодирующую полипептид RGN, может дополнительно включать последовательность, кодирующую crRNA и/или tracrRNA или crRNA и tracrRNA, объединенные для создания направляющей РНК. Последовательность (последовательности), кодирующая crRNA и/или tracrRNA, может быть функционально связана, по меньшей мере, с одной последовательностью транскрипционного контроля для экспрессии crRNA и/или tracrRNA в целевом организме или клетке-хозяине. Например, полинуклеотид, кодирующий crRNA и/или tracrRNA, может быть функционально связан с промоторной последовательностью, которая распознается РНК-полимеразой III (Pol III). Примеры подходящих промоторов Pol III включают, но ими не ограничиваются, промоторы РНК млекопитающих U6, U3, H1 и 7SL и промоторы U6 и U3 риса.

Как указано выше, конструкции экспрессии, включающие нуклеотидные последовательности, кодирующие RGN, crRNA, tracrRNA и/или sgRNA, могут быть использованы для трансформации интересующих организмов. Методы трансформации включают введение нуклеотидной конструкции в целевой организм. Под «введением» подразумевается введение нуклеотидной

конструкции в клетку-хозяина таким образом, чтобы конструкция получала доступ к внутреннему пространству клетки-хозяина. Способы, предусмотренные настоящим изобретением, не требуют конкретного введения нуклеотидной конструкции в интересующий организм, а только чтобы нуклеотидная конструкция получила доступ к внутреннему пространству по меньшей мере одной клетки-хозяина. Клетка-хозяин может быть эукариотической или прокариотической клеткой. В конкретных вариантах осуществления настоящего изобретения эукариотическая клетка-хозяин представляет собой растительную клетку, клетку млекопитающего, клетку птицы или клетку насекомого. В некоторых вариантах осуществления настоящего изобретения эукариотическая клетка, которая содержит или экспрессирует или модифицирована представленной в настоящем изобретении RGN, является клеткой человека. В некоторых вариантах осуществления настоящего изобретения эукариотическая клетка, содержит или экспрессирует представленную в настоящем изобретении RGN, или которая модифицирована представленной в настоящем изобретении RGN, является клеткой кроветворения, такой как иммунная клетка (т.е., клетка врожденной или адаптивной иммунной системы), включая, без ограничений, В-клетки, Т-клетки, натуральные клетки-киллеры (NK), плюрипотентные стволовые клетки, индуцированные плюрипотентные стволовые клетки, химерные антигенные рецепторы Т-клеток (CAR-T – chimeric antigen receptor-T), моноциты, макрофаги и дендритные клетки.

Методы введения нуклеотидных конструкций в растения и иные клетки-хозяева известные в области техники, к которой относится настоящее изобретение, включают, но ими не ограничиваются, методы стабильной трансформации, методы транзитной трансформации и методы введения нуклеотидных конструкций с помощью вируса.

В результате применения этих методов получается трансформированный организм, например, растение, включая целые растения, а также органы растений (например, листья, стебли, корни и т.д.), семена, растительные клетки, черенки, эмбрионы и их потомство. Растительные клетки могут быть дифференцированными или недифференцированными (например, каллус, клетки суспензионной культуры, протопласты, клетки листьев, клетки корней, клетки флоэмы, пыльца).

«Трансгенные организмы», или «трансформированные организмы», или «стабильно трансформированные» организмы, или клетки, или ткани относятся к организмам, которые содержат интегрированный полинуклеотид, кодирующий RGN, crRNA и/или tracrRNA по настоящему изобретению. Известно, что другие экзогенные или эндогенные последовательности нуклеиновых кислот или фрагменты ДНК также могут быть включены в клетку-хозяина. Трансформация с помощью *Agrobacterium* или методом биолистики являются двумя преимущественно используемыми подходами к трансформации растительных клеток. Однако трансформация клетки-хозяина может быть выполнена путем инфицирования, трансфекции, микроинъекции, электропорации, микропроекции, биолистики или бомбардировки частицами, с применением карбидокремниевых волокон, ультразвука, ПЭГ, совместного осаждения фосфатом кальция, методом поликатионного ДМСО, применением DEAE-декстрана, а также может быть опосредована применением вирусов, липосом и т.п. Вирус-опосредованное введение полинуклеотида, кодирующего RGN, crRNA и/или tracrRNA, включает ретровирусное, лентивирусное, аденовирусное и адено-ассоциированное вирус-опосредованное введение и экспрессию, а также использование *Caulimoviruses*, *Geminiviruses* и РНК-вирусов растений.

Протоколы трансформации, а также протоколы введения полипептидов или полинуклеотидных последовательностей в растения, могут различаться в зависимости от типа клетки-хозяина (например, клетки однодольного или двудольного растения), предназначенной для трансформации. Методы трансформации известны в данной области и включают методы, изложенные в патентах US 8575425, US 7692068, US 8802934, US 7541517, сущность каждого приведенного патента полностью включена в настоящее изобретение в виде ссылки. См. также Rakoczy-Trojanowska M., *Cell Mol Biol Lett.*, 2002, 7:849-858; Jones и соавт., *Plant Methods*, 2005, 1:5; Rivera и соавт., *Physics of Life Reviews*, 2012, 9:308-345; Bartlett и соавт., *Plant Methods*, 2008, 4:1-12; Bates GW, *Methods in Molecular Biology*, 1999, 111:359-366; Binns, *Thomashow Annual Reviews in Microbiology*, 1988, 42:575-606; Christou P, *The Plant Journal*, 1992, 2:275-281; Christou P, *Euphytica*, 1995, 85:13-27; Tzfira и соавт., *TRENDS in Genetics*, 2004, 20:375-383; Yao с соавт., *Journal of Experimental Botany*, 2006, 57:3737-3746; Zupan, Zambryski, *Plant Physiology*, 1995, 107:1041-1047; Jones и соавт., *Plant*

Methods, 2005, 1:5; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки.

Трансформация может привести к стабильному или транзиторному включению нуклеиновой кислоты в клетку. «Стабильная трансформация» означает, что нуклеотидная конструкция, введенная в клетку-хозяина, интегрируется в геном клетки-хозяина и способна наследоваться ее потомством. «Переходная трансформация» означает, что полинуклеотид вводится в клетку-хозяина, но не интегрируется в геном клетки-хозяина.

Методы трансформации хлоропластов известны в области техники, к которой относится настоящее изобретение. Например, Svab и соавт., *Proc. Natl. Acad. Sci. USA*, 1990, 87:8526-8530; Svab, Maliga, *Proc. Natl. Acad. Sci. USA*, 1993, 90:913-917; Svab, Maliga, *EMBO J.*, 1993, 12:601-606. Метод основан на доставке ДНК, содержащей селективный маркер, с помощью биолиственной пушки и нацеливании ДНК на геном пластид посредством гомологичной рекомбинации. Кроме того, трансформация пластид может быть осуществлена путем трансактивации молчащего трансгена, переносимого пластидами, путем предпочтительной для ткани экспрессии, ядерно-кодируемой и направленной на пластиды РНК-полимеразы. О такой системе было сообщено в McBride и соавт., *Proc. Natl. Acad. Sci. USA*, 1994, 91:7301-7305.

Трансформированные клетки могут быть выращены в трансгенном организме, например, растении, в соответствии с обычными способами. См., например, McCormick и соавт., *Plant Cell Reports*, 1986, 5:81-84. Затем эти растения могут быть выращены и либо опылены одним и тем же трансформированным штаммом, либо разными штаммами, с идентификацией полученного гибрида, обладающего конститутивной экспрессией желаемого фенотипического признака.

Можно вырастить два или более поколений для обеспечения стабильного поддержания и наследования экспрессии желаемого фенотипического признака, а затем собрать семена, чтобы убедиться, что экспрессия желаемого фенотипического признака достигнута. Таким образом, настоящее изобретение обеспечивает трансформированные семена (также называемые «трансгенные семена»), имеющие нуклеотидную конструкцию согласно настоящему изобретению, например, кассету экспрессии по-настоящему изобретения, стабильно включенную в геном.

В качестве альтернативы в организм могут быть введены трансформированные клетки. Эти клетки могли быть получены из организма, в котором клетки трансформируются методом *ex vivo*.

Последовательности, предусмотренные в настоящем изобретении, могут
5 быть использованы для трансформации любых видов растений, включая, но ими не ограничиваясь, однодольные и двудольные. Примеры представляющих интерес растений включают, но ими не ограничиваются, кукурузу (маис), сорго, пшеницу, подсолнечник, томаты, крестоцветные, перец, картофель, хлопок, рис, сою, сахарную свеклу, сахарный тростник, табак, ячмень, масличный рапс,
10 *Brassica* sp., люцерну, рожь, просо, сафлор, арахис, сладкий картофель, маниоку, кофе, кокос, ананас, цитрусовые, какао, чай, банан, авокадо, инжир, гуаву, манго, оливу, папайю, кешью, макадамию, миндаль, овес, овощи, декоративные и хвойные растения.

Овощи включают, но ими не ограничиваются, помидоры, салат-латук,
15 стручковую фасоль, лимскую фасоль, горох и представителей рода *Curcumis*, таких как огурец, канталупа и мускусная дыня. Декоративные растения включают, но ими не ограничиваются, азалию, гортензию, гибискус, розы, тюльпаны, нарциссы, петунии, гвоздику, пуансеттию и хризантему. Предпочтительно, растениями, для которых применимо настоящее изобретения, являются культурными растениями (например, кукуруза, сорго, пшеница,
20 подсолнечник, томаты, крестоцветные, перец, картофель, хлопок, рис, соя, сахарная свекла, сахарный тростник, табак, ячмень, масличный рапс и т.д.).

Используемый в настоящем изобретении термин «растение» включает
25 растительные клетки, протопласты клеток растений, культуры тканей растительных клеток, из которых можно регенерировать растение, каллус растения, комки клеток и растительные клетки, которые находятся в растениях или частях растения, таких как эмбрионы, пыльца, яйцеклетки, семена, листья, цветы, ветви, плоды, зерна, колосья, початки, шелуха, стебли, корни, кончики корней, пыльники и т.п. Под зерном подразумевается зрелое семечко,
30 произведенное коммерческими производителями для целей, отличных от выращивания или воспроизводства вида. Потомство, варианты и мутанты регенерированных растений также включены в рамки охвата настоящего изобретения при условии, что эти части содержат вводимые полинуклеотиды. Дополнительно предусматривают переработанный растительный продукт или

побочный продукт, который сохраняет последовательности, описанные в настоящем изобретении, например, соевую муку.

Полинуклеотиды, кодирующие RGN, crRNA, и/или tracrRNA, также могут быть использованы для трансформации любых прокариотических видов, включая, но ими не ограничиваясь, археи и бактерии (например, *Bacillus* sp., *Klebsiella* sp., *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., *Agrobacterium*, *Lactobacillus* sp.).

Полинуклеотиды, кодирующие RGN, crRNA и/или tracrRNA, могут быть использованы для трансформации любых эукариотических видов, включая, но ими не ограничиваясь, животных (например, млекопитающих, насекомых, рыб, птиц и рептилий), грибы, амёбы, водоросли и дрожжи.

Обычные методы переноса генов на вирусной и невирусной основе могут быть использованы для введения нуклеиновых кислот в клетки млекопитающих, насекомых, птиц или целевые ткани. Такие методы могут быть использованы для введения нуклеиновых кислот, кодирующих компоненты системы RGN, в культуру клеток или в организм хозяина. Невирусные векторные системы доставки включают ДНК-плазмиды, РНК (например, вектор транскрипции, описанный в настоящем изобретении), нуклеиновую кислоту в чистом виде (незащищенную) и нуклеиновую кислоту, в комплексе со средством доставки, таким как липосома. Системы доставки вирусных векторов включают ДНК- и РНК-вирусы, которые после доставки в клетку имеют либо эписомальный, либо интегрированный геном. Для обзора методов генной терапии см. публикации Anderson, *Science*, 1992, 256: 808- 813; Nabel, Feigner, *TIBTECH*, 1993, 11:211-217; Mitani, Caskey, *TIBTECH*, 1993, 11:162-166; Dillon, *TIBTECH*, 1993, 11:167-175; Miller, *Nature*, 1992, 357:455-460; Van Brunt, *Biotechnology*, 1988, 6(10): 1149-1154; Vigne, *Restorative Neurology, Neuroscience*, 1995, 8:35-36; Kremer, Perricaudet, *British Medical Bulletin*, 1995, 51(1):31-44; Haddada и соавт., в кн.: «Current Topics in Microbiology, Immunology», под ред. Doerfler, Bohm, 1995; Yu и соавт., *Gene Therapy*, 1994, 1:13-26 (1994).

Методы невирусной доставки нуклеиновых кислот включают липофекцию, нуклеофекцию, микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, поликатионные или липидные конъюгаты нуклеиновых кислот, незащищенную ДНК, искусственные вирионы и усиленные агентами

поглощение ДНК. Липофекция описана, например, в US 5049386, US 4946787 и US 4897355, а реагенты для липофекции доступны на коммерческой основе (например, продукты Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые подходят для эффективной рецептор-распознающей

5 липофекции полинуклеотидов, включают липиды Feigner, WO 91/17424; WO 91/16024. Доставка может осуществляться в клетки (например, при введении *in vitro* или *ex vivo*) или в ткани-мишени (например, при введении *in vivo*).

Приготовление комплексов липид:нуклеиновая кислота, включая целевые липосомы, такие как иммунолипидные комплексы, хорошо известно

10 специалистам в области техники, к которой относится настоящее изобретение. (См., например, Crystal, *Science*, 1995, 270:404-410; Blaese и соавт., *Cancer Gene Ther.*, 1995, 2:291-297; Behr и соавт., *Bioconjugate Chem.*, 1994, 5:382-389; Remy и соавт., *Bioconjugate Chem.*, 1994, 5:647-654 (1994); Gao и соавт., *Gene Therapy*, 1995, 2:710-722; Ahmad и соавт., *Cancer Res.*, 1992, 52:4817-4820; US 4186183,

15 US 4217344, US 4235871, US 4261975, US 4485054, US 4501728, US 4774085, US 4837028, US 4946787).

Применение РНК- или ДНК-вирусных систем для доставки нуклеиновых кислот использует преимущества высокотехнологичных разработок для нацеливания вируса на определенные клетки организма и транспортировки

20 вирусной полезной нагрузки в ядро. Вирусные векторы могут вводиться непосредственно пациентам (*in vivo*) или использоваться для обработки клеток *in vitro*, а модифицированные клетки могут вводиться пациентам (*ex vivo*).

Традиционные системы для переноса генов на основе вирусов включают ретровирусные, лентивирусные, аденовирусные, адено-ассоциированные

25 векторы и векторы вируса простого герпеса. При использовании ретровирусных, лентивирусных и адено-ассоциированных вирусных методов трансфекции генов возможна интеграция в геном хозяина, что часто приводит к длительной экспрессии введенного трансгена. Дополнительно, высокая эффективность трансдукции наблюдалась во многих различных типах клеток и тканях-мишенях.

30 Ретровирусный тропизм может быть изменен за счет включения чужеродных белков оболочки, что расширяет потенциальную популяцию клеток-мишеней. Лентивирусные векторы относятся к ретровирусным векторам, которые способны трансдуцировать или инфицировать неделящиеся клетки, которые обычно дают высокие титры вирусов. Выбор ретровирусной системы

переноса генов зависит от ткани-мишени. Ретровирусные векторы состоят из цис-действующих длинных терминальных повторов с возможностью упаковки до 6-10 т.н. чужеродной последовательности. Минимальное количество цис-действующих LTR достаточно для репликации и упаковки векторов, которые затем используются для интеграции терапевтического гена в клетку-мишень для обеспечения постоянной экспрессии трансгена. Широко используемые ретровирусные векторы включают векторы на основе вируса лейкемии мышей (MuLV), вируса лейкемии гиббона (GaLV), вируса иммунодефицита обезьян (SIV), вируса иммунодефицита человека (ВИЧ) и их комбинаций (см., например, Buchscher и соавт., *J. Viral.*, 1992, 66:2731-2739; Johann и соавт., *J. Viral.*, 1992, 66:1635-1640; Somnerfelt и соавт., *J. Viral.*, 1990, 176:58-59; Wilson и соавт., *J. Viral.*, 1989, 63:2374-2378; Miller и соавт., *J. Viral.*, 1991, 65:2220-2224; PCT/US94/05700).

В тех случаях, когда предпочтительна транзиторная экспрессия, можно использовать векторные системы на основе аденовирусов. Векторы на основе аденовирусов способны обеспечить очень высокую эффективность трансдукции во многих типах клеток и не требуют клеточного деления. С помощью таких векторов были получены высокие титры и уровни экспрессии. Такой вектор можно производить в больших количествах в относительно простой системе. Векторы адено-ассоциированного вируса («AAV») также могут быть использованы для трансдукции клеток нуклеиновыми кислотами-мишенями, например, при получении нуклеиновых кислот и пептидов *in vitro*, а также для процедур генной терапии *in vivo* и *ex vivo* (см., например, West и соавт., *Virology*, 1987, 160:38-47; US 4797368; WO 93/24641; Katin, *Human Gene Therapy*, 1994, 5:793-80; Muzyczka, *J. Clin. Invest.*, 1994, 94:1351). Конструирование рекомбинантных векторов AAV описано в ряде публикаций, включая US 5173414; Tratschin и соавт., *Mol. Cell. Biol.*, 1985, 5:3251-3260; Tratschin, и соавт., *Mol. Cell. Biol.*, 1984, 4:2072-2081; Hermonat, Muzyczka, *PNAS*, 1984, 81:6466-6470; Samulski и соавт., *J. Viral.*, 1989, 63:03822-3828. Клетки-упаковщики обычно используются для формирования вирусных частиц, способных инфицировать клетку-хозяина. К таким клеткам относятся клетки 293, которые упаковывают аденовирус, и клетки ψ J2 или клетки PA317, которые упаковывают ретровирус.

Вирусные векторы, используемые в генной терапии, обычно создают путем получения клеточной линии, которая упаковывает вектор нуклеиновой кислоты в вирусную частицу. Векторы обычно содержат минимальные вирусные последовательности, необходимые для упаковки и последующей интеграции в организм хозяина, при этом другие вирусные последовательности заменяют кассетой экспрессии для экспрессируемого полинуклеотида (полинуклеотидов). Утраченные вирусные функции обычно обеспечиваются в транс упаковочной клеточной линией. Например, векторы AAV, используемые в генной терапии, обычно содержат только последовательности ITR из генома AAV, которые необходимы для упаковки и интеграции в геном хозяина. Вирусная ДНК упакована в клеточной линии, которая содержит вспомогательную плазмиду, кодирующую другие гены AAV, а именно *rep* и *cap*, но без последовательностей ITR.

Линия клеток также может быть инфицирована аденовирусом в качестве вируса-помощника. Вирус-помощник способствует репликации вектора AAV и экспрессии генов AAV из вспомогательной плазмиды. Вспомогательная плазида не упакована в значительных количествах из-за отсутствия последовательностей ITR. Контаминацию аденовирусом можно уменьшить, например, с помощью термической обработки, к которой более чувствителен аденовирус, чем AAV. Дополнительные способы доставки нуклеиновых кислот в клетки известны специалистам в области техники, к которой относится настоящее изобретение. См., например, US 20030087817, сущность приведенного патента полностью включена в настоящее изобретение в виде ссылки.

В некоторых вариантах осуществления настоящего изобретения клетка-хозяин трансфицируется транзиторно или нетранзиторно одним или несколькими векторами, описанными в настоящем изобретении. В некоторых вариантах осуществления настоящего изобретения клетку трансфецируют в том виде, в котором она естественно встречается у субъекта. В некоторых вариантах осуществления настоящего изобретения клетка, которую трансфицируют, берется у субъекта. В некоторых вариантах осуществления настоящего изобретения клетка, которую трансфицируют, получена из клеток, взятых у субъекта, например, из клеточной линии. В некоторых вариантах осуществления настоящего изобретения клеточная линия может представлять собой клетки

млекопитающих, насекомых или птиц. Широкое разнообразие клеточных линий для культуры тканей известно в области техники, к которой относится настоящее изобретение. Примеры клеточных линий включают, но ими не ограничиваются, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLaS3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panel, PC-3, TFl, CTLL-2, CIR, Rat6, CVI, RPTE, AIO, T24, 182, A375, ARH-77, Calul, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI- 231, HB56, TIB55, lurkat, 145.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4. COS, COS-1, COS-6, COS-M6A, эпителий почки обезьяны BS-C-1, фибробласты эмбриона мыши BALB/3T3, фибробласты 3T3 Swiss, 3T3-L1, 132-d5 фетальные фибробласты человека; фибробласты мыши 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-I cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr-/-, COR-L23, COR-L23/CPR, COR-L235010, CORL23/ R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepalclc7, HL-60, HMEC, HT-29, *IY* клетки, K562 клетки, Ku812, KCL22, KGI, KY01, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCKII, MDCKII, MOR/ 0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, линии клеток OPCN/OPCT, Peer, PNT-1A/ PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, линию клеток THP1, U373, U87, U937, VCaP, линию клеток Vero, WM39, WT-49, X63, YAC-1, YAR и их трансгенные разновидности. Клеточные линии доступны из различных источников, известных специалистам в области техники, к которой относится настоящее изобретение (см., например, Коллекцию Американских типовых культур (the American Type Culture Collection – ATCC (Манассас, Вирджиния))).

В некоторых вариантах осуществления настоящего изобретения клетка, трансфицированная одним или несколькими векторами, описанными в настоящем изобретении, используется для создания новой клеточной линии, содержащей одну или более векторных последовательностей. В некоторых вариантах осуществления настоящего изобретения клетка, тразиторно трансфецированная компонентами системы RGN, как описано в настоящем

изобретении (например, путем транзиторной трансфекции одного или более векторов или трансфекции РНК), и модифицированная за счет активности системы RGN, используется для создания новой клеточной линии, включающей клетки, содержащие модификацию, но без какой-либо другой экзогенной последовательности. В некоторых вариантах осуществления настоящего изобретения клетки, транзиторно или нетранзиторно трансфицированные одним или несколькими векторами, описанными в настоящем изобретении или клеточные линии, полученные из таких клеток, используются для оценки одного или более тестируемых соединений.

В некоторых вариантах осуществления настоящего изобретения один или более векторов, описанных в настоящем изобретении, используют для получения трансгенного растения или трансгенного животного не являющегося человеком. В некоторых вариантах осуществления настоящего изобретения трансгенное животное представляет собой млекопитающее, например, мышь, крысу, хомяка, кролика, корову или свинью. В некоторых вариантах осуществления настоящего изобретения трансгенное животное представляет собой птицу, например, курицу или утку. В некоторых вариантах осуществления настоящего изобретения трансгенное животное представляет собой насекомое, например, комара или клеща.

IV. Варианты и фрагменты полипептидов и полинуклеотидов

В описании настоящего изобретения представлены активные варианты и фрагменты встречающейся в природе (т.е. дикий-тип) РНК-направляемой нуклеазы, аминокислотная последовательность которой указана как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, а также активные варианты и фрагменты природных повторов CRISPR, как последовательности, указанные как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124, а также активные варианты и фрагменты природной tracrRNA, как последовательности, указанные как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125, а также полинуклеотиды, кодирующие их.

Хотя активность варианта или фрагмента может быть изменена по сравнению с интересующим полинуклеотидом или полипептидом, вариант и фрагмент должны сохранять функциональность интересующего полинуклеотида

или полипептида. Например, вариант или фрагмент может иметь повышенную активность, пониженную активность, другой спектр активности или любое другое изменение активности по сравнению с представляющим интерес полинуклеотидом или полипептидом.

5 Фрагменты и варианты встречающихся в природе полипептидов RGN, таких как описанные в настоящем изобретении, будут сохранять специфическую для последовательности РНК-направляемую ДНК-связывающую активность. В
10 конкретных вариантах осуществления настоящего изобретения фрагменты и варианты встречающихся в природе полипептидов RGN, таких как описанные в настоящем изобретении, будут сохранять нуклеазную активность (одноцепочечную или двухцепочечную).

Фрагменты и варианты встречающихся в природе повторов CRISPR, таких как описанные в настоящем изобретении, будут сохранять способность, будучи
15 частью направляющей РНК (содержащей tracrRNA), связываться с РНК-направляемой нуклеазой (в комплексе с направляющей РНК) и направлять ее к нуклеотидной последовательности-мишени специфическим для последовательности образом.

Фрагменты и варианты встречающихся в природе tracrRNA, таких как описанные в настоящем изобретении, будут сохранять способность, будучи
20 частью направляющей РНК (содержащей CRISPR РНК), направлять РНК-направляемую нуклеазу (в комплексе с направляющей РНК) к нуклеотидной последовательности-мишени специфичным для последовательности образом.

Используемый в настоящем изобретении термин «фрагмент» относится к части полинуклеотидной или полипептидной последовательности.
25 Используемые в настоящем изобретении термины «фрагменты» или «биологически активные части» включают полинуклеотиды, содержащие достаточное количество смежных нуклеотидов для сохранения биологической активности (т.е. связывания с RGN и направления ее специфическим для последовательности образом к нуклеотидной последовательности-мишени, когда
30 он содержится в направляющей РНК). «Фрагменты» или «биологически активные части» включают полипептиды, содержащие достаточное количество смежных аминокислотных остатков для сохранения биологической активности (т.е. связывания с целевой нуклеотидной последовательностью специфичным для последовательности образом при образовании комплекса с направляющей

РНК). Фрагменты белков RGN включают те, которые короче полноразмерных последовательностей из-за использования альтернативного расположенного ниже по цепи сайта старта. Биологически активная часть белка RGN может быть полипептидом, который содержит, например, 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700 или более смежных аминокислотных остатков, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579. Такие биологически активные части могут быть получены рекомбинантными методами и оценены на специфическую для последовательности РНК-направляемую ДНК-связывающую активность. Биологически активный фрагмент последовательности повтора CRISPR может содержать по меньшей мере 8 смежных аминокислот, представленных как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124. Биологически активная часть последовательности повтора CRISPR может представлять собой полинуклеотид, который включает, например, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 или 22 смежных нуклеотида, представленных как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124. Биологически активная часть tracrRNA может представлять собой полинуклеотид, который содержит, например, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110 или более смежных нуклеотидов, представленных как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125.

В общем, в настоящем описании под «вариантами» подразумеваются в основном сходные последовательности. Для полинуклеотидов вариант включает удаление и/или добавление одного или более нуклеотидов в одном или более из внутренних сайтов нативного полинуклеотида и/или замену одного или более нуклеотидов в одном или более сайтах нативного полинуклеотида.

Используемый в описании настоящего изобретения полинуклеотид или полипептид «нативного» или «дикого типа» содержит встречающуюся в природе нуклеотидную последовательность или аминокислотную последовательность, соответственно. Для полинуклеотидов консервативные варианты включают те последовательности, которые из-за вырожденности генетического кода кодируют нативную аминокислотную последовательность интересующего гена.

Встречающиеся в природе аллельные варианты, подобные этим, могут быть идентифицированы с использованием хорошо известных методов молекулярной биологии, таких как, например, полимеразная цепная реакция (ПЦР) и методы гибридизации, описанные ниже. Варианты полинуклеотидов также включают синтетически полученные полинуклеотиды, например, полученные с использованием сайт-направленного мутагенеза, но которые все еще кодируют представляющий интерес полипептид или полинуклеотид. Как правило, варианты конкретного полинуклеотида, представленного в настоящем изобретении, будут иметь по меньшей мере, примерно на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более процентов идентичны последовательности данного конкретного полинуклеотида; процент идентичности определяют по программам выравнивания последовательности и параметрам, описанными в настоящем изобретении.

Варианты конкретного полинуклеотида, представленного в настоящем изобретении (т.е. эталонного полинуклеотида), также могут быть оценены путем сравнения процента идентичности последовательности между полипептидом, кодируемым вариантом полинуклеотида, и полипептидом, кодируемым эталонным полинуклеотидом. Процент идентичности последовательности между любыми двумя полипептидами может быть рассчитан с помощью программ выравнивания последовательности и параметров, описанных в настоящем изобретении. Если любую данную пару полинуклеотидов, описанных в настоящем изобретении, оценивают путем сравнения процентной идентичности последовательностей, общих для двух полипептидов, которые они кодируют, процентная идентичность последовательностей между двумя кодируемыми полипептидами составляет по меньшей мере примерно 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более.

В конкретных вариантах осуществления настоящего изобретения полинуклеотиды кодируют полипептид РНК-направляемой нуклеазы, включающий аминокислотную последовательность, которая по меньшей мере на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или больше идентична аминокислотной последовательности, представленной как

SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579. В некоторых вариантах осуществления настоящего изобретения варианты, представленные как SEQ ID NO: 63, сохраняют изолейцин в аминокислотном положении, соответствующем 305; валин в аминокислотном положении, соответствующем 328; лейцин в аминокислотном положении, соответствующем 366; треонин в аминокислотном положении, соответствующем 368; и валин в аминокислотном положении, соответствующем 405 в последовательности, представленной как SEQ ID NO: 63. Аминокислотное положение первой аминокислотной последовательности, «соответствующее» определенному положению второй аминокислотной последовательности, относится к положению в первой аминокислотной последовательности при оптимальном выравнивании первой и второй аминокислотных последовательностей, которое совпадает с указанным положением аминокислотного остатка во второй последовательности. В конкретных вариантах осуществления настоящего изобретения варианты, указанные как SEQ ID NO: 63, по меньшей мере на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или больший процент идентичны последовательности, представленной как SEQ ID NO: 63 за пределами тех аминокислотных остатков, которые сохраняются в последовательности, представленной как SEQ ID NO: 63 (т.е., I305, V328, L366, T368 и V405).

Биологически активный вариант полипептида RGN по настоящему изобретению, может отличаться всего примерно на 1-15 аминокислотных остатков, всего примерно на 1-10, например на 6-10, всего на 5, всего на 4, всего на 3, всего на 2 или же всего на 1 аминокислотный остаток. В конкретных вариантах осуществления настоящего изобретения полипептиды могут включать N-концевое или C-концевое отсечение, которое предполагает по меньшей мере удаление 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700 аминокислот или более от N- или C-конца полипептида.

В некоторых вариантах осуществления настоящего изобретения полинуклеотиды содержат или кодируют повтор CRISPR, состоящий из нуклеотидной последовательности, которая по меньшей мере на 40%, 45%, 50%,

55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или больший % идентична нуклеотидной последовательности, представленной как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124.

5 Представленные в настоящем изобретении полинуклеотиды могут содержать или кодировать *tracrRNA*, содержащий нуклеотидную последовательность, которая по меньшей мере на 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или больший процент идентичность
10 нуклеотидной последовательности, представленной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125.

Биологически активные варианты повтора CRISPR или *tracrRNA* по настоящему изобретению могут отличаться всего примерно на 1-15 нуклеотидов, всего примерно на 1-10, например на 6-10, всего на 5, всего на 4, всего на 3,
15 всего на 2 или всего на 1 нуклеотид. В конкретных вариантах осуществления настоящего изобретения полинуклеотиды могут включать 5' или 3' отсечение, которое может включать делецию по меньшей мере из 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 95, 100, 105, 110 нуклеотидов или более с 5' или 3' конца полинуклеотида.

20 Следует признать, что полипептиды RGN, повторы CRISPR и *tracrRNA*, представленные в настоящем изобретении, могут быть модифицированы с целью создания вариантов белков и полинуклеотидов. Изменения, разработанные человеком, могут быть внесены путем применения методов сайт-направленного мутагенеза. Кроме того, так же могут быть идентифицированы нативные, пока
25 неизвестные или еще не идентифицированные полинуклеотиды и/или полипептиды, структурно и/или функционально связанные с описанными в настоящем изобретении последовательностями, которые попадают в область охвата настоящего изобретения. Консервативные аминокислотные замены могут быть сделаны в неконсервативных областях, которые не изменяют функцию
30 белков RGN. Кроме того, могут быть внесены модификации, которые улучшают активность RGN.

Также вариативные полинуклеотиды и белки включают последовательности и белки, полученные в результате мутагенной и рекомбиногенной процедуры, такой как перетасовка ДНК. При данной процедуре один или более различных

белков RGN, представленных в настоящем изобретении (например, указанных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579), подвергается манипуляциям для создания нового белка RGN, обладающего желаемыми свойствами. Таким образом, библиотеки

5 рекомбинантных полинуклеотидов генерируют из популяции родственных последовательностей полинуклеотидов, содержащих участки последовательностей, которые имеют значительную идентичность последовательностей и могут быть гомологично рекомбинированы *in vitro* или *in vivo*. Например, используя этот подход, участки последовательности,

10 кодирующие интересующий домен, могут быть перетасованы между представленными в настоящем изобретении последовательностями RGN и другими известными генами RGN для получения нового гена, кодирующего белок с улучшенным интересующим свойством, например, увеличенным K_m в случае фермента. Стратегии для такой перетасовки ДНК известны в области

15 техники, к которой относится настоящее изобретение. См., например, Stemmer, *Proc. Natl. Acad. Sci. USA*, 1994, 91:10747-10751; Stemmer, *Nature*, 1994, 370:389-391; Cramer и соавт., *Nature Biotech.*, 1997, 15:436-438; Moore и соавт., *J. Mol. Biol.*, 1997, 272:336-347; Zhang и соавт., *Proc. Natl. Acad. Sci. USA*, 1997, 94:4504-4509; Cramer и соавт., *Nature*, 1998, 391:288-291; US 5605793, US 5837458.

20 «Перетасованная» нуклеиновая кислота представляет собой нуклеиновую кислоту, полученную с помощью процедуры перетасовки, такой как любая процедура перетасовки, изложенная в настоящем изобретении. Перетасованные нуклеиновые кислоты получают путем рекомбинации (физической или виртуальной) двух или более нуклеиновых кислот (или строк с символами

25 нуклеотидов), например, искусственным и, по желанию, рекурсивным способом. Как правило, в процессах перетасовки используют один или более этапов скрининга для выявления представляющих интерес нуклеиновых кислот; скрининг может быть выполнен до или после любого этапа рекомбинации. В некоторых (но не во всех) вариантах осуществления перестановки желательно

30 проводить несколько раундов рекомбинации перед отбором для увеличения разнообразия пула, подлежащего скринингу. Совокупный процесс рекомбинации и отбора по желанию повторяется рекурсивно. В зависимости от обстоятельств, перестановка может относиться к совокупному процессу рекомбинации и отбора

или, в другом варианте, может просто относиться к отдельным рекомбинационным стадиям совокупного процесса.

Используемый в настоящем изобретении термины «идентичность последовательности» или «идентичность» в контексте двух полинуклеотидов или полипептидных последовательностей относится к остаткам двух последовательностей, которые являются одинаковыми при выравнивании для максимального соответствия в определенном окне сравнения. Когда процент идентичности последовательности используется применительно к белкам, принимается, что положения остатков, которые не являются идентичными, часто отличаются консервативными аминокислотными заменами, когда аминокислотные остатки заменяются на другие аминокислотные остатки с аналогичными химическими свойствами (например, заряд или гидрофобность) и поэтому не изменяют функциональные свойства молекулы. Когда последовательности отличаются консервативными заменами, процент идентичности последовательности может быть урегулирован в сторону увеличения, чтобы скорректировать консервативный характер замены. О последовательностях, отличающихся такими консервативными заменами, говорят, что они имеют «сходство последовательностей» или «подобие». Способы такой корректировки хорошо известны специалистам в данной области техники. Как правило, корректировка включает оценку консервативной замены как частичного, а не полного несоответствия, тем самым увеличивая процентную идентичность последовательности. Так, например, если идентичная аминокислота оценивается в 1 балл, а неконсервативная замена оценивается в 0 баллов, то консервативная замена оценивается в баллах от 0 до 1. Оценка консервативных замен рассчитывается, например, как реализовано в программе PC/GENE (фирма Intelligenetics, Маунтин-Вью, Калифорния).

Используемый в настоящем изобретении термин «процент идентичности последовательности» означает значение, определяемое при сравнении двух оптимально выровненных последовательностей в окне сравнения, где часть полинуклеотидной последовательности может включать добавления или удаления (т.е. пробелы) по сравнению с эталонной последовательностью (которая не включает добавления или удаления) для оптимального выравнивания двух последовательностей. Процент идентичности рассчитывается путем определения количества позиций, в которых идентичное основание нуклеиновой

кислоты или остаток аминокислоты встречается в обеих последовательностях для получения количества совпадающих позиций, деления количества совпадающих позиций на общее количество позиций в окне сравнения и умножения результата на 100 для получения процента идентичности последовательности.

Если не указано иное, значения идентичности/подобия последовательности, представленные в настоящем изобретении, относятся к значению, полученному с помощью GAP Version 10 с использованием следующих параметров: % идентичности и % сходства для нуклеотидной последовательности с использованием GAP Weight of 50 и Length Weight of 3 и матрицы для оценки `nwsgapdna.cmp`; % идентичности и % сходства для аминокислотной последовательности с использованием GAP Weight of 8 и Length Weight of 2 и матрицы для оценки BLOSUM62; или любой эквивалентной программы. Под «эквивалентной программой» подразумевается любая программа сравнения последовательностей, которая для любых двух рассматриваемых последовательностей генерирует выравнивание, имеющее идентичные совпадения нуклеотидных или аминокислотных остатков и идентичную процентную идентичность последовательности при сравнении с соответствующим выравниванием, сгенерированным с помощью GAP Version 10.

Две последовательности являются «оптимально выровненными», когда они выравниваются для оценки сходства с использованием определенной матрицы аминокислотных замен (например, BLOSUM62), штрафа за существование пробелов и штрафа за расширение пробелов таким образом, чтобы получить наивысший балл, возможный для этой пары последовательностей. Матрицы аминокислотных замен и их использование для количественной оценки сходства между двумя последовательностями хорошо известны в данной области техники и описаны, например, в работах Dayhoff и соавт., (1978) «A model of evolutionary change in proteins» в кн.: «Atlas of Protein Sequence, Structure» 5(3): 345-352 (под ред. Dayhoff MO), изд-во Natl. Biomed. Res. Found., Вашингтон; Henikoff и соавт., *Proc. Natl. Acad. Sci. USA.*, 1992, 89:10915-10919. Матрица BLOSUM62 часто используется в протоколах выравнивания последовательностей в качестве матрицы замещения по умолчанию. Штраф за существование пробела налагается за введение одного аминокислотного пробела в одну из выровненных

последовательностей, а штраф за расширение пробела налагается за каждую
дополнительную пустую аминокислотную позицию, вставленную в уже
открытый пробел. Выравнивание определяется позициями аминокислот каждой
последовательности, с которых начинается и заканчивается выравнивание, и, по
5 желанию, вставкой пробела или нескольких пробелов в одну или обе
последовательности, чтобы получить максимально возможный балл. Хотя
оптимальное выравнивание и подсчет баллов могут быть выполнены вручную,
процесс облегчается при использовании компьютерного алгоритма
выравнивания, например, gapped BLAST 2.0, описанного в статье Altschul и
10 соавт., *Nucleic Acids Res.*, 1997, 25:3389-3402, и доступного для общественности
на Вебсайте Национального центра биотехнологической информации (National
Center for Biotechnology Information Website, www.ncbi.nlm.nih.gov).
Оптимальные выравнивания, включая множественные выравнивания, могут быть
подготовлены с помощью, например, Altschul и соавт., *Nucleic Acids Res.*, 1997,
15 25:3389-3402.

Что касается аминокислотной последовательности, которая оптимально
выровнена с эталонной последовательностью, аминокислотный остаток
«соответствует» положению в эталонной последовательности, с которым остаток
сопряжен при выравнивании. «Позиция» обозначается номером, который
20 последовательно идентифицирует каждую аминокислоту в эталонной
последовательности исходя из ее положения относительно N-конца. Из-за
делений, инсерций, усечений, слияний и т.д., которые должны быть учтены при
определении оптимального выравнивания, в общем случае, номер
аминокислотного остатка в тестируемой последовательности, определенный
25 простым подсчетом от N-конца, не обязательно будет таким же, как номер
соответствующей позиции в эталонной последовательности. Например, в случае,
когда в выравниваемой тестируемой последовательности имеется делеция, в
месте делеции не будет аминокислоты, соответствующей положению в
эталонной последовательности. Там, где имеется вставка в выровненную
30 эталонную последовательность, эта вставка не будет соответствовать какой-либо
аминокислотной позиции в эталонной последовательности. В случае отсечений
или слияний могут быть участки аминокислот либо в эталонной, либо в
выровненной последовательности, которые не соответствуют ни одной
аминокислоте в соответствующей последовательности.

V. Антитела

Рассматриваются антитела к полипептидам RGN или рибонуклеопротеинам, содержащим RGN полипептиды согласно настоящему изобретению, включая те, которые имеют аминокислотную последовательность, представленную как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, а так же их активные варианты или фрагменты. Способы получения антител хорошо известны в области техники, к которой относится настоящее изобретение (см., например, публикацию Harlow, Lane «Antibodies: A Laboratory Manual», изд-во Cold Spring Harbor Laboratory, Cold Spring Harbor, Нью-Йорк, 1988; US 4196265). Такие антитела могут быть использованы в наборах для обнаружения и выделения полипептидов RGN или рибонуклеопротеинов. Таким образом, в описании настоящего изобретения представлены наборы, содержащие антитела, которые специфически связываются с полипептидами или рибонуклеопротеинами, представленными в настоящем изобретении, включая, например, полипептиды, содержащие последовательность, представленную как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579.

VI. Системы и рибонуклеопротеиновые комплексы для связывания интересующей целевой последовательности и методы их получения

Настоящее изобретение описывает систему для связывания интересующей целевой последовательности-мишени, где система включает по меньшей мере одну направляющую РНК или нуклеотидную последовательность, кодирующую ее, и по меньшей мере одну РНК-направляемую нуклеазу или нуклеотидную последовательность, кодирующую ее. Направляющая РНК гибридизируется с интересующей целевой последовательностью, а также образует комплекс с полипептидом RGN, тем самым направляя полипептид RGN на связывание с целевой последовательностью. В некоторых вариантах осуществления настоящего изобретения RGN содержит аминокислотную последовательность, указанную как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579 или ее активный вариант или фрагмент. В некоторых вариантах осуществления настоящего изобретения направляющая РНК включает последовательность повтора CRISPR, содержащую нуклеотидную

последовательность, указанную как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124 или ее активный вариант или фрагмент.

В конкретных вариантах осуществления настоящего изобретения направляющая РНК включает в себя tracrRNA, содержащую нуклеотидную последовательность, указанную как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125 или ее активный вариант или фрагмент. Направляющая РНК системы может быть отдельной направляющей РНК или двунаправленной направляющей РНК. В конкретных вариантах осуществления настоящего изобретения система включает РНК-направляемую нуклеазу, гетерологичную направляющей РНК, при этом RGN и направляющая РНК не встречаются в природе в комплексе друг с другом (т.е. связанными друг с другом).

Система для связывания интересующей последовательности-мишени, представленная в настоящем изобретении, может представлять собой рибонуклеопротеиновый комплекс, который состоит по меньшей мере из одной молекулы РНК, связанной по меньшей мере с одним белком.

Рибонуклеопротеиновые комплексы, представленные в настоящем изобретении, включают по меньшей мере одну направляющую РНК в качестве РНК-компонента и РНК-направляемую нуклеазу в качестве белкового компонента. Такие рибонуклеопротеиновые комплексы могут быть очищены из клетки или организма, который естественным образом экспрессирует полипептид RGN и был сконструирован для экспрессии конкретной направляющей РНК, специфичной для интересующей последовательности-мишени. В другом варианте рибонуклеопротеиновый комплекс может быть очищен из клетки или организма, который был трансформирован полинуклеотидами, кодирующими полипептид RGN и направляющую РНК, и культивирован в условиях, позволяющих экспрессировать полипептид RGN и направляющую РНК. Таким образом, предложены методы получения полипептида RGN или рибонуклеопротеинового комплекса RGN. Такие методы включают культивирование клетки, содержащей нуклеотидную последовательность, кодирующую полипептид RGN, и в некоторых вариантах осуществления настоящего изобретения – нуклеотидную последовательность, кодирующую направляющую РНК, в условиях, в которых экспрессируется полипептид RGN (и в некоторых вариантах осуществления настоящего изобретения -направляющая

РНК). Затем полипептид RGN или рибонуклеопротеин RGN может быть очищен из лизата культивированных клеток.

Методы очистки полипептида RGN или рибонуклеопротеинового комплекса RGN из лизата биологического образца известны в области техники, к которой относится настоящее изобретение (например, эксклюзионная и/или аффинная хроматография, двумерный электрофорез (2D-PAGE), ВЭЖХ, обращено-фазовая хроматография, иммунопреципитация). В конкретных методах полипептид RGN получен рекомбинантно и содержит метку очистки для облегчения очистки, включая, но ими не ограничиваясь, глутатион-S-трансферазу (GST), хитин-связывающий белок (CBP), мальтозосвязывающий белок, тиоредоксин (TRX), поли(NANP), тандемную метку аффинной очистки (TAP), мус, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, 10xHis, биотин-карбоксильный белок-носитель (BCCP) и кальмодулин. Как правило, содержащий метку полипептид RGN или рибонуклеопротеиновый комплекс RGN очищают с помощью иммобилизованной металл-аффинной хроматографии. Следует понимать, что могут быть использованы другие аналогичные методы, известные в данной области техники, включая другие формы хроматографии или, например, иммунопреципитация, по отдельности или в комбинации.

