

(19)



**Евразийское
патентное
ведомство**

(21) **202390840** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.07.12

(22) Дата подачи заявки
2021.09.10

(51) Int. Cl. *C12N 9/80* (2006.01)
C12N 9/22 (2006.01)
C12N 15/52 (2006.01)
C12N 15/62 (2006.01)
C12N 9/78 (2006.01)

**(54) МОДИФИЦИРУЮЩИЕ ДНК ФЕРМЕНТЫ И ИХ АКТИВНЫЕ ФРАГМЕНТЫ,
ВАРИАНТЫ И СПОСОБЫ ИХ ПРИМЕНЕНИЯ**

(31) 63/077,089; 63/146,840

(32) 2020.09.11; 2021.02.08

(33) US

(86) PCT/US2021/049853

(87) WO 2022/056254 2022.03.17

(88) 2022.04.21

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

(72) Изобретатель:

Боуэн Тайсон Д., 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 описаны композиции и способы, включающие новые полипептиды дезаминазы для направленного редактирования нуклеиновых кислот. Композиции содержат полипептиды дезаминазы. Также предложены гибридные белки, содержащие ДНК-связывающий полипептид и дезаминазу по настоящему изобретению. Гибридные белки включают РНК-направляемые нуклеазы, гибридизованные с дезаминазами, необязательно в составе комплекса с направляющими РНК. Композиции также включают молекулы нуклеиновой кислоты, кодирующие дезаминазы или гибридные белки. Также предложены векторы и клетки-хозяина, содержащие молекулы нуклеиновой кислоты, кодирующие дезаминазы или гибридные белки.

A1

202390840

202390840

A1

МОДИФИЦИРУЮЩИЕ ДНК ФЕРМЕНТЫ И ИХ АКТИВНЫЕ ФРАГМЕНТЫ И ВАРИАНТЫ И СПОСОБЫ ИХ ПРИМЕНЕНИЯ

5

Перекрестные ссылки на родственные заявки

В настоящей заявке испрашивается приоритет в связи с предварительными заявками №63/077089, поданной 11 сентября 2020 г. и № 63/46840, поданной 8 февраля 2021 г., каждая из которых в полном объеме включена в данный документ.

10

Заявление в отношении перечня последовательностей

Перечень последовательностей, связанный с этой заявкой, предоставлен в формате ASCII вместо бумажной копии и в данном контексте включен в качестве ссылки в описание изобретения. Копия ASCII с названием L103438_1230WO_0108_1_SL.txt размером 1071246 байт, была создана 9 сентября 2021 г. и подана в электронном виде через EFS-Web.

15

Область изобретения

Настоящее изобретение относится к области молекулярной биологии и редактирования генов.

20

Предпосылки создания изобретения

Направленное редактирование или модификация генома быстро становится важным инструментом фундаментальных и прикладных исследований.

Первоначальные методы включали нуклеазы для генной инженерии, такие как мегануклеазы, гибридные белки цинковых пальцев или TALEN, требующие

25

создания химерных нуклеаз, содержащих сконструированные, программируемые, специфичные для последовательности ДНК-связывающие домены, специфичные для каждой конкретной последовательности-мишени.

РНК-управляемые нуклеазы (RGN), такие как (Cas)-ассоциированные белки бактериальной системы CRISPR-Cas, то есть ассоциированные с

30

кластеризованными регулярно расположенными короткими палиндромными повторами (CRISPR), позволяют направлять связывание на определенные последовательности за счет образования комплексов нуклеаз с направляющей РНК, которая специфично гибридизуется с определенной последовательностью -

мишенью. Продуцирование направляющих РНК, специфичных для мишени, является менее затратным и более эффективным способом, чем создание химерных нуклеаз для каждой последовательности-мишени. Такие нуклеазы, управляемые РНК, можно использовать для редактирования геномов при введении специфичного для последовательности двухцепочечного разрыва, который восстанавливается с помощью подверженного ошибкам негомологичного соединения концов (NHEJ) для внесения мутации в определенное место генома.

Кроме того, RGN можно использовать для направленного редактирования ДНК. Направленное редактирование последовательностей нуклеиновых кислот, например, направленное расщепление для введения специфической модификации в геномную ДНК, обеспечивает очень тонкий подход к изучению функции и экспрессии генов. RGN также можно использовать для создания химерных белков, которые используют РНК-управляемую активность RGN в сочетании с ферментом, модифицирующим ДНК, таким как дезаминаза, для направленного редактирования оснований. Направленное редактирование можно использовать для борьбы с генетическими заболеваниями у людей или для включения агрономически полезных мутаций в геномы сельскохозяйственных растений. Разработка инструментов редактирования генома обеспечивает новые подходы к терапии млекопитающих и к агробιοтехнологии на основе редактирования генов.

Краткое описание вариантов осуществления изобретения

В данном документе предлагаются композиции и способы модификации молекулы ДНК-мишени. Композиции находят применение для модификации исследуемой молекулы ДНК-мишени. Предлагаемые композиции содержат полипептиды дезаминазы. Предлагаются также гибридные белки, содержащие полипептид, связывающий молекулу нуклеиновой кислоты (например, ДНК-связывающий полипептид), и полипептид дезаминазы, и рибонуклеопротеиновые комплексы, содержащие гибридный белок, включающий РНК-управляемую нуклеазу и полипептид дезаминазы, и рибонуклеиновые кислоты. Предлагаемые композиции также включают молекулы нуклеиновой кислоты, кодирующие полипептиды дезаминазы или гибридные белки, а также векторы и клетки-хозяина, содержащие молекулы нуклеиновой кислоты. Способы, раскрытые в настоящем документе, направлены

на связывание исследуемой последовательности-мишени с исследуемой молекулой ДНК-мишенью и модификацию исследуемой молекулы ДНК-мишени.

Подробное описание настоящего изобретения

Многие модификации и другие варианты осуществления изобретений, изложенные в данном документе, представляются очевидными специалистам в данной области техники, к которой относятся указанные изобретения, с учетом принципов, представленных в предшествующих описаниях. Следовательно, следует понимать, что изобретения не должны ограничиваться конкретными раскрытыми вариантами осуществления и что модификации и другие варианты осуществления предназначены для включения в объем прилагаемой формулы изобретения. Хотя в данном документе используются конкретные термины, они используются только в общем и описательном смысле, а не в целях ограничения.

I. Обзор

В данном раскрытии предложены новые адениндезаминазы и гибридные белки, которые содержат полипептид, связывающийся с молекулой нуклеиновой кислоты, такой как ДНК-связывающий полипептид, и новый полипептид деаминазы. В некоторых вариантах осуществления ДНК