«Выделенный» или «очищенный» полипептид или его биологически активная часть в значительной степени свободны от компонентов, которые обычно сопровождают или взаимодействуют с полипептидом, находясь в его естественной среде. Таким образом, выделенный или очищенный полипептид по существу свободен от другого клеточного материала или культуральной среды, если он получен рекомбинантным методом или по сути, свободен от химических прекурсоров или других химических веществ, если он получен химическим синтезом. Белок, в значительной степени свободный от клеточного материала, включает препараты белка, содержащие менее 30%, 20%, 10%, 5% или 1% (по сухому весу) загрязняющего белка. Когда белок по настоящему изобретению или его биологически активную часть получают рекомбинантным путем, оптимальная культуральная среда представляет менее 30%, 20%, 10%, 5% или 1% (по сухому весу) химических прекурсоров или не представляющих интерес белков.

Конкретные методы для связывания и/или расщепления представляющей интерес последовательности-мишени, представленные в настоящем изобретении, включают использование собранного *in vitro* рибонуклеопротеинового комплекса RGN. Сборка рибонуклеопротеинового комплекса RGN *in vitro* может
5 быть выполнена с помощью любого метода, известного в данной области техники, в котором полипептид RGN контактирует с направляющей РНК в условиях, обеспечивающих связывание полипептида RGN с направляющей РНК. Используемый в настоящем изобретении термины «контакт», «контактирование», «контактировали» означают помещение компонентов
10 желаемой реакции вместе в условиях, подходящих для проведения-желаемой реакции. Полипептид RGN может быть очищен из биологического образца, клеточного лизата или культуральной среды, получен путем трансляции *in vitro* или химически синтезирован. Направляющая РНК может быть очищена из биологического образца, клеточного лизата или культуральной среды,
15 транскрибирована *in vitro* или синтезирована химическим путем. Полипептид RGN и направляющая РНК могут быть приведены в контакт в растворе (например, буферном солевом растворе), чтобы обеспечить сборку рибонуклеопротеинового комплекса RGN *in vitro*.

20 VII. Методы связывания, расщепления или модификации целевой последовательности

Настоящее изобретение предоставляет методы связывания, расщепления и/или модификации целевой нуклеотидной последовательности, представляющей интерес. Методы включают доставку системы, включающей по
25 меньшей мере одну направляющую РНК или полинуклеотид, кодирующий ее, и по меньшей мере один полипептид RGN или полинуклеотид, кодирующий его, к целевой последовательности или клетке, органелле или эмбриону, содержащему целевую последовательность. В некоторых вариантах осуществления настоящего изобретения RGN содержит аминокислотную последовательность SEQ ID NO: 1,
30 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579 или ее активный вариант или фрагмент. В некоторых вариантах осуществления настоящего изобретения направляющая РНК включает последовательность повтора CRISPR, которая содержит нуклеотидную последовательность SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124 или

ее активный вариант или фрагмент. В некоторых вариантах осуществления настоящего изобретения направляющая РНК включает в себя tracrRNA, которая содержит нуклеотидную последовательность, представленную как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125 или ее активный вариант или фрагмент. Направляющая РНК системы может быть отдельной направляющей РНК или двунаправленной направляющей РНК. RGN системы может быть мертвой RGN, обладать никаказной активностью или представлять собой слитый полипептид. В некоторых вариантах осуществления настоящего изобретения слитый полипептид содержит полипептид, редактирующий основания, например, цитидиндезаминазу или аденозиндезаминазу. В других вариантах осуществления настоящего изобретения слитый белок RGN содержит обратную транскриптазу. В других вариантах осуществления настоящего изобретения слитый белок RGN содержит полипептид, который рекрутирует участников функционального комплекса репарации нуклеиновых кислот, таких как участник нуклеотидной эксцизионной репарации (NER – nucleotide excision repair) или транскрипционно-связанной нуклеотидной эксцизионной репарации (Transcription Coupled-Nucleotide Excision Repair - TC-NER) (Wei и соавт., *PNAS USA*, 2015, 112(27):E3495-504; Troelstra и соавт., *Cell*, 1992, 71:939-953; Margef и соавт., *J. Mol. Biol.*, 2017, 429(9):1277-1288) как описано в патентной заявке US 62/966203, подана 27 января 2020 года; сущность патента полностью включена в настоящее изобретение в виде ссылки. В некоторых вариантах осуществления настоящего изобретения слитый белок RGN содержит CSB (van den Boom и соавт., 2004, *J. Cell Biol.*, 2004, 166(1):27-36; van Gool и соавт., *EMBO J.*, 1997, 16(19):5955-65; пример которого указан как SEQ ID NO: 608), который является участником пути TC-NER (эксцизионная репарация нуклеотидов), его функция заключается в привлечении других участников метаболического пути. В дополнительных вариантах осуществления настоящего изобретения слитый белок RGN содержит активный домен CSB, например кислотный домен CSB, который включает аминокислотные остатки 356-394 из последовательности, представленной как SEQ ID NO: 608 (Teng и соавт., *Nat. Commun.*, 2018, 9(1):4115).

В конкретных вариантах осуществления настоящего изобретения RGN и/или направляющая РНК гетерологичны клетке, органелле или эмбриону, в которые введены RGN и/или направляющая РНК (или полинуклеотид

(полинуклеотиды), кодирующие по меньшей мере одну из RGN и направляющую РНК).

В тех вариантах осуществления настоящего изобретения, где метод включает доставку полинуклеотида, кодирующего направляющую РНК и/или полипептид RGN, клетку или эмбрион затем можно культивировать в условиях, в которых экспрессируются направляющая РНК и/или полипептид RGN. В различных вариантах осуществления настоящего изобретения представленный в настоящем изобретении способ включает в себя взаимодействие последовательности-мишени с рибонуклеопротеиновым комплексом RGN.

Рибонуклеопротеиновый комплекс RGN может включать RGN, которая является нуклеазой или обладает нуклеазной активностью. В некоторых вариантах осуществления настоящего изобретения RGN из рибонуклеопротеинового комплекса представляет собой слитый полипептид, содержащий полипептид, редактирующий основания. В некоторых вариантах осуществления настоящего изобретения метод включает введение в клетку, органеллу или эмбрион, содержащий целевую последовательность, рибонуклеопротеинового комплекса RGN. Рибонуклеопротеиновый комплекс RGN может быть очищен из биологического образца, получен рекомбинантно и впоследствии очищен или собран *in vitro*, как описано в настоящем изобретении. В тех вариантах осуществления настоящего изобретения, где рибонуклеопротеиновый комплекс RGN, контактирующий с целевой последовательностью или клеточной органеллой или эмбрионом, был собран *in vitro*, способ, представленный в настоящем изобретении, может дополнительно включать сборку комплекса *in vitro* до контакта с целевой последовательностью, клеткой, органеллой или эмбрионом.

Очищенный или собранный *in vitro* рибонуклеопротеиновый комплекс RGN может быть введен в клетку, органеллу или эмбрион с помощью любого метода, известного в данной области техники, включая, но им не ограничиваясь, электропорацию. В качестве альтернативы, полипептид RGN и/или полинуклеотид, кодирующий или включающий направляющую РНК, может быть введен в клетку, органеллу или эмбрион с помощью любого метода, известного в данной области техники (например, электропорация).

После доставки или контакта с целевой последовательностью или клеткой, органеллой или эмбрионом, содержащим целевую последовательность,

направляющая РНК направляет RGN на связывание с целевой последовательностью специфическим для последовательности образом. В тех вариантах осуществления настоящего изобретения, где RGN обладает нуклеазной активностью, полипептид RGN расщепляет интересующую целевую последовательность после связывания. Впоследствии целевая последовательность может быть модифицирована с помощью эндогенных механизмов репарации, таких как негомологичное соединение концов или направленная гомологичная репарация с предоставленным донорным полинуклеотидом.

10 Методы измерения связывания полипептида RGN с целевой последовательностью известны в данной области техники и включают метод иммунопреципитации хроматина, анализ сдвига подвижности на геле, метод вытягивания (pull-down) ДНК, репортерные методы, метода захвата и обнаружения в микропланшетах. Аналогично, методы измерения расщепления или модификации целевой последовательности известны в данной области и включают методы расщепления *in vitro* или *in vivo*, где расщепление подтверждается с помощью ПЦР, секвенирования или гель-электрофореза, с или без присоединения соответствующей метки (например, радиоизотопа, флуоресцентного вещества) к целевой последовательности для облегчения обнаружения продуктов деградации. В качестве альтернативы может быть использован метод амплификации, опосредованной никующим ферментом (NTEXPAR) (см., например, Zhang и соавт., *Chem. Sci.*, 2016, 7:4951-4957). *In vivo* расщепление может быть оценено с помощью нуклеазного анализа Surveyor (Guschin и соавт., *Methods Mol. Biol.*, 2010, 649:247-256).

25 В некоторых вариантах осуществления настоящего изобретения методы включают использование одного типа RGN в комплексе с более чем одной направляющей РНК. Несколько направляющих РНК могут быть нацелены на различные области одного гена или на несколько генов.

30 В тех вариантах осуществления настоящего изобретения, где донорный полинуклеотид не предусмотрен, двухцепочечный разрыв, внесенный полипептидом RGN, может быть восстановлен с помощью метода негомологичного соединения концов (NHEJ). Поскольку суть методики NHEJ, допускает возможность ошибки, восстановление двухцепочечного разрыва может привести к модификации целевой последовательности. Используемый в

настоящем изобретении термин, «модификация» применительно к молекуле нуклеиновой кислоты означает изменение нуклеотидной последовательности молекулы нуклеиновой кислоты, которое может быть делецией, инсерцией или заменой одного или нескольких нуклеотидов или их комбинацией. Модификация целевой последовательности может привести к экспрессии измененного белкового продукта или инактивации кодирующей последовательности.

В тех вариантах осуществления настоящего изобретения, где присутствует донорный полинуклеотид, донорная последовательность в донорном полинуклеотиде может быть интегрирована в целевую нуклеотидную последовательность или заменена на нее в процессе восстановления введенного двухцепочечного разрыва, что приводит к введению экзогенной донорной последовательности. Таким образом, донорный полинуклеотид включает донорную последовательность, которую желательно внедрить в интересующую последовательность-мишень. В некоторых вариантах осуществления настоящего изобретения последовательность донора изменяет исходную последовательность нуклеотида-мишени таким образом, что вновь интегрированная последовательность донора не будет распознана и расщеплена RGN. Интеграция донорной последовательности может быть усилена включением в донорный полинуклеотид фланкирующих последовательностей, называемых в настоящем изобретении «плечи гомологии», которые имеют значительную идентичность с последовательностями, фланкирующими нуклеотида-мишени, что позволяет осуществлять процесс гомологично-направленной репарации. В некоторых вариантах осуществления настоящего изобретения «плечи гомологии» имеют длину по меньшей мере 50 пар оснований, по меньшей мере 100 пар оснований, и до 2000 пар оснований или более, и имеют по меньшей мере 90%, по меньшей мере 95% или более процентов гомологии последовательности с соответствующей последовательностью в нуклеотидной последовательности-мишени.

В тех вариантах осуществления настоящего изобретения, где полипептид RGN вводит двухцепочечные ступенчатые разрывы, донорный полинуклеотид может содержать донорную последовательность, фланкированную совместимыми выступами, что позволяет прямое лигирование донорной последовательности к расщепленной последовательности нуклеотида-мишени,

включающей выступы, с помощью негомологичного процесса репарации во время репарации двухцепочечного разрыва.

В тех вариантах осуществления настоящего изобретения, где метод включает использование RGN, которая является нисказой (т.е. способна
5 расщеплять только одну нить двухцепочечного полинуклеотида), метод может включать введение двух RGN нисказ, которые нацелены на идентичные или перекрывающиеся целевые последовательности и расщепляют разные нити полинуклеотида. Например, нисказа RGN, которая расщепляет только
положительную (+) нить двухцепочечного полинуклеотида, может быть введена
10 вместе со второй нисказой RGN, которая расщепляет только отрицательную (-) нить двухцепочечного полинуклеотида.

В некоторых вариантах осуществления настоящего изобретения, представлена методика связывания и обнаружения целевой последовательности, включающая введение в клетку, органеллу или эмбрион по меньшей мере одной
15 направляющей РНК или полинуклеотида, кодирующего ее, и по меньшей мере одного полипептида RGN или полинуклеотида, кодирующего его, экспрессию направляющей РНК и/или полипептида RGN (если введены кодирующие последовательности), при этом полипептид RGN представляет собой «мертвую» RGN, то есть не обладающую нуклеазной активностью), и дополнительно
20 включает выявляемую метку, а методика в дальнейшем предполагает детекцию данной метки. Детектируемая метка может быть присоединена к RGN в виде слитого белка (например, флуоресцентного белка) или может представлять собой небольшую молекулу, конъюгированную с полипептидом RGN или включенную в него, которая может быть детектирована визуально или другими
25 способами.

Также в настоящем изобретении представлены способы модуляции экспрессии последовательности-мишени или представляющего интерес гена при регуляции последовательности-мишени. Способы включают введение в клетку, органеллу или эмбрион по меньшей мере одной направляющей РНК или
30 кодирующего ее полинуклеотида, и по меньшей мере одного полипептида RGN или кодирующего ее полинуклеотида, экспрессирующего направляющую РНК и/или полипептид RGN (если введены кодирующие последовательности), где полипептид RGN представляет собой RGN, утратившую нуклеазную активность. В некоторых из этих вариантов осуществления настоящего изобретения

«мертвая» RGN представляет собой слитый белок, содержащий домен модулятора экспрессии (т.е. домен эпигенетической модификации, домен активации транскрипции или домен репрессора транскрипции), как описано в настоящем изобретении.

5 Представленное в настоящем изобретении описание приводит способы связывания и/или модификации целевой нуклеотидной последовательности, вызывающей интерес. Способы включают доставку системы, содержащей, по меньшей мере, одну направляющую РНК или полинуклеотид, кодирующий ее, и, по меньшей мере, один полипептид слияния, содержащий RGN согласно
10 представляемому изобретению и полипептид, редактирующий основание, например цитидиндезаминазу или аденозиндезаминазу или полинуклеотид, кодирующий полипептид слияния, к последовательности-мишени или клетке, органелле или эмбриону содержащему целевую последовательность.

15 Специалисту в области техники, к которой относится настоящее изобретение, очевидно, что любой из описанных в настоящем изобретении способов может быть использован для нацеливания на одну целевую последовательность или несколько целевых последовательностей. Таким образом, способы включают использование одного полипептида RGN в комбинации с множеством различных направляющих РНК, которые могут
20 нацеливаться на множество различных последовательностей в пределах одного гена и/или множества генов. Также в настоящем изобретении рассматриваются способы, которые предполагают введение множества различных направляющих РНК в комбинации с множеством различных полипептидов RGN. Эти направляющие РНК и направляющие РНК/RGN полипептидные системы могут
25 нацеливаться на множество различных последовательностей в пределах одного гена и/или множества генов.

Настоящее изобретение предусматривает наборы, содержащие один или несколько каких-либо элементов, описанных в вышеназванных способах и
30 составах. В некоторых вариантах осуществления настоящего изобретения набор содержит векторную систему и инструкции по использованию набора. В некоторых вариантах осуществления настоящего изобретения векторная система содержит (а) первый регуляторный элемент, функционально связанный с последовательностью ДНК, кодирующей последовательность crRNA, и один или более сайтов вставки для включения направляющей последовательности выше

кодируемой последовательности crRNA, где при экспрессии, направляющая последовательность направляет специфичное для последовательности связывание комплекса RGN с последовательностью-мишенью в эукариотической клетке, в которой комплекс RGN содержит фермент RGN, образующий комплекс с направляющим полинуклеотидом РНК; и/или (б) второй регуляторный элемент, функционально связанный с кодирующей фермент последовательностью, кодирующей указанный фермент RGN, содержащий последовательность ядерной локализации. Элементы могут быть предоставлены по отдельности или в комбинациях и могут быть предоставлены в любом подходящем контейнере, таком как флакон, бутылка или тубик.

В некоторых вариантах осуществления настоящего изобретения набор включает инструкции на одном или нескольких языках. В некоторых вариантах осуществления настоящего изобретения набор включает один или более реагентов для применения в процессе, использующем один или несколько элементов, описанных в настоящем изобретении. Реагенты могут быть предоставлены в любом подходящем контейнере. Например, набор может включать один или несколько реакционных буферов или буферов для хранения. Реагенты могут быть предоставлены в виде, пригодном для использования в конкретном анализе или в форме, требующей добавления одного или нескольких других компонентов перед использованием (например, в виде концентрата или лиофилизированной формы). Буфер может быть любым буфером, включая, но ими не ограничиваясь, натрий карбонатный буфер, натрий бикарбонатный буфер, боратный буфер, Трис-буфер, буфер MOPS, буфер HEPES и их комбинации. В некоторых вариантах осуществления настоящего изобретения буфер является щелочным. В некоторых вариантах осуществления настоящего изобретения буфер имеет рН от примерно 7 до примерно 10.

В некоторых вариантах осуществления настоящего изобретения набор включает один или несколько олигонуклеотидов, соответствующих направляющей последовательности для вставки в вектор таким образом, чтобы функционально связать направляющую последовательность и регуляторный элемент. В некоторых вариантах осуществления настоящего изобретения набор включает матричный полинуклеотид для гомологичной рекомбинации. В одном из вариантов осуществления настоящего изобретения предусматривают способы использования одного или нескольких элементов системы RGN. Система RGN

по настоящему изобретению обеспечивает эффективное средство для модификации целевого полинуклеотида. Система RGN по настоящему изобретению, имеет широкий спектр применения, включая модификацию (например, делецию, инсерцию, транслокацию, инактивацию, активацию, редактирование оснований) целевого полинуклеотида во множестве типов 5 клеток. Таким образом, система RGN по настоящему изобретению имеет широкий спектр применения, например, в генной терапии, скрининге лекарств, диагностике заболеваний и прогнозировании. Например, система RGN или комплекс RGN, включает фермент RGN, соединенный с направляющей 10 последовательностью, гибридизованной с целевой последовательностью в составе целевого полинуклеотида.

VIII. Целевые полинуклеотиды

В одном аспекте осуществления настоящего изобретения изобретение, 15 представленное в настоящем изобретении, предусматривает методы модификации целевого полинуклеотида в эукариотической клетке, методы могут быть *in vivo*, *ex vivo* или *in vitro*. В некоторых вариантах осуществления настоящего изобретения методы включает отбор образца клетки или популяции 20 клеток от человека, животного или растения (включая микроводоросли) и модификацию клетки или клеток. Культивирование может происходить на любой стадии *ex vivo*. Клетка или клетки могут быть даже повторно введены в животное или растение (включая микроводоросли).

Используя естественную вариабельность, селекционеры комбинируют наиболее полезные гены для получения желательных свойств, таких как 25 урожайность, качество продукции, однородность, выносливость, устойчивость к вредителям. Эти желательные качества также включают рост, требования к продолжительности светового дня, температурные требования, дату начала цветения или репродуктивного развития, содержание жирных кислот, устойчивость к насекомым, заболеваниям, нематодам, грибам, гербицидам, 30 устойчивость к различным факторам окружающей среды, включая засуху, жару, влажность, холод, ветер и неблагоприятные почвенные условия, включая высокую засоленность почвы. Источники этих полезных генов включают местные или зарубежные сорта, поддерживаемые из поколения в поколение сорта, родственные дикорастущие растения и индуцированные мутации,

например, обработанные мутагенами растительные материалы. Используя настоящее изобретение, селекционеры получают новый инструмент для индуцирования мутаций. Соответственно, специалист в данной области техники может проанализировать геном на наличие источников полезных генов и
5 использовать настоящее изобретение применительно к вариантам с желаемыми характеристиками или признаками для индуцирования роста полезных генов с большей точностью, чем предыдущие мутагенные агенты, и, следовательно, ускорить и улучшить программы селекции растений.

Целевой полинуклеотид системы RGN может быть любым
10 полинуклеотидом, эндогенным или экзогенным для эукариотической клетки. Например, целевой полинуклеотид может быть полинуклеотидом, находящимся в ядре эукариотической клетки. Целевой полинуклеотид может быть последовательностью, кодирующей генный продукт (например, белок) или некодирующей последовательностью (например, регуляторный полинуклеотид
15 или нежелательная ДНК). Не опираясь на какую-либо теорию, в настоящем изобретении предполагают, что целевая последовательность должна быть связана с мотивом, прилегающим к протоспейсеру (PAM – Protospacer Adjacent Motif), то есть короткой последовательностью, распознаваемой системой RGN. Точные требования к последовательности и длине PAM отличаются в
20 зависимости от используемой RGN, но обычно PAM представляет собой последовательность из 2-5 пар оснований, примыкающих к протоспейсеру (то есть целевой последовательности).

Целевой полинуклеотид системы RGN может включать ряд генов и полинуклеотидов, связанных с заболеванием, а также генов и полинуклеотидов,
25 связанных с сигнальными биохимическими путями. Примеры целевых полинуклеотидов включают последовательность, связанную с сигнальным биохимическим путем, например, ген или полинуклеотид, связанный с сигнальным биохимическим путем. Примеры целевых полинуклеотидов включают ген или полинуклеотид, ассоциированный с заболеванием.
30 Используемый в настоящем изобретении термин «ассоциированный с заболеванием» ген или полинуклеотид, относится к любому гену или полинуклеотиду, который продуцирует продукты транскрипции или трансляции на аномальном уровне или в аномальной форме в клетках, полученных из тканей, пораженных заболеванием, по сравнению с контрольными тканями или

здоровыми клетками без заболевания. Это может быть ген, который экспрессируется на аномально высоком уровне; это может быть ген, который экспрессируется на аномально низком уровне, где измененная экспрессия коррелирует с возникновением и/или прогрессированием заболевания.

5 Используемый в настоящем изобретении термин ген, ассоциированный с заболеванием, также относится к гену, обладающему мутацией (мутациями) или генетической вариацией, которая непосредственно ответственна или находится в состоянии неравновесного сцепления гена (генов), ответственного за этиологию заболевания (например, вызывающая мутация). Продукты транскрипции или
10 трансляции могут быть известными или неизвестными, и могут находиться на нормальном или аномальном уровне. В некоторых вариантах осуществления настоящего изобретения заболевание может быть заболеванием животного. В некоторых вариантах осуществления настоящего изобретения заболевание может быть заболеванием птиц. В некоторых вариантах осуществления
15 настоящего изобретения заболевание может быть заболеванием млекопитающих. В некоторых вариантах осуществления настоящего изобретения заболевание может быть заболеванием человека. Примеры генов и полинуклеотидов, ассоциированных с заболеваниями человека, доступны из баз данных Института генетической медицины Маккьюсика-Натанса, Университет Джонса Хопкинса
20 (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University) (Балтимор, штат Мэриленд) и Национального центра биотехнологической информации, Национальной медицинской библиотеки (National Center for Biotechnology Information, National Library of Medicine) (Бетесд, штат Мэриленд), которые доступны по интернету.

25 Хотя системы RGN особенно полезны благодаря своей относительной простоте нацеливания на представляющие интерес геномные последовательности, все еще остается вопрос о том, что RGN может сделать для устранения вызывающей мутации. Один из подходов заключается в получении
30 слитого белка между RGN (предпочтительно неактивным или никасным вариантом RGN) и ферментом редактирования оснований или активным доменом фермента редактирования оснований, например, цитидиндезаминазой или аденозиндезаминазой (US 9840699, сущность приведенного патента полностью включена в настоящее изобретение в виде ссылки). В некоторых вариантах осуществления настоящего изобретения изложенные методы

предполагают контакт молекулы ДНК со (а) слитым белком, содержащим RGN по настоящему изобретению и полипептид, редактирующий основания, такой как дезаминаза; и (б) gRNA, нацеливающей слитый белок (а) на целевую нуклеотидную последовательность цепи ДНК; при этом молекула ДНК контактирует со слитым белком и gRNA в количестве и в условиях, эффективных для дезаминирования нуклеотидных оснований. В некоторых вариантах осуществления настоящего изобретения, последовательность ДНК-мишени включает последовательность, ассоциированную с заболеванием или расстройством, и в которой дезаминирование нуклеотидных оснований приводит к последовательности, которая не связана с заболеванием или расстройством. В некоторых вариантах осуществления настоящего изобретения, целевая последовательность ДНК находится в аллеле культурного растения, при этом конкретный аллель интересующего признака приводит к получению растения с меньшей агрономической ценностью. В результате дезаминирования нуклеотидных оснований образуется аллель, который улучшает признак и повышает агрономическую ценность растения.

В некоторых вариантах осуществления настоящего изобретения последовательность ДНК включает точечную мутацию $T \rightarrow C$ или $A \rightarrow G$, связанную с заболеванием или расстройством, и где дезаминирование мутантного основания C или G приводит к последовательности, которая не связана с заболеванием или расстройством. В некоторых вариантах осуществления, дезаминирование исправляет точечную мутацию в последовательности, связанную с заболеванием или расстройством.

В некоторых вариантах осуществления настоящего изобретения, последовательность, связанная с заболеванием или расстройством, кодирует белок, и при этом дезаминирование вводит стоп-кодон в последовательность, связанную с заболеванием или расстройством, что приводит к отсечению кодируемого белка. В некоторых вариантах осуществления настоящего изобретения контакт осуществляют *in vivo* у субъекта восприимчивого к наличию или с диагностированным заболеванием или расстройством. В некоторых вариантах осуществления настоящего изобретения заболевание или расстройство представляет собой заболевание, связанное с точечной мутацией или мутацией одного основания в геноме. В некоторых вариантах осуществления настоящего изобретения заболевание представляет собой

генетическое заболевание, злокачественную опухоль, метаболическое заболевание или лизосомную болезнь накопления.

IX. Фармацевтические составы и способы лечения

5 В этом разделе представлены фармацевтические составы, содержащие описанные в настоящем изобретении полипептиды RGN и их активные варианты, и фрагменты, а также кодирующие их полинуклеотиды, описанные в настоящем изобретении gRNA или кодирующие их полинуклеотиды, описанные в настоящем изобретении системы или клетки, содержащие любые из
10 полипептидов RGN или кодирующих RGN полинуклеотиды, gRNA или кодирующие gRNA полинуклеотиды или системы RGN, а так же фармацевтически приемлемый носитель.

Фармацевтический состав – это композиция, которая используется для профилактики, снижения интенсивности, излечения или иного лечения
15 заданного состояния или заболевания, и которая содержит активный ингредиент (т.е. полипептиды RGN, кодирующие RGN полинуклеотиды, gRNA, кодирующие gRNA полинуклеотиды, системы RGN или клетки, содержащие любой из вышеперечисленных компонентов) и фармацевтически приемлемый носитель.

Используемый в настоящем изобретении термин «фармацевтически приемлемый носитель» относится к материалу, который не вызывает
20 значительного раздражения в организме и не отменяет активность и свойства активного ингредиента (т.е. полипептидов RGN, полинуклеотидов, кодирующих RGN, gRNA, полинуклеотидов, кодирующих gRNA, систем RGN или клеток, содержащих любой из вышеперечисленных компонентов). Фармацевтически приемлемые носители должны обладать достаточно высокой чистотой и
25 достаточно низкой токсичностью, чтобы сделать их пригодными для введения субъекту, проходящему лечение. Носитель может быть инертным или обладать фармацевтическими преимуществами. В некоторых вариантах осуществления настоящего изобретения фармацевтически приемлемый носитель содержит один
30 или более совместимых твердых или жидких наполнителей, разбавителей или инкапсулирующих веществ, которые подходят для введения человеку или позвоночному животному. В некоторых вариантах фармацевтически приемлемый носитель не является природным. В некоторых вариантах

осуществления настоящего изобретения фармацевтически приемлемый носитель и активный ингредиент не встречаются вместе в природе.

Фармацевтические составы, используемые в предусмотренных в настоящем изобретении способах, могут быть приготовлены с использованием подходящих носителей, вспомогательных веществ и других агентов, которые обеспечивают подходящий перенос, доставку, переносимость и т.п. Множество подходящих составов известно специалистам в данной области техники. См., например, кн.: Remington, «The Science, Practice of Pharmacy», (21-е изд. 2005 г.). Подходящие составы включают, например, порошки, пасты, мази, желе, воска, масла, липиды, липидные (катионные или анионные) везикулы (такие как везикулы LIPOFECTIN), липидные наночастицы, конъюгаты ДНК, безводные абсорбирующие пасты, эмульсии масло-в-воде и вода-в-масле, эмульсии карбовакс (полиэтиленгликоли различной молекулярной массы), полутвердые гели и полутвердые смеси, содержащие карбовакс. Фармацевтические составы для перорального или парентерального применения могут быть приготовлены в виде лекарственных форм в единичной дозе, подходящей для дозировки активных ингредиентов. Такие лекарственные формы в единичной дозе включают, например, таблетки, драже, капсулы, инъекции (ампулы), суппозитории и т.д.

В некоторых вариантах осуществления настоящего изобретения, когда клетки, включающие или модифицированные представленными в настоящем изобретении RGN, gRNA, системами RGN или кодирующими их полинуклеотидами, вводятся субъекту, клетки вводятся в виде суспензии с фармацевтически приемлемым носителем. Специалисты в данной области техники понимают, что фармацевтически приемлемый носитель для использования в клеточном составе не будет включать буферы, соединения, агенты криоконсервации, консерванты или другие агенты в количествах, которые существенно нарушают жизнеспособность клеток, доставляемых субъекту. Состав, содержащий клетки, может включать, например, осмотические буферы, которые позволяют поддерживать целостность клеточной мембраны, и, по желанию, питательные вещества для поддержания жизнеспособности клеток или улучшения приживания при введении. Такие составы и суспензии известны специалистам в области техники, к которой относится настоящее изобретение

и/или могут быть адаптированы для использования с клетками, описанными в настоящем изобретении, с помощью обычных экспериментов.

Клеточный состав также может быть эмульгирован или представлен в виде липосомной композиции при условии, что процедура эмульгирования не оказывает отрицательного влияния на жизнеспособность клеток. Клетки и любой другой активный ингредиент могут быть смешаны с наполнителями, которые являются фармацевтически приемлемыми и совместимыми с активным ингредиентом, в количествах, подходящих для использования в терапевтических способах, описанных в настоящем изобретении.

Дополнительные агенты, включенные в клеточную композицию, могут содержать фармацевтически приемлемые соли содержащихся в ней компонентов. Фармацевтически приемлемые соли включают соли присоединения кислоты (образованные со свободными аминогруппами полипептида), которые образуются с неорганическими кислотами, такими как, например, соляная или фосфорная кислоты, или с органическими кислотами, такими как уксусная, винная, миндальная и подобные. Соли, образованные свободными карбоксильными группами, также могут быть получены из неорганических оснований, таких как, например, гидроксид натрия, калия, аммония, кальция или трехвалентного железа, и таких органических оснований, как изопропиламин, триметиламин, 2-(этиламино)этанол, гистидин, прокаин и т.п.

Физиологически переносимые и фармацевтически приемлемые носители хорошо известны в области техники, к которой относится настоящее изобретение. Примеры жидких носителей представляют собой стерильные водные растворы, которые не содержат никаких материалов в дополнение к активным ингредиентам или содержат буфер, такой как фосфат натрия при физиологическом значении pH, физиологический солевой раствор или оба, такие как физиологический раствор с фосфатным буфером. Кроме того, водные носители могут содержать более одной буферной соли, а также такие соли, как хлорид натрия и калия, декстрозу, полиэтиленгликоль и другие растворители. Жидкие составы также могут содержать жидкие фазы в дополнение к воде и без нее. Примерами таких дополнительных жидких фаз являются глицерин, растительные масла, такие как хлопковое масло, и водно-масляные эмульсии. Количество активного соединения, используемого в клеточных составах,

которое эффективно при лечении конкретного расстройства или состояния, может зависеть от природы расстройства или состояния и может быть определено стандартными клиническими методами.

Представленные в настоящем изобретении полипептиды RGN, направляющие РНК, системы RGN или полинуклеотиды, кодирующие их, могут быть соединены с фармацевтически приемлемыми вспомогательными веществами, такими как носители, растворители, стабилизаторы, адъюванты, разбавители и т.д., в зависимости от конкретного способа введения и лекарственной формы. В некоторых вариантах осуществления настоящего изобретения эти фармацевтические составы разрабатываются для достижения физиологически совместимого рН, который варьируется от рН примерно 3 до рН примерно 11, от рН примерно 3 до рН примерно 7, в зависимости от формулы и способа введения. В некоторых вариантах осуществления настоящего изобретения рН может быть отрегулирован в диапазоне от рН примерно 5,0 до рН примерно 8. В некоторых вариантах осуществления настоящего изобретения составы могут включать терапевтически эффективное количество по меньшей мере одного соединения, описанного в настоящем изобретении, вместе с одним или более фармацевтически приемлемым эксципиентом. В некоторых вариантах осуществления настоящего изобретения составы включают комбинацию описанных в изобретении соединений или второй активный ингредиент, полезный для лечения или предотвращения роста бактерий (в качестве примера, но ими не ограничиваясь, аскорбиновую кислоту, антибактериальные или антимикробные агенты) или включают комбинацию реагентов представленных в настоящем изобретении.

Подходящие вспомогательные вещества включают, например, молекулы-носители, которые включают крупные, медленно метаболизируемые макромолекулы, такие как белки, полисахариды, полимолочные кислоты, полигликолевые кислоты, полимерные аминокислоты, сополимеры аминокислот и неактивные вирусные частицы. Другие примеры вспомогательных веществ включают антиоксиданты (в качестве примера, но ею не ограничиваясь, аскорбиновую кислоту), хелатообразующие агенты (в качестве примера, но ею не ограничиваясь, ЭДТА), углеводы (в качестве примера, но ими не ограничиваясь, декстрин, гидроксилалкилцеллюлозу и гидроксилалкилметилцеллюлозу), стеариновую кислоту, жидкости (в качестве

примера, но ими не ограничиваясь, масла, воду, физиологический раствор, глицерин и этанол), смачивающие или эмульгирующие агенты, вещества, повышающие рН, и т.п.

В некоторых вариантах осуществления настоящего изобретения составы
5 поставляются в контейнерах для единичных доз или нескольких доз, например, в герметичных ампулах и флаконах, и могут храниться в сублимированном (лиофилизированном) состоянии, требующем добавления стерильного жидкого носителя, например, физиологического раствора, воды для инъекций, полужидкой пены или геля непосредственно перед использованием. Растворы и
10 суспензии для внутривенных инъекций могут быть приготовлены из стерильных порошков, гранул и таблеток описанного ранее типа. В некоторых вариантах осуществления настоящего изобретения активный ингредиент растворяют в забуференном жидком растворе, который замораживают в контейнере с единичной дозой или несколькими дозами, а затем размораживают для инъекции
15 или хранят/стабилизируют в холодильнике до использования.

Терапевтический агент (агенты) может содержаться в системах контролируемого высвобождения биологических объектов. Для продления действия лекарства часто желательно замедлить всасывание лекарства при
20 подкожном, интратекальном или внутримышечном введении. Это может быть достигнуто путем использования жидкой суспензии кристаллического или аморфного материала с плохой растворимостью в воде. Скорость всасывания препарата зависит от скорости его растворения, которая, в свою очередь, может зависеть от размера кристаллов и кристаллической формы. В качестве
25 альтернативы отсроченное всасывание парентерально вводимой лекарственной формы достигается путем растворения или суспендирования лекарственного средства в масляном средстве для введения. В некоторых вариантах осуществления настоящего изобретения использование имплантата длительного устойчивого высвобождения может быть особенно подходящим для лечения хронических заболеваний. Имплантаты длительного устойчивого
30 высвобождения хорошо известны специалистам в области техники, к которой относится настоящее изобретение.

В настоящем изобретении представлены методы лечения заболевания у субъекта, нуждающегося в этом. Методы включают введение субъекту, нуждающемуся в лечении, эффективного количества описанного в настоящем

изобретении полипептида RGN или его активного варианта или фрагмента или полинуклеотида, кодирующего его, описанного в настоящем изобретении gRNA или полинуклеотида, кодирующего его, описанной в настоящем изобретении системы RGN или модифицированных клеток или клеток содержащих любую из
5 вышеназванных композиции.

В некоторых вариантах осуществления настоящего изобретения лечение включает редактирование генов *in vivo* путем введения описанного в настоящем изобретении полипептида RGN, gRNA или системы RGN или полинуклеотида (ов), кодирующего их. В некоторых вариантах осуществления настоящего
10 изобретения лечение включает редактирование генов *ex vivo*, при котором клетки генетически модифицируют *ex vivo* с помощью описанного в настоящем изобретении полипептида RGN, gRNA или системы RGN или полинуклеотида (ов), их кодирующих, затем модифицированные клетки вводят субъекту. В некоторых вариантах осуществления настоящего изобретения, генетически
15 модифицированные клетки берут от субъекта, которому затем вводят модифицированные клетки, пересаженные клетки упоминаются здесь как аутологичные. В некоторых вариантах осуществления настоящего изобретения генетически модифицированные клетки берут от другого субъекта (т.е. донора) в пределах того же вида, что и субъект, которому вводят модифицированные
20 клетки (т.е. реципиент), и пересаженные клетки упоминаются здесь как аллогенные. В некоторых примерах, описанных в настоящем изобретении, клетки могут быть размножены в культуре перед введением субъекту, нуждающемуся в этом.

В некоторых вариантах осуществления настоящего изобретения
25 заболевание, подлежащее лечению с помощью описанных в настоящем изобретении композиций, является заболеванием, которое можно лечить с помощью иммунотерапии, такой как технология использования химерного антигенного рецептора (Chimeric Antigen Receptor- CAR) Т-клетки. Такие заболевания включают, но ими не ограничиваются, злокачественные опухолевые
30 образования. В некоторых вариантах осуществления настоящего изобретения заболевание, подлежащее лечению с помощью представленных в настоящей работе композиций, связано с причинно-следственной мутацией. Используемый в настоящем изобретении термин, «вызывающая мутация» относится к конкретному нуклеотиду, нуклеотидам или последовательности нуклеотидов в

геноме, которая вносит вклад в тяжесть заболевания или наличие заболевания или расстройства у субъекта. Коррекция вызывающей мутации приводит к улучшению по меньшей мере одного симптома, возникающего в результате заболевания или расстройства. В некоторых вариантах вызывающая мутация находится рядом с сайтом РАМ, распознаваемым RGN по настоящему изобретению. Вызывающая мутация может быть исправлена с помощью представленной в настоящем изобретении RGN или химерного полипептида, содержащего представленную в настоящем изобретении RGN и полипептид, редактирующий основание (т.е. редактор оснований). Примеры заболеваний, связанных с вызывающей мутацией, включают, но ими не ограничиваются, муковисцидоз, синдром Херлера, атаксию Фридрейха, болезнь Гентингтона и серповидно клеточную анемию. В некоторых вариантах осуществления настоящего изобретения заболевание, подлежащее лечению с помощью представленных в настоящем изобретении RGN, является заболеванием из табл. 11. Дополнительные, но ими не ограничивающиеся, примеры генов и мутаций, связанных с заболеваниями, доступны из баз Института генетической медицины МакКюсика-Натанса, Университет Джонса Хопкинса (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University) (Балтимор, штат Мэриленд) и Национального центра биотехнологической информации, Национальной медицинской библиотеки (National Center for Biotechnology Information, National Library of Medicine) (Бетесд, штат Мэриленд), базы данных доступны через Интернет.

Используемый в настоящем изобретении термин «лечение» или «лечащий» или «смягчающий» или «улучшающий» используется взаимозаменяемо. Эти термины относятся к подходу для получения полезных или желаемых результатов, включая, но ими не ограничиваясь, терапевтическую пользу и/или профилактическую пользу. Под терапевтической пользой понимается любое терапевтически значимое улучшение или воздействие на одно или несколько заболеваний, состояний или симптомов, находящихся на лечении. С профилактической целью составы могут вводиться субъекту, подверженному риску развития конкретного заболевания, состояния или симптома или субъекту, сообщающему об одном или более физиологических симптомах заболевания, даже если заболевание, состояние или симптом, возможно, еще не проявились.

Используемый в настоящем изобретении термин «эффективное количество» или «терапевтически эффективное количество» относится к количеству агента, которое является достаточным для достижения полезных или желаемых результатов. Терапевтически эффективное количество может варьироваться в зависимости от одного или более параметров: субъекта и болезненного состояния, подлежащего лечению, веса и возраста субъекта, тяжести болезненного состояния, способа введения и т.п., которые могут быть легко определены специалистом в данной области техники, к которой относится настоящее изобретение. Конкретная доза может варьировать в зависимости от одного или более параметров: конкретного выбранного средства, режима дозирования, которого следует придерживаться, вводится ли средство в комбинации с другими соединениями, времени введения и системы, в которой осуществляется доставка соединения.

Используемый в настоящем изобретении термин «введение» относится к введению активного ингредиента субъекту способом или путем, который приводит, по меньшей мере, к частичной локализации введенного активного ингредиента в желаемом месте, таком как место повреждения или восстановления, таким образом, что достигается желаемый эффект (или эффекты). В тех вариантах осуществления настоящего изобретения, в которых вводят клетки, клетки могут вводиться любым подходящим путем, который приводит к доставке в желаемое место субъекта, в котором, по меньшей мере часть имплантированных клеток или компонентов клеток остается жизнеспособной. Период жизнеспособности клеток после введения субъекту может составлять от нескольких часов, например, двадцати четырех часов, до нескольких дней, до нескольких лет или даже продолжительности жизни пациента, т.е. долгосрочного приживания. Например, в некоторых аспектах, описанных в настоящем изобретении эффективное количество фоторецепторных клеток или клеток-предшественников сетчатки вводят системным путем, таким как внутрибрюшинное или внутривенное введение.

В некоторых вариантах осуществления настоящего изобретения введение предполагает введение посредством вирусной доставки. В некоторых вариантах осуществления настоящего изобретения введение предполагает введение посредством электропорации. В некоторых вариантах осуществления настоящего изобретения введение предполагает введение посредством доставки

наночастицами. В некоторых вариантах осуществления настоящего изобретения введение предполагает введение посредством доставки липосомами. Любой эффективный способ введения может быть использован для введения эффективного количества фармацевтического состава, описанного в настоящем изобретении. В некоторых вариантах осуществления настоящего изобретения введение предполагает введение нижеперечисленными методами, а именно: внутривенно, подкожно, внутримышечно, перорально, ректально, с помощью аэрозоля, парентерально, офтальмологически, легочно, трансдермально, вагинально, орально, назально и путем местного введения или любой комбинации этих методов. В некоторых вариантах осуществления настоящего изобретения для доставки клеток используется введение путем инъекции или инфузии.

Используемый в настоящем изобретении термин «субъект» относится к любому индивидууму, для которого желательны диагностика, лечение или терапия. В некоторых вариантах осуществления настоящего изобретения субъектом является животное. В некоторых вариантах осуществления настоящего изобретения субъект является млекопитающим. В некоторых вариантах осуществления настоящего изобретения субъектом является человек.

Эффективность лечения может быть определена квалифицированным клиницистом. Однако лечение считается «эффективным лечением», если какой-либо один или все признаки или симптомы заболевания или расстройства изменены благоприятным образом (например, уменьшены по меньшей мере на 10%) или улучшены или ослаблены другие клинически признанные симптомы или маркеры заболевания. Эффективность также может быть измерена на том основании, что состояние пациента не ухудшается, что подтверждается при госпитализации или необходимости медицинского вмешательства (например, прогрессирование заболевания остановлено или, по крайней мере, замедлено). Методы измерения этих показателей известны специалистам в данной области техники. Лечение включает в себя: (1) подавление заболевания, например, прекращение или замедление прогрессирования симптомов; или (2) облегчение заболевания, например, регресс симптомов; и (3) предотвращение или уменьшение вероятности развития симптомов.

А. Модификация вызывающих мутаций с помощью редактирования оснований

Примером генетически наследуемого заболевания, которое может быть скорректировано с помощью подхода, основанного на использовании слитого белка RGN-редактор оснований по настоящему изобретению, является синдром Гурлера. Синдром Гурлера, также известный как мукополисахаридоз 1 (МПС-1), является результатом дефицита α -L-идуронидазы (IDUA), что приводит к лизосомальной болезни накопления, характеризующейся на молекулярном уровне накоплением дерматансульфата и гепарансульфата в лизосомах. Это заболевание обычно является наследственным генетическим обусловленным, вызванным мутациями в гене IDUA, кодирующем α -L-идуронидазу. Распространенными мутациями IDUA являются W402X и Q70X, обе нонсенс-мутации приводят к преждевременной остановке трансляции. Такие мутации хорошо поддаются точному редактированию генома (PGE), поскольку реверсия одного нуклеотида, например, с помощью редактирования основания, восстанавливает кодирующую последовательность дикого типа и приводит к экспрессии белка под контролем эндогенных регуляторных механизмов генетического локуса. Кроме того, поскольку известно, что гетерозиготные формы протекают бессимптомно, PGE терапия, направленная на одну из этих мутаций, была бы полезна для значительной части пациентов с этим заболеванием, поскольку необходимо корректировать только один из мутировавших аллелей (Bunge и соавт., *Hum. Mol. Genet.*, 1994, 3(6): 861-866; сущность публикации полностью включена в настоящее изобретение в виде ссылки).

Современные методы лечения синдрома Гурлера включают заместительную ферментную терапию и трансплантацию костного мозга (Vellodi и соавт., *Arch. Dis. Child.*, 1997, 76(2): 92-99; Peters и соавт., *Blood*, 1998, 91(7): 2601-2608; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки). Хотя заместительная ферментная терапия оказала значительное влияние на выживаемость и качество жизни пациентов с синдромом Гурлера, этот подход требует дорогостоящих и трудоемких еженедельных инфузий. Другие методы лечения синдрома Гурлера включают доставку гена IDUA на векторе экспрессии или вставку гена в высоко-экспрессируемый локус, такой как локус сывороточного альбумина (US 9956247

сущность патента полностью включена в настоящее изобретение в виде ссылки). Однако эти методы не восстанавливают исходный локус IDUA до правильной кодирующей последовательности. Стратегия редактирования генома имеет ряд преимуществ, в частности, регуляция экспрессии генов будет контролироваться естественными механизмами, присутствующими у здоровых людей. Кроме того, при редактировании оснований не требуется вызывать двухцепочечные разрывы ДНК, что может привести к большим хромосомным перестройкам, гибели клеток или онкогенезу из-за нарушения механизмов подавления опухоли. Общая стратегия может быть направлена на использование слитых белков RGN-редактор оснований, представленных в настоящем изобретении, для нацеливания и исправления определенных мутаций, вызывающих заболевания в геноме человека. Следует понимать, для лечения заболеваний, которые могут быть скорректированы путем редактирования оснований, также могут быть применены аналогичные методы. Кроме того, следует отметить, что аналогичные методы нацеливания на мутации, вызывающие заболевание, мутации у других видов, в частности, обычных домашних животных или сельскохозяйственных животных, также могут быть применены с использованием RGN, представленных в настоящем изобретении. Обычные домашние животные и сельскохозяйственные животные включают собак, кошек, лошадей, свиней, коров, овец, кур, ослов, змей, хорьков и рыбу, включая лосося и креветок.

Б. Изменение причинно-следственных мутаций путем целевого удаления RGN, представленные в настоящем изобретении, также могут быть полезны в терапевтических подходах к лечению такого заболевания человека, в котором причинно-следственная мутация является более сложной. Например, некоторые заболевания, такие как атаксия Фридрейха и болезнь Гентингтона, являются результатом значительного увеличения числа повторов трехнуклеотидного мотива в определенной области гена, что влияет на способность экспрессируемого белка функционировать или быть экспрессируемым. Атаксия Фридрейха (FRDA) - аутосомно-рецессивное заболевание, приводящее к прогрессирующей дегенерации нервной ткани в спинном мозге. Пониженный уровень белка фратаксина (FXN) в митохондриях вызывает окислительные повреждения и приводит к дефициту железа на клеточном уровне. Пониженная

экспрессия FXN связана с экспансией триплета GAA в интроне 1 гена FXN соматической и зародышевой линии. У пациентов с FRDA повтор GAA часто состоит из более чем 70, иногда даже более чем 1000 (чаще всего 600-900) триплетов, в то время как у здоровых людей таких повторов примерно 40 или меньше (Pandolfo и соавт., *Handbook of Clinical Neurology*, 2012 103: 275-294; Campuzano и соавт., *Science*, 1996, 271: 1423-1427; Pandolfo, *Adv. Exp. Med. Biol.*, 2002, 516: 99-118; сущность каждой публикации полностью включена в настоящее изобретение в виде ссылки).

Расширение последовательности тринуклеотидных повторов, вызывающее атаксию Фридрейха (FRDA – Friedreich’s Ataxia), происходит в определенном генетическом локусе в гене FXN, называемом областью нестабильности FRDA. РНК-направляемые нуклеазы (RGN) могут быть использованы для удаления области нестабильности в клетках пациентов с FRDA. Этот подход требует 1) RGN и направляющей РНК, последовательность которой может быть запрограммирована на нацеливание на аллель в геноме человека; и 2) подхода к доставке RGN и направляющей последовательности РНК. Многие нуклеазы, используемые для редактирования генома, такие как широко используемая нуклеаза Cas9 из *S. pyogenes* (SpCas9), слишком велики для упаковки в адено-ассоциированные вирусные (AAV – adeno-associated viral) векторы, особенно если учесть длину гена SpCas9 и направляющей РНК, в дополнение к другим генетическим элементам, необходимым для функциональной экспрессии кассеты. Это усложняет подход с использованием SpCas9.

Некоторые РНК-направляемые нуклеазы настоящего изобретения хорошо подходят для упаковки в вектор AAV вместе с направляющей РНК. Упаковка двух направляющих РНК, вероятно, потребует второго вектора, но этот подход все равно выгодно отличается от того, что потребуются от более крупной нуклеазы, такой как SpCas9, которая может потребовать разделения белковой последовательности между двумя векторами. Настоящее изобретение охватывает стратегию с использованием RGN по настоящему изобретению, в которой удаляют область геномной нестабильности. Такая стратегия применима к другим заболеваниям и расстройствам, имеющим аналогичную генетическую основу, таким как болезнь Гентингтона. Аналогичные стратегии с использованием RGN, представленной в настоящем изобретении, также могут быть применимы к аналогичным заболеваниям и расстройствам у животных, имеющих

агрономическое или экономическое значение, включая собак, кошек, лошадей, свиней, коров, овец, кур, ослов, змей, хорьков и рыб, включая лосося и креветок.

В. Изменение вызывающих мутаций путем направленного мутагенеза

5 RGN, представленные в настоящем изобретении, также могут быть использованы для введения разрушительных мутаций, которые могут привести к благоприятному эффекту. Генетические дефекты в генах, кодирующих гемоглобин, в частности цепь бета-глобина (ген HBB), могут быть причиной некоторых заболеваний, известных как гемоглобинопатия, включая серповидно
10 клеточную анемию и талассемию.

У взрослых людей гемоглобин представляет собой гетеротетрамер, состоящий из двух альфа (α)-подобных глобиновых цепей и двух бета (β)-подобных глобиновых цепей и 4 групп гема. У взрослых тетрамер $\alpha_2\beta_2$ называется гемоглобином А (HbA) или нормальным гемоглобином взрослого
15 человека. Как правило, альфа- и бета-глобиновые цепи синтезируются в соотношении примерно 1:1, и это соотношение, по-видимому, является критически с точки зрения стабилизации гемоглобина и красных кровяных телец (RBC – red blood cells). У развивающегося плода вырабатывается другая форма гемоглобина, фетальный гемоглобин (HbF), который имеет более высокое
20 сродство к кислороду, чем гемоглобин А, что позволяет доставлять кислород в организм ребенка через кровотоки матери. Фетальный гемоглобин также содержит две α -глобиновые цепи, но вместо взрослых β -глобиновых цепей он имеет две фетальные гамма (γ)-глобиновые цепи (т.е. фетальный гемоглобин – это $\alpha_2\gamma_2$). Регуляция перехода от производства гамма-глобина к производству
25 бета-глобина довольно сложна, и в основном включает в себя понижающую регуляцию транскрипции гамма-глобина с одновременной повышающей регуляцией транскрипции бета-глобина. Примерно на 30й неделе беременности синтез гамма-глобина у плода начинает снижаться, в то время как производство бета-глобина увеличивается. Примерно к 10 месяцам жизни новорожденного
30 гемоглобин почти полностью состоит из $\alpha_2\beta_2$, хотя некоторое количество HbF сохраняется до взрослого возраста (примерно 1-3% от общего количества гемоглобина). У большинства пациентов с гемоглобинопатией присутствуют гены, кодирующие гамма-глобин, но их экспрессия относительно низкая из-за нормальной репрессии генов, происходящей после родов, как описано выше.

Серповидно клеточная анемия вызывается мутацией V6E в гене β -глобина (HBB) (на уровне ДНК GAG в GTG), полученный в результате гемоглобин называется «гемоглобин S» или «HbS». В условиях пониженного содержания кислорода молекулы HbS агрегируют и образуют волокнистые преципитаты. Эти агрегаты вызывают аномалию или «серповидность» красных кровяных телец, что приводит к потере гибкости клеток. Серповидные красные кровяные тельца больше не могут проникать в капилляры, что может привести к вазоокклюзивному кризу у пациентов с серповидно клеточной анемией. Кроме того, серповидные красные кровяные тельца более хрупкие, чем нормальные красные кровяные тельца, и склонны к гемолизу, что в конечном итоге приводит к анемии у пациента.

Лечение и ведение пациентов с серповидно клеточной анемией – это дело всей жизни, включающее антибиотикотерапию, обезболивание и переливание крови во время острых эпизодов. Одним из подходов к лечению является применение гидроксимочевины, которая оказывает свое лечебное действие частично за счет увеличения выработки гамма-глобина. Однако долгосрочные побочные эффекты хронической терапии гидроксимочевинной до сих пор неизвестны, лечение дает нежелательные побочные эффекты и может иметь различную эффективность у разных пациентов. Несмотря на повышение эффективности лечения серповидно клеточной анемии, продолжительность жизни пациентов по-прежнему составляет всего лишь примерно 50 лет, а сопутствующие заболевания оказывают глубокое влияние на качество жизни пациента.

Талассемия (альфа-талассемия и бета-талассемия) также является заболеванием, связанными с гемоглобином, и обычно связано с уменьшением экспрессии глобиновых цепей. Это может произойти из-за мутаций в регуляторных областях генов или из-за мутации последовательности, кодирующей глобин, которые приводят к снижению экспрессии или снижению уровней или функциональности глобина. Лечение талассемии обычно включает переливание крови и хелатную терапию железом. Трансплантация костного мозга также используется для лечения людей с тяжелой талассемией, если удастся найти подходящего донора, но эта процедура может быть сопряжена со значительными рисками.

Один из подходов, который был предложен для лечения как серповидно-клеточной анемии (SCD – sickle cell disease) так и бета-талассемии, заключается в увеличении экспрессии гамма-глобина таким образом, чтобы HbF функционально заменял аберрантный взрослый гемоглобин. Как упоминалось выше, лечение больных серповидно-клеточной анемией гидроксимочевиной считается успешным отчасти благодаря опосредованному ею эффекту увеличения экспрессии гамма-глобина (DeSimone, *Proc. Nat'l. Acad. Sci. USA*, 1982, 79(14):4428-31; Ley, и соавт., *N. Engl. J. Medicine*, 1982, 307: 1469-1475; Ley и соавт., *Blood*, 1983, 62: 370-380; Constantoulakis и соавт., *Blood*, 1988, 72(6):1961-1967; сущность каждой приведенной публикации полностью включена в настоящее изобретение в виде ссылки). Повышение экспрессии HbF включает идентификацию генов, продукты которых играют роль в регуляции экспрессии гамма-глобина. Одним из таких генов является BCL11A. BCL11A кодирует белок, содержащий структурный мотив «цинковые пальцы», который экспрессируется во взрослых эритроидных клетках-предшественниках, понижающая регуляция его экспрессии приводит к увеличению экспрессии гамма-глобина (Sankaran и соавт., *Science*, 2008, 322: 1839, сущность публикации полностью включена в настоящее изобретение в виде ссылки). Было предложено использовать ингибирующую РНК, направленную на ген BCL11A (например, см. US 2011/0182867, сущность патента полностью включена в настоящее изобретение в виде ссылки), но эта технология имеет несколько потенциальных недостатков, в том числе заключающихся в том, что полный нокаут может быть не достигнут, доставка таких РНК может быть проблематичной, и кроме того, РНК должны присутствовать постоянно, что требует многократного лечения в течение жизни.

RGN, представленные в настоящем изобретении, могут быть использованы для нацеливания на энхансерную область BCL11A, для того, чтобы нарушить экспрессию BCL11A и тем самым увеличить экспрессию гамма-глобина. Это целевое нарушение может быть достигнуто путем негомологичного концевоего соединения (NHEJ), при котором RGN, представленные в настоящем изобретении, нацеливаются на определенную последовательность в области энхансера BCL11A и делают двухцепочечный разрыв, а клеточный механизм восстанавливает разрыв, обычно одновременно внося разрушительные мутации. Подобно тому, как это описано для других заболеваний, RGN, представленные в

настоящем изобретении, могут иметь преимущества перед другими известными RGN из-за их относительно небольшого размера, что позволяет упаковывать кассеты экспрессии для RGN и ее направляющей РНК в один вектор AAV для доставки *in vivo*. Подобные стратегии с использованием RGN, представленные в
5 настоящем изобретении, могут быть также применимы к аналогичным заболеваниям и расстройствам, как у людей, так и у животных, имеющих агрономическое или экономическое значение.

X. Клетки, содержащие полинуклеотидную генетическую модификацию

10 Настоящее изобретение представляет клетки и организмы, содержащие интересующую последовательность-мишень, которая была модифицирована с использованием технологического процесса, опосредованного RGN, crRNA и/или tracrRNA, как описано в настоящем изобретении. В некоторых вариантах осуществления настоящего изобретения RGN содержит аминокислотную
15 последовательность SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579 или ее активный вариант или фрагмент. В некоторых вариантах осуществления настоящего изобретения направляющая РНК содержит последовательность повтора CRISPR, содержащую нуклеотидную последовательность, указанную как, SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57,
20 64, 71, 77, 84, 90, 97, 104, 111, 118 или 124 или ее активный вариант или фрагмент. В конкретных вариантах осуществления настоящего изобретения направляющая РНК содержит tracrRNA, содержащую нуклеотидную последовательность SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125 или ее активный вариант или фрагмент. Направляющая
25 РНК системы может быть отдельной направляющей РНК или двойной направляющей РНК.

Модифицированные клетки могут быть эукариотическими (например, клетки млекопитающих, растений, насекомых) или прокариотическими. Также предусмотрены органеллы и эмбрионы, содержащие по меньшей мере одну
30 нуклеотидную последовательность, которая была модифицирована с использованием технологического процесса, опосредованного RGN, crRNA и/или tracrRNA, как описано в настоящем изобретении. Генетически модифицированные клетки, организмы, органеллы и эмбрионы могут быть

гетерозиготными или гомозиготными по модифицированной нуклеотидной последовательности.

Хромосомная модификация клетки, организма, органеллы или эмбриона может привести к изменению экспрессии (повышающая или понижающая регуляция), инактивации или экспрессии измененного белкового продукта или интегрированной последовательности. В тех вариантах осуществления настоящего изобретения, в которых хромосомная модификация приводит либо к инактивации гена, либо к экспрессии нефункционального белкового продукта, генетически модифицированная клетка, организм, органелла или эмбрион называются «нокаутированными». Фенотип нокаута может быть результатом делеционной мутации (т.е. удаления по меньшей мере одного нуклеотида), инсерционной мутации (т.е. вставки по меньшей мере одного нуклеотида) или нонсенс мутации (т.е. замены по меньшей мере одного нуклеотида таким образом, что вводится стоп-кодон).

Альтернативно, хромосомная модификация клетки, организма, органеллы или эмбриона может вызвать «сбой», который является результатом хромосомной интеграции нуклеотидной последовательности, кодирующей белок. В некоторых из этих вариантов осуществления настоящего изобретения кодирующая последовательность интегрирована в хромосому таким образом, что хромосомная последовательность, кодирующая белок дикого типа, инактивирована, но экзогенно введенный белок экспрессируется.

В других вариантах осуществления настоящего изобретения хромосомная модификация приводит к получению варианта белкового продукта. Экспрессированный вариант белкового продукта может иметь, по меньшей мере одну аминокислотную замену и/или добавление или удаление по меньшей мере одной аминокислоты. Вариант белкового продукта, кодируемый измененной хромосомной последовательностью, может проявлять измененные характеристики или активности по сравнению с белком дикого типа, включая, но ими не ограничиваясь, измененную ферментативную активность или субстратную специфичность.

В других вариантах осуществления настоящего изобретения хромосомная модификация может приводить к измененному паттерну экспрессии белка. В качестве примера, но ими не ограничиваясь, хромосомные изменения в регуляторных областях, контролирующей экспрессию белкового продукта,

могут приводить к сверхэкспрессии или понижающей регуляции белкового продукта или измененному паттерну тканевой или временной экспрессии.

Модифицированные клетки могут быть выращены в организм, такой как, например, растение, в соответствии с обычными способами. Смотри, например, 5 McCormick и соавт., *Plant Cell Reports*, 1986, 5:81-84. Затем эти растения могут быть выращены, и либо опылены тем же модифицированным штаммом, либо разными штаммами, таким образом, полученный гибрид будет иметь генетическую модификацию. Настоящее изобретение обеспечивает генетически модифицированные семена. Потомство, варианты и мутанты регенерированных 10 растений также включены в объем настоящего изобретения, при условии, что эти части содержат генетическую модификацию. Дополнительно предлагается обработанный растительный продукт или побочный продукт, который сохраняет генетическую модификацию, в том числе, например, соевая мука.

Методы, представленные в настоящем изобретении, могут быть 15 использованы для модификации любого вида растений, включая, но ими не ограничиваясь, однодольные и двудольные растения. Примеры представляющих интерес растений включают, но ими не ограничиваются, кукурузу (маис), сорго, пшеницу, подсолнечник, томаты, крестоцветные, перец, картофель, хлопок, рис, сою, сахарную свеклу, сахарный тростник, табак, ячмень, масличный рапс, 20 *Brassica sp.*, люцерну, рожь, просо, сафлор, арахис, сладкий картофель, маниоку, кофе, кокос, ананас, цитрусовые, кокосовую пальму, чай, банан, авокадо, инжир, гуаву, манго, оливу, папайю, кешью, макадамию, миндаль, овес, овощи, декоративные и хвойные растения.

Овощи включают, но ими не ограничиваются, помидоры, салат-латук, 25 стручковую фасоль, лимскую фасоль, горох и представителей рода *Curcumis*, таких как огурец, канталупа и мускусная дыня. Декоративные растения включают, но ими не ограничиваются, азалию, гортензию, гибискус, розы, тюльпаны, нарциссы, петунии, гвоздику, пуансеттию и хризантему. Предпочтительно, растениями, на которые ориентировано настоящее 30 изобретения, являются культурные растения (например, кукуруза, сорго, пшеница, подсолнечник, томаты, крестоцветные, перец, картофель, хлопок, рис, соя, сахарная свекла, сахарный тростник, табак, ячмень, масличный рапс и т.д.).

Методы, представленные в настоящем изобретении, также могут быть использованы для генетической модификации любого прокариотического вида,

включая, но ими не ограничиваясь, археи и бактерии (например, *Bacillus* sp., *Klebsiella* sp. *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., *Agrobacterium*, *Lactobacillus* sp.).

5 Методы, предусмотренные в настоящем изобретении, могут быть использованы для генетической модификации любого эукариотического вида или его клеток, включая, но ими не ограничиваясь, животные (например, млекопитающие, насекомые, рыбы, птицы и рептилии), грибы, амёбы, водоросли и дрожжи. В некоторых вариантах осуществления настоящего изобретения
10 клетки, которые модифицируют с помощью раскрытых в настоящем изобретении методов, включают клетки кроветворного происхождения, такие как клетки иммунной системы, включая, но ими не ограничиваясь, В-клетки, Т-клетки, природные клетки-киллеры (NK), плюрипотентные стволовые клетки, индуцированные плюрипотентные стволовые клетки, химерные антигенные
15 рецепторные Т-клетки (CAR-T - Chimeric Antigen Receptor T), моноциты, макрофаги и дендритные клетки.

 Клетки, которые были модифицированы, могут быть введены в организм. Эти клетки могут происходить из того же организма (например, человека) в случае аутологичной клеточной трансплантации, когда клетки модифицируются
20 *ex vivo*. В качестве альтернативы клетки могут происходить из другого организма того же вида (например, другого человека) в случае аллогенной клеточной трансплантации.

 В настоящем изобретении употребление существительного в единственном числе может подразумевать также множественное число. Например,
25 «полипептид» означает один или более полипептидов.

 Все публикации и патентные заявки, упомянутые в описании настоящего изобретения, указывают на уровень специалистов в области техники, к которой относится представляемое изобретение. Все публикации и патентные заявки включены в настоящее описание в виде ссылки, как если бы каждая отдельная
30 публикация или патентная заявка были специально и индивидуально указаны для включения в виде ссылки.

 Хотя настоящее изобретение довольно подробно описано и снабжено примерами для ясности понимания, очевидно, что в пределах прилагаемых

вариантов осуществления настоящего изобретения могут практиковаться определенные изменения и модификации.

5 Варианты осуществления настоящего изобретения включают, но ими не ограничиваются:

1. Молекулу нуклеиновой кислоты, содержащую полинуклеотид, кодирующий полипептид РНК-направляемой нуклеазы (RGN), где указанный полинуклеотид содержит нуклеотидную последовательность, кодирующую полипептид RGN, включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

15 где указанный полипептид RGN способен связывать последовательность ДНК-мишени специфическим для РНК-направляющей последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной последовательностью ДНК-мишени;

 где указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

20 2. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 1, в котором указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

25 3. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 1, в котором указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

30 4. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 1, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении

аминокислоты 368, и валин в положении аминокислоты 405 последовательности, представленной как SEQ ID NO: 63.

5. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-4, где указанный полипептид RGN способен
5 расщеплять указанную последовательность ДНК-мишени при связывании.

6. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 5, где указанный полипептид RGN способен вызывать двухцепочечный разрыв.

7. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 5, где указанный полипептид RGN способен вызывать
10 одноцепочечный разрыв.

8. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-4, где указанный полипептид RGN не обладает нуклеазной активностью или является никазой.

15 9. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-8, где полипептид RGN функционально слит с полипептидом, редактирующим основания.

10. Молекулу нуклеиновой кислоты по варианту осуществлению настоящего изобретения 9, где полипептид, редактирующий основания,
20 представляет собой дезаминазу.

11. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 10, где дезаминаза представляет собой цитидиндезаминазу или адениндезаминазу.

25 12. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-11, где полипептид RGN включает один или более сигналов ядерной локализации.

13. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-12, где полипептид RGN кодон-оптимизирован для экспрессии в эукариотической клетке.

30 14. Молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-13, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающему к протоспейсеру (PAM).

15. Вектор, включающий молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-14.

16. Вектор по варианту осуществления настоящего изобретения 15, дополнительно включающий по меньшей мере одну нуклеотидную последовательность, кодирующую указанную gRNA, способную гибридизироваться с указанной последовательностью ДНК-мишени.

17. Вектор по варианту осуществления настоящего изобретения 16, где направляющая РНК выбрана из группы, состоящей из:

а) направляющей РНК, включающей:

10 i) CRISPR РНК, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1;

б) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 8;

в) направляющей РНК, включающей::

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 15;

г) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 22;

 д) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 31;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 29;

 е) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 37; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 38;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 36;

 ж) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 44; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 45;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 43;

з) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 51; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 52;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 50;

и) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 58;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 56;

к) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 64; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 65;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

л) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 71; и

5 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 72;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 70;

м) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 77; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 76;

н) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 85;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83;

о) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 90; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 91;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 89;

п) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 96;

 р) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 103;

 с) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 110;

 т) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 117;

у) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 123; и

ф) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83.

18. Вектор по варианту осуществления настоящего изобретения 16, где направляющая РНК выбрана из группы, состоящей из:

а) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 1;

б) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 9; и

5 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 10;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 8;

10 в) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 16; и

15 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 17;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 15;

г) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 23; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 24;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 22;

д) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 30; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 31;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 29;

е) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 37; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 38;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 36;

 ж) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 43;

 з) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 51; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 52;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 50;

 и) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 57; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 58;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 56;

к) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 64; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 65;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

л) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 71; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 72;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 70;

м) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 77; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 76;

н) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84; и

5 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 85;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83;

о) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 90; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 91;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 89;

п) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 97; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 98;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 96;

р) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 104; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 103;

с) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 110;

т) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 117;

у) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 123; и

ф) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83.

19. Вектор по варианту осуществления настоящего изобретения 16, где направляющая РНК выбрана из группы, состоящей из:

а) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 1;

б) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 10;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 8;

в) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 15;

г) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23; и

5 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 22;

10 д) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 30; и

15 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 31;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 29;

е) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 36;

ж) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 44; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 45;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 43;

з) направляющей РНК, включающей:

5 i) CRISPR РНК, которая на 100% идентична последовательности, представленной как SEQ ID NO: 51; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 52;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 50;

 и) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 57; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 58;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 56;

 к) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 64; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 65;

 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

30 л) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как ; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 70;

м) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 77; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 76;

 н) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 85;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 83;

 о) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 90; и

 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 91;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 89;

 п) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 97; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 98;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 96;

р) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 104; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 105;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 103;

с) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 111; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 112;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 110;

т) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 118; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 119;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 117;

у) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124; и

5 ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 123; и

ф) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 83.

20. Вектор по любому из вариантов осуществления настоящего изобретения 16-19, где указанная gRNA представляет собой отдельную направляющую РНК.

20 21. Вектор по любому из вариантов осуществления настоящего изобретения 16-19, где указанная gRNA представляет собой двойную направляющую РНК.

22. Клетку, включающую молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-14, или вектор по любому из вариантов осуществления настоящего изобретения 15-21.

25 23. Способ получения полипептида RGN, включающий культивирование клетки по варианту осуществления настоящего изобретения 22 в условиях, в которых экспрессируется полипептид RGN.

24. Способ получения полипептида RGN, включающий введение в клетку гетерологичной молекулы нуклеиновой кислоты, включающей нуклеотидную последовательность, кодирующую полипептид РНК-направляемой нуклеазы (RGN), включающей аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

где указанный полипептид RGN способен связывать последовательность ДНК-мишени специфическим для РНК направляющей последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной целевой ДНК-мишени;

5 и культивирование указанной клетки в условиях, в которых экспрессируется полипептид RGN.

25. Способ осуществления настоящего изобретения по варианту 24, в котором указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 10 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

26. Способ осуществления настоящего изобретения по варианту 24, в котором указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 15 96, 103, 110, 117, 123 и 570-579.

27. Способ осуществления настоящего изобретения по варианту 24, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности в последовательности SEQ ID NO: 63.

28. Способ по любому из вариантов осуществления настоящего изобретения 23-27, дополнительно включающий очистку указанного полипептида RGN.

29. Способ по любому из вариантов осуществления настоящего изобретения 23-27, где указанная клетка дополнительно экспрессирует одну или несколько направляющих РНК, способных связываться с указанным полипептидом RGN для образования рибонуклеопротеинового комплекса RGN.

30. Способ осуществления настоящего изобретения по варианту 29, дополнительно включающий очистку указанного рибонуклеопротеинового комплекса RGN.

31. Выделенный полипептид РНК-направляемой нуклеазы (RGN), где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579; и

где указанный полипептид RGN способен связывать последовательность ДНК-мишени специфическим для РНК-направляющей последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной целевой ДНК-мишенью.

32. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 31, в котором указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

33. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 31, в котором указанный полипептид RGN включает аминокислотную последовательность, который на 100% идентичен любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

34. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 31, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности, представленной как SEQ ID NO: 63.

35. Выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-34, где указанный полипептид RGN способен расщеплять указанную последовательность ДНК-мишени при связывании.

36. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 35, где расщепление указанным полипептидом RGN приводит к образованию двухцепочечного разрыва.

37. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 35, где расщепление указанным полипептидом RGN приводит к образованию одноцепочечного разрыва.

5 38. Выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-34, где полипептид RGN является неактивной нуклеазой или никазой.

39. Выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-38, где полипептид RGN функционально слит с полипептидом, редактирующим основания.

10 40. Выделенный полипептид RGN по варианту осуществления настоящего изобретения 39, где полипептид, редактирующий основания, является дезаминазой.

15 41. Выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-40, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM).

42. Выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-41, где полипептид RGN включает один или более сигналов ядерной локализации.

20 43. Молекулу нуклеиновой кислоты, включающую полинуклеотид, кодирующий CRISPR РНК (сгRNA), где указанная сгRNA включает спейсерную последовательность и последовательность повтора CRISPR, где указанная последовательность повтора CRISPR включает нуклеотидную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51,
25 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 и 124;

где направляющая РНК, включающая:

а) указанную сгRNA и

б) транс-активирующую CRISPR РНК (tracrRNA), гибридизованную с указанной последовательностью повторов CRISPR указанной сгRNA,

30 способна гибридизоваться с последовательностью ДНК-мишени специфическим для последовательности образом через спейсерную последовательность указанной сгRNA, когда указанная направляющая РНК связана с полипептидом РНК-направляемой нуклеазы (RGN), и

где указанный полинуклеотид, кодирующий crRNA, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

44. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 43, в котором указанная последовательность повтора CRISPR
5 включает нуклеотидную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 и 124.

45. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 43, где указанная последовательность повтора CRISPR включает
10 нуклеотидную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 2, 9, 16, 23, 30, 37, 44, 51, 57, 64, 71, 77, 84, 90, 97, 104, 111, 118 и 124.

46. Вектор, включающий молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 43-45.

15 47. Вектор по варианту осуществления настоящего изобретения 46, где указанный вектор дополнительно включает полинуклеотид, кодирующий указанную tracrRNA.

48. Вектор по варианту осуществления настоящего изобретения 47, где указанная tracrRNA выбрана из группы, состоящей из:

20 а) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 3, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2;

25 б) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9;

30 в) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16;

г) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24, где указанная

н) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 85, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84;

5 о) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 91, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 90;

10 п) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97;

15 р) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104;

20 с) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111;

т) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118; и

25 у) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124.

30 49. Вектор по варианту осуществления настоящего изобретения 47, где указанная tracrRNA выбрана из группы, состоящей из:

а) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2;

у) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124.

5 50. Вектор по варианту осуществления настоящего изобретения 47, где указанная tracrRNA выбрана из группы, состоящей из:

а) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
10 NO: 2;

б) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 10, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 9;

15 в) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 16;

20 г) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 23;

25 д) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 31, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 30;

30 е) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 37;

ж) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 45, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID
NO: 44;

з) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 52, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 51;

5 и) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 58, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 57;

10 к) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 65, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 64;

15 л) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 71;

20 м) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 77;

н) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 85, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 84;

25 о) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 91, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 90;

30 п) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 98, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 97;

р) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 105, где указанная последовательность повтора

CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 104;

5 с) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 112, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 111;

10 т) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 119, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 118; и

15 у) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 124.

51 51. Вектор по любому из вариантов осуществления настоящего изобретения 47-50, где полинуклеотид, кодирующий указанную crRNA, и полинуклеотид, кодирующий указанную tracrRNA, функционально связаны с одним промотором и кодируются как одна направляющая РНК.

20 52. Вектор по любому из вариантов осуществления настоящего изобретения 47-50, где указанный полинуклеотид, кодирующий указанную crRNA, и указанный полинуклеотид, кодирующий указанную tracrRNA, функционально связаны с разными промоторами.

25 53. Вектор по любому из вариантов осуществления настоящего изобретения 46-52, где указанный вектор дополнительно включает полинуклеотид, кодирующий указанный полипептид RGN.

54. Вектор по варианту осуществления настоящего изобретения 53, где указанный полипептид RGN выбран из группы, состоящей из:

30 а) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 1, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3;

 б) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 8, где указанная последовательность повтора

и) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 56, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO:58;

к) полипептида RGN, который по меньшей мере на 90% идентичен любой из последовательностей SEQ ID NO: 63 и 570-579, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 64, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 65;

л) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 70, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 71, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO:72;

м) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 76, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 77, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO:78;

н) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 83, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 85;

о) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 89, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 90, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 91;

п) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 96, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности,

представленной как SEQ ID NO: 97, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98;

5 р) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 103, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105;

10 с) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 110, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112;

15 т) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 117, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119;

20 у) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 123, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125; и

25 ф) полипептида RGN, который по меньшей мере на 90% идентичен последовательности SEQ ID NO: 83, где указанная последовательность повтора CRISPR по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78;

55. Вектор по варианту осуществления настоящего изобретения 53, где указанный полипептид RGN выбран из группы, состоящей из:

30 а) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 1, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3;

CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 97, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 98;

5 р) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 103, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 104, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105;

10 с) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 110, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112;

15 т) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 117, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119;

20 у) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 123, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125;

25 ф) полипептида RGN, который по меньшей мере на 95% идентичен последовательности SEQ ID NO: 83, где указанная последовательность повтора CRISPR по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78.

30 56. Вектор по варианту осуществления настоящего изобретения 53, где указанный полипептид RGN выбран из группы, состоящей из:

а) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 1, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 3;

б) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 8, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 9, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 10;

5 в) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 15, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 16, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 17;

10 г) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 22, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 23, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 24;

15 д) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 29, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 30, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 31;

20 е) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 36, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 37, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 38;

25 ж) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 43, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 44, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 45;

30 з) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 50, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 51, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 52;

tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 98;

р) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 103, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 104, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 105;

с) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 110, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 111, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 112;

т) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 117, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 119;

у) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 123, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 125;

ф) полипептида RGN, который на 100% идентичен последовательности SEQ ID NO: 83, где указанная последовательность повтора CRISPR на 100% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 78.

57. Молекулу нуклеиновой кислоты, включающую полинуклеотид, кодирующий транс-активирующую CRISPR РНК (tracrRNA), включающую нуклеотидную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125;

в которой направляющая РНК, включающая:

а) указанную tracrRNA и

б) crRNA, включающую спейсерную последовательность и последовательность повтора CRISPR, где указанная tracrRNA гибридизируется с последовательностью повтора CRISPR указанной crRNA,

способна гибридизироваться с последовательностью ДНК-мишени специфическим для последовательности образом через спейсерную последовательность указанной crRNA, когда указанная направляющая РНК связана с полипептидом РНК-направляемой нуклеазы (RGN), и

где указанный полинуклеотид, кодирующий tracrRNA, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

58. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 57, где указанная tracrRNA включает нуклеотидную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125.

59. Молекулу нуклеиновой кислоты по варианту осуществления настоящего изобретения 57, где указанная tracrRNA включает нуклеотидную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3, 10, 17, 24, 31, 38, 45, 52, 58, 65, 72, 78, 85, 91, 98, 105, 112, 119 или 125.

60. Вектор, включающий молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 57-59.

61. Вектор по варианту осуществления настоящего изобретения 60, где указанный вектор дополнительно включает полинуклеотид, кодирующий указанную crRNA.

62. Вектор по варианту осуществления настоящего изобретения 61, где указанная crRNA включает последовательность повтора CRISPR, выбранную из группы, состоящей из:

а) последовательности повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, при этом указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3;

б) последовательности повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, где указанная

указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125.

63. Вектор по варианту осуществления настоящего изобретения 61, где указанная crRNA включает последовательность повтора CRISPR, выбранную из группы, состоящей из:

а) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3.

б) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 9, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 10.

в) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 16, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 17.

г) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 23, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 24.

д) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 30, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 31;

е) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 37, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 38;

ж) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44, где указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45;

указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105;

с) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, где

5 указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112;

т) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, где

10 указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119;

у) последовательности повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, где

указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125.

15 64. Вектор по варианту осуществления 61, где указанная crRNA включает последовательность повтора CRISPR, выбранную из группы, состоящей из:

а) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID
20 NO: 3;

б) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID
NO: 10;

25 в) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID
NO: 17;

г) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID
30 NO: 24;

д) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 30, при этом указанная

tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 31;

5 е) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 38;

10 ж) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 44, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 45; з) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 51, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 52;

15 и) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 57, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 58;

20 к) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 64, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 65;

25 л) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 71, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 72;

м) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 77, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 78;

30 н) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 85;

о) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 90, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 91;

5 п) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 97, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 98;

10 р) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 104, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 105;

15 с) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 111, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 112;

20 т) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 118, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 119;

у) последовательности повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124, при этом указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 125.

25 65. Вектор по любому из вариантов осуществления настоящего изобретения 61-64, где указанный полинуклеотид, кодирующий указанную sgRNA, и указанный полинуклеотид, кодирующий указанную tracrRNA, функционально связаны с одним и тем же промотором и кодируются как одна направляющая РНК.

30 66. Вектор по любому из вариантов осуществления настоящего изобретения 61-64, где указанный полинуклеотид, кодирующий указанную sgRNA, и указанный полинуклеотид, кодирующий указанную tracrRNA, функционально связаны с разными промоторами.

67. Вектор по любому из вариантов осуществления настоящего изобретения 60-66, где указанный вектор дополнительно включает полинуклеотид, кодирующий указанный полипептид RGN.

5 68. Вектор по варианту осуществления настоящего изобретения 67, где указанный полипептид RGN выбран из группы, состоящей из:

а) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 1, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3;

б) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 8, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10;

в) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 15, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17;

г) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 22, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24;

д) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 29, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30, и указанная

с) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 110, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111, и
5 указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112;

т) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 117, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90%
10 идентична последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119;

у) полипептида RGN, который по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 123, где указанная crRNA
15 включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125;

ф) полипептида RGN, который по меньшей мере на 90% идентичен
20 последовательности, представленной как SEQ ID NO: 83, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78.

25 69. Вектор по варианту осуществления настоящего изобретения 67, где указанный полипептид RGN выбран из группы, состоящей из:

а) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 1, где указанная crRNA
30 включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3;

б) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 8, где указанная crRNA

tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 85;

5 о) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 89, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 90, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 91;

10 п) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 96, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 97, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 98;

15 р) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 103, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 104, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105;

20 с) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 110, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112;

25 т) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 117, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119;

у) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 123, где указанная crRNA

включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125;

5 ф) полипептида RGN, который по меньшей мере на 95% идентичен последовательности, представленной как SEQ ID NO: 83, где указанная crRNA включает последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA по меньшей мере на 95% идентична последовательности,
10 представленной как SEQ ID NO: 78;

70. Вектор по варианту осуществления настоящего изобретения 67, где указанный полипептид RGN выбран из группы, состоящей из:

а) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 1, где указанная crRNA включает
15 последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 3;

б) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 8, где указанная crRNA включает
20 последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 10;

в) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 15, где указанная crRNA включает
25 последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 17;

г) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 22, где указанная crRNA включает
30 последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 24;

д) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 29, где указанная crRNA включает

последовательности, представленной как SEQ ID NO: 118, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 119;

у) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 123, где указанная crRNA включает последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 125;

ф) полипептида RGN, который на 100% идентичен последовательности, представленной как SEQ ID NO: 83, где указанная crRNA включает последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и указанная tracrRNA на 100% идентична последовательности, представленной как SEQ ID NO: 78.

71. Систему для связывания последовательности ДНК-мишени молекулы ДНК, причем указанная система включает:

а) одну или несколько направляющих РНК, способных гибридизоваться с указанной целевой последовательностью ДНК или один или несколько полинуклеотидов, включающих одну или несколько нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК (gRNA); и

б) полипептид РНК-направляемой нуклеазы (RGN), включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579, или полинуклеотид, включающий нуклеотидную последовательность, кодирующую полипептид RGN;

где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК, и указанная нуклеотидная последовательность, кодирующая полипептид RGN, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности;

где одна или несколько направляющих РНК способны гибридизоваться с целевой последовательностью ДНК, и

где одна или несколько направляющих РНК способны образовывать комплекс с полипептидом RGN, чтобы направить указанный полипептид RGN на связывание с целевой последовательностью ДНК молекулы ДНК.

5 72. Систему для связывания последовательности ДНК-мишени молекулы ДНК, причем указанная система включает:

а) одну или несколько направляющих РНК, способных гибридизироваться с указанной целевой последовательностью ДНК или один или несколько полинуклеотидов, включающих одну или несколько нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК (gRNA); и

10 б) полипептид РНК-направляемой нуклеазы (RGN), включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

15 где одна или более направляющих РНК способны гибридизироваться с целевой последовательностью ДНК, и

где одна или несколько направляющих РНК способны образовывать комплекс с полипептидом RGN, чтобы направить полипептид RGN на связывание с указанной целевой последовательностью ДНК молекулы ДНК.

20 73. Систему по варианту осуществления настоящего изобретения 71 или 72, где по меньшей мере одна из указанных последовательностей нуклеотидов, кодирующих одну или несколько направляющих РНК, функционально связана с промотором, гетерологичным указанной последовательности нуклеотидов.

25 74. Систему по любому из вариантов осуществления настоящего изобретения 71-73, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

30 75. Систему по любому из вариантов осуществления настоящего изобретения 71-73, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, указанных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

76. Систему по любому из вариантов осуществления настоящего изобретения 71-73, где указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368 и валин в положении аминокислоты 405 в SEQ ID NO: 63.

77. Систему по любому из вариантов осуществления настоящего изобретения 71-76, в которой указанный полипептид RGN и указанная одна или несколько направляющих РНК не встречаются в природе в комплексе друг с другом.

78. Систему по любому из вариантов осуществления настоящего изобретения 71-77, где указанная последовательность целевой ДНК представляет собой эукариотическую последовательность целевой ДНК.

79. Систему по любому из вариантов осуществления настоящего изобретения 71-78, где указанная gRNA представляет собой отдельную направляющую РНК (sgRNA).

80. Систему по любому из вариантов осуществления настоящего изобретения 71-78, где указанная gRNA представляет собой двойную направляющую РНК.

81. Система по любому из вариантов осуществления настоящего изобретения 71-80, где указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по

меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 8;

5 в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 15;

10 г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 22;

20 д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 29;

25 е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 36;

30 ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 45, где указанный

полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 43;

5 з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 50;

10 и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 56;

15 к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

20 л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 70;

25 м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 90% идентична

последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 76;

5 н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 90, и tracrRNA, которая по меньшей мере на 90% идентична
15 последовательности, представленной как SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ
25 ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105, где указанный
30 полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ

ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83.

82. Систему по любому из вариантов осуществления настоящего изобретения 71-80, где указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по

меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 1;

5 б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 8;

10 в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 15;

15 г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 22;

20 д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 29;

30 е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 38, где указанный

полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 36;

5 ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 43;

10 з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 50;

15 и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 56;

20 к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

25 л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 95% идентична

последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 70;

5 м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична
15 последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 95, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по
25 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по
меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 95% идентична
последовательности, представленной как SEQ ID NO: 98, где указанный
30 полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по
меньшей мере на 95% идентична последовательности, представленной как SEQ

ID NO: 104, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 103;

5 с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 110;

10 т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 117;

15 у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 123;

20 ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83;

30

83. Систему по любому из вариантов осуществления настоящего изобретения 71-80, где указанная gRNA выбрана из группы, состоящей из:

- а) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 1;
- б) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 8;
- в) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 15;
- г) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 22;
- д) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 29;
- е) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37, и

tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 36;

5 ж) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 43;

10 з) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 50;

15 и) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 56;

20 к) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

25 л) gRNA, включающей последовательность повтора CRISPR, которая на 30 100% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 70;

последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 110;

5 т) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 117;

10 у) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 123;

15 ф) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 83.

20 84. Систему по любому из вариантов осуществления настоящего изобретения 71-83, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM).

25 85. Систему по любому из вариантов осуществления настоящего изобретения 71-84, где целевая последовательность ДНК находится внутри клетки.

86. Систему по варианту осуществления настоящего изобретения 85, где клетка представляет собой эукариотическую клетку.

30 87. Систему по варианту осуществления настоящего изобретения 86, где эукариотическая клетка представляет собой растительную клетку.

88. Систему по варианту осуществления настоящего изобретения 86, где эукариотическая клетка является клеткой млекопитающего.

89. Систему по варианту осуществления настоящего изобретения 88, где указанная клетка млекопитающего является клеткой человека.

90. Систему по варианту осуществления настоящего изобретения 89, где указанная человеческая клетка является иммунной клеткой.

91. Систему по варианту осуществления настоящего изобретения 90, где указанная иммунная клетка является стволовой клеткой.

5 92. Систему по варианту осуществления настоящего изобретения 91, где стволовая клетка является индуцированной плюрипотентной стволовой клеткой.

93. Систему по варианту осуществления настоящего изобретения 86, где эукариотическая клетка представляет собой клетку насекомого.

10 94. Систему по варианту осуществления настоящего изобретения 85, где клетка представляет собой прокариотическую клетку.

95. Систему по любому из вариантов осуществления настоящего изобретения 71-94, где при транскрипции одна или несколько направляющих РНК способны гибридизоваться с целевой последовательностью ДНК, и направляющая РНК способна образовывать комплекс с полипептидом RGN для прямого расщепления последовательности ДНК.

96. Систему по варианту осуществления настоящего изобретения 95, где расщепление приводит к образованию двухцепочечного разрыва.

97. Систему по варианту осуществления настоящего изобретения 95, где расщепление приводит к одноцепочечному разрыву.

20 98. Систему по любому из вариантов осуществления настоящего изобретения 71-94, где указанный полипептид RGN не обладает нуклеазной активностью или представляет собой никазу.

25 99. Систему по любому из вариантов осуществления настоящего изобретения 71-98, где полипептид RGN функционально связан с полипептидом, редактирующим основания.

100. Систему по варианту осуществления настоящего изобретения 99, где полипептид, редактирующий основания, представляет собой дезаминазу.

101. Систему по варианту осуществления настоящего изобретения 100, где дезаминаза представляет собой цитидиндезаминазу или адениндезаминазу.

30 102. Систему по любому из вариантов осуществления настоящего изобретения 71-101, где полипептид RGN включает один или более сигналов ядерной локализации.

103. Систему по любому из вариантов осуществления настоящего изобретения 71-102, где полипептид RGN кодон-оптимизирован для экспрессии в эукариотической клетке.

5 104. Систему по любому из вариантов осуществления настоящего изобретения 71-103, где нуклеотидные последовательности, кодирующие одну или более направляющих РНК, и нуклеотидная последовательность, кодирующая полипептид RGN, расположены на одном векторе.

10 105. Систему по любому из вариантов осуществления воплощения настоящего изобретения 71-104, где указанная система дополнительно включает один или несколько донорных полинуклеотидов или одну или несколько нуклеотидных последовательностей, кодирующих один или несколько донорных полинуклеотидов.

15 106. Фармацевтический состав, включающий молекулу нуклеиновой кислоты по любому из вариантов осуществления настоящего изобретения 1-14, 43-45 и 57-59, вектор по любому из вариантов осуществления настоящего изобретения 15-21, 46-56 и 60-70, клетку по варианту осуществления настоящего изобретения 22, выделенный полипептид RGN по любому из вариантов осуществления настоящего изобретения 31-42 или систему по любому из вариантов осуществления настоящего изобретения 71-105, и фармацевтически приемлемый носитель.

20 107. Способ связывания последовательности ДНК-мишени молекулы ДНК, включающий доставку системы по любому из вариантов осуществления настоящего изобретения 71-105 к указанной последовательности ДНК-мишени или клетке, включающей целевую последовательность ДНК.

25 108. Способ по варианту осуществления настоящего изобретения 107, в котором указанный полипептид RGN или указанная направляющая РНК дополнительно включает выявляемую метку, что позволяет обнаружить указанную последовательность ДНК-мишени.

30 109. Способ по варианту осуществления настоящего изобретения 107, в котором указанная направляющая РНК или указанный полипептид RGN дополнительно включает модулятор экспрессии, тем самым модулируя экспрессию указанной последовательности ДНК-мишени или гена, находящегося под транскрипционным контролем указанной последовательности ДНК-мишени.

110. Способ расщепления или модификации последовательности ДНК-мишени молекулы ДНК, включающий доставку системы по любому из вариантов осуществления настоящего изобретения 71-105 к последовательности ДНК-мишени или клетке, включающей молекулу ДНК, и расщепление или
5 модификацию указанной последовательности ДНК-мишени.

111. Способ по варианту осуществления настоящего изобретения 110, в котором модифицированная последовательность ДНК-мишени включает инсерцию гетерологичной ДНК в последовательность ДНК-мишени.

112. Способ по варианту осуществления настоящего изобретения 110, где
10 указанная модифицированная последовательность ДНК-мишени включает делецию по меньшей мере одного нуклеотида из последовательности ДНК-мишени.

113. Способ по варианту осуществления настоящего изобретения 110, в котором указанная модифицированная последовательность ДНК-мишени
15 включает мутацию по меньшей мере одного нуклеотида в последовательности ДНК-мишени.

114. Способ связывания последовательности ДНК-мишени молекулы ДНК, включающий:

а) сборку рибонуклеотидного комплекса РНК-направляемой нуклеазы
20 (RGN) *in vitro* путем объединения:

i) одной или нескольких направляющих РНК, способных
гибридизироваться с последовательностью ДНК-мишени; и

ii) полипептида RGN, включающего аминокислотную последовательность,
25 которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

в условиях, подходящих для образования рибонуклеотидного комплекса RGN; и

б) контактирование указанной последовательности ДНК-мишени или
30 клетки, содержащей указанную последовательность ДНК-мишени, с собранным *in vitro* рибонуклеотидным комплексом RGN;

где одна или несколько направляющих РНК гибридизируются с последовательностью ДНК-мишени, тем самым направляя указанный полипептид RGN для связывания с последовательностью ДНК-мишени.

115. Способ по варианту осуществления настоящего изобретения 114, где указанный полипептид RGN или указанная направляющая РНК дополнительно включают выявляемую метку, что позволяет обнаружить указанную последовательность ДНК-мишени.

5 116. Способ по варианту осуществления настоящего изобретения 114, в котором указанная направляющая РНК или указанный полипептид RGN дополнительно включает модулятор экспрессии, что позволяет модулировать экспрессию указанной последовательности ДНК-мишени.

10 117. Способ расщепления и/или модификации последовательности ДНК-мишени в молекуле ДНК, включающий контакт молекулы ДНК с:

а) полипептидом РНК-направляемой нуклеазы (RGN), где RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123, и 570-579; и

15 б) одной или несколькими направляющими РНК, способными нацеливать RGN на последовательность ДНК-мишени;

где одна или несколько направляющих РНК гибридизуются с целевой последовательностью ДНК, тем самым направляя указанный полипептид RGN для связывания с указанной целевой последовательностью ДНК и расщеплять и/или модифицировать указанную последовательности ДНК-мишени.

20 118. Способ по варианту осуществления настоящего изобретения 117, в котором расщепление указанным полипептидом RGN приводит к образованию двухцепочечного разрыва.

25 119. Способ по варианту осуществления настоящего изобретения 117, в котором расщепление указанным полипептидом RGN приводит к одноцепочечному разрыву.

30 120. Способ по варианту осуществления настоящего изобретения 117, где указанный полипептид RGN не обладает нуклеазной активностью или не является нисказой и функционально слит с полипептидом, редактирующим основания.

121. Способ по варианту осуществления настоящего изобретения 120, где полипептид, редактирующий основания, представляет собой дезаминазу.

122. Способ по варианту осуществления настоящего изобретения 121, где дезаминаза представляет собой цитидиндезаминазу или адениндезаминазу.

123. Способ по варианту осуществления настоящего изобретения 117, в котором указанная модифицированная последовательность ДНК-мишени включает инсерцию гетерологичной ДНК в последовательность ДНК-мишени.

5 124. Способ по варианту осуществления настоящего изобретения 117, где указанная модифицированная последовательность ДНК-мишени включает делецию по меньшей мере одного нуклеотида из последовательности ДНК-мишени.

10 125. Способ по варианту осуществления настоящего изобретения 117, где указанная модифицированная последовательность ДНК-мишени включает мутацию по меньшей мере одного нуклеотида в последовательности ДНК-мишени.

126. Способ любому из вариантов осуществления настоящего изобретения 114-125, где указанная последовательность ДНК-мишени расположена рядом с мотивом, примыкающим к протоспейсеру (РАМ).

15 127. Способ по любому из вариантов осуществления настоящего изобретения 114-126, где указанная последовательность ДНК-мишени является эукариотической последовательностью ДНК-мишени.

20 128. Способ по любому из вариантов осуществления настоящего изобретения 114-127, где указанная gRNA является отдельная направляющей РНК (sgRNA).

129. Способ по любому из вариантов осуществления настоящего изобретения 114-127, где указанная gRNA представляет собой двойную направляющую РНК.

25 130. Способ по любому из вариантов осуществления настоящего изобретения 114-129, где указанная RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

30 131. Способ по любому из вариантов осуществления настоящего изобретения 114-129, где указанный RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

132. Способ по любому из вариантов осуществления настоящего изобретения 114-129, где указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности, представленной как SEQ ID NO: 63.

133. Способ по любому из вариантов осуществления настоящего изобретения 114-129, где:

а) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3;

б) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 8, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10;

в) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 15, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17;

г) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 22, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24;

д) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 29, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 31;

е) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 36, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 38;

ж) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 43, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 45;

з) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 50, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 52;

и) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 56, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 58;

к) указанная RGN по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90%

последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105;

5 с) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 110, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как
10 SEQ ID NO: 112;

 т) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 117, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая
15 по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119;

 у) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая
20 по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125; или

 ф) указанная RGN по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по
25 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78;

134. Способ по любому из вариантов осуществления настоящего
30 изобретения 114-129, где:

 а) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 1, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по

последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45;

5 з) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 50, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ
10 ID NO: 52;

 и) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 56, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по
15 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 58;

 к) указанная RGN по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95%
20 идентична последовательности SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 65;

 л) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 70, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична
25 последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 72;

 м) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 76, указанная направляющая РНК включает
30 последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78;

по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119;

5 у) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125; или

10 ф) указанная RGN по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78;

15 135. Способ по любому из вариантов осуществления настоящего изобретения 114-129, где:

а) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 1, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, на 100% идентична последовательности, представленной как SEQ ID NO: 3;

20 б) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 8, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 10;

25 в) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 15, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17;

30 г) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 22, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности,

представленной как SEQ ID NO: 23, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24;

д) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 29, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 31;

е) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 36, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38;

ж) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 43, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 45;

з) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 50, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 52;

и) указанная RGN по на 100% идентична последовательности, представленной как SEQ ID NO: 56, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 58;

к) указанная RGN на 100% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 64, и tracrRNA, на 100% идентична последовательности SEQ ID NO: 65;

л) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 70, указанная направляющая РНК включает последовательность

повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72;

5 м) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 76, указанная направляющая РНК включает последовательность повтора crRNA, на 100% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78;

10 н) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, на 100% идентична последовательности, представленной как SEQ ID NO: 85;

15 о) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 89, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 100, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 91;

20 п) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 96, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 98;

25 р) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 103, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 105;

30 с) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 110, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 112;

т) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 117, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 119;

5 у) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125; или

ф) указанная RGN на 100% идентична последовательности, представленной как SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78;

136. Способ по любому из вариантов осуществления настоящего изобретения 107-135, где последовательность ДНК-мишени находится внутри клетки.

20 137. Способ по варианту осуществления настоящего изобретения 136, где клетка представляет собой эукариотическую клетку.

138. Способ по варианту осуществления настоящего изобретения 137, где эукариотическая клетка представляет собой растительную клетку.

25 139. Способ по варианту осуществления настоящего изобретения 137, где эукариотическая клетка является клеткой млекопитающего.

140. Способ по варианту осуществления настоящего изобретения 139, где указанная клетка млекопитающего является клеткой человека.

141. Способ по варианту осуществления настоящего изобретения 140, где указанная человеческая клетка является иммунной клеткой.

30 142. Способ по варианту осуществления настоящего изобретения 141, где указанная иммунная клетка является стволовой клеткой.

143. Способ по варианту осуществления настоящего изобретения 142, где указанная стволовая клетка является индуцированной плюрипотентной стволовой клеткой.

144. Способ по варианту осуществления настоящего изобретения 137, где эукариотическая клетка представляет собой клетку насекомого.

145. Способ по варианту осуществления настоящего изобретения 136, где клетка является прокариотической клеткой.

5 146. Способ по любому из вариантов осуществления настоящего изобретения 136-145, дополнительно включающий культивирование клетки в условиях, в которых экспрессируется полипептид RGN и расщепляется последовательность ДНК-мишени для получения молекулы ДНК, содержащей модифицированную последовательность ДНК; и отбор клетки, содержащей
10 модифицированную последовательность ДНК-мишени.

147. Клетку, включающая модифицированную последовательность ДНК-мишени в соответствии со способом по варианту осуществления настоящего изобретения 146.

15 148. Клетку по варианту осуществления настоящего изобретения 147, где клетка является эукариотической клеткой.

149. Клетку по варианту осуществления настоящего изобретения 148, где эукариотическая клетка представляет собой растительную клетку.

150. Растение, включающее клетку по варианту осуществления настоящего изобретения 149.

20 151. Семя, включающее клетку по варианту осуществления настоящего изобретения 149.

152. Клетку по варианту осуществления настоящего изобретения 148, где эукариотическая клетка представляет собой клетку млекопитающего.

25 153. Клетку по варианту осуществления настоящего изобретения 152, где указанная клетка млекопитающего является клеткой человека.

154. Клетку по варианту осуществления настоящего изобретения 153, где указанная клетка человека является иммунной клеткой.

155. Клетку по варианту осуществления настоящего изобретения 154, где указанная иммунная клетка является стволовой клеткой.

30 156. Клетку по варианту осуществления настоящего изобретения 155, где указанная стволовая клетка является индуцированной плюрипотентной стволовой клеткой.

157. Клетку по варианту осуществления настоящего изобретения 148, где эукариотическая клетка представляет собой клетку насекомого.

158. Клетку по варианту осуществления настоящего изобретения 147, где клетка является прокариотической клеткой.

159. Фармацевтический состав, включающий клетку по любому из вариантов осуществления настоящего изобретения 148 и 152-156 и фармацевтически приемлемый носитель.

160. Способ получения генетически модифицированной клетки с исправлением мутации, вызывающей генетически наследуемое заболевание, указанный способ включает введение в клетку:

а) полипептида РНК-направляемой нуклеазы (RGN), где полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, или полинуклеотида, кодирующего указанный полипептид RGN, причем указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором для обеспечения экспрессии полипептида RGN в клетке; и б) направляющей РНК (gRNA) или полинуклеотида, кодирующего указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, функционально связан с промотором для обеспечения экспрессии gRNA в клетке,

причем RGN и gRNA нацелены на геномное местоположение вызывающей мутации и изменяют геномную последовательность для удаления вызывающей мутации.

161. Способ по варианту осуществления настоящего изобретения 160, в котором RGN не проявляет активности нуклеазы или нисказы и слита с полипептидом, обладающим активностью по редактированию оснований.

162. Способ по варианту осуществления настоящего изобретения 161, где полипептид, редактирующий основания, представляет собой дезаминазу.

163. Способ по варианту осуществления настоящего изобретения 162, где полипептид, редактирующий основания, представляет собой цитидиндезаминазу или адениндезаминазу.

164. Способ по любому из вариантов осуществления настоящего изобретения 160-163, где генетически наследуемое заболевание вызвано однонуклеотидным полиморфизмом.

165. Способ по любому из вариантов осуществления настоящего изобретения 160-163, где генетически наследуемое заболевание представляет собой синдром Гурлера.

5 166. Способ по любому из вариантов осуществления настоящего изобретения 160-163, где gRNA дополнительно включает спейсерную последовательность, которая нацелена на ближайшую область, расположенную к корректируемому однонуклеотидному полиморфизму.

10 167. Способ получения генетически модифицированной клетки с делецией в вызывающей заболевание области геномной нестабильности, указанный способ включает введение в клетку:

а) полипептида РНК-направляемой нуклеазы (RGN), где полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-15 579, или полинуклеотида, кодирующего указанный полипептид RGN, причем указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором для обеспечения экспрессии полипептида RGN в клетке; и

б) первой направляющей РНК (gRNA) или полинуклеотида, кодирующего указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, 20 функционально связан с промотором для обеспечения экспрессии gRNA в клетке, а также где gRNA включает спейсерную последовательность, которая нацелена на 5'-конец области геномной нестабильности; и

в) второй направляющей РНК (gRNA) или полинуклеотида, кодирующего указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, 25 функционально связан с промотором для обеспечения экспрессии gRNA в клетке, а также где указанная вторая gRNA включает спейсерную последовательность, которая направлена на 3'-конец области геномной нестабильности;

причем RGN и две gRNA нацелены на область геномной нестабильности, и 30 по меньшей мере часть области геномной нестабильности удалена.

168. Способ по варианту осуществления настоящего изобретения 167, где генетически наследуемое заболевание представляет собой атаксию Фридрейха или болезнь Гентингтона.

169. Способ по варианту осуществления настоящего изобретения 167, где первая gRNA дополнительно включает спейсерную последовательность, которая нацелена на область внутри или непосредственно у области геномной нестабильности.

5 170. Способ по варианту осуществления настоящего изобретения 169, где вторая gRNA дополнительно включает спейсерную последовательность, которая нацелена на область внутри или непосредственно у области геномной нестабильности.

10 171. Способ по любому из вариантов осуществления настоящего изобретения 160-170, где указанный полипептид RGN по меньшей мере на 95% идентичен любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579.

15 172. Способ по любому из вариантов осуществления настоящего изобретения 160-170, в котором указанный полипептид RGN на 100% идентичен любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579.

20 173. Способ по любому из вариантов осуществления настоящего изобретения 160-170, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности, представленной как SEQ ID NO: 63.

25 174. Способ по любому из вариантов осуществления настоящего изобретения 160-170, где указанная gRNA, указанная первая gRNA, указанная вторая gRNA или указанная первая gRNA и указанная вторая gRNA выбраны из gRNA, выбранной из группы, состоящей из:

30 а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 8;

в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 15;

г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 22;

д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 29;

е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по

меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 36;

ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 43;

з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 50;

и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 56;

к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 72, где указанный

полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 70;

5 м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ
10 ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 85, где указанный
15 полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ
20 ID NO: 90, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 89;

25 п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по
30 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична

последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 103;

5 с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 110;

 т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 90% идентична
15 последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 117;

 у) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по
25 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 123;

 ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична
30 последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83.

175. Способ по любому из вариантов осуществления настоящего изобретения 160-170, где указанная gRNA, указанная первая gRNA, указанная

вторая gRNA или указанная первая gRNA и указанная вторая gRNA выбрана из gRNA, выбранной из группы, состоящей из:

- а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 1;
- б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 8;
- в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 15;
- г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 22;
- д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 31, где указанный

полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 29;

5 е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 36;

10 ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 43;

15 з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 50;

20 и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 56;

25 к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 95% идентична

последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

5 л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 70;

м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична
15 последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ
25 ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 95, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 91, где указанный
30 полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ

ID NO: 97, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83.

176. Способ по любому из вариантов осуществления 160-170, где указанная gRNA, указанная первая gRNA, указанная вторая gRNA или указанная первая gRNA и указанная вторая gRNA выбраны из gRNA, выбранной из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 8;

в) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 15;

г) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную

tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

5 л) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности,
10 представленной как SEQ ID NO: 70;

м) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную
15 последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как
20 SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 90, и tracrRNA, которая на 100% идентична последовательности, представленной как
25 SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая на
30 100% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 96;

5 р) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 103;

10 с) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 110;

15 т) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 117;

20 у) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 123;

25 ф) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 83.

30 177. Способ по любому из вариантов осуществления настоящего изобретения 160-176, где клетка представляет собой клетку животного.

178. Способ по любому из вариантов осуществления настоящего изобретения 177, где животная клетка представляет собой клетку млекопитающего.

5 179. Способ по варианту осуществления настоящего изобретения 177, где клетка получена из собаки, кошки, мыши, крысы, кролика, лошади, коровы, свиньи или человека.

10 180. Способ получения генетически модифицированной кроветворной прогениторной клетки млекопитающего со сниженной экспрессией мРНК и белка BCL11A, метод включает введение в изолированную кроветворную прогениторную клетку человека:

а) полипептида РНК-направляемой нуклеазы (RGN), где полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой последовательности, представленной как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579, или 15 полинуклеотида, кодирующего указанный полипептид RGN, причем указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором для обеспечения экспрессии полипептида RGN в клетке; и

б) направляющую РНК (gRNA) или полинуклеотид, кодирующий 20 указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, функционально связан с промотором для обеспечения экспрессии gRNA в клетке,

25 посредством чего RGN и gRNA экспрессируются в клетке и расщепляются в области энхансера BCL11A, что приводит к генетической модификации кроветворной прогениторной клетки человека и снижению экспрессии мРНК и/или белка BCL11A.

181. Способ по варианту осуществления настоящего изобретения 180, в котором указанный полипептид RGN по меньшей мере на 95% идентичен любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579.

30 182. Способ по варианту осуществления настоящего изобретения 180, в котором указанный полипептид RGN на 100% идентичен любой из последовательностей, представленных как SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 или 570-579.

183. Способ по варианту осуществления настоящего изобретения 180, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности, представленной как SEQ ID NO: 63, и имеет изолейцин в аминокислотном положении 305, валин в аминокислотном положении 328, лейцин в аминокислотном положении 366, треонин в аминокислотном положении 368, и валин в аминокислотном положении 405 последовательности, представленной как SEQ ID NO: 63.

184. Способ по варианту осуществления настоящего изобретения 180, где указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 8;

в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 15;

г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 24, где указанный

полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 22;

5 д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 29;

10 е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 36;

15 ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 43;

20 з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 50;

25 и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 90% идентична

последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 56;

5 к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 90% идентична
15 последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 70;

м) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ
25 ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична
30 последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ

ID NO: 90, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 83.

185. Способ по варианту осуществления настоящего изобретения настоящего изобретения 180, в котором указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 8;

в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 95% идентична

последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 15;

5 г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по
10 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 22;

д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 95% идентична
15 последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 29;

е) gRNA, включающей последовательность повтора CRISPR, которая по
20 меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ
25 ID NO: 36;

ж) gRNA, включающей последовательность повтора CRISPR, которая по
меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 45, где указанный
30 полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 43;

з) gRNA, включающей последовательность повтора CRISPR, которая по
меньшей мере на 95% идентична последовательности, представленной как SEQ

ID NO: 51, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 50;

и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 56;

к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 70;

м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 95, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по

меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 83.

186. Способ по варианту осуществления настоящего изобретения 180, где указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 2, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 9, и tracrRNA, которая на 100% идентична последовательности, представленной как

SEQ ID NO: 10, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 8;

5 в) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 16, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 15;

10 г) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 23, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 22;

15 д) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 30, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 29;

20 е) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 36;

30 ж) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 44, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 43;

з) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 51, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 50;

и) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 57, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 56;

к) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 64, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

л) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 70;

м) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 77, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 76;

н) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную

tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 123;

5 ф) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности,
10 представленной как SEQ ID NO: 83.

187. Способ по любому из вариантов осуществления 180-186, где gRNA дополнительно включает спейсерную последовательность, которая нацелена на область внутри или близлежащую области энхансера BCL11A.

188. Способ лечения заболевания, включающий введение нуждающемуся в
15 лечении субъекту эффективного количества фармацевтической композиции по варианту осуществления настоящего изобретения 106 или 159.

189. Способ по варианту осуществления настоящего изобретения 188, где
указанное заболевание связано с вызывающей мутацией, и указанное
эффективное количество указанной фармацевтической композиции корректирует
20 указанную вызывающую мутацию.

190. Применение молекулы нуклеиновой кислоты по любому из вариантов
осуществления настоящего изобретения 1-14, 43-45 и 57-59, вектора по любому
из вариантов осуществления настоящего изобретения 15-21, 46-56 и 60-70,
клеток по любому из вариантов осуществления настоящего изобретения 22, 147,
25 148 и 152-156, выделенного полипептида RGN по любому из вариантов
осуществления настоящего изобретения 31-42 или системы по любому из
вариантов осуществления настоящего изобретения 71-105 для лечения
заболевания у субъекта.

191. Применение по варианту осуществления настоящего изобретения 190,
30 где указанное заболевание связано с вызывающей мутацией, и указанное
лечение включает корректирование указанной вызывающей мутации.

192. Применение молекулы нуклеиновой кислоты по любому из вариантов
осуществления настоящего изобретения 1-14, 43-45 и 57-59, вектора по любому
из вариантов осуществления настоящего изобретения 15-21, 46-56 и 60-70,

клетки по любому из вариантов осуществления настоящего изобретения 22, 147, 148 и 152-156, выделенного полипептида RGN по любому из вариантов осуществления настоящего изобретения 31-42 или системы по любому из вариантов осуществления настоящего изобретения 71-105 для получения лекарственного средства, полезного для лечения заболевания.

193. Применение по варианту осуществления настоящего изобретения 192, где указанное заболевание связано с вызывающей мутацией, и эффективное количество указанного лекарственного средства корректирует указанную вызывающую мутацию.

Следующие примеры применения настоящего изобретения приведены в качестве иллюстрации осуществления настоящего изобретения, но они не ограничивают рамки его охвата.

Примеры

Пример 1. Идентификация РНК-направляемых нуклеаз

Идентифицировано 19 различных CRISPR-ассоциированных РНК-направляемых нуклеаз (RGN), описанных в табл.1. В табл. 1 приведено название каждой RGN, ее аминокислотная последовательность, источник, из которого она получена, и процессированные последовательности crRNA и tracrRNA (методы идентификации см. в примере 2). В табл. 1 представлена общая последовательность одиночной направляющей РНК (sgRNA), где поли-N указывает на расположение спейсерной последовательности, которая определяет последовательность нуклеиновой кислоты-мишени sgRNA. Для систем RGN APG06622, APG02787 и APG06248 консервативная последовательность в основании стержня-шпильки tracrRNA представляет собой UNANNA (SEQ ID NO: 129). Для APG06007, APG09344 и APG07991 последовательность в том же местоположении представляет собой UNANNA (SEQ ID NO: 130). Для APG02874, APG03850 и APG07553 последовательность в том же местоположении представляет собой UNANNA (SEQ ID NO №: 131). Для систем RGN APG03031, APG09208, APG05586, APG08770, APG03021, APG06015, APG01868 и APG02998 консервативная последовательность в основании стержня шпильки tracrRNA представляет собой UNANNC (SEQ ID NO: 132). Для

APG08167 и APG01604 последовательность в том же местоположении представляет собой CNANNC (SEQ ID NO: 133).

Таблица 1. Краткое описание идентификаторов SEQ ID NO и CRISPR-связанных систем

RGN ID	SEQ ID NO.	Источник	Последовательность повтора crRNA (SEQ ID NO:)	tracrRNA (SEQ ID NO:)	Каркас sgRNA (SEQ ID NO:)
APG06622	1	<i>Pedobacter</i> sp.	2	3	4
APG02787	8	<i>Chitinophaga</i> sp.	9	10	11
APG06248	15	<i>Mucilaginibacter</i> sp.	16	17	18
APG06007	22	<i>Acidovorax</i> sp.	23	24	25
APG02874	29	<i>Bacillus</i> sp.	30	31	32
APG03850	36	<i>Bacillus</i> sp.	37	38	39
APG07553	43	<i>Bacillus</i> sp.	44	45	46
APG03031	50	<i>Chryseobacterium</i> sp.	51	52	53
APG09208	56	<i>Bacillus</i> sp.	57	58	59
APG05586	63	<i>Enterococcus</i> sp.	64	65	66
APG08770	70	<i>Enterococcus</i> sp.	71	72	73
APG08167	76	<i>Staphylococcus</i> sp.	77	78	79
APG01604	83	<i>Staphylococcus</i> sp.	84	85	86
APG03021	89	<i>Streptococcus</i> sp.	90	91	92
APG06015	96	<i>Pediococcus</i> sp.	97	98	99
APG09344	103	<i>Weissella</i> sp.	104	105	106
APG07991	110	<i>Enterococcus</i> sp.	111	112	113
APG01868	117	<i>Enterococcus</i> sp.	118	119	120
APG02998	123	<i>Enterococcus</i> sp.	124	125	126

5

Пример 2. Идентификация направляющей РНК и конструирование sgRNA

Культуры бактерий, которые изначально экспрессируют представленную РНК-направляемую нуклеазную систему, выращивают до середины логарифмической фазы (ОП600 ~0,600), гранулируют и подвергают мгновенной заморозке. РНК выделяют из гранул с использованием набора для выделения mirVana miRNA Isolation Kit (Life Technologies, Карлсбад, Калифорния), а библиотеки для секвенирования выделенной РНК готовили с использованием набора NEBNext Small RNA Library Prep kit (NEB, Беверли, Массачусетс). Библиотечный препарат фракционируют с использованием 6% полиакриламидного геля для улавливания фрагментов РНК размером менее 200

15

нуклеотидов для обнаружения crRNA и tracrRNA, соответственно. Глубокое секвенирование (спаренный конец 75 п.н.) было выполнено на Next Seq 500 (High Output kit) поставщик (фирма MoGene, Сент-Луис, Миссури). Прочтения обрабатывают с использованием процедуры тримминга с помощью программы Cutadapt и соотносят с эталонными геномами с помощью программного обеспечения Bowtie2. Пользовательский конвейер действий RNAseq был написан на языке python для обнаружения транскриптов crRNA и tracrRNA. Границы обработанной crRNA определяют по охвату последовательностей массивом нативных спейсерных повторов. Анти-повторяющаяся часть tracrRNA идентифицирована с использованием разрешающих параметров BLASTn. Глубина секвенирования РНК подтвердила границы обработанной tracrRNA путем идентификации транскрипта, содержащего анти-повтор. Ручную обработку РНК выполняют с использованием системы прогнозирования вторичной структуры NUPACK - программного обеспечения для сворачивания РНК. Кассеты sgRNA получают синтезом ДНК и, как правило, в направлении 5' > 3' следующим образом: последовательность спейсера 20-30 п.н., функционально связанная на своем 3'-конце с обработанной частью повтора crRNA, функционально связанной с некомплементарным линкером 4 п.н. (AAAG; SEQ ID NO: 249), функционально связанным на своем 3'-конце с обработанной tracrRNA. Другие некомплементарные линкеры размером 4 п.н. также могут быть использованы.

Для анализа *in vitro* sgRNA синтезируют путём транскрипции кассет sgRNA *in vitro* с помощью набора для синтеза gRNA GeneArt™ Precision (фирма ThermoFisher). Обработанные последовательности crRNA и tracrRNA для каждого из полипептидов RGN идентифицируют и приводят в табл. 1. Смотрите ниже sgRNA, сконструированные для PAM библиотек 1 и 2.

Пример 3. Определение требований к PAM для каждой RGN

Требования к PAM для каждой RGN определяют с использованием метода истощения PAM, по сути адаптированного из публикаций Kleinstiver и соавт., *Nature.*, 2015, 523:481-485 и Zetsche и соавт., *Cell.*, 2015, 163:759-771. Вкратце, две библиотеки плазмид (L1 и L2) создают в каркасе pUC18 (ampR), каждая из них содержит отдельную последовательность протоспейсера (мишени) 30 п.н., окруженную 8 случайными нуклеотидами (т.е. областью PAM). Целевая

последовательность и фланкирующая область PAM библиотеки 1 и библиотеки 2 для каждого RGN приведены в табл. 2.

Библиотеки отдельно электропоруют в клетки *E. coli* BL21(DE3), содержащие векторы экспрессии pRSF-1b, включающие RGN по настоящему изобретению (кодон-оптимизированные для *E. coli*) вместе с родственной sgRNA, содержащей спейсерную последовательность, соответствующую протоспейсеру в L1 или L2. Количество библиотечной плазмиды, использованное в реакции трансформации, берут в достаточном количестве для получения $>10^6$ КОЕ. RGN и sgRNA находятся под контролем промоторов T7 в каркасе pRSF-1b. Реакция трансформации протекает в течение 1 ч, после чего разводят средой LB, содержащей карбенициллин и канамицин, и выращивают в течение ночи. На следующий день смесь разводят в самоиндуцирующейся среде Overnight Instant Express™ TB (фирма Millipore Sigma), чтобы обеспечить экспрессию RGN и sgRNA, и выращивают в течение дополнительных 4 или 20 ч, после чего клетки разделяют и с помощью набора Mini-prep (фирма Qiagen, Джермантаун, Мэриленд) выделяют плазмидную ДНК. В присутствии соответствующей sgRNA плазмиды, содержащие PAM, который распознается RGN, будут расщепляться, что приводит к их удалению из популяции. Плазмиды, содержащие PAM, которые не распознаются RGN или которые трансформируются в бактерии, не содержащие соответствующей sgRNA, выживут и реплицируются. Участки PAM и протоспейсера нерасщепленных плазмид являются ПЦР-амплифицированными и подготовленными для секвенирования в соответствии с опубликованными протоколами (руководство по подготовке 16s-метагеномной библиотеки 15044223B, фирма Illumina, Сан-Диего, Калифорния). Глубокое секвенирование (режим одностороннего секвенирование, 75 п.н.) осуществляют на генном анализаторе MiSeq (фирма Illumina) поставщиком (фирма MoGene, Сент-Луис, Миссури). Обычно на ампликон получают 1-4 млн. прочтений. Участки PAM выделяют, подсчитывают и нормализуют по общему количеству считываний для каждого образца. PAM, которые приводят к расщеплению плазмиды, идентифицируют исходя из их недостаточной представленности в сравнении с контролем (т.е. когда библиотеку трансформируют в *E. coli*, содержащую RGN, но не имеющую соответствующей sgRNA). Для представления требования к PAM для новой RGN, коэффициенты истощения (частота в образце/ частота в контроле) для всех

последовательностей в рассматриваемой области преобразуют в значения обогащения с помощью преобразования $-\log_2$. У удовлетворяющих требованиям РАМ значения обогащения $>2,3$ (что соответствует коэффициентам истощения $< \sim 0,2$). РАМ, в обеих библиотеках, превышающие этот порог, отбирают и используют для создания веб-логотипов последовательностей, которые, например, можно сгенерированы с помощью веб-сервиса в интернете, известного как «weblogo». Последовательности РАМ идентифицируют и представляют, если в наиболее обогащенных РАМ наблюдается последовательная закономерность. Консенсусный РАМ (имеющий коэффициент обогащения (EF – enrichment factor) $>2,3$) для каждой RGN, приведенной в табл. 2. Ориентация РАМ также указана в табл. 2.

Таблица 2: Определение РАМ или РАМ-подобного мотива

RGN ID	sgRNA L1 (SEQ ID NO.)	sgRNA L2 (SEQ ID NO.)	РАМ (SEQ ID NO.)	Ориентация РАМ
APG06622	5	6	7	5'-мишень-РАМ-3'
APG02787	12	13	14	5'-мишень-РАМ-3'
APG06248	19	20	21	5'-мишень-РАМ-3'
APG06007	26	27	28	5'-мишень-РАМ-3'
APG02874	33	34	35	5'-мишень-РАМ-3'
APG03850	40	41	42	5'-мишень-РАМ-3'
APG07553	47	48	49	5'-мишень-РАМ-3'
APG03031	54	55	35	5'-мишень-РАМ-3'
APG09208	60	61	62	5'-мишень-РАМ-3'
APG05586	67	68	69	5'-мишень-РАМ-3'
APG08770	74	75	69	5'-мишень-РАМ-3'
APG08167	80	81	82	5'-мишень-РАМ-3'
APG01604	87	88	82	5'-мишень-РАМ-3'
APG03021	93	94	95	5'-мишень-РАМ-3'
APG06015	100	101	102	5'-мишень-РАМ-3'
APG09344	107	108	109	5'-мишень-РАМ-3'
APG07991	114	115	116	5'-мишень-РАМ-3'
APG01868	121	122	116	5'-мишень-РАМ-3'
APG02998	127	128	116	5'-мишень-РАМ-3'

15 Пример 4. Демонстрация активности редактирования генов в клетках млекопитающих

Кассеты экспрессии RGN получают и интродуцируют в векторы для экспрессии в млекопитающих. Каждую RGN кодон-оптимизируют для экспрессии в человеке (SEQ ID NO: 134-152) и функционально сливают с 5'-конца с последовательностью ядерной локализации SV40 (NLS; SEQ ID NO: 251)

и метками 3xFLAG (SEQ ID NO: 252), и функционально сливают на 3'-конце с последовательностью нуклеоплазмина NLS (SEQ ID NO: 253). Используют две копии последовательности NLS, функционально слитых в тандем. Каждая экспрессионная кассета находится под контролем промотора цитомегаловируса (CMV) (SEQ ID NO: 258). В данной области техники известно, что энхасер транскрипции CMB (SEQ ID NO: 259) может быть включен в конструкции, содержащие промотор CMV. Конструкции экспрессии направляющей РНК, кодирующие по одной gRNA, каждая под контролем промотора U6 РНК-полимеразы III человека (SEQ ID NO: 260), получают и интродуцируют в вектор pTwist High Copy Amp. Сиквенсы последовательностей-мишеней для каждой направляющей приведены в табл. 3.

Некоторые из описанных выше в настоящем изобретении конструкций вводят в клетки млекопитающих. За один день до трансфекции 1×10^5 клеток HEK293T (фирма Sigma) высевают в 24-луночные планшеты в модифицированную среду Eagle Dulbecco (DMEM – Dulbecco's modified Eagle medium) плюс 10 об.% фетальной сыворотки теленка (фирма Gibco) и 1% пенициллин-стрептомицина (фирма Gibco). На следующий день, когда слияние клеток достигает 50-60%, 500 нг плазмиды экспрессии RGN плюс 500 нг плазмиды экспрессии отдельной gRNA совместно трансфицируют, используя 1,5 мкл липофектамина 3000 (фирма Thermo Scientific) на лунку, следуя инструкциям производителя. После 48 ч роста собирают общую геномную ДНК с использованием набора для выделения геномной ДНК (фирма Machery-Nagel) в соответствии с инструкциями производителя.

Для определения скорости редактирования для каждой RGN для каждой геномной мишени анализируют общую геномную ДНК. Сначала получают олигонуклеотиды, которые используют для ПЦР-амплификации и последующего анализа амплифицированного геномного сайта-мишени. Используемые олигонуклеотидные последовательности перечислены в табл. 4.

Все ПЦР проводят с использованием 10 мкл 2X Master Mix Phusion High-Fidelity ДНК-полимеразы (фирма Thermo Scientific) в реакции объемом 20 мкл, включающей 0,5 мкм каждого праймера. Большие области генома, охватывающие каждый целевой ген, сначала амплифицируют с использованием праймеров PCR№1, используя программу: 98°C, 1 мин; 30 циклов [98°C, 10 сек; 62°C, 15 сек; 72°C, 5 мин]; 72°C, 5 мин; 12°C, постоянно. Один микролитр этой

ПЦР смеси затем дополнительно амплифицируют с использованием праймеров, специфичных для каждой направляющей РНК (праймеры для PCR№2), используя программу: 98°C, 1 мин; 35 циклов [98°C, 10 сек; 67°C, 15 сек; 72°C, 30 сек]; 72°C, 5 мин; 12°C, постоянно. Праймеры для ПЦР №2 включают последовательности выступов адаптера транспозазы Nextera Read 1 и Read 2 для секвенирования в Illumina.

Для RGN APG02874, APG03850 и APG09208 методы реализуют согласно описанному выше в настоящем изобретении. Несколько различных генов в геноме человека нацелены на направляемое РНК расщепление. Эти локусы включены в приведённую ниже табл. 3 вместе со ссылками на sgRNA SEQ ID NO. Также показан процент инсерций и делеций (indel), который является показателем активности RGN.

Таблица 3. Целевые последовательности и последовательности sgRNA для направляющих РНК, используемых для тестирования активности редактирования генов в клетках млекопитающих

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG02874, APG09208	RelA	SGN000973, SGN000778	153	188, 206
APG02874	RelA	SGN000974	154	189
APG02874	RelA	SGN000975, SGN000780	155	190, 208
APG02874, APG09208	AurkB	SGN000976, SGN000775	156	191, 204
APG02874, APG09208	AurkB	SGN000977, SGN000776	157	192, 205
APG02874	AurkB	SGN000978	158	193
APG02874	VEGFA	SGN000979	159	194
APG02874	VEGFA	SGN000981	160	195
APG03850	RelA	SGN000982	161	196
APG03850	RelA	SGN000983	162	197
APG03850	RelA	SGN000984	163	198
APG03850	AurkB	SGN000985	164	199
APG03850	AurkB	SGN000986	165	200
APG03850	AurkB	SGN000987	166	201
APG03850	VEGFA	SGN000988	167	202
APG03850	VEGFA	SGN000990	168	203
APG09208	RelA	SGN000779	169	207
APG09208	AurkB	SGN000793	170	209
APG09208	AurkB	SGN000794	171	210
APG05586	TRA	SGN001163	553	557
APG05586	TRA	SGN001164	554	558
APG05586	VEGFA	SGN001165	555	559
APG05586	VEGFA	SGN001166	556	560
APG09208	RelA	SGN000778	153	206
APG09208	AurkB	SGN000793	609	811

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG05586, APG08770, APG09298	EMX1	SGN001159	610	812
APG05586, APG08770, APG09298	TRA	SGN001162	611	813
APG05586, APG08770, APG09298	TRA	SGN001163	612	814
APG05586, APG08770, APG09298	TRA	SGN001164	613	815
APG05586, APG08770, APG09298	VEGFA	SGN001165	614	816
APG05586, APG08770, APG09298	VEGFA	SGN001166	615	817
APG05586, APG09298	VEGFA	SGN001167	616	818
APG09208	RelA	SGN001213	617	819
APG08167, APG01604	VEGFA	SGN001245	618	820
APG08167, APG01604	VEGFA	SGN001246	619	821
APG08167, APG01604	VEGFA	SGN001247	620	822
APG08167, APG01604	RelA	SGN001248	621	823
APG08167, APG01604	RelA	SGN001249	622	824
APG08167, APG01604	RelA	SGN001250	623	825
APG08167, APG01604	AurkB	SGN001251	624	826
APG08167, APG01604	AurkB	SGN001252	625	827
APG08167, APG01604	AurkB	SGN001253	626	828
APG07991	RelA	SGN001312	627	829
APG07991	RelA	SGN001313	628	830
APG07991	RelA	SGN001314	629	831
APG01868	RelA	SGN001315	630	832
APG01868	RelA	SGN001316	631	833
APG01868	RelA	SGN001317	632	834
APG02998	RelA	SGN001318	633	835
APG02998	RelA	SGN001319	634	836
APG02998	RelA	SGN001320	635	837
APG09344	RelA	SGN001321	636	838
APG06015	RelA	SGN001322	637	839
APG03021	RelA	SGN001323	638	840
APG09344	RelA	SGN001324	639	841
APG06015	RelA	SGN001325	640	842
APG03021	RelA	SGN001326	641	843
APG06015	RelA	SGN001327	642	844
APG03021	RelA	SGN001328	643	845

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG09344	RelA	SGN001329	644	846
APG03021	TRA	SGN001330	645	847
APG03021	TRA	SGN001331	646	848
APG03021	TRA	SGN001332	647	849
APG06015	TRA	SGN001333	648	850
APG06015	TRA	SGN001334	649	851
APG06015	TRA	SGN001335	650	852
APG09344	TRA	SGN001336	651	853
APG09344	TRA	SGN001337	652	854
APG09344	TRA	SGN001338	653	855
APG07991	TRA	SGN001339	654	856
APG07991	TRA	SGN001340	655	857
APG07991	TRA	SGN001341	656	858
APG01868	TRA	SGN001342	657	859
APG01868	TRA	SGN001343	658	860
APG01868	TRA	SGN001344	659	861
APG01868	TRA	SGN001692	660	862
APG02998	TRA	SGN001345	661	863
APG02998	TRA	SGN001346	662	864
APG02998	TRA	SGN001347	663	865
APG03021	HAO1	SGN001348	664	866
APG03021	HAO1	SGN001349	665	867
APG03021	HAO1	SGN001350	666	868
APG06015	HAO1	SGN001351	667	869
APG06015	HAO1	SGN001352	668	870
APG06015	HAO1	SGN001353	669	871
APG09344	HAO1	SGN001354	670	872
APG09344	HAO1	SGN001355	671	873
APG09344	HAO1	SGN001356	672	874
APG07991	HAO1	SGN001357	673	875
APG07991	HAO1	SGN001358	674	876
APG07991	HAO1	SGN001359	675	877
APG01868	HAO1	SGN001360	676	878
APG01868	HAO1	SGN001361	677	879
APG01868	HAO1	SGN001362	678	880
APG02998	HAO1	SGN001363	679	881
APG02998	HAO1	SGN001364	680	882
APG02998	HAO1	SGN001365	681	883
APG05586, APG09298	TRA	SGN001371	682	884
APG05586, APG09298	TRA	SGN001372	683	885
APG05586, APG09298	TRA	SGN001373	684	886
APG05586, APG09298	TRA	SGN001374	685	887
APG05586, APG09298	TRA	SGN001375	686	888
APG05586, APG09298	TRA	SGN001376	687	889
APG05586, APG09298	TRA	SGN001377	688	890
APG05586, APG09298	TRA	SGN001378	689	891
APG05586,	TRA	SGN001379	690	892

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG09298				
APG05586, APG09298	TRA	SGN001380	691	893
APG05586, APG09298	TRA	SGN001381	692	894
APG05586, APG09298	TRA	SGN001382	693	895
APG05586, APG09298	B2M	SGN001383	694	896
APG05586, APG09298	B2M	SGN001384	695	897
APG05586, APG09298	B2M	SGN001385	696	898
APG05586, APG09298	B2M	SGN001386	697	899
APG05586, APG09298	B2M	SGN001387	698	900
APG05586, APG09298	B2M	SGN001388	699	901
APG05586, APG09298	B2M	SGN001389	700	902
APG05586, APG09298	B2M	SGN001390	701	903
APG05586, APG09298	B2M	SGN001391	702	904
APG05586, APG09298	B2M	SGN001392	703	905
APG05586, APG09298	B2M	SGN001393	704	906
APG05586, APG09298	B2M	SGN001394	705	907
APG05586, APG09298	LDHA	SGN001395	706	908
APG05586, APG09298	LDHA	SGN001396	707	909
APG05586, APG09298	LDHA	SGN001397	708	910
APG05586, APG09298	LDHA	SGN001399	709	911
APG05586, APG09298	LDHA	SGN001400	710	912
APG05586, APG09298	LDHA	SGN001401	711	913
APG05586, APG09298	LDHA	SGN001402	712	914
APG05586, APG09298	LDHA	SGN001403	713	915
APG05586, APG09298	LDHA	SGN001404	714	916
APG05586, APG09298	LDHA	SGN001405	715	917
APG05586, APG09298	HAO1	SGN001406	716	918
APG05586, APG09298	HAO1	SGN001407	717	919
APG05586,	HAO1	SGN001408	718	920

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG09298				
APG05586, APG09298	HAO1	SGN001409	719	921
APG05586, APG09298	HAO1	SGN001410	720	922
APG05586, APG09298	HAO1	SGN001411	721	923
APG05586, APG09298	HAO1	SGN001412	722	924
APG05586, APG09298	HAO1	SGN001413	723	925
APG05586, APG09298	HAO1	SGN001414	724	926
APG05586, APG09298	HAO1	SGN001415	725	927
APG05586, APG09298	HAO1	SGN001416	726	928
APG01604	B2M	SGN001592	727	929
APG01604	B2M	SGN001593	728	930
APG01604	B2M	SGN001594	729	931
APG01604	B2M	SGN001595	730	932
APG01604	B2M	SGN001596	731	933
APG01604	B2M	SGN001597	732	934
APG01604	B2M	SGN001598	733	935
APG01604	B2M	SGN001599	734	936
APG01604	B2M	SGN001600	735	937
APG01604	B2M	SGN001601	736	938
APG01604	B2M	SGN001602	737	939
APG01604	B2M	SGN001603	738	940
APG01604	HAO1	SGN001616	739	941
APG01604	HAO1	SGN001617	740	942
APG01604	HAO1	SGN001618	741	943
APG01604	HAO1	SGN001619	742	944
APG01604	HAO1	SGN001620	743	945
APG01604	HAO1	SGN001621	744	946
APG01604	HAO1	SGN001622	745	947
APG01604	HAO1	SGN001623	746	948
APG01604	HAO1	SGN001624	747	949
APG01604	HAO1	SGN001625	748	950
APG01604	HAO1	SGN001626	749	951
APG01604	HAO1	SGN001627	750	952
APG01604	LDHA	SGN001640	751	953
APG01604	LDHA	SGN001641	752	954
APG01604	LDHA	SGN001642	753	955
APG01604	LDHA	SGN001643	754	956
APG01604	LDHA	SGN001644	755	957
APG01604	LDHA	SGN001645	756	958
APG01604	LDHA	SGN001646	757	959
APG01604	LDHA	SGN001647	758	960
APG01604	LDHA	SGN001648	759	961
APG01604	LDHA	SGN001649	760	962
APG01604	LDHA	SGN001650	761	963
APG01604	LDHA	SGN001651	762	964
APG01604	TRA	SGN001664	763	965
APG01604	TRA	SGN001665	764	966

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG01604	TRA	SGN001666	765	967
APG01604	TRA	SGN001667	766	968
APG01604	TRA	SGN001668	767	969
APG01604	TRA	SGN001669	768	970
APG01604	TRA	SGN001670	769	971
APG01604	TRA	SGN001671	770	972
APG01604	TRA	SGN001672	771	973
APG01604	TRA	SGN001673	772	974
APG01604	TRA	SGN001674	773	975
APG01604	TRA	SGN001675	774	976
APG01868	TRA	SGN001684	775	977
APG01868	TRA	SGN001685	776	978
APG01868	TRA	SGN001686	777	979
APG01868	TRA	SGN001687	778	980
APG01868	TRA	SGN001688	779	981
APG01868	TRA	SGN001689	780	982
APG01868	TRA	SGN001690	781	983
APG01868	TRA	SGN001691	782	984
APG01868	EMX1	SGN001697	783	985
APG01868	EMX1	SGN001698	784	986
APG01868	EMX1	SGN001699	785	987
APG01868	EMX1	SGN001700	786	988
APG01868	EMX1	SGN001701	787	989
APG01868	EMX1	SGN001702	788	990
APG01868	EMX1	SGN001703	789	991
APG01868	EMX1	SGN001704	1176	992
APG01868	EMX1	SGN001705	1177	993
APG01868	EMX1	SGN001706	1178	994
APG01868	EMX1	SGN001707	1179	995
APG01868	EMX1	SGN001708	1180	996
APG01868	EMX1	SGN001709	1181	997
APG01868	EMX1	SGN001710	1182	998
APG01868	EMX1	SGN001711	1183	999
APG01868	HAO1	SGN001713	1184	1000
APG01868	HAO1	SGN001714	1185	1001
APG01868	HAO1	SGN001715	790	1002
APG01868	HAO1	SGN001716	791	1003
APG01868	HAO1	SGN001717	792	1004
APG01868	HAO1	SGN001718	793	1005
APG01868	HAO1	SGN001719	794	1006
APG01868	HAO1	SGN001721	795	1007
APG01868	HAO1	SGN001722	796	1008
APG01868	HAO1	SGN001723	797	1009
APG01868	HAO1	SGN001724	798	1010
APG01868	LDHA	SGN001725	799	1011
APG01868	LDHA	SGN001726	800	1012
APG01868	LDHA	SGN001727	801	1013
APG01868	LDHA	SGN001728	802	1014
APG01868	LDHA	SGN001729	803	1015
APG01868	LDHA	SGN001730	804	1016
APG01868	LDHA	SGN001731	805	1017
APG01868	LDHA	SGN001732	806	1018
APG01868	LDHA	SGN001733	807	1019
APG01868	LDHA	SGN001734	808	1020
APG01868	LDHA	SGN001735	809	1021

RGN ID	Ген	ID направляющих	Целевая последовательность (SEQ ID NO.)	sgRNA (SEQ ID NO.)
APG01868	VEGFA	SGN001785	810	1022

Таблица 4: Олигонуклеотиды для определения активности редактирования генов в клетках млекопитающих

Наименование	Последовательность праймеров	SEQ ID NO
SGN000977 FWD	CTTGTAGCTGGAGGTCCATC	172
SGN000977 REV	TGTTGGCAAATCTAGTCTCG	173
SGN000978 FWD	ACATTTGACGAGCAGCGAA	174
SGN000978 REV	GGCCCTGGAGAGGTTTTAA	175
SGN000979, SGN000982, SGN000988, SGN000990 FWD	ACACAGCTTCCCGTTCTCAG	176
SGN000979, SGN000982, SGN000988, SGN000990 REV	ATCACCCAGCTTCCCTGTG	177
SGN000981 FWD	GGCGTCGCACTGAAACTTTT	178
SGN000981 REV	AGTTCATGGTTTCGGAGGCC	179
SGN000983 FWD	CGACCAAACAAGTGCAAAGG	180
SGN000983 REV	GGGTTGTTGTTGGTCTGGAT	181
SGN000775, SGN000776, SGN000793, SGN000794, SGN000976, SGN000985, SGN000986, SGN000987 FWD	ACTGCCATGGGAAGAAGGTG	182
SGN000775, SGN000776, SGN000793, SGN000794, SGN000976, SGN000985, SGN000986, SGN000987 REV	ACAATTCTCCTGCCTCAGCC	183
SGN000778, SGN000973, SGN000984 FWD	TGGCCCCTATGTGGAGATCA	184
SGN000778, SGN000973, SGN000984 REV	GGCAGAGCTCAGCCTCATAG	185
SGN000779, SGN000780, SGN000974, SGN000975 FWD	ATATCCCCACTTCCCCTGCT	186
SGN000779, SGN000780, SGN000974, SGN000975 REV	CACCTCAAGGACAGCTCTGG	187
SGN001163, SGN001164 FWD	TTGATAGCTTGTGCCTGTCC	561
SGN001163, SGN001164 REV	AGAGTCTCTCAGCTGGTACA	562
SGN001165, SGN001166 FWD	GCGACAGGGGCAAAGTGAGT	563
SGN001165, SGN001166 REV	CTAGCACTTCTCGCGGCTCC	564
SGN001382 FWD	AACTCATGCCTGCTGCTCTT	1023
SGN001382 REV	CAGTCTCACGCAGTCACTCA	1024
SGN001371, SGN001372, SGN001373, SGN001689, SGN001690 FWD	AACTGAGGCGGCTGAAATGA	1025
SGN001371, SGN001372, SGN001373, SGN001689, SGN001690 REV	TGGGACATGCAAGCCCATAA	1026
SGN001730 FWD	AAGATGTTGACATGCTCTTCC	1027
SGN001730 REV	TATGCAGTCAAAGCCTCA	1028
SGN001353, SGN001356, SGN001359, SGN001362, SGN001365, SGN001406, SGN001407, SGN001412, SGN001413, SGN001414, SGN001622, SGN001623, SGN001624, SGN001625, SGN001626, SGN001627 FWD	AAGTCATTTGCTTGTTTGGA	1029
SGN001353, SGN001356, SGN001359, SGN001362, SGN001365, SGN001406, SGN001407, SGN001412, SGN001413, SGN001414, SGN001622, SGN001623, SGN001624, SGN001625, SGN001626, SGN001627 REV	TGGTGCATTCAGAGAAGGAG	1030

Наименование	Последовательность праймеров	SEQ ID NO
N001413,SGN001414,SGN001622,SGN001623,SGN001624,SGN001625,SGN001626,SGN001627 REV		
SGN001375,SGN001376,SGN001377 FWD	AATGAAGCCAGGCAAGAGCA	1031
SGN001375,SGN001376,SGN001377 REV	CTGTGCAAACCCAGGCTAGA	1032
SGN001251,SGN001252 FWD	ACATTTGACGAGCAGCGAA	1033
SGN001251,SGN001252 REV	GGCCCTGGAGAGGTTTTAA	1034
SGN001380 FWD	ACCCGGCCTGCTTTTCTTAA	1035
SGN001380 REV	GGCAGCGAGGCATACATAGT	1036
SGN001731,SGN001732,SGN001733,SGN001734,SGN001735 FWD	ACCCTGCTTTTTCTGCCTTT	1037
SGN001731,SGN001732,SGN001733,SGN001734,SGN001735 REV	CAGGCCTAATGGACATTAATCC T	1038
SGN001374,SGN001688 FWD	ACTACTAAGGGGCCATCT	1039
SGN001374,SGN001688 REV	CAGGAGGAGGATTCGGAACC	1040
SGN001726,SGN001727,SGN001728,SGN001729 FWD	AGGAAAATGAATCACAATTACT	1041
SGN001726,SGN001727,SGN001728,SGN001729 REV	GTGCGAAAGGGCAAGATTCT	1042
SGN001397 FWD	AGGCCTTCAACTCTCTTTTGGC A	1043
SGN001397 REV	GGATGGGGTCAAGGTATGGGC	1044
SGN001213,SGN001248,SGN001249,SGN001321,SGN001322,SGN001323 FWD	ATGACATTCAGGCCACAGTG	1045
SGN001213,SGN001248,SGN001249,SGN001321,SGN001322,SGN001323 REV	CTTCCTCCTATTCAGGCCCA	1046
SGN001725 FWD	CAGCTTTTGAAATGGGGTGC	1047
SGN001725 REV	CAACAAATGGAGACCATCTGGA	1048
SGN001381 FWD	CAGTATTCTAAGGACGCCAGAA A	1049
SGN001381 REV	GCACTTTGGGAGGCTGAA	1050
SGN001390,SGN001391,SGN001392,SGN001393,SGN001394,SGN001592,SGN001593,SGN001594,SGN001595,SGN001596,SGN001597,SGN001598,SGN001599,SGN001600,SGN001601,SGN001602,SGN001603 FWD	CGGGCATTCTGAAGCTG	1051
SGN001390,SGN001391,SGN001392,SGN001393,SGN001394,SGN001592,SGN001593,SGN001594,SGN001595,SGN001596,SGN001597,SGN001598,SGN001599,SGN001600,SGN001601,SGN001602,SGN001603 REV	GTAGGCCAAAGGTCTCCCC	1052
SGN001383,SGN001384,SGN001385,SGN001386,SGN001387,SGN001388,SGN001389 FWD	CTTGACACCAAGTTAGCCCC	1053
SGN001383,SGN001384,SGN001385,SGN001386,SGN001387,SGN001388,SGN001389 REV	TCATACACAACCTTCAGCAGC	1054
SGN000793,SGN001253 FWD	CTTGTAGCTGGAGGTCCATC	1055
SGN000793,SGN001253 REV	TGTTGGCAAATCTAGTCTCG	1056
SGN001379 FWD	GAGCAGCTGAGTCAATGATAGT	1057
SGN001379 REV	GGAGAGATCTGGAGGGAACCTT A	1058
SGN001165,SGN001166,SGN001245 FWD	GCAAAGTGAGTGACCTGCTT	1059
SGN001165,SGN001166,SGN001245 REV	GAGCTAGCACTTCTCGCG	1060

Наименование	Последовательность праймеров	SEQ ID NO
SGN001738 FWD	GCTGTTTGGGAGGTCAGAAA	1061
SGN001738 REV	GAATATTGAAGGGGGCAGGG	1062
SGN001167,SGN001246,SGN001247,SGN001785 FWD	GGACACTTCCCAAAGGACC	1063
SGN001167,SGN001246,SGN001247,SGN001785 REV	CACGTCCTCACTCTCGAAGA	1064
SGN001399,SGN001402,SGN001403,SGN001646,SGN001647,SGN001648,SGN001649,SGN001650,SGN001651 FWD	GGCCTTCACTCTTACAGACCC	1065
SGN001399,SGN001402,SGN001403,SGN001646,SGN001647,SGN001648,SGN001649,SGN001650,SGN001651 REV	GGATGGGGTCAAGGTATGGGC	1066
SGN001162,SGN001163,SGN001164,SGN001330,SGN001331,SGN001332,SGN001333,SGN001334,SGN001335,SGN001336,SGN001337,SGN001338,SGN001339,SGN001340,SGN001341,SGN001342,SGN001343,SGN001344,SGN001345,SGN001346,SGN001347,SGN001664,SGN001665,SGN001666,SGN001667,SGN001668,SGN001669,SGN001670,SGN001671,SGN001672,SGN001673,SGN001674,SGN001675,SGN001684,SGN001685,SGN001686,SGN001687,SGN001691,SGN001692 FWD	GGGCAAAGAGGGAAATGAGA	1067
SGN001162,SGN001163,SGN001164,SGN001330,SGN001331,SGN001332,SGN001333,SGN001334,SGN001335,SGN001336,SGN001337,SGN001338,SGN001339,SGN001340,SGN001341,SGN001342,SGN001343,SGN001344,SGN001345,SGN001346,SGN001347,SGN001664,SGN001665,SGN001666,SGN001667,SGN001668,SGN001669,SGN001670,SGN001671,SGN001672,SGN001673,SGN001674,SGN001675,SGN001684,SGN001685,SGN001686,SGN001687,SGN001691,SGN001692 REV	GAACCTGGCCATTCCTGAAG	1068
SGN001714,SGN001722 FWD	GGTTTTTGGAGGTGGAGTTGA	1069
SGN001714,SGN001722 REV	CCCCCTAACCAAGTGAAAAGA	1070
SGN001716,SGN001717,SGN001721 FWD	TAGATAAATGAGCAGTGAACAGCC	1071
SGN001716,SGN001717,SGN001721 REV	TCCACAAAGGATCACAAAGTCA	1072
SGN001313,SGN001316,SGN001319 FWD	TGAGAGACAGTGGGACAGAC	1073
SGN001313,SGN001316,SGN001319 REV	AGTCCTAGAGGAGGCAGAAC	1074
SGN001702,SGN001703,SGN001707,SGN001709,SGN001710,SGN001711 FWD	TGAGTCCGAGCAGAAGAAGA	1075
SGN001702,SGN001703,SGN001707,SGN001709,SGN001710,SGN001711 REV	GGAGATTGGAGACACGGAGA	1076
SGN001723 FWD	TGGAGGTGGAGTTGAATAACA	1077
SGN001723 REV	CTTCTCCCCCTAACCAAGTG	1078
SGN001395,SGN001396 FWD	TGGATCTCCAACATGGCAGCC	1079
SGN001395,SGN001396 REV	CGGAAGGCTAAGGAGGGAGGA	1080
SGN000778,SGN001250,SGN001312,SGN001314,SGN001315,SGN001317,SGN001318,SGN001320,SGN001324,SGN001325,SGN001326,SGN001327,SGN001328,SGN001329 FWD	TGGCCCCCTATGTGGAGATCA	1081
SGN000778,SGN001250,SGN001312,SGN001314,SGN001315,SGN001317,SGN001318,SGN001320,SGN001324,SGN001325,SGN001326,SGN001327,SGN001328,SGN001329 REV	GGCAGAGCTCAGCCTCATAG	1082

Наименование	Последовательность праймеров	SEQ ID NO
001328,SGN001329 REV		
SGN001159,SGN001697,SGN001698,SGN001699,SGN001700,SGN001701,SGN001704,SGN001705,SGN001706,SGN001708 FWD	TGTTAGACCCATGGGAGCAG	1083
SGN001159,SGN001697,SGN001698,SGN001699,SGN001700,SGN001701,SGN001704,SGN001705,SGN001706,SGN001708 REV	GTTGCCACCCCTAGTCATT	1084
SGN001400,SGN001401,SGN001404,SGN001405,SGN001640,SGN001641,SGN001642,SGN001643,SGN001644,SGN001645 FWD	TTCCACGCTAAGGTATGGGCC	1085
SGN001400,SGN001401,SGN001404,SGN001405,SGN001640,SGN001641,SGN001642,SGN001643,SGN001644,SGN001645 REV	GCCAACAGCACCAACCCCAA	1086
SGN001348,SGN001349,SGN001350,SGN001351,SGN001352,SGN001354,SGN001355,SGN001357,SGN001358,SGN001360,SGN001361,SGN001363,SGN001364,SGN001408,SGN001409,SGN001410,SGN001411,SGN001415,SGN001416,SGN001616,SGN001617,SGN001618,SGN001619,SGN001620,SGN001621,SGN001713,SGN001715,SGN001718,SGN001719,SGN001724 FWD	TTGCTCACTTGATGTAAGCAA	1087
SGN001348,SGN001349,SGN001350,SGN001351,SGN001352,SGN001354,SGN001355,SGN001357,SGN001358,SGN001360,SGN001361,SGN001363,SGN001364,SGN001408,SGN001409,SGN001410,SGN001411,SGN001415,SGN001416,SGN001616,SGN001617,SGN001618,SGN001619,SGN001620,SGN001621,SGN001713,SGN001715,SGN001718,SGN001719,SGN001724 REV	TTTTGGTACGGTCTTTGTGT	1088

Очищенную геномную ДНК подвергают ПЦР№1 и ПЦР№2 как указано выше. После второй ПЦР амплификации ДНК очищают с помощью набора для очистки ПЦР (фирма Zymo) в соответствии с инструкциями производителя и элюируют в воду. 200-500 нг очищенного продукта после ПЦР№ 2 смешивают с 2 мкл 10X NEB буфера 2 и водой в 20 мкл реакции и отжигают для образования гетеродуплексов ДНК, используя программу: 95°C, 5 мин; 95-85°C, охлаждение со скоростью 2°C/сек; 85-25°C, охлаждение со скоростью 0,1°C/сек.; 12°C, постоянно. После отжига отбирают 5 мкл ДНК для контроля без фермента и добавляют 1 мкл эндонуклеазы I T7 (NEB) после чего инкубируют реакцию при 37°C в течение 1 ч. После инкубации добавляют 5x FlashGel красителя (фирма Lonza) и анализируют по 5 мкл каждой реакционной смеси и контрольных образцов в 2,2% агарозном геле FlashGel (фирма Lonza) методом геле-электрофореза. После визуализации геля процент негомологичного концевое соединения (NHEJ – non-homologous end joining) определяют, используя следующее уравнение: процент событий NHEJ = 100 x [1-(расщепленная 1-

фракция) ($1/2$)], где (расщепленную фракцию) определяют как: (плотность расщепленных продуктов)/(плотность расщепленных продуктов + нерасщепленная исходная группа).

Для некоторых образцов для анализа результатов после экспрессии в клетках млекопитающих используют SURVEYOR®. Клетки инкубируют при 37°C в течение 72 ч после трансфекции перед извлечением геномной ДНК. Геномную ДНК экстрагируют с использованием раствора для извлечения ДНК QuickExtract (фирма Epicentre) в соответствии с протоколом производителя. Геномную область, фланкирующую сайт-мишень RGN, амплифицируют методом ПЦР, продукты очищают с использованием QiaQuick Spin Column (фирма Qiagen) в соответствии с протоколом производителя. 200-500 нг очищенных ПЦР-продуктов смешивают с 1 мкл 10×Taq ДНК-полимеразным ПЦР-буфером (фирма Enzymatics) и высокоочищенной водой до конечного объема 10 мкл и подвергают процессу повторного отжига, чтобы обеспечить образование гетеродуплекса, используя программу: 95°C в течение 10 мин, от 95°C до 85°C изменение -2°C/сек, от 85°C до 25°C изменение - 0,25°C/сек и выдержка при 25°C в течение 1 мин.

После повторного отжига продукты обрабатывают нуклеазой SURVEYOR® и энхансером S SURVEYOR® (фирма Integrated DNA Technologies) в соответствии с протоколом производителя и анализируют в 4-20% полиакриламидном геле Novex TBE (фирма Life Technologies). Гели окрашивают ДНК-красителем SYBR Gold (фирма Life Technologies) в течение 10 мин и получают изображение с помощью системы визуализации Gel Doc (фирма Bio-rad). Количественная оценка основана на относительных интенсивностях полос. Процент Indel определяют по формуле, $100 \times (1 - (1 - (b+c)/(a+b+c))^{1/2})$, где a - интегрированная интенсивность нерасщепленного ПЦР продукта, b и c – интегрированные интенсивности каждого продукта расщепления.

Дополнительно, продукты ПЦР№2, содержащие выступающие последовательности Illumina, подвергают библиотечной подготовке в соответствии с протоколом библиотеки метагеномного секвенирования Illumina 16S. Глубокое секвенирование выполняют на платформе Illumina Mi-Seq поставщиком услуг (фирма MOGene). Как правило, на ампликон генерируется 200 000 парных прочтений по 250 п.н. (2 x 100 000 прочтений). Прочтения анализируют с использованием программы CRISPResso (Pinello и соавт., *Nature*

Biotech, 2016, 34:695-697) для расчета степени редактирования. Выравнивания выходных данных обработаны вручную, чтобы подтвердить сайты инсерций и делеций, а также идентифицировать сайты микрогомологии в местах рекомбинации. Показатели редактирования приведены в табл. 5. Все

5 эксперименты проводятся в клетках человека. «Целевая последовательность (последовательность-мишень)» – это последовательность внутри гена-мишени. Для каждой последовательности-мишени направляющая РНК включает комплементарную последовательность-мишень РНК и соответствующую sgRNA в зависимости от используемой RGN. Выбранная разбивка экспериментов по 10 направляющим РНК показана в табл. 6.1-6.3.

Таблица 5. Активность RGN в клетках млекопитающих.

RGN	Направляющий ID	Ген-мишень	Общая степень редактирования	Доля делеций в образце	Доля инсерций в образце
APG02874	SGN000973	RelA	Нет данных		
APG02874	SGN000974	RelA	0,10%	37,23%	62,77%
APG02874	SGN000975	RelA	Нет данных		
APG02874	SGN000976	AurkB	0,56%	35,17%	64,84%
APG02874	SGN000977	AurkB	0,30%	76,34%	23,65%
APG02874	SGN000978	AurkB	Нет данных		
APG02874	SGN000979	VEGFA	0,17%	13,95%	86,05%
APG02874	SGN000981	VEGFA	Нет данных		
APG03850	SGN000982	RelA	0,19%	40,55%	59,45%
APG03850	SGN000983	RelA	Нет данных		
APG03850	SGN000984	RelA	0,05%	72,17%	27,84%
APG03850	SGN000985	AurkB	0,37%	66,62%	33,37%
APG03850	SGN000986	AurkB	Нет данных		
APG03850	SGN000987	AurkB	Нет данных		
APG03850	SGN000988	VEGFA	Нет данных		
APG03850	SGN000990	VEGFA	Нет данных		
APG09208	SGN000775	AurkB	Нет данных		
APG09208	SGN000776	AurkB	Нет данных		
APG09208	SGN000778	RelA	Нет данных		
APG09208	SGN000779	RelA	Нет данных		
APG09208	SGN000780	RelA	Нет данных		
APG09208	SGN000793	AurkB	1,18%	59,97%	40,02%
APG09208	SGN000794	AurkB	Нет данных		
APG05586	SGN001163	TRA	23,49%	95,03%	5,41%
APG05586	SGN001164	TRA	1,37%	95,65%	4,36%
APG05586	SGN001165	VEGFA	65,59%	98,58%	2,03%
APG05586	SGN001166	VEGFA	10,48%	94,98%	5,02%
APG06015	SGN001322	RelA	Нет данных		
APG06015	SGN001325	RelA	Нет данных		
APG06015	SGN001327	RelA	Нет данных		
APG06015	SGN001333	TRA	Нет данных		
APG06015	SGN001334	TRA	0,03	100%	0%
APG06015	SGN001335	TRA	Нет данных		
APG06015	SGN001351	HAO1	Нет данных		
APG06015	SGN001352	HAO1	Нет данных		
APG06015	SGN001353	HAO1	Нет данных		

RGN	Направляющий ID	Ген-мишень	Общая степень редактирования	Доля делеций в образце	Доля инсерций в образце
APG09344	SGN001321	RelA	0,02	0%	100%
APG09344	SGN001324	RelA	Нет данных		
APG09344	SGN001329	RelA	Нет данных		
APG09344	SGN001336	TRA	Нет данных		
APG09344	SGN001337	TRA	Нет данных		
APG09344	SGN001338	TRA	Нет данных		
APG09344	SGN001354	HAO1	Нет данных		
APG09344	SGN001355	HAO1	Нет данных		
APG09344	SGN001356	HAO1	Нет данных		
APG07991	SGN001312	RelA	2,16	56,61%	43,39%
APG07991	SGN001313	RelA	2,64	80,6%	19,39%
APG07991	SGN001314	RelA	2,49	32,48%	67,51%
APG07991	SGN001339	TRA	6,75	76,47%	27,42%
APG07991	SGN001340	TRA	6,66	69,19%	33,89%
APG07991	SGN001341	TRA	2,84	60,63%	39,38%
APG07991	SGN001357	HAO1	13,34	79,66%	21,31%
APG07991	SGN001358	HAO1	0,05	60,78%	39,22%
APG07991	SGN001359	HAO1	0,21	80,53%	19,48%
APG01868	SGN001315	RelA	10,78	64,07%	36,36%
APG01868	SGN001316	RelA	6,26	87,09%	13,73%
APG01868	SGN001317	RelA	19,42	62,57%	38,11%
APG01868	SGN001342	TRA	4,66	39,32%	60,68%
APG01868	SGN001343	TRA	41,39	60,13%	41,39%
APG01868	SGN001344	TRA	5,42	60,46%	39,53%
APG01868	SGN001360	HAO1	0,38	69,28%	30,72%
APG01868	SGN001361	HAO1	9,48	84,92%	17,07%
APG01868	SGN001362	HAO1	0,32	28,97%	71,03%
APG02998	SGN001318	RelA	1,13	72,78%	27,22%
APG02998	SGN001319	RelA	2,55	91,76%	8,22%
APG02998	SGN001320	RelA	3,48	43,97%	56,04%
APG02998	SGN001345	TRA	2,01	79,92%	26,06%
APG02998	SGN001346	TRA	5,87	40,14%	61,02%
APG02998	SGN001347	TRA	2,26	59,08%	40,89%
APG02998	SGN001363	HAO1	0,4	16,34%	83,66%
APG02998	SGN001364	HAO1	Нет данных		
APG02998	SGN001365	HAO1	1,14	65,98%	34,02%
APG09208	SGN000778	RelA	0,26	0%	100,00%
APG09208	SGN000793	AurkB	1,64	50,94%	49,06%
APG09208	SGN001213	RelA	0,05	100%	0,00%
APG05586	SGN001159	EMX1	40,13	83,86%	16,13%
APG05586	SGN001162	TRA	46,12	86,07%	14,45%
APG05586	SGN001163	TRA	43,77	93,38%	7,58%
APG05586	SGN001164	TRA	8	95,23%	4,80%
APG05586	SGN001165	VEGFA	65,59	98,58%	2,03%
APG05586	SGN001166	VEGFA	10,48	94,98%	5,02%
APG05586	SGN001167	VEGFA	43,8	92,97%	7,35%
APG08770	SGN001159	EMX1	11,9	80,98%	20,13%
APG08770	SGN001162	TRA	14,57	90,2%	10,24%
APG08770	SGN001163	TRA	12,78	95,47%	5,49%
APG08770	SGN001164	TRA	3,09	92,84%	9,00%
APG08770	SGN001165	VEGFA	23,12	92,75%	8,00%
APG08770	SGN001166	VEGFA	10,52	93,45%	6,92%
APG08167	SGN001245	VEGFA	Нет данных		
APG08167	SGN001246	VEGFA	Нет данных		
APG08167	SGN001247	VEGFA	Нет данных		

RGN	Направляющий ID	Ген-мишень	Общая степень редактирования	Доля делеций в образце	Доля инсерций в образце
APG08167	SGN001248	RelA	Нет данных		
APG08167	SGN001249	RelA	Нет данных		
APG08167	SGN001250	RelA	Нет данных		
APG08167	SGN001251	AurkB	Нет данных		
APG08167	SGN001252	AurkB	Нет данных		
APG08167	SGN001253	AurkB	Нет данных		
APG01604	SGN001245	VEGFA	69,13	96,82%	7,23%
APG01604	SGN001246	VEGFA	4,57	79,07%	24,78%
APG01604	SGN001247	VEGFA	18,49	96,17%	5,09%
APG01604	SGN001248	RelA	17,04	94,78%	5,61%
APG01604	SGN001249	RelA	5,53	87,88%	14,96%
APG01604	SGN001250	RelA	21,18	81,7%	19,19%
APG01604	SGN001251	AurkB	8,38	84,67%	15,34%
APG01604	SGN001252	AurkB	24,74	90,49%	10,22%
APG01604	SGN001253	AurkB	0,32	86,44%	13,56%
APG03021	SGN001323	RelA	13,73	87,67%	12,31%
APG03021	SGN001326	RelA	1,03	78,9%	21,11%
APG03021	SGN001328	RelA	7,12	92,18%	9,68%
APG03021	SGN001330	TRA	Нет данных		
APG03021	SGN001331	TRA	0,45	23,15%	76,85%
APG03021	SGN001332	TRA	0,25	49,44%	50,57%
APG03021	SGN001348	HAO1	0,08	30,16%	69,84%
APG03021	SGN001349	HAO1	0,03	58,33%	41,67%
APG03021	SGN001350	HAO1	0,36	70,53%	29,48%

Конкретные инсерции и делеции для соответствующих направляющих показаны в табл. 6.1-6.3. В этих таблицах последовательность-мишень обозначена жирными заглавными буквами. 8-нуклеотидные области РАМ дважды подчеркнуты, а основные выявленные нуклеотиды выделены жирным шрифтом. Инсерции обозначаются строчными буквами. Делеции обозначаются тире (---). Местоположение INDEL вычисляют от ближайшего конца РАМ последовательности-мишени, причем конец является нулевым местоположением. Местоположение является положительным (+), если находится со стороны мишени на конце; местоположение является отрицательным (-), если находится на стороне РАМ конца.

Таблица 6.1. Конкретные инсерции и делеции для направляющей 977 с использованием RGN APG02874

Направляющая SGN000977 (SEQ ID NO: 157)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
GGAGAGGTTTTTAATGGCCCAGCCTCACACCCAGG	218651	99,7				
GGAGAGGTTTTTAATGGCCCAG----ACACCCAGG	161	0,07	24,88	Делеция	-1	4
GGAGAGGTTTTTAATGGCCCAGtacCCTCACACCCAGG	153	0,07	23,65	Инсерция	+3	3
GGAGAGGTTTTTAATGGCCCAG-----G	137	0,06	21,17	Делеция	-9	12
GGAGAGGTTTTTAATGGCCCAGC-----ACCCAGG	130	0,06	20,09	Делеция	-3	5
GGAGAGGTTT-AATGGCCCAGCCTCACACCCAGG	33	0,02	5,1	Делеция	+13	1
GGAGAGGTTTTTAATGGCC-AGCCTCACACCCAGG	33	0,02	5,1	Делеция	+5	1

Таблица 6.2. Конкретные инсерции и делеции для направляющей 985 с использованием RGN APG03850

5

Направляющая SGN000985 (SEQ ID NO: 164)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
AGGCTGGGCCATTAAAACCTCTCCAGGGGCCGTG	218367	99,63				
AGGCTGGGCCATTAAAACCTCTCCaAGGGGCCGTG	268	0,12	33,37	Инсерция	3	1
AGGCTGGGCCATTAAAACCTCT-CAGGGGCCGTG	204	0,09	25,4	Делеция	4	1
AGGCTGGGCCATTAAAACCTCT-----CCGTG	140	0,06	17,43	Делеция	-2	7
AGGCTGGGCCATTAAAACCTCTCC-GGGGCCGTG	137	0,06	17,06	Делеция	2	1
AGGCT-GGCCATTAAAACCTCTCCAGGGGCCGTG	28	0,01	3,49	Делеция	21	1
C----- ----- -----T	26	0,01	3,24	Делеция	-39	88

Таблица 6.3. Конкретные вставки и удаления для направляющей 793 с использованием RGN APG09208

Направляющая SGN000793 (SEQ ID NO: 170)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
<u>AGGTTTTAATGGCCAGCCTCACACCCAGGCTTG</u>	169578	98,82				
C----- ----- -----G	471	0,27	23,25	Делеция	-17	80
<u>TGGAGAGGTTTTAATGGCCAGCCTCACACCC</u> <u>AGGCTGGCCTCCC</u>	398	0,23	19,64	Инсерция	+3	1
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACAC</u> --- -----G	190	0,11	9,38	Делеция	-15	18
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACAC</u> - -----TCCC	133	0,08	6,56	Делеция	-10	12
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACACC</u> <u>-AGGCTGGCCTCCC</u>	110	0,06	5,43	Делеция	0	1
<u>CTGGAGAGGTTTTAATGGCCAG</u> ----- ----GTCTGGCCTCCC	106	0,06	5,23	Делеция	-2	12
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACAGg</u> <u>CCCAGGCTGGCCTCCC</u>	102	0,06	5,03	Инсерция	+3	2
<u>CTGGAGAGGTTTTAATGGCCAGCC</u> ----- <u>-CAGGCTGGCCTCCC</u>	92	0,05	4,54	Делеция	+1	7
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACACt</u> Gtttgacctggagccactctctgcaccccgct gacc <u>CCCAGGCTGGCCTCCC</u>	61	0,04	3,01	Инсерция	+3	38
C----- ----- -----TCCC	50	0,03	2,47	Делеция	-10	44
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACACa</u> ctaggtgtattataagaatcttataaac <u>CC</u> <u>CAGGCTGGCCTCCC</u>	48	0,03	2,37	Инсерция	+3	29
<u>CTGGAGAGGTTTTAATGGCCAGCCTCACAC</u> Ccagctttcgttcgcaactcgagtggaaga ttggacttgctg <u>CCCAGGCTGGCCTCCC</u>	39	0,02	1,92	Инсерция	+3	43
<u>CTGGAGAGGTTTTAATGGCCAGCCTC</u> --- -----CC	36	0,02	1,78	Делеция	-12	18
<u>CTGGAGAGGTTTTAATGTCCCAGCCTCACAC</u> Gcactgttcacgtggctgatcatacactga tcacgtgattgatcat <u>CCCAGGCTGGCC</u> TCCC	34	0,02	1,68	Инсерция	+3	46
<u>CTGGAGAGGTTTTAATGGCCCA</u> ----- -----T	27	0,02	1,33	Делеция	-23	34

Направляющая SGN001166 (SEQ ID NO: 556)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
CGCGCGGACCACGGCTCCTCCGAAG----- ----- ----- ----- -----G	615	0,28	2,66	Делеция	140	- 13 7
CGCGCGGACCACGG----- ----- ----- ----- -----T	610	0,28	2,64	Делеция	132	- 11 8
CGCGCGGACCACGGCTCCTCC----CGAGAACA GCCCAGAAG	551	0,25	2,39	Делеция	4	3
CGCGCGGACCACGGCTCCTCCGA--CGAGAACA GCCCAGAAG	497	0,23	2,15	Делеция	2	3
GCGCGGACcACGGCTCCTCCGAAGTCGAGAACA GCCCAGAAG	438	0,2	1,9	Делеция	1	3
CGCGCGGACCACGGCTCCTCCGAAG-GAGAACA GCCCAGAAG	381	0,17	1,65	Делеция	1	2
G-----gaGGCgg----- -----G	334	0,15	1,45	Делеция	49	- 13
CGCGCGGACCACGGCTCCTCCGAA----- -----A	328	0,15	1,42	Делеция	27	- 23
CGCGGACCACGGCTCCTCCGAAGGagGAGAACA GCCCAGAAG	326	0,15	1,41	Инсерция	2	2
CGCGCGGACCACGGCTCCTCCGAAGC----- --AGAAG	321	0,15	1,39	Делеция	11	-9
CGCGCGGACCACGGCTCCTCCGAAGAG----- -----G	301	0,14	1,3	Делеция	18	- 17
CGCGCGGACCACGGCTCCTCCGAAG----- ----AAG	290	0,13	1,26	Делеция	14	- 11
CGCGCGGACCACGGCTCC----- GAGAACAGCCCAGAAG	266	0,12	1,15	Делеция	8	2
CGCGCGGACCACGGCTCCTC-----GAGAACA GCCCAGAAG	265	0,12	1,15	Делеция	6	2
CGCGCGGACCACGGCTC----- GAGAACAGCCCAGAAG	232	0,11	1,01	Делеция	9	2
GCGCGGACCACGGCTCCTCCGAAGcCGAGAACAGC CCAGAAG	217	0,1	0,94	Инсерция	1	2
CGCGCGGACCACGGCTCC-----CGAGAACA GCCCAGAAG	199	0,09	0,86	Делеция	7	3
CGCGCGGACCACGGCTCCTCCGAAG----- -----T	192	0,09	0,83	Делеция	17	- 14
CGCGCGGACCACGGCTCCTCC----- ---CAGAAG	177	0,08	0,77	Делеция	15	-8
CGCGCGGACCACGGC-----GAGAACA GCCCAGAAG	158	0,07	0,68	Делеция	11	2
CGCGCGGACCACGGCTCCTCC----- ----- ----- -----T	149	0,07	0,65	Делеция	97	- 90
CGCGCGGACCACGGC-----CGAGAACA GCCCAGAAG	147	0,07	0,64	Делеция	10	3

Направляющая SGN001166 (SEQ ID NO: 556)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
CGCGCGGACCACGGCTCC----- ---CAGAAG	131	0,06	0,57	Делеция	18	-8
CGCGCGGACCACGGCTCCTCCGAAGC--- ---CAGAAG	127	0,06	0,55	Делеция	10	-8
CGCGCGGACCACGGCTCCTCCGAAG----- ----- -----G	124	0,06	0,54	Делеция	53	-50
CGCGCGGACCACGGCCCCCTCCGA-----GAACA GCCCAGAAG	113	0,05	0,49	Делеция	5	0
CGCGCGGACCACGGCTCCTCC----- --AGAAG	112	0,05	0,49	Делеция	16	-9
CGCGCGGACCACGGCTCC----- AGCCCAGAAG	104	0,05	0,45	Делеция	14	-4
CGCGCGGA----- GAACAGCCCAGAAG	102	0,05	0,44	Делеция	20	0
CGCGCGGACCACGGC----- CCAGAAG	101	0,05	0,44	Делеция	20	-7
CGCGCG----- AGAACAGCCCAGAAG	98	0,04	0,42	Делеция	21	1
CGCGCGGACCACGGCTCCTCCG---CGAGAACA GCCCAGAAG	97	0,04	0,42	Делеция	3	3
CGCGCGGACCACGGCTCCTCCG----- ----- ----- -G	94	0,04	0,41	Делеция	78	-72
CGCGCGGACCACGGCTCCTCCGAAG---GAACA GCCCAGAAG	91	0,04	0,39	Делеция	3	0
CGCGCGGACCACGGCTCCTCCGAAG----- ----AGAAG	90	0,04	0,39	Делеция	12	-9
CGCGC----- --CCAGAAG	80	0,04	0,35	Делеция	30	-7
GCGCGGACCACGGCTCCTCCGAAGCCAGAACA GCCCAGAAG	74	0,03	0,32	Инсерция	1	1
C----- -----AGAAG	73	0,03	0,32	Делеция	56	-9
CGCGCGGACCACGGCTCCTCCGAAG----- ----- -----A	67	0,03	0,29	Делеция	47	-44
CGCGCGGACCACGGCTCCTCCGAA-CGAGAACA GCCCAGACG	66	0,03	0,29	Делеция	1	3
CGCGCGGACCACGGCTC----- ----AGAAG	63	0,03	0,27	Делеция	20	-9
CGCGCGGACC----- ----CAGAAG	63	0,03	0,27	Делеция	26	-8
CGCGCGGACCACGGCTCCTCCG----- ----- ----- ----- ----- -----T	62	0,03	0,27	Делеция	179	-173
C----- -----AGAAG	60	0,03	0,26	Делеция	51	-9

Направляющая SGN001166 (SEQ ID NO: 556)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
CGCGCGGACCACGGCTCCTCC-----A GCCCAGAAG	59	0,03	0,26	Делеция	11	-4
CGCGCGGACCAC----- ----AGAAG	56	0,03	0,24	Делеция	25	-9
CGCGCGGAC-----GAGAACA GCCCAGAAG	55	0,02	0,24	Делеция	17	2
CGCGCGGACCACGGCTCCTCT----CGAGAACA GCCCAGAAG	54	0,02	0,23	Делеция	4	3
CGCGCGGACCACGGCTCCTCCGAAG----- GCCCAGAAG	54	0,02	0,23	Делеция	8	-5
G----- -----AGAACAGCCCAGAAG	53	0,02	0,23	Делеция	39	1
CGCGCGGACCA----- -----GAAG	52	0,02	0,23	Делеция	27	-10
CGCGCGGACCACGGCTCC----- ----AGAAG	52	0,02	0,23	Делеция	19	-9
CGCGCGGACCACGGCTCCTCCG----GAGAACA GCCCAGAAG	50	0,02	0,22	Делеция	4	2
CGCGCGGACCACGGCTCC----- -CCCAGAAG	48	0,02	0,21	Делеция	16	-6
CGCGCGGACCACGGCTCCTCCGA----- -----GAAG	48	0,02	0,21	Делеция	15	-10
CGCGCGGA-----ACA GCCCAGAAG	47	0,02	0,2	Делеция	22	-2
CGCGCGGACCACGGCTCCTCCGAAG-----ACA GCCCAGAAG	46	0,02	0,2	Делеция	5	-2
CGCGCGGACCACGGCTCCT-----GCGAGAACA GCCCAGAAG	45	0,02	0,19	Делеция	5	4
CGCGCGGACC-----CGAGAACA GCCCAGAAG	45	0,02	0,19	Делеция	15	3
CGCG-----AGAACA GCCCAGAAG	44	0,02	0,19	Делеция	23	1
CGCGCGGACCA-----GAACA GCCCAGAAG	44	0,02	0,19	Делеция	17	0
CG-----A GCCCAGAAG	44	0,02	0,19	Делеция	30	-4
CGCGCGGACCACGGCTCCTCCGA-----GAACA GCCCAGACG	44	0,02	0,19	Делеция	5	0
CGCGCGGACCACGGCTCCTCGGA-----GAACA GCCCAGAAG	42	0,02	0,18	Делеция	5	0
CGCGCGGACCACGGCTCCTCCTC-CGAGAACA GCCCAGAAG	42	0,02	0,18	Делеция	2	3
CGCGCGG-----GGAAGAACA GCCCAGAAG	41	0,02	0,18	Делеция	17	4
CGCGCGGACCACGGCTCCTCCGAAG----- -----GAAG	39	0,02	0,17	Делеция	13	-10
CGCGCGGACC-----GAGAACA GCCCAGAAG	39	0,02	0,17	Делеция	16	2
CGCGCGGACCACGGCTCCTCCGAAG----- -----GA-----G	39	0,02	0,17	Делеция	17	-16
GCGGACCACGGCTCCTCCGAAGaagCGAGAACA GCCCAGAAG	38	0,02	0,16	Инсерция	3	3

Направляющая SGN001166 (SEQ ID NO: 556)	# прочтений	% прочтений	% INDEL	Тип	Локализация INDEL	Размер
GGACCACGGCTCCTCCGAAGGAGGagGAGAACA GCCCAGAAG	38	0,02	0,16	Инсерция	5	2
CGCGCGGACCAC----- -----AAG	38	0,02	0,16	Делеция	27	-11
CGCGCGGACCACGGCTCCTCCGAAGC-----A GCCCAGAAG	38	0,02	0,16	Делеция	6	-4
T----- -----T	37	0,02	0,16	Делеция	47	-15
CGCGCGGACCACGGCTGTTCTGA-----GAACA GCCCAGAAG	37	0,02	0,16	Делеция	5	0
C----- CGAGAACAGCCCAGAAG	36	0,02	0,16	Делеция	31	3
CGCGCGGACCACGGC----- -----GAAG	36	0,02	0,16	Делеция	23	-10
CGCGCGGA----- -----AG	35	0,02	0,15	Делеция	32	-12
CGCGCGGACCACGG-----AACAA GCCCAGAAG	34	0,02	0,15	Делеция	15	-1
CGCGCGGACCACGGCTC-----AACAA GCCCAGAAG	34	0,02	0,15	Делеция	12	-1
CGCGCGGACCACGGCCCCCTCC----CGAGAACA GCCCAGAAG	33	0,01	0,14	Делеция	4	3
G----- -----A	33	0,01	0,14	Делеция	57	-21
CGCGCGGACCACGGCTCCTCCGA-GCGAGAACA GCCCAGAAG	32	0,01	0,14	Делеция	1	4
CGCGCGGACCACGTT-----CGAGAACA GCCCAGAAG	32	0,01	0,14	Делеция	10	3
CGCGA-----AGAACA GCCCAGAAG	31	0,01	0,13	Делеция	22	1
CGCGCGGACCACGG----- -----AAG	31	0,01	0,13	Делеция	25	-11
CGCG-----AACAA GCCC-----G	30	0,01	0,13	Делеция	36	-20
CGCGCGGACCACGGCTCCTCCGAAG----AACAA GCCCAGCAG	30	0,01	0,13	Делеция	4	-1
CGCGCGGACCACGGCTCCT-----GAACA GCCCAGAAG	30	0,01	0,13	Делеция	9	0
CGCGGACCACGGCTCCTCCGAAGcGCGAGAACA GCCCAGAAG	30	0,01	0,13	Инсерция	2	1
CG----- AGAACAGCCCAGAAG	29	0,01	0,13	Делеция	25	1
CGCGCGGACCACGGCCCCCTCCGAAG---- AACAGCCCAGAAG	29	0,01	0,13	Делеция	4	-1
CGCGCGGACCAC----- AGCCCAGAAG	29	0,01	0,13	Делеция	20	-4
CGCGCGGACCACGG----- ACAGCCCAGAAG	29	0,01	0,13	Делеция	16	-2
CGCGCGGACCACGGCTCCT----- -----AGAAG	28	0,01	0,12	Делеция	17	-9

Надежность нескольких нуклеаз была проанализирована путем тестирования их способности редактировать множество различных сайтов-мишеней в нескольких генах, На расширенной панели направляющих протестировано более 40 мишеней, Все протестированные белки показали надежное редактирование на различных сайтах. Результаты приведены в табл. 7.

Таблица 7. Надежность выбранных РНК-направляемых нуклеаз

RGN	Ген	Номер SGN	Степень общего редактирования (%)
APG01868	TRA	SGN001684	42,62
		SGN001685	1,43
		SGN001686	8,19
		SGN001687	8,83
		SGN001688	0,21
		SGN001689	0
		SGN001690	0
		SGN001691	20,86
	EMX1	SGN001697	0,93
		SGN001698	0
		SGN001699	18,84
		SGN001700	21,8
		SGN001701	39,94
		SGN001702	2,16
		SGN001703	0,79
		SGN001704	3,49
		SGN001705	14,49
		SGN001706	1,04
		SGN001707	2,51
		SGN001708	1,75
		SGN001709	3,62
	HAO1	SGN001710	0
		SGN001711	13,39
		SGN001713	22,29
		SGN001714	0,67
		SGN001715	0,23
		SGN001716	7,1
		SGN001717	0,68
		SGN001718	0,21
		SGN001719	0
SGN001721		0,17	
SGN001722		1,36	
SGN001723	24,16		
SGN001724	0,2		

RGN	Ген	Номер SGN	Степень общего редактирования (%)
	LDHA	SGN001725	4,92
		SGN001726	0
		SGN001727	1,32
		SGN001728	0,28
		SGN001729	1,94
		SGN001730	5,2
		SGN001731	11,28
		SGN001732	6,04
		SGN001733	1,12
		SGN001734	2,11
		SGN001735	7
APG05586	TRA	SGN001371	30,52
		SGN001372	26,01
		SGN001373	26,67
		SGN001374	22,94
		SGN001375	21,17
		SGN001376	0
		SGN001377	0
		SGN001378	30,04
		SGN001379	39,13
		SGN001380	4,93
		SGN001381	17,14
		SGN001382	5,6
	B2M	SGN001383	36,23
		SGN001384	44,28
		SGN001385	6,33
		SGN001386	5,55
		SGN001387	48,71
		SGN001388	31,78
		SGN001389	41
		SGN001390	48,15
		SGN001391	46,19
		SGN001392	37,22
		SGN001393	31,62
	LDHA	SGN001394	29,72
		SGN001396	32,88
		SGN001397	41,45
		SGN001399	43,28
		SGN001400	2,52
		SGN001401	37,52
		SGN001402	0,37
	SGN001403	53,13	

RGN	Ген	Номер SGN	Степень общего редактирования (%)
		SGN001404	44,06
		SGN001405	1,46
	HAO1	SGN001406	21,98
		SGN001407	9,13
		SGN001408	25,06
		SGN001409	43,81
		SGN001410	37,6
		SGN001411	40,75
		SGN001412	18,2
		SGN001413	28,44
		SGN001414	29,39
		SGN001415	0,39
		SGN001416	43,59
		APG01604	B2M
SGN001593	22,7		
SGN001594	23,6		
SGN001595	33,23		
SGN001596	20,88		
SGN001597	1,26		
SGN001598	26,7		
SGN001599	9,41		
SGN001600	28,88		
SGN001601	7,39		
SGN001602	27,8		
SGN001603	6,93		
HAO1	SGN001616		
	SGN001617		2,55
	SGN001618		0,6
	SGN001619		6,29
	SGN001620		7,92
	SGN001621		13,03
	SGN001622		5,32
	SGN001623		18,58
	SGN001624		20,03
	SGN001625		1,65
LDHA	SGN001626		0,18
	SGN001627		0,73
	SGN001640		3,75
	SGN001641		1,13
	SGN001642		14,2
	SGN001643		12,16
	SGN001644	4,9	

RGN	Ген	Номер SGN	Степень общего редактирования (%)
		SGN001645	4,78
		SGN001646	0,74
		SGN001647	2,89
		SGN001648	0
		SGN001649	0,15
		SGN001650	4,4
		SGN001651	7,26
	TRA	SGN001664	0,44
		SGN001665	3,4
		SGN001666	17,24
		SGN001667	2,33
		SGN001668	0,3
		SGN001669	13,68
		SGN001670	0,54
		SGN001671	5,09
		SGN001672	0,22
		SGN001673	12,22
		SGN001674	17,33
		SGN001675	11,69
APG09298	B2M	SGN001383	22,25
		SGN001384	9,5
		SGN001385	2,37
		SGN001386	0,4
		SGN001387	38,03
		SGN001388	17,66
		SGN001389	42,03
		SGN001390	42,88
		SGN001391	16,31
		SGN001392	16,5
		SGN001393	6,27
	SGN001394	17,15	
	HAO1	SGN001406	3,46
		SGN001407	1,87
		SGN001408	10,71
		SGN001409	18,79
		SGN001410	15,09
		SGN001411	14,81
		SGN001412	0,63
		SGN001413	5,92
		SGN001414	5,24
		SGN001415	0,03
SGN001416		15,41	

RGN	Ген	Номер SGN	Степень общего редактирования (%)
	LDHA	SGN001395	3,15
		SGN001396	28,53
		SGN001397	11,69
		SGN001399	23,92
		SGN001400	0,89
		SGN001401	33,97
		SGN001402	0
		SGN001403	25,62
		SGN001404	16,96
		SGN001405	0,11
	TRA	SGN001162	44,26
		SGN001163	42,06
		SGN001164	8,1
		SGN001371	21,05
		SGN001372	3,95
		SGN001373	7,8
		SGN001374	9,04
		SGN001375	2,94
		SGN001376	0
		SGN001377	0
		SGN001378	22,14
		SGN001379	27,43
		SGN001380	0,16
SGN001381	14,18		
SGN001382	2,78		

Пример 5. Белковая инженерия APG05586

Консервативные и переменные остатки APG05586 идентифицируют путем сравнения APG05586 с близкородственными гомологами, включая APG08770 (SEQ ID NO: 70), APG09882 (указан как SEQ ID NO: 568 и описан в PCT/US2020/045759, сущность приведенного патента полностью включена в настоящее изобретение в виде ссылки), и APG01658 (указан как SEQ ID NO: 569 и описан в PCT/US2020/045759). Чтобы идентифицировать критические аминокислотные остатки в белке APG05586 конструируют несколько вариантов, содержащих мутации в неконсервативных местоположениях. В общей сложности 132 остатка изменены в десяти модифицированных вариантах RGN APG05586 (SEQ ID NO: 570-579). Затем эти десять вариантов APG05586 и дикий тип APG05586 анализируют на активность в клетках млекопитающих. Используя

направляющую РНК для APG05586 (SEQ ID NO: 66), RGN тестируют на активность по шести целевым местоположениям генома (табл. 8), следуя методам, описанным в примере 4. Кодон-оптимизированные кодирующие последовательности для млекопитающих для каждого варианта представлены в виде SEQ ID NO: 580-589. 5' и 3' нуклеотидные последовательности праймеров для амплификации, используемые для обнаружения активности редактирования генов, также представлены в табл. 8. Степени редактирования приведены в табл. 9.

Таблица 8. Последовательности-мишени

ID направляющей	Ген	Последовательность-мишень (SEQ ID NO:)	sgRNA (SEQ ID NO:)	5' праймер для амплификации	3' праймер для амплификации
SGN001159	EMX1	590	591	592	593
SGN001162	TRA	594	595	596	597
SGN001163	TRA	598	599	596	597
SGN001164	TRA	600	601	596	597
SGN001165	VEGFA	602	603	604	605
SGN001166	VEGFA	606	607	604	605

10

Таблица 9. Степень редактирования для APG05586 и ее вариантов

RGN ID	RGN SEQ ID NO:	SGN001159	SGN001162	SGN001163	SGN001164	SGN001165	SGN001166
APG09298	570	33,40%	44,26%	42,06%	8,10%	33,94%	46,19%
APG06251	571	0%	43,59%	48,82%	5,90%	25,07%	35,08%
APG03066	572	0%	0%	0%	0%	0%	0%
APG01560	573	15,12%	16,38%	17,59%	2,52%	44,99%	20,68%
APG02777	574	0%	0,08%	0%	0%	0%	0%
APG05761	575	24,25%	21,95%	21,86%	10,45%	23,72%	20,17%
APG02479	576	5,61%	9,25%	11,00%	1,45%	4,53%	6,34%
APG08385	577	31,16%	39,95%	35,42%	2,37%	31,41%	26,65%
APG09217	578	29,43%	41,12%	44,71%	4,79%	38,16%	31,76%
APG06657	579	36,81%	44,20%	41,54%	4,44%	16,29%	21,86%
APG05586	63	40,13%	46,12%	43,77%	8,00%	96,76%	77,38%

Относительная активность вариантов позволяет предположить, какие участки в белке толерантны к мутациям. Табл. 10, приведенная ниже, представляет собой сводную информацию об активности мутантного образца RGN с разным количеством введенных мутаций по сравнению с APG005586. «-»

15

означает отсутствие активности; «+» означает 1-15% редактирования по крайней мере в 4 из 6 мишеней; «++» означает 10-25% редактирования по крайней мере в 4 из 6 мишеней; и «+++» - 20-50% редактирования по крайней мере в 4 из 6 мишеней.

5 Таблица 10. Суммирование изменений в белковой инженерии и степени редактирования

Белок	Результат активности	Число мутаций
APG09298	+++	10
APG06251	+++	12
APG03066	-	87
APG01560	++	10
APG02777	-	37
APG05761	++	10
APG02479	+	12
APG08385	+++	12
APG09217	+++	14
APG06657	+++	14
APG05586	+++	0

Варианты APG03066 и APG02777 содержат слишком много мутаций, чтобы идентифицировать специфические остатки, важные для функционирования, однако низкая активность этих вариантов указывает на то, что обширные изменения в мостиковой спирали и домене распознавания в этом белке недопустимы. Все остальные варианты содержат 14 или менее мутаций, что позволяло идентифицировать специфические остатки, важные для активности. Основываясь на этих результатах, несколько остатков были идентифицированы как важные для функционирования белка. Мутации I305L, V328A, L366I, T368S и V405A приводят к снижению активности в анализируемых вариантах. Предполагают, что все эти мутации находятся в домене узнавания белка. Снижение активности для APG01560 является результатом множественных изменений, не локализованных в определенной области внутри белка.

20 Пример 6. Идентификация мишеней, связанных с заболеваниями

База данных клинических вариантов была получена из базы данных NCBI ClinVar, которая доступна через Интернет на веб-сайте NCBI ClinVar. В этом списке были идентифицированы патогенные однонуклеотидные полиморфизмы (Single Nucleotide Polymorphisms - SNP). Используя информацию о геномном локусе, были идентифицированы мишени CRIGvSPR в области, перекрывающей и окружающей каждый SNP, в табл. 11 приведены SNP, которые могут быть скорректированы за счет использования комбинации редактора оснований и

RGN по настоящему изобретению для нацеливания на вызывающую мутацию (Causal Mutation – «Casl Mut»). В табл. 11 указан только один вариант обозначения каждого из перечисленных заболеваний, «№ RS» соответствует регистрационному номеру RS в базе данных SNP на веб-сайте NCBI, ID аллелей соответствует номеру доступа вызывающего аллеля, кроме того, номер доступа хромосомы также предоставляет справочную информацию о присоединении, которую можно найти на веб-сайте NCBI. В табл. 11 также представлена информация о последовательности генома-мишени, соответствующей RGN, для каждого заболевания. Информация о последовательности-мишени также обеспечивает последовательность протоспейсера для получения необходимой sgRNA для соответствующей RGN по настоящему изобретению.

Таблица 11. Связанные с заболеванием мишени для RGN

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Болезнь Штаргардта I	1800553	APG06622	C>T	22927	NC_000001,10,NC_000001,11	ABCA4	261
Болезнь Штаргардта I	1800728	APG06622	A>G	98777	NC_000001,10,NC_000001,11	ABCA4	262
Болезнь Штаргардта I	61751374	APG06622	G>A	22933	NC_000001,10,NC_000001,11	ABCA4	263
Болезнь Штаргардта I	61750641	APG02874, APG03031, APG09208	G>A	105317	NC_000001,10,NC_000001,11	ABCA4	264
Болезнь Штаргардта I	1800553	APG02787	G>A	22927	NC_000001,10,NC_000001,11	ABCA4	265
Болезнь Штаргардта I	1800553	APG06007	G>A	22927	NC_000001,10,NC_000001,11	ABCA4	266
Болезнь Штаргардта I	1800553	APG03850	G>A	22927	NC_000001,10,NC_000001,11	ABCA4	267
Болезнь Штаргардта I	1800553	APG05586	G>A	22927	NC_000001,10,NC_000001,11	ABCA4	268
Болезнь Штаргардта I	1800553	APG08167, APG01604	C>T	22927	NC_000001,10,NC_000001,11	ABCA4	269
Семейный гиперинсулинизм	1,51E+08	APG02874, APG03031, APG09208	G>A	24127	NC_000011,9,NC_000011,10	ABCC8	270
Дефицит ацил-КоА дегидрогеназы жирных кислот с очень длинной углеродной цепью	1,14E+08	APG06622	T>C	33877	NC_000017,10,NC_000017,11	ACADVL	271
Дефицит ацил-КоА дегидрогеназы жирных кислот с очень длинной углеродной цепью	1,14E+08	APG06248	T>C	33877	NC_000017,10,NC_000017,11	ACADVL	272
Дефицит ацил-КоА дегидрогеназы жирных кислот с очень длинной углеродной цепью	3,7E+08	APG02874, APG03031, APG09208	G>A	98197	NC_000017,10,NC_000017,11	ACADVL	273
Синдром Барайцера – Уинтера I	2,82E+08	APG02874, APG03031, APG09208	G>A	38553	NC_000007,13,NC_000007,14	ACTB	274

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Тяжелый иммунодефицит из-за дефицита ADA	1,22E+08	APG02874, APG03031, APG09208	G>A	16996	NC_000020,10,NC_000020,11	ADA	275
Тяжелый иммунодефицит из-за дефицита ADA	1,22E+08	APG02874, APG03031, APG09208	T>C	17004	NC_000020,10,NC_000020,11	ADA	276
Первичная гипероксалурия	1,22E+08	APG06622	G>A	38436	NC_000002,11,NC_000002,12	AGXT	277
Врожденное нарушение гликозилирования	28939378	APG06622	C>T	19763	NC_000016,9,NC_000016,10	ALG1	278
Гипофосфатазия	1,22E+08	APG06622	G>A	28709	NC_000001,10,NC_000001,11	ALPL	279
Гипофосфатазия	1,22E+08	APG02874, APG03031, APG09208	G>A	28709	NC_000001,10,NC_000001,11	ALPL	280
Колоректальный рак	1,38E+08	APG06622	C>T	15837	NC_000005,9,NC_000005,10	APC	281
Метахроматическая лейкодиетрофия	80338815	APG06622	C>T	18090	NC_000022,10,NC_000022,11	ARSA	282
Болезнь Вильсона	1,94E+08	APG02874, APG03031, APG09208	G>A	44393	NC_000013,10,NC_000013,11	ATP7B	283
Кардио-лицевой-кожный синдром	1,8E+08	APG06622	T>C	29012	NC_000007,13,NC_000007,14	BRAF	284
Кардио-лицевой-кожный синдром	1,8E+08	APG06248	T>C	29012	NC_000007,13,NC_000007,14	BRAF	285
Рак молочной железы и/или яичников	41293455	APG06622	G>A	32714	NC_000017,10,NC_000017,11	BRCA1	286
Рак молочной железы и/или яичников	41293465	APG06622	G>A	70268	NC_000017,10,NC_000017,11	BRCA1	287
Рак молочной железы и/или яичников	55770810	APG06622	G>A	70063	NC_000017,10,NC_000017,11	BRCA1	288
Рак молочной железы и/или яичников	62625307	APG06622	G>A	69596	NC_000017,10,NC_000017,11	BRCA1	289
Рак молочной железы и/или яичников	62625308	APG06622	G>A	32710	NC_000017,10,NC_000017,11	BRCA1	290
Рак молочной железы и/или яичников	80356962	APG06622	C>T	70247	NC_000017,10,NC_000017,11	BRCA1	291

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Рак молочной железы и/или яичников	41293455	APG02787	C>T	32714	NC_000017,10,NC_000017,11	BRCA1	292
Рак молочной железы и/или яичников	41293455	APG06007	C>T	32714	NC_000017,10,NC_000017,11	BRCA1	293
Рак молочной железы и/или яичников	41293455	APG03850	C>T	32714	NC_000017,10,NC_000017,11	BRCA1	294
Рак молочной железы и/или яичников	41293455	APG05586	C>T	32714	NC_000017,10,NC_000017,11	BRCA1	295
Рак молочной железы и/или яичников	41293455	APG08167, APG01604	G>A	32714	NC_000017,10,NC_000017,11	BRCA1	296
Рак молочной железы и/или яичников	80356962	APG08167, APG01604	C>T	70247	NC_000017,10,NC_000017,11	BRCA1	297
Рак молочной железы и/или яичников	80358163	APG08167, APG01604	T>C	46006	NC_000017,10,NC_000017,11	BRCA1	298
Рак молочной железы и/или яичников	45580035	APG06622	C>T	67431	NC_000013,10,NC_000013,11	BRCA2	299
Рак молочной железы и/или яичников	80359212	APG06622	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	300
Рак молочной железы и/или яичников	80359003	APG02874, APG03031, APG09208	G>A	67069	NC_000013,10,NC_000013,11	BRCA2	301
Рак молочной железы и/или яичников	80359004	APG02874, APG03031, APG09208	G>A	46672	NC_000013,10,NC_000013,11	BRCA2	302
Рак молочной железы и/или яичников	80359212	APG02787	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	303
Рак молочной железы и/или яичников	80359212	APG06007	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	304
Рак молочной железы и/или яичников	80359212	APG03850	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	305
Рак молочной железы и/или яичников	80359212	APG05586	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	306
Рак молочной железы и/или яичников	80359212	APG08167, APG01604	C>T	67494	NC_000013,10,NC_000013,11	BRCA2	307
Рак молочной железы и/или яичников	80359071	APG08167, APG01604	G>A	67203	NC_000013,10,NC_000013,11	BRCA2	308

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Расстройства, связанные с CAPN3	1,21E+08	APG02874, APG03031, APG09208	G>A	32661	NC_000015,9,NC_000015,10	CAPN3;POMT1	309
Дефицит CBS	5742905	APG06622	A>G	15159	NC_000021,8,NC_000021,9	CBS	310
Дефицит CBS	5742905	APG02874, APG03031, APG09208	T>C	15159	NC_000021,8,NC_000021,9	CBS	311
Дефицит CBS	1,22E+08	APG02874, APG03031, APG09208	G>A	15156	NC_000021,8,NC_000021,9	CBS	312
Муковисцидоз	77010898	APG06622	G>A	22168	NC_000007,13,NC_000007,14	CFTR	313
Муковисцидоз	75096551	APG02874, APG03031, APG09208	G>A	33858	NC_000007,13,NC_000007,14	CFTR	314
Муковисцидоз	75527207	APG02787	G>A	22159	NC_000007,13,NC_000007,14	CFTR	315
Муковисцидоз	75527207	APG06007	G>A	22159	NC_000007,13,NC_000007,14	CFTR	316
Муковисцидоз	75527207	APG03850	G>A	22159	NC_000007,13,NC_000007,14	CFTR	317
Муковисцидоз	75527207	APG05586	G>A	22159	NC_000007,13,NC_000007,14	CFTR	318
Муковисцидоз	78655421	APG02787	G>A	22148	NC_000007,13,NC_000007,14	CFTR	319
Муковисцидоз	78655421	APG06007	G>A	22148	NC_000007,13,NC_000007,14	CFTR	320
Муковисцидоз	78655421	APG03850	G>A	22148	NC_000007,13,NC_000007,14	CFTR	321
Муковисцидоз	78655421	APG05586	G>A	22148	NC_000007,13,NC_000007,14	CFTR	322
Муковисцидоз	75527207	APG08167, APG01604	G>A	22159	NC_000007,13,NC_000007,14	CFTR	323
Врожденная миотония	80356701	APG02874, APG03031, APG09208	T>C	33902	NC_000007,13,NC_000007,14	CLCN1	324

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Несовершенный остеогенез, I тип	72645321	APG02874, APG03031, APG09208	G>A	414022	NC_000017,10,NC_000017,11	COL1A1	325
Несовершенный остеогенез, I тип	72645321	APG02874, APG03031, APG09208	G>A	414022	NC_000017,10,NC_000017,11	COL1A1	326
Синдром Альпорта I типа, X-сцепленный рецессивный	1,05E+08	APG02874, APG03031, APG09208	G>A	35796	NC_000023,10,NC_000023,11	COL4A5	327
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG06622	C>T	23992	NC_000001,10,NC_000001,11	CPT2	328
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG06248	C>T	23992	NC_000001,10,NC_000001,11	CPT2	329
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG06007	C>T	23992	NC_000001,10,NC_000001,11	CPT2	330
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG03850	C>T	23992	NC_000001,10,NC_000001,11	CPT2	331
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG05586	C>T	23992	NC_000001,10,NC_000001,11	CPT2	332
Дефицит карнитинпальмитоилтрансферазы II	74315294	APG08167, APG01604	C>T	23992	NC_000001,10,NC_000001,11	CPT2	333
Дефицит бета-гидроксилазы дофамина	74853476	APG02874, APG03031, APG09208	T>C	16789	NC_000009,11,NC_000009,12	DBH	334
Врожденная микроцефалия	11555217	APG06622	C>T	34125	NC_000011,9,NC_000011,10	DHCR7	335
Синдром Смита -Лемли - Опица	80338853	APG06622	G>A	21822	NC_000011,9,NC_000011,10	DHCR7	336
Синдром Смита -Лемли - Опица	80338857	APG02874, APG03031, APG09208	G>A	34128	NC_000011,9,NC_000011,10	DHCR7	337
Синдром Смита -Лемли - Опица	11555217	APG06007	G>A	34125	NC_000011,9,NC_000011,10	DHCR7	338

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Синдром Смита -Лемли - Опица	11555217	APG03850	G>A	34125	NC_000011,9,NC_000011,10	DHCR7	339
Синдром Смита -Лемли - Опица	11555217	APG05586	G>A	34125	NC_000011,9,NC_000011,10	DHCR7	340
Семейная дизавтономия	1,11E+08	APG06248	A>G	21124	NC_000009,11,NC_000009,12	ELP1	341
Гипертирозинемия	80338901	APG06622	G>A	26909	NC_000015,9,NC_000015,10	FAH	342
Гипертирозинемия	80338901	APG06248	G>A	26909	NC_000015,9,NC_000015,10	FAH	343
Анемия Фанкони	1,05E+08	APG06622	G>A	27086	NC_000009,11,NC_000009,12	FANCC	344
Синдром Марфана	3,98E+08	APG06622	C>T	51454	NC_000015,9,NC_000015,10	FBN1	345
Синдром Марфана	7,28E+08	APG06622	A>G	175979	NC_000015,9,NC_000015,10	FBN1	346
Синдром Марфана	1,38E+08	APG02874, APG03031, APG09208	G>A	31496	NC_000015,9,NC_000015,10	FBN1	347
Синдром Марфана	1,38E+08	APG02874, APG03031, APG09208	G>A	31496	NC_000015,9,NC_000015,10	FBN1	348
Синдром Марфана	3,88E+08	APG02874, APG03031, APG09208	G>A	38652	NC_000015,9,NC_000015,10	FBN1	349
FGFR3-связанные расстройства	1,22E+08	APG06622	C>T	31371	NC_000004,11,NC_000004,12	FGFR3	350
Болезнь накопления гликогена, тип IA	1801175	APG08167, APG01604	C>T	27037	NC_000017,10,NC_000017,11	G6PC	351
Болезнь накопления гликогена, тип II	3,98E+08	APG02874, APG03031, APG09208	G>A	415590	NC_000017,10,NC_000017,11	GAA	352
Дефицит UDP-глюкозо-гексозо-1-фосфатуридилтрансферазы	75391579	APG06622	A>G	18653	NC_000009,11,NC_000009,12	GALT	353
Глутаровая ацидурия, 1 тип	1,21E+08	APG02874, APG03031, APG09208	G>A	17127	NC_000019,9,NC_000019,10	GCDH	354

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
X-сцепленная глухота	76434661	APG06622	C>T	53916	NC_000013,10,NC_000013,11	GJB2	355
X-сцепленная глухота	80338945	APG06622	A>G	32055	NC_000013,10,NC_000013,11	GJB2	356
X-сцепленная глухота	80338945	APG02874, APG03031, APG09208	T>C	32055	NC_000013,10,NC_000013,11	GJB2	357
X-сцепленная глухота	80338945	APG02874, APG03031, APG09208	T>C	32055	NC_000013,10,NC_000013,11	GJB2	358
X-сцепленная глухота	1,11E+08	APG02874, APG03031, APG09208	G>A	53902	NC_000013,10,NC_000013,11	GJB2	359
X-сцепленная глухота	1,05E+08	APG06007	G>A	32041	NC_000013,10,NC_000013,11	GJB2	360
X-сцепленная глухота	1,05E+08	APG03850	G>A	32041	NC_000013,10,NC_000013,11	GJB2	361
X-сцепленная глухота	1,05E+08	APG08167, APG01604	C>T	32041	NC_000013,10,NC_000013,11	GJB2	362
X-сцепленная глухота	80338945	APG08167, APG01604	A>G	32055	NC_000013,10,NC_000013,11	GJB2	363
Миопатия с телами включения 2	28937594	APG06622	A>G	21064	NC_000009,11,NC_000009,12	GNE	364
Миопатия с телами включения 2	28937594	APG06248	A>G	21064	NC_000009,11,NC_000009,12	GNE	365
бета Талассемия	33930165	APG02874, APG03031, APG09208	G>A	30165	NC_000011,9,NC_000011,10	HBV	366
Мукополисахаридоз, I тип	1,22E+08	APG06622	G>A	26947	NC_000004,11,NC_000004,12	IDUA	367
Мукополисахаридоз, I тип	1,22E+08	APG06622	C>T	26948	NC_000004,11,NC_000004,12	IDUA	368
Мукополисахаридоз, I тип	1,22E+08	APG02874, APG03031, APG09208	G>A	26947	NC_000004,11,NC_000004,12	IDUA	369

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Врожденный синдром удлиненного интервала QT	1,99E+08	APG02874, APG03031, APG09208	G>A	67758	NC_000011,9,NC_000011,10	KCNQ1	370
Врожденный синдром удлиненного интервала QT	1,99E+08	APG02874, APG03031, APG09208	T>C	67776	NC_000011,9,NC_000011,10	KCNQ1	371
Миопатия с телами включения 2	28942080	APG06622	G>A	18735	NC_000019,9,NC_000019,10	LDLR	372
Миопатия с телами включения 2	1,22E+08	APG06622	C>T	18725	NC_000019,9,NC_000019,10	LDLR	373
бета Талассемия	1,38E+08	APG06622	G>A	171217	NC_000019,9,NC_000019,10	LDLR	374
Мукополисахаридоз, I тип	7,46E+08	APG06622	C>T	228192	NC_000019,9,NC_000019,10	LDLR	375
Мукополисахаридоз, I тип	7,66E+08	APG06622	G>A	228162	NC_000019,10,NC_000019,9	LDLR	376
Мукополисахаридоз, I тип	7,69E+08	APG06622	G>A	228176	NC_000019,10,NC_000019,9	LDLR	377
Врожденный синдром удлиненного интервала QT	7,46E+08	APG06248	C>T	228192	NC_000019,9,NC_000019,10	LDLR	378
Врожденный синдром удлиненного интервала QT	7,69E+08	APG06248	G>A	228176	NC_000019,10,NC_000019,9	LDLR	379
Миопатия с телами включения 2	3,76E+08	APG02874, APG03031, APG09208	G>A	198012	NC_000019,10,NC_000019,9	LDLR	380
Миопатия с телами включения 2	7,69E+08	APG02874, APG03031, APG09208	G>A	228176	NC_000019,10,NC_000019,9	LDLR	381

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
бета Талассемия	7,75E+08	APG02874, APG03031, APG09208	T>C	228197	NC_000019,9,NC_000019,10	LDLR	382
Мукополисахаридоз, I тип	7,76E+08	APG02874, APG03031, APG09208	G>A	246116	NC_000019,9,NC_000019,10	LDLR	383
Мукополисахаридоз, I тип	8,79E+08	APG02874, APG03031, APG09208	T>C	246008	NC_000019,10,NC_000019,9	LDLR	384
Мукополисахаридоз, I тип	1,38E+08	APG02787	G>A	171217	NC_000019,9,NC_000019,10	LDLR	385
Врожденный синдром удлиненного интервала QT	1,38E+08	APG06007	G>A	171217	NC_000019,9,NC_000019,10	LDLR	386
Врожденный синдром удлиненного интервала QT	1,38E+08	APG03850	G>A	171217	NC_000019,9,NC_000019,10	LDLR	387
Миопатия с телами включения 2	1,38E+08	APG05586	G>A	171217	NC_000019,9,NC_000019,10	LDLR	388
Миопатия с телами включения 2	1,38E+08	APG08167, APG01604	G>A	171217	NC_000019,9,NC_000019,10	LDLR	389
Бета-талассемия	1,22E+08	APG08167, APG01604	C>T	18725	NC_000019,9,NC_000019,10	LDLR	390
Кардио-лицевой-кожный синдром	1,22E+08	APG06622	A>G	28390	NC_000015,9,NC_000015,10	MAP2K1	391
MECP2-связанные расстройства	28934906	APG06622	G>A	26850	NC_000023,10,NC_000023,11	MECP2	392
MECP2-связанные расстройства	28935468	APG06622	G>A	26863	NC_000023,10,NC_000023,11	MECP2	393
MECP2-связанные расстройства	61749721	APG06622	G>A	26868	NC_000023,10,NC_000023,11	MECP2	394
MECP2-связанные расстройства	61750240	APG06622	G>A	26854	NC_000023,10,NC_000023,11	MECP2	395
MECP2-связанные расстройства	28935468	APG06248	G>A	26863	NC_000023,10,NC_000023,11	MECP2	396
MECP2-связанные расстройства	28934906	APG06007	C>T	26850	NC_000023,10,NC_000023,11	MECP2	397

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
MECP2-связанные расстройства	28934906	APG03850	C>T	26850	NC_000023,10,NC_000023,11	MECP2	398
MECP2-связанные расстройства	61750240	APG02787	C>T	26854	NC_000023,10,NC_000023,11	MECP2	399
MECP2-связанные расстройства	61750240	APG06007	C>T	26854	NC_000023,10,NC_000023,11	MECP2	400
MECP2-связанные расстройства	61750240	APG03850	C>T	26854	NC_000023,10,NC_000023,11	MECP2	401
MECP2-связанные расстройства	61750240	APG05586	C>T	26854	NC_000023,10,NC_000023,11	MECP2	402
Синдром Ангельмана	61751362	APG02787	C>T	26858	NC_000023,10,NC_000023,11	MECP2	403
Синдром Ангельмана	61751362	APG06007	C>T	26858	NC_000023,10,NC_000023,11	MECP2	404
Синдром Ангельмана	61751362	APG03850	C>T	26858	NC_000023,10,NC_000023,11	MECP2	405
Синдром Ангельмана	61751362	APG05586	C>T	26858	NC_000023,10,NC_000023,11	MECP2	406
Синдром Ангельмана	28934906	APG08167, APG01604	G>A	26850	NC_000023,10,NC_000023,11	MECP2	407
Семейная средиземноморская лихорадка	28940579	APG06622	A>G	17579	NC_000016,9,NC_000016,10	MEFV	408
Семейная средиземноморская лихорадка	61752717	APG06622	T>C	17577	NC_000016,9,NC_000016,10	MEFV	409
Семейная средиземноморская лихорадка	1,05E+08	APG06622	C>T	17588	NC_000016,9,NC_000016,10	MEFV	410
Семейная средиземноморская лихорадка	61752717	APG06248	T>C	17577	NC_000016,9,NC_000016,10	MEFV	411
Семейная средиземноморская лихорадка	1,05E+08	APG06248	C>T	17588	NC_000016,9,NC_000016,10	MEFV	412
Семейная средиземноморская лихорадка	1,05E+08	APG02874, APG03031, APG09208	G>A	17588	NC_000016,9,NC_000016,10	MEFV	413
Семейная средиземноморская лихорадка	28940579	APG08167, APG01604	A>G	17579	NC_000016,9,NC_000016,10	MEFV	414
Болезнь Шарко-Мари-Тута, тип 2	28940293	APG02874, APG03031, APG09208	T>C	17309	NC_000001,10,NC_000001,11	MFN2	415
Наследственный неполипозный рак толстой кишки	63751657	APG02874, APG03031, APG09208	G>A	95331	NC_000003,11,NC_000003,12	MLH1	416

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Наследственный неполипозный рак толстой кишки	63751711	APG02874, APG03031, APG09208	G>A	95792	NC_000003,11,NC_000003,12	MLH1	417
Метилмалоновая ацидемия	1,22E+08	APG06622	C>T	16462	NC_000001,10,NC_000001,11	MMACHC	418
Метилмалоновая ацидемия	1,22E+08	APG06248	C>T	16462	NC_000001,10,NC_000001,11	MMACHC	419
Метилмалоновая ацидемия	1,22E+08	APG02874, APG03031, APG09208	G>A	16464	NC_000001,10,NC_000001,11	MMACHC	420
Наследственный синдром, предрасполагающий к раку	63750636	APG06622	C>T	96378	NC_000002,11,NC_000002,12	MSH2	421
Наследственный синдром, предрасполагающий к раку	63749843	APG06622	C>T	94826	NC_000002,11,NC_000002,12	MSH6	422
Наследственный синдром, предрасполагающий к раку	7,86E+08	APG06622	C>T	181998	NC_000002,12,NC_000002,11	MSH6	423
Наследственный синдром, предрасполагающий к раку	63750741	APG02874, APG03031, APG09208	T>C	94663	NC_000002,11,NC_000002,12	MSH6	424
Наследственный синдром, предрасполагающий к раку	63749843	APG02787	C>T	94826	NC_000002,11,NC_000002,12	MSH6	425
Наследственный синдром, предрасполагающий к раку	63749843	APG06007	C>T	94826	NC_000002,11,NC_000002,12	MSH6	426
Наследственный синдром, предрасполагающий к раку	63749843	APG03850	C>T	94826	NC_000002,11,NC_000002,12	MSH6	427
Наследственный синдром, предрасполагающий к раку	63749843	APG05586	C>T	94826	NC_000002,11,NC_000002,12	MSH6	428
Наследственный синдром, предрасполагающий к раку	63751017	APG02787	C>T	94786	NC_000002,11,NC_000002,12	MSH6	429
Наследственный синдром, предрасполагающий к раку	63751017	APG06007	C>T	94786	NC_000002,11,NC_000002,12	MSH6	430

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Наследственный синдром, предрасполагающий к раку	63751017	APG03850	C>T	94786	NC_000002,11,NC_000002,12	MSH6	431
Наследственный синдром, предрасполагающий к раку	63751017	APG05586	C>T	94786	NC_000002,11,NC_000002,12	MSH6	432
Наследственный синдром, предрасполагающий к раку	63751017	APG08167, APG01604	C>T	94786	NC_000002,11,NC_000002,12	MSH6	433
MUTYH-ассоциированный полипоз	34612342	APG06622	T>C	20332	NC_000001,10,NC_000001,11	MUTYH	434
MUTYH-ассоциированный полипоз	36053993	APG06622	C>T	20333	NC_000001,10,NC_000001,11	MUTYH	435
MUTYH-ассоциированный полипоз	36053993	APG02787	G>A	20333	NC_000001,10,NC_000001,11	MUTYH	436
MUTYH-ассоциированный полипоз	36053993	APG06007	G>A	20333	NC_000001,10,NC_000001,11	MUTYH	437
MUTYH-ассоциированный полипоз	36053993	APG03850	G>A	20333	NC_000001,10,NC_000001,11	MUTYH	438
MUTYH-ассоциированный полипоз	36053993	APG05586	G>A	20333	NC_000001,10,NC_000001,11	MUTYH	439
MUTYH-ассоциированный полипоз	36053993	APG08167, APG01604	C>T	20333	NC_000001,10,NC_000001,11	MUTYH	440
MUTYH-ассоциированный полипоз	34612342	APG08167, APG01604	T>C	20332	NC_000001,10,NC_000001,11	MUTYH	441
Гипериммуноглобулинемия D с приступами лихорадки	28934897	APG06622	G>A	26968	NC_000012,11,NC_000012,12	MVK	442
MYBPC3-связанные расстройства	2E+08	APG06622	C>T	174776	NC_000011,9,NC_000011,10	MYBPC3	443
MYBPC3-связанные расстройства	3,88E+08	APG06622	G>A	45725	NC_000011,9,NC_000011,10	MYBPC3	444
MYBPC3-связанные расстройства	3,98E+08	APG06622	C>T	51962	NC_000011,9,NC_000011,10	MYBPC3	445
MYBPC3-связанные расстройства	2E+08	APG06248	C>T	174776	NC_000011,9,NC_000011,10	MYBPC3	446
MYBPC3-связанные расстройства	1,88E+08	APG02874, APG03031, APG09208	T>C	45267	NC_000011,9,NC_000011,10	MYBPC3	447
MYBPC3-связанные расстройства	2E+08	APG02874, APG03031, APG09208	G>A	174776	NC_000011,9,NC_000011,10	MYBPC3	448

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
МУВРС3-связанные расстройства	3,98E+08	APG02874, APG03031, APG09208	G>A	51820	NC_000011,9,NC_000011,10	МУВРС3	449
МУВРС3-связанные расстройства	3,98E+08	APG02874, APG03031, APG09208	G>A	51962	NC_000011,9,NC_000011,10	МУВРС3	450
Кардиомиопатия	3218716	APG06622	C>T	52071	NC_000014,8,NC_000014,9	МҮН7	451
Кардиомиопатия	36211715	APG06622	C>T	29159	NC_000014,8,NC_000014,9	МҮН7	452
Кардиомиопатия	1,22E+08	APG06622	G>A	29128	NC_000014,8,NC_000014,9	МҮН7	453
Кардиомиопатия	3,72E+08	APG06622	C>T	52045	NC_000014,8,NC_000014,9	МҮН7	454
Кардиомиопатия	3,98E+08	APG02874, APG03031, APG09208	T>C	52276	NC_000014,8,NC_000014,9	МҮН7	455
Кардиомиопатия	3,72E+08	APG06007	G>A	52045	NC_000014,8,NC_000014,9	МҮН7	456
Кардиомиопатия	3,72E+08	APG03850	G>A	52045	NC_000014,8,NC_000014,9	МҮН7	457
Аутосомно-рецессивная глухота	1,11E+08	APG02874, APG03031, APG09208	G>A	52388	NC_000011,9,NC_000011,10	МҮО7А	458
Врожденные генетические заболевания	80358259	APG06622	A>G	18006	NC_000018,9,NC_000018,10	NPС1	459
Врожденные генетические заболевания	80358259	APG02874, APG03031, APG09208	T>C	18006	NC_000018,9,NC_000018,10	NPС1	460
Врожденные генетические заболевания	1,2E+08	APG02874, APG03031, APG09208	G>A	18010	NC_000018,9,NC_000018,10	NPС1	461
Фенилкетонурия	5030851	APG06622	G>A	15628	NC_000012,11,NC_000012,12	РАН	462
Фенилкетонурия	5030858	APG06622	G>A	15616	NC_000012,11,NC_000012,12	РАН	463
Гиперфенилаланинемия	5030860	APG06622	T>C	15632	NC_000012,11,NC_000012,12	РАН	464

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Фенилкетонурия	62516101	APG02874, APG03031, APG09208	G>A	15658	NC_000012,11,NC_000012,12	PAH	465
Гиперфенилаланинемия, не фенилкетонурия	62644499	APG02874, APG03031, APG09208	G>A	15656	NC_000012,11,NC_000012,12	PAH	466
Фенилкетонурия	5030858	APG02787	C>T	15616	NC_000012,11,NC_000012,12	PAH	467
Фенилкетонурия	5030858	APG06007	C>T	15616	NC_000012,11,NC_000012,12	PAH	468
Фенилкетонурия	5030858	APG03850	C>T	15616	NC_000012,11,NC_000012,12	PAH	469
Фенилкетонурия	5030858	APG05586	C>T	15616	NC_000012,11,NC_000012,12	PAH	470
Гиперфенилаланинемия, не фенилкетонурия	5030860	APG08167, APG01604	T>C	15632	NC_000012,11,NC_000012,12	PAH	471
Гиперфенилаланинемия, не фенилкетонурия	62642937	APG08167, APG01604	G>A	15667	NC_000012,11,NC_000012,12	PAH	472
Семейный рак молочной железы	1,8E+08	APG06622	G>A	132139	NC_000016,10,NC_000016,9	PALB2	473
Семейный рак молочной железы	1,8E+08	APG02874, APG03031, APG09208	G>A	132185	NC_000016,10,NC_000016,9	PALB2	474
Нарушение биогенеза пероксисом 1B	61750420	APG08167, APG01604	C>T	22555	NC_000007,13,NC_000007,14	PEX1	475
Иммунодефицит 14	3,98E+08	APG06622	G>A	94255	NC_000001,10,NC_000001,11	PIK3CD	476
Поликистозная дисплазия почек	1,38E+08	APG06622	G>A	19147	NC_000006,11,NC_000006,12	PKHD1	477
Синдром гликопротеинов с дефицитом углеводов, I тип	28936415	APG02787	G>A	22745	NC_000016,9,NC_000016,10	PMM2	478
Синдром гликопротеинов с дефицитом углеводов, I тип	28936415	APG06007	G>A	22745	NC_000016,9,NC_000016,10	PMM2	479
Синдром гликопротеинов с дефицитом углеводов, I тип	28936415	APG03850	G>A	22745	NC_000016,9,NC_000016,10	PMM2	480

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Синдром гликопротеинов с дефицитом углеводов, I тип	28936415	APG05586	G>A	22745	NC_000016,9,NC_000016,10	PMM2	481
Синдром гликопротеинов с дефицитом углеводов, I тип	28936415	APG08167, APG01604	G>A	22745	NC_000016,9,NC_000016,10	PMM2	482
POLG-связанное состояние	1,14E+08	APG06622	C>T	28535	NC_000015,9,NC_000015,10	POLG	483
POLG-связанное состояние	1,14E+08	APG06622	C>T	28541	NC_000015,9,NC_000015,10	POLG	484
POLG-связанное состояние	1,14E+08	APG02874, APG03031, APG09208	G>A	28535	NC_000015,9,NC_000015,10	POLG	485
POLG-связанное состояние	1,14E+08	APG06007	G>A	28535	NC_000015,9,NC_000015,10	POLG	486
POLG-связанное состояние	1,14E+08	APG03850	G>A	28535	NC_000015,9,NC_000015,10	POLG	487
POLG-связанное состояние	1,14E+08	APG08167, APG01604	C>T	28541	NC_000015,9,NC_000015,10	POLG	488
Цероидный липофусциноз нейрональный 1	1,38E+08	APG06622	G>A	23943	NC_000001,10,NC_000001,11	PPT1	489
Цероидный липофусциноз нейрональный 1	1,38E+08	APG02787	C>T	23943	NC_000001,10,NC_000001,11	PPT1	490
Цероидный липофусциноз нейрональный 1	1,38E+08	APG06007	C>T	23943	NC_000001,10,NC_000001,11	PPT1	491
Цероидный липофусциноз нейрональный 1	1,38E+08	APG03850	C>T	23943	NC_000001,10,NC_000001,11	PPT1	492
Цероидный липофусциноз нейрональный 1	1,38E+08	APG05586	C>T	23943	NC_000001,10,NC_000001,11	PPT1	493
Цероидный липофусциноз нейрональный 1	1,38E+08	APG08167, APG01604	G>A	23943	NC_000001,10,NC_000001,11	PPT1	494
Кардиомиопатия	1,22E+08	APG06622	C>T	21885	NC_000007,13,NC_000007,14	PRKAG2	495
Синдром Каудена	1,22E+08	APG06622	C>T	22852	NC_000010,10,NC_000010,11	PTEN	496
PTPN11-связанное расстройство	1,22E+08	APG06622	C>T	28370	NC_000012,11,NC_000012,12	PTPN11	497

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Лимфома В-лимфобластного лейкоза, без подтипа ICD-O	1,22E+08	APG06622	A>G	28372	NC_000012,11,NC_000012,12	PTPN11	498
PTPN11-связанное расстройство	3,98E+08	APG06622	A>G	49032	NC_000012,11,NC_000012,12	PTPN11	499
PTPN11-связанное расстройство	1,22E+08	APG06248	C>T	28370	NC_000012,11,NC_000012,12	PTPN11	500
PTPN11-связанное расстройство	1,22E+08	APG06248	A>G	28379	NC_000012,11,NC_000012,12	PTPN11	501
PTPN11-связанное расстройство	28933386	APG02787	A>G	28365	NC_000012,11,NC_000012,12	PTPN11	502
PTPN11-связанное расстройство	28933386	APG06007	A>G	28365	NC_000012,11,NC_000012,12	PTPN11	503
PTPN11-связанное расстройство	28933386	APG03850	A>G	28365	NC_000012,11,NC_000012,12	PTPN11	504
PTPN11-связанное расстройство	28933386	APG05586	A>G	28365	NC_000012,11,NC_000012,12	PTPN11	505
PTPN11-связанное расстройство	1,22E+08	APG02787	A>G	28372	NC_000012,11,NC_000012,12	PTPN11	506
PTPN11-связанное расстройство	1,22E+08	APG06007	A>G	28372	NC_000012,11,NC_000012,12	PTPN11	507
PTPN11-связанное расстройство	1,22E+08	APG03850	A>G	28372	NC_000012,11,NC_000012,12	PTPN11	508
PTPN11-связанное расстройство	1,22E+08	APG05586	A>G	28372	NC_000012,11,NC_000012,12	PTPN11	509
PTPN11-связанное расстройство	28933386	APG08167, APG01604	A>G	28365	NC_000012,11,NC_000012,12	PTPN11	510
Болезнь накопления гликогена	1,17E+08	APG06622	G>A	17337	NC_000011,9,NC_000011,10	PYGM	511
Болезнь накопления гликогена	1,17E+08	APG06248	G>A	17337	NC_000011,9,NC_000011,10	PYGM	512
Рак молочной железы и яичников, семейный 4	3,88E+08	APG06622	G>A	39241	NC_000017,10,NC_000017,11	RAD51D	513
Дилатационная кардиомиопатия IDD	2,68E+08	APG02874, APG03031, APG09208	G>A	15310	NC_000010,10,NC_000010,11	RBM20	514
RET-связанные расстройства	74799832	APG06622	T>C	28958	NC_000010,10,NC_000010,11	RET	515
RET-связанные расстройства	74799832	APG02874, APG03031, APG09208	T>C	28958	NC_000010,10,NC_000010,11	RET	516
RYR1-связанные расстройства	1,18E+08	APG06622	C>T	28003	NC_000019,9,NC_000019,10	RYR1	517

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
RYR1-связанные расстройства	2,01E+08	APG06622	C>T	169564	NC_000019,9,NC_000019,10	RYR1	518
RYR1-связанные расстройства	1,18E+08	APG02874, APG03031, APG09208	G>A	76888	NC_000019,9,NC_000019,10	RYR1	519
RYR1-связанные расстройства	1,18E+08	APG02874, APG03031, APG09208	G>A	76888	NC_000019,9,NC_000019,10	RYR1	520
RYR1-связанные расстройства	1,18E+08	APG02874, APG03031, APG09208	G>A	76835	NC_000019,9,NC_000019,10	RYR1	521
RYR1-связанные расстройства	1,18E+08	APG02874, APG03031, APG09208	T>C	28014	NC_000019,9,NC_000019,10	RYR1	522
Синдром Швахмана	1,14E+08	APG06622	A>G	18235	NC_000007,13,NC_000007,14	SBDS	523
Синдром Бругада	1,38E+08	APG06622	C>T	24416	NC_000003,11,NC_000003,12	SCN5A	524
Синдром Бругада	28937316	APG02874, APG03031, APG09208	G>A	24408	NC_000003,11,NC_000003,12	SCN5A	525
Синдром Бругада	45546039	APG02874, APG03031, APG09208	G>A	48043	NC_000003,11,NC_000003,12	SCN5A	526
Синдром Бругада	72549410	APG02874, APG03031, APG09208	G>A	78547	NC_000003,11,NC_000003,12	SCN5A	527
Синдром Коудена 3	80338844	APG06622	C>T	21935	NC_000011,9,NC_000011,10	SDHD	528
Синдром Коудена 3	80338844	APG06248	C>T	21935	NC_000011,9,NC_000011,10	SDHD	529
Дефицит альфа-1-антитрипсина	28929474	APG06622	C>T	33006	NC_000014,8,NC_000014,9	SERPINA1	530
Дефицит альфа-1-антитрипсина	28929474	APG02787	G>A	33006	NC_000014,8,NC_000014,9	SERPINA1	531
Дефицит альфа-1-антитрипсина	28929474	APG03850	G>A	33006	NC_000014,8,NC_000014,9	SERPINA1	532

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Дефицит альфа-1-антитрипсина	28929474	APG05586	G>A	33006	NC_000014,8,NC_000014,9	SERPINA1	533
Дефицит альфа-1-антитрипсина	28929474	APG08167, APG01604	C>T	33006	NC_000014,8,NC_000014,9	SERPINA1	534
Конечностно-поясная мышечная дистрофия, тип 2D	28933693	APG06622	C>T	24476	NC_000017,10,NC_000017,11	SGCA	535
Мукополисахаридоз, MPS-III-A	1,05E+08	APG06622	C>T	20146	NC_000017,10,NC_000017,11	SGSH	536
Мукополисахаридоз, MPS-III-A	1,05E+08	APG02874, APG03031, APG09208	G>A	20146	NC_000017,10,NC_000017,11	SGSH	537
Синдром Нунана	2,68E+08	APG02787	A>G	21860	NC_000010,10,NC_000010,11	SHOC2	538
Синдром Нунана	2,68E+08	APG06007	A>G	21860	NC_000010,10,NC_000010,11	SHOC2	539
Синдром Нунана	2,68E+08	APG03850	A>G	21860	NC_000010,10,NC_000010,11	SHOC2	540
Синдром Нунана	2,68E+08	APG05586	A>G	21860	NC_000010,10,NC_000010,11	SHOC2	541
Синдром Нунана	2,68E+08	APG08167, APG01604	A>G	21860	NC_000010,10,NC_000010,11	SHOC2	542
SLC26A2-связанные расстройства	1,05E+08	APG06622	C>T	19128	NC_000005,9,NC_000005,10	SLC26A2	543
SLC26A2-связанные расстройства	1,11E+08	APG02874, APG03031, APG09208	G>A	52666	NC_000007,13,NC_000007,14	SLC26A4	544
Семейная гипертрофическая кардиомиопатия 2	7,28E+08	APG02874, APG03031, APG09208	G>A	172354	NC_000001,10,NC_000001,11	TNNT2	545
Болезнь Шарко-Мари-Тута, тип 2C	3,98E+08	APG02874, APG03031, APG09208	G>A	48018	NC_000012,11,NC_000012,12	TRPV4	546
Фокальная кортикальная дисплазия, тип II	28934872	APG06622	G>A	27436	NC_000016,9,NC_000016,10	TSC2	547
Амилоидогенный транстретиновый амилоидоз	76992529	APG06622	G>A	28465	NC_000018,9,NC_000018,10	TTR	548

Болезнь	№ RS	RGN	Вызывающие мутации	ID аллелей	Обозначения хромосом	Обозначения генов	Мишень (SEQ ID NO.)
Амилоидогенный транстретиновый амилоидоз	76992529	APG06248	G>A	28465	NC_000018,9,NC_000018,10	TTR	549
Кожно-глазной альбинизм	1,05E+08	APG06622	C>T	18816	NC_000011,9,NC_000011,10	TYR	550
Кожно-глазной альбинизм	1,22E+08	APG02874, APG03031, APG09208	G>A	18814	NC_000011,9,NC_000011,10	TYR	551
Болезнь Виллебранда	41276738	APG06622	C>T	15335	NC_000012,11,NC_000012,12	VWF	552

Пример 7. Нацеливание на мутации, ответственные за синдром Гурлера

Ниже описывают возможное лечение синдрома Гурлера, также называемого MPS-1, с использованием системы направляемого РНК редактирования оснований, которая исправляет мутацию, ответственную за синдром Гурлера, у 5 значительной части пациентов с этим заболеванием. В этом подходе используют слитый белок редактирования оснований, который направляется РНК и может быть упакован в один вектор AAV для доставки в широкий спектр типов тканей. В зависимости от точных регуляторных элементов и используемого домена редактора оснований также может быть возможно сконструировать один вектор, 10 который кодирует как слитый белок редактирования оснований, так и одну направляющую РНК для нацеливания на связанный с заболеванием локус.

Пример 7.1. Идентификация RGN с идеальным PAM

Генетическое заболевание МПС-1 представляет собой лизосомную болезнь накопления, характеризующуюся на молекулярном уровне накоплением 15 дерматансульфата и гепарансульфата в лизосомах. Это заболевание, как правило, является наследственным генетическим заболеванием, вызванным мутациями в гене *IDUA* (последовательность гена обозначена NCBI NG_008103.1), который кодирует α -L-идуронидазу. Заболевание является результатом дефицита α -L-идуронидазы. Наиболее распространенными 20 мутациями *IDUA*, обнаруженными в исследованиях людей североευропейского происхождения, являются W402X и Q70X, обе бессмысленные мутации, приводящие к преждевременной терминации трансляции (Bunge с соавт., *Hum. Mol. Genet.* 1994, 3(6): 861- 866, публикация включена в настоящее описание в качестве ссылки). Реверсия одного нуклеотида восстановит кодирующую 25 последовательность дикого типа и приведет к экспрессии белка, контролируемой эндогенными регуляторными механизмами генетического локуса.

Мутация W402X гена *Idua* человека является причиной высокой доли случаев MPS-1H. Редакторы оснований могут нацеливаться на узкое окно последовательности относительно сайта связывания протоспейсерного 30 компонента направляющей РНК, и, таким образом, присутствие последовательности PAM на определенном расстоянии от целевого локуса имеет важное значение для успеха стратегии. Учитывая ограничения, заключающиеся в том, что мутация-мишень должна находиться на открытой нецелевой цепи (NTS – non-target strand) во время взаимодействия белка, редактирующего

основания, и что отпечаток домена RGN может блокировать доступ к области рядом с PAM, считается, что доступный локус быть в 10-30 п.н. от PAM. Во избежание редактирования и мутагенеза других близлежащих аденозиновых оснований в этом окне проверяют различные линкеры. Идеальное окно составляет 12-16 п.н. от PAM.

Последовательность PAM, совместимая с APG02874, APG09208 и APG05586, хорошо видна в генетическом локусе. Эти нуклеазы имеют последовательность PAM 5'-nnnnCC-3' (SEQ ID NO: 35), 5'-nnnnC-3' (SEQ ID NO: 62) и 5'-nnRYA-3' (SEQ ID NO: 69), соответственно, и имеют компактный размер, что потенциально позволяет осуществлять доставку с помощью одного вектора AAV. Этот подход к доставке дает множество преимуществ по сравнению с другими, такими как доступ к широкому спектру тканей (печень, мышцы, ЦНС) и хорошо проявившиеся профиль безопасности и технологии производства.

Для Cas9 из *S. pyogenes* (SpyCas9) требуется PAM-последовательность NGG (SEQ ID NO: 256), которая присутствует рядом с локусом W402X, но размер SpyCas9 препятствует упаковке в единый вектор AAV и, таким образом, отказывается от вышеупомянутых преимуществ этого подхода. Хотя может использоваться стратегия двойной доставки (например, публикация Ryu с соавт., *Nat. Biotechnol.*, 2018, 36(6): 536-539, включенная в настоящее описание в виде ссылки), это значительно усложнит и удорожит получение. Кроме того, доставка двойного вирусного вектора значительно снижает эффективность коррекции генов, поскольку для успешного редактирования в данной клетке требуется заражение обоими векторами и сборка слитого белка в клетке.

Обычно используемый ортолог Cas9 из *S. aureus* (SauCas9) значительно меньше по размеру по сравнению со SpyCas9, но имеет более сложные требования к PAM — NGRRT (SEQ ID NO: 257). Эта последовательность находится за пределами диапазона, который, как ожидается, будет полезен для базового редактирования вызывающего локуса.

Пример 7.2. Слитые конструкции RGN и последовательностей sgRNA

Последовательность ДНК, кодирующую слитый белок со следующими доменами, получают с использованием стандартных методов молекулярной биологии: 1) домен RGN с мутациями, инактивирующими активность по расщеплению ДНК («мертвые» или «никазы»); 2) аденозиндезаминаза,

пригодная для редактирования оснований. Все конструкции, описанные в таблице ниже, содержат слитый белок с активным доменом редактирования оснований, в данном примере ADAT (SEQ ID NO: 211), функционально слитый с N-концом мертвой RGN APG02874 (SEQ ID NO: 214), APG09208 (SEQ ID NO: 216) и APG005586 (SEQ ID NO: 567). Также можно использовать другие аденозиндеаминазы, применимые для редактирования оснований ДНК (см., например, заявку PCT PCT/US2019/068079). В данной области техники известно, что слитый белок также может быть получен с ферментом, редактирующим основания на С-конце RGN. Кроме того, RGN и редактор оснований слитого белка обычно разделены линкерной аминокислотной последовательностью. В данной области техники известно, что длина стандартных линкеров находится в диапазоне от 15 до 30 аминокислот.

Кроме того, в данной области техники известно, что некоторые слитые белки между RGN и ферментом редактирования оснований могут также содержать по меньшей мере один домен ингибитора урацилгликозилазы (UGI) (SEQ ID NO: 212), который может повышать эффективность редактирования оснований (патент US 10167457, включенный в настоящее изобретение в виде ссылки). Таким образом, слитый белок может содержать RGN APG02874, APG09208 или их вариант, аденозиндезаминазу и необязательно по меньшей мере один UGI.

Таблица 12. Конструкции для РНК-нацеленного редактирования оснований

SEQ ID NO:	Конструкция	RGN	Мертвая нуклеаза (D) или никаза (N)	Редактор оснований
213	Nuc-ADAT-линкер-d APG02874 -линкер-SV40	APG02874	D	ADAT
215	Nuc-ADAT-линкер-d APG09208-линкер-SV40	APG09208	D	ADAT
565	Nuc-ADAT-линкер-d APG05586-линкер-SV40	APG05586	D	ADAT

Доступные сайты редактирования RGN определяются последовательностью PAM. При объединении RGN с доменом редактирования оснований целевой остаток для редактирования должен находиться на нецелевой цепи (NTS – non-target strand), поскольку NTS является одноцепочечной, а RGN связана с локусом. Оценка ряда нуклеаз и соответствующих направляющих РНК позволяет выбрать наиболее подходящий инструмент редактирования генов для данного конкретного локуса. Несколько потенциальных последовательностей PAM, на

которые могут быть нацелены описанные выше конструкции в гене *Idua* человека, находятся вблизи мутантного нуклеотида, ответственного за мутацию W402X. Также продуцируется последовательность, кодирующая транскрипт направляющей РНК, которая содержит 1) «спейсер», комплементарный некодирующей цепи ДНК в локусе, связанным с болезнью; и 2) последовательность РНК, необходимая для ассоциации направляющей РНК с RGN. Такая sgRNA может быть закодирована, например, SEQ ID NO: 217 для системы RGN APG02874, SEQ ID NO: 218 для системы RGN APG09208 или SEQ ID NO: 566 для системы RGN APG005586. Эти молекулы sgRNA и подобные sgRNA, которые могут быть разработаны специалистом в данной области, могут быть оценены на предмет их эффективности в направлении указанных выше редакторов оснований к интересующему локусу.

Пример 7.3. Анализ активности в клетках пациентов с болезнью Гурлера

Для проверки стратегии генотипирования и оценки описанных выше конструкций используют фибробласты пациентов с болезнью Гурлера. Конструируют вектор, содержащий соответствующие промоторы по цепи выше последовательности, кодирующей белок слияния, и последовательность, кодирующую sgRNA, для их экспрессии в клетках человека, подобно векторам, описанным в примере 4. Признано, что могут применяться промоторы и другие элементы ДНК (например, энхансеры, или терминаторы), которые либо известны высокими уровнями экспрессии в клетках человека, либо могут специфически хорошо экспрессироваться в клетках фибробластов. Вектор трансфицируют в фибробласты с использованием стандартных методов, например метод трансфекции, аналогичной описанному в примере 4. В качестве альтернативы можно использовать электропорацию. Клетки культивируют в течение 1-3 дней. Геномную ДНК (gDNA) выделяют с использованием стандартных методик. Эффективность редактирования определяется путем проведения анализа генотипирования qPCR и/или секвенирования следующего поколения на очищенной gDNA, как описано ниже.

В анализе Taqman™ qPCR используют зонды, специфичные для аллеля дикого типа и мутантного аллеля. Эти зонды несут флуорофоры, которые разрешаются по их спектральным свойствам возбуждения и/или испускания с использованием прибора для qPCR. Набор для генотипирования, содержащий праймеры и зонды для ПЦР, которые имеются в продаже (например, Thermo

Fisher Taqman™ SNP genotyping assay ID C__27862753_10 для SNP ID rs121965019), или разработать. Пример разработанного набора праймеров и зонд показан в табл. 13.

Таблица 13. Праймеры и зонды для RT-PCR

Описание	Последовательность	SEQ ID NO:
Forward-праймер для амплификации	5'-GACTCCTTCACCAAG-3'	219
Reverse-праймер для амплификации	5'-GTAGATCAGCACCG-3'	220
Зонд дикого типа	5'-CTCTGGGCCGAAGT-3'	221
Зонд W402X	5'-CTCTAGGCCGAAGT-3'	222

5

После эксперимента по редактированию gDNA подвергают анализу методом количественной ПЦР (qPCR) с использованием стандартных методов, праймеров и зондов, описанных выше. Ожидаемые результаты показаны в табл.

14. Эту систему *in vitro* можно использовать для целесообразной оценки конструкций и выбора конструкции с высокой эффективностью редактирования для дальнейших исследований. Системы будут оцениваться в сравнении с клетками с мутацией W402X и без нее и предпочтительно с некоторыми клетками, гетерозиготными по этой мутации. Значения Ct можно сравнивать либо с эталонным геном, либо с полной амплификации локуса с использованием красителя, такого как Sybr green.

15

Таблица 14. Ожидаемые результаты ПЦР

Генотип	Трансфекция с редактором оснований	Ожидаемый результат ПЦР
Idua ^{WT/WT}	Нет	Гомозиготный дикий тип
Idua ^{WT/W402X}	Нет	Гетерозиготный: 50% дикого типа, 50% W402X
Idua ^{W402X/W402X}	Нет	Гомозиготный W402X
Idua ^{W402X/W402X}	Да	Вариабельный

Ткани также можно анализировать с помощью метода секвенирования следующего поколения. Можно использовать сайты связывания праймеров, такие как показанные ниже (табл. 15), или другие подходящие сайты связывания праймеров, которые могут быть идентифицированы специалистом в данной области. После ПЦР-амплификации продукты, содержащие выступающие последовательности Illumina Nextera XT, подвергаются подготовке библиотеки в соответствии с протоколом Illumina 16S Metagenomic Sequencing Library.

20

25 Глубокое секвенирование выполняется на платформе Illumina Mi-Seq. Как

правило, на ампликон генерируются 200000 считываний парных концов по 250 п.н. (2 x 100000 считываний). Прочтения анализируют с использованием CRISPResso (Pinello с соавт., 2016) для расчета степени редактирования.

5 Выравнивание выходных данных проводится вручную для подтверждения сайтов инсерций и делеций, а также для выявления сайтов микрогомологии в сайтах рекомбинации.

Таблица 15. Сайты связывания праймеров методом секвенирование следующего поколения (NGS – Next Generation Sequencing)

Направление	Последовательность	SEQ ID NO:
Forward	5'-ACTTCCTCCAGCC-3'	223
Reverse	5'-GAACCCCGGCTTA-3'	224

10 Вестерн-блоттинг клеточного лизата трансфицированных клеток и контрольных клеток с использованием анти-IDUA антитела проводят для проверки экспрессии полноразмерного белка, а анализ активности фермента на клеточном лизате с использованием субстрата 4-метилумбеллиферил-α-L-идуронида подтверждает, что фермент каталитически активен (публикация

15 Норвуд с соавт., *Clin.Chim. Acta*, 1979, 92(2): 257-265, включена в настоящее описание в качестве ссылки). Эти эксперименты проводят в сравнении с исходной клеточной линией IduaW402X/W402X (без трансфекции), клеточной линией IduaW402X/W402X, трансфицированной конструкцией редактирования основания и случайной направляющей последовательностью, и клеточной

20 линией, экспрессирующей IDUA дикого типа.

Пример 7.4. Подтверждение возможности лечения заболевания на мышинной модели

Для проверки эффективности данного терапевтического подхода используют модель мыши с нонсенс-мутацией в аналогичной аминокислоте.

25 Мышиная линия несет мутацию W392X в своем гене *Idua* (идентификатор гена: 15932), которая соответствует гомологичной мутации при синдроме Гурлера у пациентов (публикация Bunge с соавт., *Hum. Mol. Genet.* 1994, 3(6): 861-866, включена в настоящее описание в качестве ссылки). Этот локус содержит последовательность, отличную от последовательности нуклеотидов у человека, в

30 которой отсутствует последовательность РАМ, необходимая для коррекции с помощью редакторов оснований, описанных в предыдущих примерах, и, таким образом, требуется разработка отдельного слитого белка для выполнения

коррекции нуклеотидов. Облегчение заболевания у такого животного может подтвердить терапевтический подход к коррекции мутации в тканях, доступных вектору доставки генов.

Мыши, гомозиготные по этой мутации, обнаруживают ряд фенотипических характеристик, сходных с пациентами с синдромом Гурлера. Слитый белок редактирования оснований-RGN, описанный выше (табл. 12), вместе с направляющей последовательностью РНК встраивают в вектор экспрессии, который обеспечивает экспрессию белка и транскрипцию РНК у мышей. Дизайн исследования показан ниже в табл. 16. Исследование включает группы, которым вводят высокую дозу вектора экспрессии, включающего слитый белок редактирования оснований и направляющую РНК последовательность, низкую дозу того же вектора экспрессии, контрольным образцом которого является мышь, обработанная вектором экспрессии, который не содержит слитый белок для редактирования оснований или направляющую РНК, и второй контроль, который представляет собой мышь дикого типа, обработанную тем же пустым вектором.

Таблица 16. Эксперимент по редактированию оснований на примере мышинной модели

Группа	Штаммы мышей	N	Обработка
1	Idua-W392X ¹	≥ 5	Низкая доза вектора
2	Idua-W392X	≥ 5	Высокая доза вектора
3	Idua-W392X	≥ 5	Растворитель
4	129/Sv (WT)	5	Растворитель

Конечные точки оценки включают массу тела, экскрецию GAG с мочой, ферментативную активность IDUA в сыворотке, активность IDUA в интересующих тканях, патологию тканей, генотипирование интересующих тканей для подтверждения коррекции SNP, а также поведенческую и неврологическую оценку. Поскольку некоторые конечные точки являются конечными, перед окончанием исследования могут быть добавлены дополнительные группы для оценки, например, тканевой патологии и тканевой активности IDUA. Дополнительные примеры конечных точек можно найти в опубликованных работах, устанавливающих модели животных с синдромом Гурлера (Shull с соавт., *Proc. Natl. Acad. Sci. U.S.A.*, 1994, (26): 12937-12941;

Wang с соавт., *Mol. Genet. Metab.*, 2010, 99(1): 62-71; Hartung с соавт., с соавт., *Mol. Ther.*, 2004, 9(6): 866-875; Liu с соавт., *Mol. Ther.*, 2005, 11(1): 35-47; Clarke с соавт., *Hum. Mol. Genet.* 1997, 6(4): 503-511; все указанные публикации включены в настоящее изобретение в виде ссылки).

5 Один возможный вектор доставки использует аденоассоциированный вирус (AAV). Получают вектор, включающий последовательность, кодирующую слитый белок редактора оснований и dRGN (например, Nuc-ADAT-линкер-dAPG19748-линкер-SV40, как описано выше), которому предшествуют энхансер CMV (SEQ ID NO: 259) и промотор (SEQ ID NO: 258), или другую подходящую комбинацию энхансера и промотора), необязательно последовательность Козака и функционально слитую на 3'-конце с последовательностью терминатора и последовательностью полиаденилирования, такой как минимальная последовательность, описанная в публикации Levitt N., Briggs D., Gil A., Proudfoot N.J. *Genes Dev.* 1989, 3 (7), 1019–1025. Вектор может дополнительно
10 содержать кассету экспрессии, кодирующую одну направляющую РНК, функционально связанную на своем 5'-конце с промотором U6 человека (SEQ ID NO: 260) или другим промотором, подходящим для выработки малых некодирующих РНК, и дополнительно содержащую инвертированную последовательности концевых повторов (ITR – inverted terminal repeat), которые
15 необходимы и хорошо известны в данной области техники для упаковки в капсид AAV. Выработка и вирусная упаковка осуществляются стандартными способами, такими как описанные в патенте US 9587250, включенном в настоящее изобретение в качестве ссылки.
20

 Другие возможные вирусные векторы включают аденовирусные и
25 лентивирусные векторы, которые обычно используются и могут содержать сходные элементы с различными возможностями упаковки и потребностями. Также могут использоваться невирусные методы доставки, такие как мРНК и sgRNA, инкапсулированные липидными наночастицами (Cullis P. R. и Allen T. M., *Adv. Drug Deliv. Rev.*, 2013, 65(1): 36-48; Finn с соавт., *Cell Rep*, 2018, 22(9): 2227-2235; обе публикации включены в настоящее изобретение в виде ссылок),
30 гидродинамическая инъекция плазмидной ДНК (публикация Suda T. и Liu D., *Mol. Ther.*, 2007, 15(12): 2063-2069, включена в настоящее изобретение в виде ссылки), или рибонуклеопротеиновые комплексы sgRNA и связанные с наночастицами золота (Lee K. с соавт., *Nat. Biomed. Eng.* 2017, 1 (11), 889–90).

Пример 7.5. Коррекция заболевания на мышинной модели с гуманизированным локусом

Чтобы оценить эффективность конструкции, идентичной редактору оснований, которая будет использоваться для терапии человека, необходима модель мыши, в которой нуклеотиды рядом с W392 изменены, чтобы соответствовать последовательности у людей вокруг W402. Это может быть достигнуто с помощью различных методов, включая использование RGN и HDR матрицы для вырезания и замены локуса в эмбрионах мыши.

Из-за высокой степени консервативности аминокислот большинство нуклеотидов в локусе мыши могут быть изменены на нуклеотиды последовательности человека с молчащими мутациями, как показано в табл. 17. Единственные изменения оснований, приводящие к изменению кодирующей последовательности в полученном сконструированном геноме мыши, происходят после введенного стоп-кодона.

Таблица 17. Нуклеотидные мутации для создания гуманизованного локуса мыши

Признак	Человек (W402X)		Мышь (W392X)		Гуманизованная мышь	
	Нуклеотид (SEQ ID NO: 225)	Код аминокислот	Нуклеотид (SEQ ID NO: 226)	Код аминокислот	Нуклеотид (SEQ ID NO: 227)	Код аминокислот
Протоспейсер	G	E	A	G	G	G
	G	E	G	E	G	E
	A		A		A	
	G		A		G	
	C	Q	C	Q	C	Q
	A		A		A	
	G		A		G	
	C	L	C	L	C	L
	T		T		T	
	C		C		C	
	T	STOP	T	STOP	T	STOP
	A		A		A	
	G		G		G	
	G	A	G	A	G	A
	C		C		C	
	C		A		C	
	G	E	G	E	G	E
	A		A		A	
	A		G		A	
	G	V	G	V	G	V
T	T		T			
G	C		G			
T	S	T	S	T	S	
C		C		C		
G		A		G		
РАМ,	C	Q	A	K	C	Q

	Человек (W402X)		Мышь (W392X)		Гуманизированная мышь	
некритич- ный	A		A		A	
	G		G		G	
	G		G		G	
PAM, критичный	C	A	C	A	C	A
	C		T		C	

После создания этой линии мышей проводят эксперименты, аналогичные описанным в примере 7.4.

Пример 8. Нацеливание на мутации, ответственные за атаксию Фридрейха

5 Распространение последовательности тринуклеотидных повторов, вызывающая атаксию Фридрейха (FRDA), происходит в определенном генетическом локусе в гене FXN, называемом областью нестабильности FRDA. Для вырезания области нестабильности в клетках пациентов с FRDA можно использовать РНК-направляемые нуклеазы (RGN). Этот подход требует 1) RGN
10 и последовательности направляющей РНК, которые можно запрограммировать для нацеливания на аллель в геноме человека; и 2) подход к доставке для RGN и направляющей последовательности. Многие нуклеазы, используемые для редактирования генома, такие как обычно используемая нуклеаза Cas9 из *S. pyogenes* (SpCas9), слишком велики, чтобы их можно было упаковать в векторы
15 аденоассоциированных вирусов (AAV), особенно с учетом длины гена SpCas9 и направляющей РНК в дополнение к другим генетическим элементам, необходимым для функциональной экспрессии кассет. Такой подход с использованием SpCas9 маловероятен.

20 Компактные РНК-направляемые нуклеазы по настоящему изобретению, например, APG03850, безусловно хорошо подходят для вырезания области нестабильности FRDA. Для APG03850 требуется PAM, который находится в непосредственной близости от области нестабильности FRDA. Кроме того, APG03850 может быть упакован в вектор AAV вместе с направляющей РНК. Для упаковки двух направляющих РНК, вероятно, потребуются второй вектор, но
25 этот подход все же выгодно отличается от того, что потребовалось бы для более крупной нуклеазы, такой как SpCas9, которая потребовала бы расщепления последовательности белка между двумя векторами.

30 В табл. 18 показано расположение геномных последовательностей-мишеней, подходящих для нацеливания APG03850 на 5'- и 3'-фланги области нестабильности FRDA, а также последовательности sgRNA для геномных

мишеней. Оказавшись в локусе, RGN может вырезать область нестабильности FA. Иссечение области можно проверить с помощью секвенирования Illumina локуса.

Таблица 18, Геномные последовательности-мишени для систем RGN

5 Genomic

№ направляющей	Расположение относительно области нестабильности FRDA	Геномная последовательность-мишень (SEQ ID NO.)	sgRNA (SEQ ID NO.)
1	5'	228	229
2	5'	230	231
3	3'	232	233
4	3'	234	235

Пример 9. Нацеливание на мутации, ответственные за серповидно-клеточную анемию

10 Направленные последовательности в области энхансера BCL11A (SEQ ID NO: 236) могут обеспечивать механизм повышения фетального гемоглобулина (HbF) для излечения или облегчения симптомов серповидно-клеточной анемии. Например, полногеномные ассоциативные исследования идентифицировали набор генетических вариаций в BCL11A, которые связаны с повышенными уровнями HbF. Эти вариации представляют собой набор SNP, обнаруженных в

15 некодирующих областях BCL11A, которые функционируют как стадийно-специфические, клонально-ограниченные энхансерные области. Дальнейшее исследование показало, что этот энхансер BCL11A необходим в эритроидных клетках для экспрессии BCL11A (Bauer с соавт., *Science*, 2013, 343:253-257; публикация включена в настоящее описание в качестве ссылки). Энхансерная

20 область была обнаружена в интроне 2 гена BCL11A, и в интроне 2 были идентифицированы три области гиперчувствительности к ДНКазе I (часто свидетельствующие о состоянии хроматина, связанного с регуляторным потенциалом). Эти три области были идентифицированы как «+62», «+58» и «+55» в соответствии с расстоянием в килобазах от сайта начала транскрипции

25 BCL11A. Этих энхансерных областей примерно 350 (+55); 550 (+58); и 350 (+62) нуклеотидов в длину (Bauer с соавт., 2013).

Пример 9.1. Выявление предпочтительных систем RGN

В настоящем примере описывают потенциальное лечение бета-гемоглобинопатий с использованием системы RGN, которая нарушает

связывание BCL11A с его сайтом связывания в локусе HBB, который является геном, ответственным за образование бета-глобина в гемоглобине взрослых. Этот подход использует метод NHEJ, который более эффективен в клетках млекопитающих. Кроме того, в этом подходе используют нуклеазу достаточно

5 малого размера, которую можно упаковать в единый вектор AAV для доставки *in vivo*.

Мотив энхансера GATA1 в энхансерной области BCL11A человека (SEQ ID NO: 236) является идеальной мишенью для разрушения с использованием РНК-

10 направляемых нуклеаз (RGN) для снижения экспрессии BCL11A с

одновременной повторной экспрессией HbF в эритроцитах взрослого человека (Wu с соавт., *Nat Med*, 2019, 387:2554). Несколько последовательностей PAM, совместимых с APG03850 или APG09208, легко обнаруживают в генетическом локусе, окружающем данный сайт GATA1. Эти нуклеазы имеют

15 последовательность PAM 5'-nnnnG-3' (SEQ ID NO: 42) и 5'-nnnnC-3' (SEQ ID NO: 62), соответственно, и имеют компактный размер, что потенциально позволяет осуществлять их доставку наряду с соответствующей направляющей РНК в одном AAV или аденовирусном векторе. Этот подход к доставке дает множество преимуществ по сравнению с другими, такими как доступ к гемопоэтическим стволовым клеткам, а также хорошо зарекомендовавший себя профиль

20 безопасности и методы производства.

Обычно используемая нуклеаза Cas9 из *S. pyogenes* (SpyCas9) требует последовательности PAM 5'-NGG-3' (SEQ ID NO: 256), некоторые из которых присутствуют рядом с мотивом GATA1. Однако размер SpyCas9 не позволяет упаковывать его в единый вектор AAV или аденовируса и, таким образом,

25 отказывается от вышеупомянутых преимуществ этого подхода. Хотя можно использовать стратегию двойной доставки, это значительно усложнит и удорожит получение. Кроме того, доставка двойного вирусного вектора значительно снижает эффективность коррекции генов, поскольку для успешного редактирования в данной клетке требуется заражение обоими векторами.

30 Получают экспрессионную кассету, кодирующую APG03850 или APG09208, кодон-оптимизированные для человека, аналогичную описанным в примере 4. Также получают кассеты экспрессии, которые экспрессируют направляющие РНК для RGN APG03850 или APG09208. Эти направляющие РНК включают: 1) последовательность протоспейсера, комплементарную либо

некодирующей, либо кодирующей цепи ДНК в энхансерном локусе BCL11A (последовательность-мишень) и 2) последовательность РНК, необходимую для ассоциации направляющей РНК с RGN. Поскольку несколько потенциальных последовательностей РАМ для нацеливания с помощью APG03850 или APG09208 окружают мотив энхансера BCL11A GATA1, получают несколько потенциальных направляющих РНК-конструкций для определения наилучшей последовательности протоспейсера, которая вызывает надежное расщепление и опосредованное NHEJ нарушение последовательности энхансера BCL11A GATA1. Геномные последовательности-мишени, представленные в табл. 19, оценивают для направления RGN в этот локус с использованием sgRNA.

Таблица 19. Последовательности-мишени для BCL11A GATA1 локуса энхансера с использованием APG03850 или APG09208

Направление	RGN	Геномные последовательности-мишени (SEQ ID NO:)	sgRNA (SEQ ID NO:)
1	APG03850	237	238
2	APG03850	239	240
3	APG03850	241	242
4	APG09208	243	244
5	APG09208	245	246
6	APG09208	247	248

Для оценки эффективности, с которой APG03850 или APG09208 генерирует инсерции или делеции, которые разрушают область энхансера BCL11A, используют клеточные линии человека, такие как клетки эмбриональной почки человека (клетки НЕК – human embryonic kidney). Получают ДНК-вектор, содержащий кассету экспрессии RGN (например, как описано в примере 4). Также получают отдельный вектор, содержащий кассету экспрессии, содержащую кодирующую последовательность для последовательности направляющей РНК из табл. 16. Такая кассета экспрессии может дополнительно содержать промотор РНК-полимеразы III U6 человека (SEQ ID NO: 260), как описано в примере 4. В качестве альтернативы можно использовать один вектор, содержащий кассеты экспрессии как RGN, так и направляющей РНК. Вектор вводят в клетки НЕК с использованием стандартных методов, таких как описанные в примере 4, и клетки культивируют в течение 1-3 дней. После этого периода культивирования выделяют геномную ДНК и определяют частоту

инсерций и делеций с использованием расщепления эндонуклеазой I T7 и/или прямого секвенирования ДНК, как описано в примере 4.

Область ДНК, охватывающую целевую область BCL11A, амплифицируют с помощью ПЦР с праймерами, содержащими выступающие последовательности Illumina Nextera XT. Эти ПЦР-ампликоны либо исследуют на образование NHEJ с помощью расщепления эндонуклеазой I T7, либо подвергают подготовке библиотеки в соответствии с протоколом библиотеки метагеномного секвенирования Illumina 16S или аналогичной подготовке библиотеки для секвенирования следующего поколения (NGS). После глубокого секвенирования сгенерированные прочтения анализируют с помощью CRISPResso для расчета степени редактирования. Выходные выравнивания проверяют вручную для подтверждения сайтов инсерции и делеции. Этот анализ выявляет предпочтительную RGN и соответствующую предпочтительную направляющую РНК (sgRNA). Анализ может привести к тому, что и APG03850, и APG09208 будут одинаково предпочтительными. Кроме того, анализ может определить, что существует более одной предпочтительной направляющей РНК или что все геномные последовательности-мишени в табл. 16 в равной степени предпочтительны.

Пример 9.2. Анализ экспрессии фетального гемоглобина

В этом примере, сгенерированные с помощью APG03850 или APG09208 инсерции или делеции, нарушающие энхансерную область BCL11A, анализируют на экспрессию фетального гемоглобина. Используют CD34+ гемопоэтические стволовые клетки (HSC – hematopoietic stem cells) здоровых доноров человека. Эти HSC культивируют и в них вводят вектор (векторы), содержащий кассеты экспрессии, включающие кодирующие области предпочтительной RGN и предпочтительной sgRNA, с использованием способов, аналогичных описанным в примере 8.1. После электропорации эти клетки дифференцируют *in vitro* в эритроциты с использованием установленных протоколов (например, из публикации Giarratana с соавт., *Nat Biotechnology*, 2004, 23:69-74, включенных в настоящее описание в качестве ссылки). Экспрессию HbF затем измеряют с помощью вестерн-блоттинга с антителом против HbF человека или количественно определяют с помощью высокоэффективной жидкостной хроматографии (ВЭЖХ). Ожидают, что успешное разрушение локуса энхансера BCL11A приведет к увеличению

продукции HbF по сравнению с HSC, подвергнутыми электропорации только с RGN, но без направляющей.

Пример 9.3. Анализ пониженного образования серповидно клеточных эритроцитов

5 В этом примере, сгенерированные с помощью APG03850 или APG09208 инсерции или делеции, нарушающие энхансерную область BCL11A, анализируют на снижение образования серповидно клеточных эритроцитов. Используют донорские CD34+ гемопоэтические стволовые клетки (HSC) от пациентов с серповидно-клеточной анемией. Эти HSC культивируют и в них
10 вводят вектор (векторы), содержащий кассеты экспрессии, включающие кодирующие области предпочтительной RGN и предпочтительной sgRNA, с использованием способов, аналогичных описанным в примере 8.1. После электропорации эти клетки дифференцируют *in vitro* в эритроциты с использованием установленных протоколов (например, из публикации Giarratana с соавт., *Nat Biotechnology*, 2004, 23:69-74). Экспрессию HbF затем измеряют с помощью вестерн-блоттинга с антителом против HbF человека или количественно определяют с помощью высокоэффективной жидкостной хроматографии (ВЭЖХ). Ожидают, что успешное разрушение локуса энхансера BCL11A приведет к увеличению продукции HbF по сравнению с HSC,
15 подвергнутыми электропорации только с RGN, но без направляющей.

Образование серповидных клеток индуцируется в этих дифференцированных эритроцитах добавлением метабисульфита. Количество серповидных и нормальных эритроцитов подсчитывают с помощью микроскопа. Ожидают, что количество серповидных клеток меньше в образцах клетках,
25 обработанных APG03850 или APG09208 плюс sgRNA, чем в образцах необработанных клеток или обработанных только RGN.

Пример 9.4. Подтверждение лечения заболевания на мышинной модели

Для оценки эффективности использования APG03850 или APG09208 для разрушения локуса BCL11A используют подходящие гуманизированные
30 мышинные модели серповидно клеточной анемии. Кассеты экспрессии, кодирующие предпочтительную RGN и предпочтительную sgRNA, упаковывают в векторы AAV или аденовирусные векторы. В частности, аденовирус типа Ad5/35 эффективен для HSC. Выбирают подходящую мышинную модель, содержащую гуманизированный локус HBB с аллелями серповидных клеток,

например B6; FVB-Tg(LCR-HBA2, LCR-HBB*E26K)53Hhb/J или B6,Cg-Hbatm1Paz Hbbtm1Tow Tg(HBA-HBBs). 41Paz/HhbJ. Этим мышам вводят гранулоцитарный колониестимулирующий фактор отдельно или в комбинации с плериксафором для мобилизации HSC в кровяном русле. Затем внутривенно вводят AAV или аденовирусы, несущие плазмиду с RGN и направляющей, и мышам дают выздороветь в течение недели. Кровь, полученную от этих мышей, тестируют в анализе серповидных клеток *in vitro* с использованием метабисульфита, и за мышами наблюдают длительное время для контроля показателей смертности и гемопоэтической функции. Ожидают, что лечение AAV или аденовирусами, несущими RGN и направляющую РНК, уменьшит серповидность, смертность и улучшит гемопоэтическую функцию по сравнению с мышами, обработанными вирусами, лишенными обеих кассет экспрессии, или вирусами, несущими только одну кассету экспрессии RGN.

Пример 10. Тестирование разных форматов доставки

Чтобы определить, способны ли RGN доставляться в различных форматах, доставку нуклеофекции мРНК и RNP тестируют с первичными Т-клетками. Очищенные CD3⁺ Т-клетки или мононуклеарные клетки периферической крови (PBMC – peripheral blood mononuclear cell) размораживают, активируют с помощью гранул CD3/CD28 (фирма ThermoFisher) в течение 3 дней, затем подвергают нуклеофекции с использованием устройства Lonza 4D-Nucleofector X и полосок Nucleocuvette. Набор P3 Primary Cell используют для доставки как мРНК, так и RNP. Клетки трансфицируют с использованием программ EO-115 и EH-115 для доставки мРНК и RNP, соответственно. Клетки культивируют в среде для размножения Т-клеток CTS OpTimizer (фирма ThermoFisher), содержащей IL-2, IL-7 и IL-15 (фирма Miltenyi Biotec), в течение 4 дней после нуклеофекции перед сбором материала с использованием набора для выделения геномной ДНК Nucleospin Tissue (фирма Machery Nagel).

Ампликоны, окружающие сайты редактирования, формируют с помощью ПЦР с использованием праймеров, указанных в табл. 4, и подвергают секвенированию методом NGS с использованием платформы Illumina Nextera с секвенированием парных концов 2x250 п.н. в соответствии с методом, описанным в примере 4.

Таблица 20. Доставка RGN в первичные T-клетки с помощью mRNA и RNP

RGN	Способ доставки	SGN	Редактирование, %
APG01604	мРНК, нуклеофекция	SGN001671	1,18
APG01604	мРНК, нуклеофекция	SGN001673	22,7
APG01604	RNP, нуклеофекция	SGN001673	0,56
APG01604	мРНК, нуклеофекция	SGN001674	0,63
APG01604	RNP, нуклеофекция	SGN001674	0,77
APG01868	мРНК, нуклеофекция	SGN001684	77,85
APG01868	RNP, нуклеофекция	SGN001684	83,73
APG01868	мРНК, нуклеофекция	SGN001691	82,53
APG01868	RNP, нуклеофекция	SGN001691	80,7
APG01868	мРНК, нуклеофекция	SGN001692	76,84
APG01868	RNP, нуклеофекция	SGN001692	86,06
APG01868	мРНК, нуклеофекция	SGN001785	0,22
APG01868	RNP, нуклеофекция	SGN001785	0,43

5 Доставка как мРНК, так и RNP показала успешное редактирование с помощью APG01868, APG01868 и с помощью RNP в нескольких геномных мишенях. Доставка RNP была ограничена при использовании APG01604, но показала одинаковую степень редактирования при доставке мРНК.

Пример 11. Тестирование редактора оснований

10 Чтобы определить, могут ли APG09298 и APG01604 выполнять редактирование оснований цитозина в клетках млекопитающих, цитозиндезаминазу функционально сливают с никазой версии каждой RGN для получения слитого белка. Идентифицируют остатки, предположительно дезактивирующие домен RuvC в RGN APG09298 и APG01604, и RGN модифицируют в варианты никазы. Никазный вариант RGN упоминают в настоящем изобретении как «nRGN». Следует учитывать, что любой вариант
15 никазы RGN может быть использован для получения слитого белка по настоящему изобретению.

20 Нуклеотидные последовательности дезаминазы и nRGN, оптимизированные для экспрессии у млекопитающих, синтезируют в виде слитых белков с N-концевой меткой ядерной локализации и клонируют в экспрессирующей плазмиде pTwist CMV (фирма Twist Biosciences). Каждый слитый белок содержит, начиная с amino-конца, NLS SV40 (SEQ ID NO: 251), функционально связанную на C-конце с меткой 3X FLAG (SEQ ID NO: 252), функционально

связанную на С-конце с дезаминазой (APG05840), функционально связанной на С-конце с пептидным линкером, функционально связанным на С-конце с nRGN (nAPG09298 или nAPG01604), функционально связанной на С-конце с пептидным линкером, функционально связанным на С-конце с белком, стабилизирующим урацил (USP2, указан как SEQ ID NO: 1089), наконец, функционально связанный на С-конце с нуклеоплазмином NLS (SEQ ID NO: 253). Аминокислотная последовательность слитых белков APG05840-nAPG09298-USP2 и APG05840-nAPG01604-USP2 представлена как SEQ ID NO: 1090 и 1091, соответственно.

Также были получены плазмиды экспрессии, содержащие кассету экспрессии, кодирующую sgRNA. Геномные последовательности-мишени человека и последовательности sgRNA для направления слитых белков к геномным мишеням указаны в табл. 3, а праймеры для амплификации геномной области перечислены в табл. 4. Те же способы, что и в примере 4, для доставки плазмид в клетки млекопитающих и секвенирования ампликона используют для проверки возможностей редактирования оснований у данных RGN, когда они были присоединены к цитозиндезаминазе.

Таблица 21. Оценка степеней рецензирования оснований для каждой исследованной RGN

Конструкция	Мишень	Мутированные прочтения, %
APG05840-nAPG09298-USP2	SGN001159	6,96
APG05840-nAPG09298-USP2	SGN001159	9,91
APG05840-nAPG09298-USP2	SGN001162	21,59
APG05840-nAPG09298-USP2	SGN001162	27,05
APG05840-nAPG09298-USP2	SGN001163	17,85
APG05840-nAPG09298-USP2	SGN001163	20,74
APG05840-nAPG09298-USP2	SGN001164	4,74
APG05840-nAPG09298-USP2	SGN001164	7,89
APG05840-nAPG09298-USP2	SGN001165	31,99
APG05840-nAPG09298-USP2	SGN001165	62,47
APG05840-nAPG09298-USP2	SGN001166	23,49
APG05840-nAPG09298-USP2	SGN001166	29,56
APG05840-nAPG01604-USP2	SGN001245	9,67
APG05840-nAPG01604-USP2	SGN001246	1,4
APG05840-nAPG01604-USP2	SGN001249	1,72
APG05840-nAPG01604-USP2	SGN001250	13,41
APG05840-nAPG01604-USP2	SGN001251	0,87

Конструкция	Мишень	Мутированные прочтения, %
APG05840-nAPG01604-USP2	SGN001252	4,28

Пример 12. Нецелевой анализ (вне мишени)

Для оценки специфичности нуклеаз было определено нецелевое редактирование в потенциальных сайтах, идентифицированных с помощью биоинформатики. Потенциальные нецелевые сайты для APG01604 идентифицируют по мишеням с менее чем пятью несовпадениями в последовательности-мишени и по крайней мере с одним совпадением остатка в последовательности PAM.

Для проверки специфичности и нецелевого редактирования APG01604 используют те же методы, что и описанные в примере 4 для доставки плазмиды в клетки млекопитающих и секвенирования ампликона. Из того же эксперимента, в котором SGN в табл. 22 проверены на целевое редактирование, потенциальные нецелевые местоположения в таблице 23 проанализированы на возможное редактирование. Праймеры в табл. 24 используют для амплификации потенциальных нецелевых сайтов с последовательностью, сходной с целевым сайтом, для поиска нецелевого редактирования.

Таблица 22. SGN, используемые для поиска нецелевого редактирования

SGN	Ген	Forward-праймер в мишени (SEQ ID NO:)	Reverse-праймер в мишени (SEQ ID NO:)
SGN001675	TRA	CCCTTGTCCATCACTGGCAT (1096)	ACCAAAGCTGCCCTTACCTG (1097)
SGN001594	B2M	CCTTAATGTGCCTCCAGCCT (1092)	AGGAGAGACTCACGCTGGAT (1093)
SGN001674	TRA	CCCTTGTCCATCACTGGCAT (1094)	ACCAAAGCTGCCCTTACCTG (1095)

Таблица 23. Анализ последовательностей вне мишени

SGN	Номер вне мишени	Последовательность локуса вне мишени	SEQ ID NO:
SGN001594	1594-2	AGCACAGCTAAGGCCTAAAGTTGAA	1098
SGN001594	1594-3	AGCACAGCTAAGGCACCGGATTGGA	1099
SGN001594	1594-4	AGCACAGCCAAGGCCAAGGGCTGAC	1100
SGN001594	1594-5	AGCAGAGCTAAGGCCAAGGCAGGTG	1101
SGN001594	1594-6	GGCACAGCTAAGGCCAGCAGTGGCC	1102
SGN001674	1674-2	GTCTCTGAGCTGGTACATGGCAGAG	1103
SGN001674	1674-3	GTCTTTTAGCTGGTACACGTGTGTC	1104
SGN001675	1675-2	AGAAGATTTGTCACTGGATTCTGAG	1105
SGN001675	1675-3	AGGACACTTGTCACTGGATTTAGGA	1106

SGN	Номер вне мишени	Последовательность локуса вне мишени	SEQ ID NO:
SGN001675	1675-4	AGGACACTTGTCACTGGATTTAGGA	1107
SGN001675	1675-5	TGAGGACTTGTCACTGGATTCAGGG	1108
SGN001675	1675-6	AGGAGACTTTTCACTGGATTTAGGG	1109
SGN001675	1675-7	TGGACACTTGTCACTGGATTTAGGG	1110
SGN001675	1675-8	CACAGACTTGTCACTGGATGTGGGG	1111
SGN001675	1675-9	TACAGACATGTCACTGGATCTGGAA	1112
SGN001675	1675-10	AACACACTTGTCACTGGATTTAGGG	1113
SGN001675	1675-11	TACAGACTGGTCACTGGATGCTGGT	1114
SGN001675	1675-12	AGGACACTTGTCACTGGATTTAGCA	1115
SGN001675	1675-13	CTAAGACTTGTTACTGGATTGTGTG	1116
SGN001675	1675-14	CTCAGACTGGTCACTGGATAATGTA	1117

Таблица 24. Праймеры, применяемые для амплификации областей вне мишени

Описание	Последовательности праймеров	SEQ ID NO:
1594-2 FWD	CCAGAAGCCAGCAGATGACA	1118
1594-2 REV	GAGTGGTGGGCTCTCAATCC	1119
1594-3 FWD	AGCAACACCCATCCAAAGGTT	1120
1594-3 REV	GGGAGTGATGATAATGCGGG	1121
1594-4 FWD	TGGACTAGAGAGGGTTGGGG	1122
1594-4 REV	TCTCTTTCCACGAGCAGCAG	1123
1594-5 FWD	GCCTTTGACCTTCCCAGATT	1124
1594-5 REV	ACCATTGGAAAGGTGGATGC	1125
1594-6 FWD	GGCTTCAGGCTTTCCTCTGT	1126
1594-6 REV	AGCATGCTGGCCTAAAGTGA	1127
1674-2 FWD	ACCATTGGTCTGCTCAGGTG	1128
1674-2 REV	CCAAAACCTGCAGTGGCTTC	1129
1674-3 FWD	GGAGAGGAACTGGGCATGAG	1130
1674-3 REV	TCCGTCTCTCCTAGGTCTGC	1131
1675-2 FWD	CCATGACTGGCCCTTCTGTT	1132
1675-2 REV	GGGTAGAGTACATGGCGACG	1133
1675-3 FWD	CCCTCCCACCAGAAGCTCTA	1134
1675-3 REV	GATAAGAGGCCCAAGGACCG	1135
1675-4 FWD	CACTTGACACGTGAGCCTCT	1136
1675-4 REV	CCTCTTTCAGCCTCTGGTGG	1137
1675-5 FWD	GACAGACTTGGTTCTGCCCT	1138
1675-5 REV	CCTCCCCTCCTCCTAGCTT	1139
1675-6 FWD	TGCTGTAGTGGGTCTGAACAG	1140
1675-6 REV	TTTACCATGCTGGCTAGGC	1141
1675-7 FWD	CCCAGCCACATGGAACCTAAG	1142

Описание	Последовательности праймеров	SEQ ID NO:
1675-7 REV	GGACCCTAGGAGTTCCTTGT	1143
1675-8 FWD	GCAGGTTGATAGGGAAGAGC	1144
1675-8 REV	TCATCTCCCAGCTGATGACA	1145
1675-9 FWD	CCACCTGTTGCACAAATCCG	1146
1675-9 REV	TCCTCCCCTGGGAACATGAT	1147
1675-10 FWD	GAGGTA CTGGGAGTGGGGAT	1148
1675-10 REV	CTTCCTGCCTCTTCCAGCTC	1149
1675-11 FWD	GCACAGCTTTTGT CATGGGG	1150
1675-11 REV	GGGGATGAGAAAACAGAGCCA	1151
1675-12 FWD	CCATGCCCCATTCTGAAGGT	1152
1675-12 REV	TTCCCCTCCTCTAGACTGCC	1153
1675-13 FWD	CTTGAGCCCAGGAGTTTGAG	1154
1675-13 REV	TGCATTCTTGGGATGACCTC	1155
1675-14 FWD	ACTGCTTTTCCCTGGACACA	1156
1675-14 REV	GGTCACAAGTCCCCTGGTC	1157

При использовании доставки плазмиды не было обнаружено заметного нецелевого редактирования для двух из трех протестированных направляющих. Один нецелевой сайт для SGN001675, 1675-8, показал 2% редактирования в нецелевом локусе. В прочтениях наблюдают дупликацию 132 п.н. в геномной области в месте разреза NGS. Два нецелевых сайта, 1675-7 и 1675-13, не удалось секвенировать из-за плохой амплификации праймеров.

Таблица 25. Редактирование вне мишени для SGN001594 и APG08167

Обозначение вне мишени	Общее число несоответствий в мишени	Общее число несоответствий в семени	Общее число несоответствий в PAM	Редактирование, %
SGN001594	0	0	3	42,79
1594-2	2	2	3	0
1594-3	2	2	3	0,01
1594-4	2	1	2	0
1594-5	3	2	2	0
1594-6	2	1	1	0

Таблица 26. Редактирование вне мишени для SGN001675 и APG08167

Обозначение вне мишени	Общее число несоответствий в мишени	Общее число несоответствий в семени	Общее число несоответствий в PAM	Редактирование, %
SGN001675	0	0	3	33,115
1675-8	2	1	3	2,02

Обозначение вне мишени	Общее число несоответствий в мишени	Общее число несоответствий в семени	Общее число несоответствий в РАМ	Редактирование, %
1675-5	4	0	3	0,06
1675-10	3	1	3	0,03
1675-12	4	0	2	0,02
1675-6	4	0	3	0
1675-2	4	0	3	0
1675-3	4	0	3	0
1675-4	4	0	3	0
1675-9	3	1	3	0
1675-11	2	0	2	0
1675-14	4	1	2	0
1675-13	4	0	2	Итог секвенирования
1675-7	4	0	3	Итог секвенирования

Таблица 27 Редактирование вне мишени для SGN001674 и APG08167

Обозначение вне мишени	Общее число несоответствий в мишени	Общее число несоответствий в семени	Общее число несоответствий в РАМ	Редактирование, %
SGN001674	0	0	3	52,17
1674-2	2	1	3	0
1674-3	2	0	1	0,03

Пример 13. Тестирование вариантов каркаса направляющей РНК

5 APG08167 и APG01604 примерно на 62,72% идентичны, но распознают один и тот же РАМ, имеют одну и ту же crRNA и имеют близкородственную последовательность tracrRNA. Из-за схожести последовательностей направляющей РНК все данные, ранее полученные для APG01604, используют как основу tracrRNA APG08167. Эти исследования были проведены, чтобы

10 подтвердить, что белок APG01604 более активен с tracrRNA, закодированной в геноме APG08167 (нативная tracrRNA APG08167). Различные длины спейсеров в crRNA также тестируют, чтобы определить, предпочитает ли белок APG01604 спейсерную последовательность из 20 или 25 пар оснований.

15 Для этого были созданы синтетические crRNA с шестью различными целевыми последовательностями, которые содержат мишень из 20 или 25 пар

оснований. Различные crRNA комбинаторно комбинируют с синтетической tracrRNA из генома APG08167 или APG01604, формируют RNP, и комплексы RNP подвергают нуклеофекции в клетки НЕК293Т. Стандартные методы в примере 4 используют для определения степени редактирования в клетках.

5 Таблица 28. Последовательности crRNA

Ген	Спейсер	Длина спейсера	SEQ ID NO
B2M	A	20	1158
	B	20	1159
	C	20	1160
	D	20	1161
	A	25	1162
	B	25	1163
	C	25	1164
	D	25	1165
TRA	E	20	1166
	F	20	1167
	E	25	1168
	F	25	1169

Таблица 29. Последовательности tracrRNA

tracrRNA	SEQ ID NO
APG08167 tracrRNA	1170
APG01604 tracrRNA	1171

Таблица 30. Последовательности праймеров для секвенирования

Праймер	SEQ ID NO:
TRAC Левый праймер	1170
TRAC Правый праймер	1171
B2M Левый праймер	1172
B2M Правый праймер	1173

10

Таблица 31. Результаты редактирования каркаса APG01604 и испытание длины спейсера

Ген	Спейсер	Длина спейсера	Каркас	Эффективность редактирования, %
B2M	A	20	APG01604 tr	0,21
	B	20	APG01604 tr	0,08
	C	20	APG01604 tr	0,27
	D	20	APG01604 tr	0,17
	A	25	APG01604 tr	0,15
	B	25	APG01604 tr	0
	C	25	APG01604 tr	1,33
	D	25	APG01604 tr	0,31
	A	20	APG08167 tr	6,58
	B	20	APG08167 tr	1,75
	C	20	APG08167 tr	6,99
	D	20	APG08167 tr	64,6

Ген	Спейсер	Длина спейсера	Каркас	Эффективность редактирования, %
	A	25	APG08167 tr	22,7
	B	25	APG08167 tr	2,36
	C	25	APG08167 tr	52,51
	D	25	APG08167 tr	82,58
TRA	E	20	APG01604 tr	0,05
	F	20	APG01604 tr	0,01
	E	25	APG01604 tr	0,02
	F	25	APG01604 tr	0,01
	E	20	APG08167 tr	0
	F	20	APG08167 tr	0,29
	E	25	APG08167 tr	2,64
	F	25	APG08167 tr	1,13

Скорость редактирования в клетках млекопитающих демонстрирует отсутствие надежного редактирования при использовании нативной последовательности *tracrRNA* APG01604. RNP APG01604 показывает гораздо более высокие степени редактирования при использовании последовательности *tracrRNA* APG08167. Кроме того, при сравнении одной и той же мишени спейсер длиной в 25 п.о. показывает более высокое редактирование в мишенях A, C, D, E и F, чем *crRNA*, которые имеют длину спейсера в 20 п.о. Эти результаты в совокупности демонстрируют, что *tracrRNA* APG08167 проявляет более высокий уровень редактирования, чем *tracrRNA* APG01604. Кроме того, RNP APG01604 лучше работает со спейсерной последовательностью из 25 п.н., чем со спейсерной последовательностью из 20 п.о.

ПЕРВОНАЧАЛЬНАЯ ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Молекула нуклеиновой кислоты, содержащая полинуклеотид, который кодирует полипептид РНК-направляемой нуклеазы (RGN), где указанный полинуклеотид содержит нуклеотидную последовательность, кодирующую полипептид RGN, включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579, где указанный полипептид RGN способен связывать последовательность ДНК-мишени специфическим для РНК-направляющей последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной последовательностью ДНК-мишени, где указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

2. Молекула нуклеиновой кислоты по п. 1, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

3. Молекула нуклеиновой кислоты по п. 1, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

4. Молекула нуклеиновой кислоты по п. 1, где указанный полипептид RGN по меньшей мере на 90% идентичен последовательности SEQ ID NO: 63 и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности SEQ ID NO: 63.

5. Молекула нуклеиновой кислоты по любому из п.п. 1-4, где указанный полипептид RGN способен расщеплять указанную последовательность ДНК-мишени при связывании.

5 6. Молекула нуклеиновой кислоты по п. 5, где указанный полипептид RGN способен вызывать двухцепочечный разрыв.

7. Молекула нуклеиновой кислоты по п. 5, где указанный полипептид RGN способен вызывать одноцепочечный разрыв.

10

8. Молекула нуклеиновой кислоты по любому из п.п. 1-4, где указанный полипептид RGN не обладает нуклеазной активностью или является никазой.

15 9. Молекула нуклеиновой кислоты по любому из п.п. 1-8, где полипептид RGN функционально слит с полипептидом, редактирующим основания.

10. Молекула нуклеиновой кислоты по п. 9, где полипептид, редактирующий основания, представляет собой дезаминазу.

20 11. Молекула нуклеиновой кислоты по любому из п.п. 1-10, где полипептид RGN включает один или несколько сигналов ядерной локализации.

12. Молекула нуклеиновой кислоты по любому из п.п. 1-11, где полипептид RGN кодон-оптимизирован для экспрессии в эукариотической клетке.

25

13. Молекула нуклеиновой кислоты по любому из п.п. 1-12, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсейсеру (PAM).

30 14. Вектор, включающий молекулу нуклеиновой кислоты по любому из п.п. 1-13.

15. Вектор по п. 14, дополнительно включающий по меньшей мере одну нуклеотидную последовательность, кодирующую указанную gRNA, способную гибридизироваться с указанной последовательностью ДНК-мишени.

5 16. Вектор по п. 15, где направляющая РНК выбрана из группы, состоящей из:

а) направляющей РНК, включающей:

i) CRISPR РНК, содержащую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 2; и

10 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 3;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 1;

15 б) направляющей РНК, включающей:

i) CRISPR РНК, содержащую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 9; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 10;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 8;

в) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 16;

и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 17;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 15;

г) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 23;
и

5 ii) tcrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 24;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 22;

10 д) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 30;
и

15 ii) tcrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 31;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 29;

е) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 37;
и

ii) tcrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 38;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 36;

ж) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 44;
и

ii) tcrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 45;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 43;

з) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 51; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 52;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 50;

 и) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 57; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 58;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 56;

 к) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 64; и

 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 65;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 63 и 570-579;

 л) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 71; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 72;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 70;

м) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 77;

и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 76;

н) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84;

и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 85;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 83;

о) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 90;

и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 91;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 89;

п) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 97; и

5 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности, представленной как SEQ ID NO: 98;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 96;

10 р) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 104; и

15 ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 105;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 103;

с) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 111; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 112;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 110;

т) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 118; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 119;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 117;

у) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 124; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 125;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 123;

ф) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84; и

ii) tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 83.

17. Вектор по п. 15, где направляющая РНК выбрана из группы, состоящей из:

25 а) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 2; и

30 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 3;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 1;

б) направляющей РНК, включающей:

и) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 9;
и

5 ii) tcracRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 10;

 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 8;

10 в) направляющей РНК, включающей:

и) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 16;
и

15 ii) tcracRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 17;

 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 15;

г) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 23;
и

ii) tcracRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 24;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 22;

д) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 30;
и

ii) tcracRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 31;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 29;

е) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 37; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 38;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 36;

 ж) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 44; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 45;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 43;

 з) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 51; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 52;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 50;

 и) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 57; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 58;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 56;

к) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 64;

и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности, представленной как SEQ ID NO: 65;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

л) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 71;

и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 72;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 70;

м) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 77;

и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 76;

н) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 85;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 83;

10 о) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 90; и

15 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 91;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 89;

п) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 97; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 98;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 96;

р) направляющей РНК, включающей:

30 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 104; и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 105;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 103;

с) направляющей РНК, включающей:

5 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 111; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 112;

10 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 110;

 т) направляющей РНК, включающей:

15 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 118; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 119;

20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 117;

 у) направляющей РНК, включающей:

25 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 124; и

 ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 125;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 123;

 ф) направляющей РНК, включающей:

 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84;
и

ii) tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 83.

18. Вектор по п. 15, где направляющая РНК выбрана из группы, состоящей из:

а) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 2; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 3;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 1;

б) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 9; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 10;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 8;

25 в) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 16; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 17;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 15;

г) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 23; и

5 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 24;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 22;

д) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 30; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 31;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 29;

е) направляющей РНК, включающей:

20 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности, представленной как SEQ ID NO: 37; и

ii) tracrRNA, которая на 100% идентична последовательности, представленной как SEQ ID NO: 38;

25 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности, представленной как SEQ ID NO: 36;

ж) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 44; и

30 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 45;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 43;

з) направляющей РНК, включающей:

i) CRISPR РНК, которая на 100% идентична последовательности SEQ ID NO: 51; и

5 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 52;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 50;

и) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 57; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 58;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 56;

к) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 64; и

20 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 65;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 63 и 570-579;

25 л) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 71; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 72;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 70;

м) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 77; и

5 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 76;

н) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 84; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 85;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 83;

о) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 90; и

20 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 91;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 89;

25 п) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 97; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 98;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 96;

р) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 104; и

5 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 105;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 103;

с) направляющей РНК, включающей:

10 i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 111; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 112;

15 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 110;

т) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 118; и

20 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 119;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 117;

25 у) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 124; и

ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 125;

30 где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 123; и

ф) направляющей РНК, включающей:

i) CRISPR РНК, включающую последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 84; и

5 ii) tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78;

где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 83.

10 19. Вектор по любому из п.п. 15-18, где указанная gRNA представляет собой одиночную направляющую РНК.

20. Вектор по любому из п.п. 15-18, где указанная gRNA представляет собой двойную направляющую РНК.

15 21. Клетка, включающая молекулу нуклеиновой кислоты по любому из п.п. 1-13 или вектор по любому из п.п. 14-20.

20 22. Способ получения полипептида RGN, включающий культивирование клетки по п. 21 в условиях, в которых экспрессируется полипептид RGN.

25 23. Способ получения полипептида RGN, включающий введение в клетку гетерологичной молекулы нуклеиновой кислоты, содержащей нуклеотидную последовательность, кодирующую полипептид РНК-направляемой нуклеазы (RGN), включающей аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

30 где указанный полипептид RGN способен связывать последовательность ДНК-мишени специфическим для РНК направляющей последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной ДНК-мишенью;

и культивирование указанной клетки в условиях, в которых экспрессируется полипептид RGN.

24. Способ по п. 23, в котором указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

5

25. Способ по п. 23, в котором указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

10

26. Способ по п. 23, в котором указанный полипептид RGN по меньшей мере на 90% идентичен последовательности SEQ ID NO: 63 и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 в последовательности SEQ ID NO: 63.

15

27. Способ по п.п. 22-26, дополнительно включающий очистку указанного полипептида RGN.

20

28. Способ по любому из п.п. 22-26, где указанная клетка дополнительно экспрессирует одну или несколько направляющих РНК, способных связываться с указанным полипептидом RGN для образования рибонуклеопротеинового комплекса RGN.

25

29. Способ по п. 28, дополнительно включающий очистку указанного рибонуклеопротеинового комплекса RGN.

30

30. Выделенный полипептид РНК-направляемой нуклеазы (RGN), где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579; и

где указанный полипептид RGN способен связывать последовательность ДНК-мишени молекулы ДНК специфическим для РНК-направляющей

последовательности образом при связывании с направляющей РНК (gRNA), способной к гибридизации с указанной последовательностью ДНК-мишени.

5 31. Выделенный полипептид RGN по п. 30, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

10 32. Выделенный полипептид RGN по п. 30, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

15 33. Выделенный полипептид RGN по п. 30, где указанный полипептид RGN по меньшей мере на 90% идентичен последовательности SEQ ID NO: 63 и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности SEQ ID NO: 63.

20 34. Выделенный полипептид RGN по любому из п.п. 30-33, где указанный полипептид RGN способен расщеплять указанную последовательность ДНК-мишени при связывании.

25 35. Выделенный полипептид RGN по п. 34, где расщепление указанным полипептидом RGN приводит к образованию двухцепочечного разрыва.

30 36. Выделенный полипептид RGN по п. 34, где расщепление указанным полипептидом RGN приводит к образованию одноцепочечного разрыва.

37. Выделенный полипептид RGN по любому из п.п. 30-33, где указанный полипептид RGN является неактивной нуклеазой или никазой.

38. Выделенный полипептид RGN по любому из п.п. 30-37, где полипептид RGN функционально слит с полипептидом, редактирующим основания.

5 39. Выделенный полипептид RGN по п. 38, где полипептид, редактирующий основания, является дезаминазой.

40. Выделенный полипептид RGN по любому из п.п. 30-39, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM).

10 41. Выделенный полипептид RGN по любому из п.п. 30-40, где полипептид RGN включает один или более сигналов ядерной локализации.

15 42. Система для связывания последовательности ДНК-мишени молекулы ДНК, где указанная система включает:

а) одну или несколько направляющих РНК, способных гибридизироваться с указанной последовательностью ДНК-мишени, или один или несколько полинуклеотидов, включающих одну или несколько нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК (gRNA); и

20 б) полипептид РНК-направляемой нуклеазы (RGN), включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579, или полинуклеотид, включающий нуклеотидную последовательность, кодирующую полипептид RGN;

25 где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК, и указанная нуклеотидная последовательность, кодирующая полипептид RGN, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности;

30 где одна или несколько направляющих РНК способны гибридизироваться с последовательностью ДНК-мишени, и

где одна или несколько направляющих РНК способны образовывать комплекс с полипептидом RGN, чтобы направить указанный полипептид RGN для связывания с последовательностью ДНК-мишени молекулы ДНК.

5 43. Система для связывания последовательности ДНК-мишени молекулы ДНК, где указанная система включает:

а) одну или несколько направляющих РНК, способных гибридизироваться с указанной последовательностью ДНК-мишени, или один или несколько полинуклеотидов, включающих одну или несколько нуклеотидных последовательностей, кодирующих одну или несколько направляющих РНК (gRNA); и

10 б) полипептид РНК-направляемой нуклеазы (RGN), включающий аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 15 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

где одна или более направляющих РНК способны гибридизироваться с последовательностью ДНК-мишени, и

где одна или несколько направляющих РНК способны образовывать комплекс с полипептидом RGN, чтобы направлять указанный полипептид RGN на связывание с указанной последовательностью ДНК-мишени молекулы ДНК.

20 44. Система по п.п. 42 или 43, где по меньшей мере одна из указанных последовательностей нуклеотидов, кодирующих одну или несколько направляющих РНК, функционально связана с промотором, гетерологичным указанной последовательности нуклеотидов.

45. Система по любому из п.п. 42-44, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 30 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

46. Система по любому из п.п. 42-44, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична

любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

5 47. Система по любому из п.п. 42-44, где указанный полипептид RGN по меньшей мере на 95% идентичен последовательности SEQ ID NO: 63 и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368 и валин в положении аминокислоты 405 в SEQ ID NO: 63.

10 48. Система по любому из п.п. 42-47, в которой указанный полипептид RGN и указанная одна или несколько направляющих РНК не встречаются в природе в комплексе друг с другом.

15 49. Система по любому из п.п. 42-48, в которой указанная последовательность ДНК-мишени представляет собой эукариотическую последовательность ДНК-мишени.

20 50. Система по любому из п.п. 42-49, в которой указанная gRNA представляет собой одиночную направляющую РНК (sgRNA).

51. Система по любому из п.п. 42-49, в которой указанная gRNA представляет собой двойную направляющую РНК.

25 52. Система по любому из п.п. 42-51, где указанная gRNA выбрана из группы, состоящей из:

30 а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 3, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 10,

где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 8;

5 в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 15;

10 г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 23, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 22;

15 д) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 30, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 31, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 29;

20 е) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 37, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 38, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 36;

30 ж) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 44, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 45, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 43;

з) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 51, и

tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 50;

5 и) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 56;

10 к) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

15 л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 70;

20 м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 76;

25 н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 90, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную

последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84, и
5 tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 83.

10 53. Система по любому из п.п. 42-51, где указанная gRNA выбрана из группы, состоящей из:

а) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 3,
15 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 1;

б) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 10,
20 где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 8;

в) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 16, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ
25 ID NO: 17, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 15;

г) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 23, и
30 tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 24, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 22;

последовательность, которая по меньшей мере на 95% идентична любой из последовательностей, представленных как SEQ: 63 и 570-579;

л) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 71, и
5 tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 70;

10 м) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 76;

15 н) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична
20 последовательности SEQ ID NO: 83;

о) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 95, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную
25 последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ
30 ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 104, и

tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 103;

5 с) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 110;

10 т) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 117;

15 у) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 123;

20 ф) gRNA, включающей последовательность повтора CRISPR, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 83.

30 54. Система по любому из п.п. 42-51, где указанная gRNA выбрана из группы, состоящей из:

 а) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 2, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 3, где указанный полипептид

100% идентична последовательности SEQ ID NO: 52, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 50;

5 и) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 57, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 58, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 56;

10 к) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 64, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 65, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей, представленных как SEQ ID NO: 63 и 570-579;

15 л) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 72, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 70;

20 м) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 76;

25 н) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 85, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 83;

30 о) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 90, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 91, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 89;

п) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 98, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 96;

р) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 104, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 105, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 103;

с) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 112, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 110;

т) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 119, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 117;

у) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 125, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 123;

ф) gRNA, включающей последовательность повтора CRISPR, которая на 100% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78, где указанный полипептид RGN включает аминокислотную последовательность, которая на 100% идентична последовательности SEQ ID NO: 83.

55. Система по любому из п.п. 42-54, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM).

56. Система по любому из п.п. 42-55, где последовательность ДНК-мишени находится внутри клетки.

5 57. Система по п. 56, где клетка представляет собой эукариотическую клетку.

58. Система по п. 57, где эукариотическая клетка представляет собой растительную клетку.

10 59. Система по п. 57, где эукариотическая клетка представляет собой клетку млекопитающего.

15 60. Система по п. 57, где эукариотическая клетка представляет собой клетку насекомого.

61. Система по п. 56, где клетка представляет собой прокариотическую клетку.

20 62. Система по любому из п.п. 42-61, где при транскрипции одна или несколько направляющих РНК способны гибридизироваться с последовательностью ДНК-мишени, и направляющая РНК способна образовывать комплекс с полипептидом RGN для прямого расщепления последовательности ДНК.

25 63. Система по п. 62, где расщепление приводит к образованию двухцепочечного разрыва.

30 64. Система по п. 62, где расщепление приводит к образованию одноцепочечного разрыва.

65. Система по любому из п.п. 42-61, где указанный полипептид RGN не обладает нуклеазной активностью или представляет собой никазу.

66. Система по любому из п.п. 42-65, где полипептид RGN функционально связан с полипептидом, редактирующим основания.

5 67. Система по п. 66, где полипептид, редактирующий основания, представляет собой дезаминазу.

68. Система по любому из п.п. 42-67, где полипептид RGN включает один или несколько сигналов ядерной локализации.

10 69. Система по любому из п.п. 42-68, где полипептид RGN кодон-оптимизирован для экспрессии в эукариотической клетке.

15 70. Система по любому из п.п. 42-69, где нуклеотидные последовательности, кодирующие одну или более направляющих РНК, и нуклеотидная последовательность, кодирующая полипептид RGN, расположены в одном векторе.

20 71. Система по любому п.п. 42-70, где указанная система дополнительно включает один или несколько донорных полинуклеотидов или одну или несколько нуклеотидных последовательностей, кодирующих один или несколько донорных полинуклеотидов.

25 72. Фармацевтический состав, включающий молекулу нуклеиновой кислоты по любому из п.п. 1-13, вектор по любому из п.п. 14-20, клетку по п. 21, выделенный полипептид RGN по любому из п.п. 30-41 или систему по любому из п.п. 42-71, и фармацевтически приемлемый носитель.

30 73. Способ связывания последовательности ДНК-мишени молекулы ДНК, включающий доставку системы по любому из п.п. 42-71 к указанной последовательности ДНК-мишени или клетке, включающей последовательность ДНК-мишени.

74. Способ по п. 73, в котором указанный полипептид RGN или указанная направляющая РНК дополнительно включает выявляемую метку, что позволяет обнаружить указанную последовательность ДНК-мишени.

5 75. Способ по п. 73, в котором указанная направляющая РНК или указанный полипептид RGN дополнительно включает модулятор экспрессии, тем самым модулируя экспрессию указанной последовательности ДНК-мишени или гена, находящихся под транскрипционным контролем указанной последовательности ДНК-мишени.

10 76. Способ расщепления или модификации последовательности ДНК-мишени молекулы ДНК, включающий доставку системы по любому из п.п. 42-71 к указанной последовательности ДНК-мишени или клетке, включающей молекулу ДНК, и расщепление или модификация указанной последовательности ДНК-мишени.

15 77. Способ по п. 76, в котором указанная модифицированная последовательность ДНК-мишени включает инсерцию гетерологичной ДНК в последовательность ДНК-мишени.

20 78. Способ по п. 76, где указанная модифицированная последовательность ДНК-мишени включает делецию по меньшей мере одного нуклеотида из последовательности ДНК-мишени.

25 79. Способ по п. 76, в котором указанная модифицированная последовательность ДНК-мишени включает мутацию по меньшей мере одного нуклеотида в последовательности ДНК-мишени.

30 80. Способ связывания последовательности ДНК-мишени молекулы ДНК, включающий:

а) сборку рибонуклеотидного комплекса РНК-направляемой нуклеазы (RGN) *in vitro* путем объединения:

і) одной или нескольких направляющих РНК, способных гибридизоваться с последовательностью ДНК-мишени; и

ii) полипептида RGN, включающего аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579;

5 в условиях, подходящих для образования рибонуклеотидного комплекса RGN; и

б) контактирование указанной последовательности ДНК-мишени или клетки, содержащей указанную последовательность ДНК-мишени, с собранным *in vitro* рибонуклеотидным комплексом RGN;

10 где одна или несколько направляющих РНК гибридизируются с последовательностью ДНК-мишени, тем самым направляя указанный полипептид RGN для связывания с последовательностью ДНК-мишени.

81. Способ по п. 80, в котором указанный полипептид RGN или указанная
15 направляющая РНК дополнительно включает выявляемую метку, что позволяет обнаружить указанную последовательность ДНК-мишени.

82. Способ по п. 80, в котором указанная направляющая РНК или
указанный полипептид RGN дополнительно включает модулятор экспрессии,
20 тем самым модулируя экспрессию указанной последовательности ДНК-мишени.

83. Способ расщепления и/или модификации последовательности ДНК-
мишени в молекуле ДНК, включающий контактирование молекулы ДНК с:

а) полипептидом РНК-направляемой нуклеазы (RGN), где RGN включает
25 аминокислотную последовательность, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123, и 570-579; и

б) одной или несколькими направляющими РНК, способными нацеливать
указанную RGN на последовательность ДНК-мишени;

30 где одна или несколько направляющих РНК гибридизируются с последовательностью ДНК-мишени, тем самым направляя указанный полипептид RGN для связывания с указанной последовательностью ДНК-мишени и расщепления и/или модификации указанной последовательности ДНК-мишени.

84. Способ по п. 83, в котором расщепление указанным полипептидом RGN приводит к образованию двухцепочечного разрыва.

5 85. Способ по п. 83, в котором расщепление указанным полипептидом RGN приводит к образованию одноцепочечного разрыва.

10 86. Способ по п. 83, где указанный полипептид RGN не обладает нуклеазной активностью или является нисказой и функционально слит с полипептидом, редактирующим основания.

87. Способ по п. 86, где полипептид, редактирующий основания, представляет собой дезаминазу.

15 88. Способ по п. 83, в котором указанная модифицированная последовательность ДНК-мишени включает инсерцию гетерологичной ДНК в последовательность ДНК-мишени

20 89. Способ по п. 83, где указанная модифицированная последовательность ДНК-мишени включает делецию по меньшей мере одного нуклеотида из последовательности ДНК-мишени.

25 90. Способ по п. 83, где указанная модифицированная последовательность ДНК-мишени включает мутацию по меньшей мере одного нуклеотида в последовательности ДНК-мишени.

91. Способ любому из п.п. 80-90, где указанная последовательность ДНК-мишени расположена рядом с мотивом, примыкающим к протоспейсеру (PAM).

30 92. Способ по любому из п.п. 80-91, где указанная последовательность ДНК-мишени является эукариотической последовательностью ДНК-мишени.

93. Способ по любому из п.п. 80-92, где указанная gRNA является одиночной направляющей РНК (sgRNA).

94. Способ по любому из п.п. 80-92, где указанная gRNA является двойной направляющей РНК.

5 95. Способ по любому из п.п. 80-94, где указанная RGN включает аминокислотную последовательность, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

10 96. Способ по любому из п.п. 80-94, где указанный RGN включает аминокислотную последовательность, которая на 100% идентична любой из последовательностей SEQ ID NO: 1, 8, 15, 22, 29, 36, 43, 50, 56, 63, 70, 76, 83, 89, 96, 103, 110, 117, 123 и 570-579.

15 97. Способ по любому из п.п. 80-94, где указанный полипептид RGN по меньшей мере на 90% идентичен последовательности SEQ ID NO: 63 и имеет изолейцин в положении аминокислоты 305, валин в положении аминокислоты 328, лейцин в положении аминокислоты 366, треонин в положении аминокислоты 368, и валин в положении аминокислоты 405 последовательности
20 SEQ ID NO: 63.

98. Способ по любому из п.п. 80-94, где:

а) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 1, указанная направляющая РНК включает последовательность
25 повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 3;

б) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 8, указанная направляющая РНК включает последовательность
30 повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 10;

в) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 15, указанная направляющая РНК включает последовательность

к) указанная RGN по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 65;

5

л) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 70, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 72;

10

м) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 76, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78;

15

н) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 85;

20

о) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 89, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 90, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 91;

25

п) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 96, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 98;

30

р) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 103, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности

SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 105;

5 с) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 110, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 112;

10 т) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 117, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 119;

15 у) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 125; или

20 ф) указанная RGN по меньшей мере на 90% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 78.

99. Способ по любому из п.п. 80-94, где:

25 а) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 1, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 2, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 3;

30 б) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 8, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 9, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 10;

SEQ ID NO: 57, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 58;

к) указанная RGN по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 64, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 65;

л) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 70, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 71, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 72;

м) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 76, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78;

н) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 85;

о) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 89, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 90, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 91;

п) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 96, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 98;

р) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 103, указанная направляющая РНК включает последовательность

повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 104, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 105;

5 с) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 110, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 112;

10 т) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 117, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 119;

15 у) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 125; или

20 ф) указанная RGN по меньшей мере на 95% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 78.

25 100. Способ по любому из п.п. 80-94, где:

а) указанная RGN на 100% идентична последовательности SEQ ID NO: 1, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 2, и tracrRNA, на 100% идентична последовательности SEQ ID NO: 3;

30 б) указанная RGN на 100% идентична последовательности SEQ ID NO: 8, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 9, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 10;

в) указанная RGN на 100% идентична последовательности SEQ ID NO: 15, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 16, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 17;

5 г) указанная RGN на 100% идентична последовательности SEQ ID NO: 22, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 23, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 24;

10 д) указанная RGN на 100% идентична последовательности SEQ ID NO: 29, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 30, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 31;

15 е) указанная RGN на 100% идентична последовательности SEQ ID NO: 36, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 37, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 38;

20 ж) указанная RGN на 100% идентична последовательности SEQ ID NO: 43, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 44, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 45;

з) указанная RGN на 100% идентична последовательности SEQ ID NO: 50, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 51, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 52;

25 и) указанная RGN по на 100% идентична последовательности SEQ ID NO: 56, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 57, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 58;

30 к) указанная RGN на 100% идентична любой из последовательностей SEQ ID NO: 63 и 570-579, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 64, и tracrRNA, на 100% идентична последовательности SEQ ID NO: 65;

л) указанная RGN на 100% идентична последовательности SEQ ID NO: 70, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 71, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 72;

5 м) указанная RGN на 100% идентична последовательности SEQ ID NO: 76, указанная направляющая РНК включает последовательность повтора crRNA, на 100% идентична последовательности SEQ ID NO: 77, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78;

10 н) указанная RGN на 100% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 84, и tracrRNA, на 100% идентична последовательности SEQ ID NO: 85;

15 о) указанная RGN на 100% идентична последовательности SEQ ID NO: 89, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 90, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 91;

20 п) указанная RGN на 100% идентична последовательности SEQ ID NO: 96, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 97, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 98;

р) указанная RGN на 100% идентична последовательности SEQ ID NO: 103, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 104, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 105;

25 с) указанная RGN на 100% идентична последовательности SEQ ID NO: 110, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 111, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 112;

30 т) указанная RGN на 100% идентична последовательности SEQ ID NO: 117, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 118, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 119;

у) указанная RGN на 100% идентична последовательности SEQ ID NO: 123, указанная направляющая РНК включает последовательность повтора crRNA,

которая на 100% идентична последовательности SEQ ID NO: 124, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 125; или

5 ф) указанная RGN на 100% идентична последовательности SEQ ID NO: 83, указанная направляющая РНК включает последовательность повтора crRNA, которая на 100% идентична последовательности SEQ ID NO: 84, и tracrRNA, которая на 100% идентична последовательности SEQ ID NO: 78.

10 101. Способ по любому из п.п. 73-100, где последовательность ДНК-мишени находится внутри клетки.

102. Способ по п. 101, где клетка является эукариотической клеткой.

15 103. Способ по п. 102, где эукариотическая клетка является клеткой растения.

104. Способ по п. 102, где эукариотическая клетка является клеткой млекопитающего.

20 105. Способ по п. 102, где эукариотическая клетка является клеткой насекомого.

106. Способ по п. 101, где клетка является прокариотической клеткой.

25 107. Способ по любому из п.п. 101-106, дополнительно включающий культивирование клетки в условиях, в которых полипептид RGN экспрессируется и расщепляет последовательность ДНК-мишени с получением молекулы ДНК, содержащей модифицированную последовательность ДНК; и отбор клетки, содержащей указанную модифицированную последовательность ДНК-мишени.

30 108. Клетка, содержащая модифицированную последовательность ДНК-мишени в соответствии со способом по п.107.

109. Клетка по п. 108, где клетка является эукариотической клеткой.

110. Клетка по п. 109, где эукариотическая клетка является клеткой растения.

5 111. Растение, включающее клетку по п. 110.

112. Семя, включающее клетку по п. 110.

10 113. Клетка по п. 109, где эукариотическая клетка является клеткой млекопитающего.

114. Клетка по п. 113, где указанная клетка млекопитающего является клеткой человека.

15 115. Клетка по п. 114, где указанная клетка человека является иммунной клеткой.

116. Клетка по п. 115, где указанная иммунная клетка является стволовой клеткой.

20

117. Клетка по п. 116, где указанная стволовая клетка является индуцированной мультипотентной стволовой клеткой.

25 118. Клетка по п. 109, где эукариотическая клетка является клеткой насекомого.

119. Клетка по п. 108, где клетка является прокариотической клеткой.

30 120. Фармацевтический состав, включающий клетку по любому из п.п. 108, 109, 113-117 и фармацевтически приемлемый носитель.

121. Способ лечения заболевания, включающий введение субъекту, нуждающемуся в этом, эффективного количества фармацевтического состава по п. 72 или п. 120.

122. Способ по п. 121, где указанное заболевание связано с вызванной мутацией, и указанное эффективное количество указанного фармацевтического состава корректирует указанную вызванную мутацию.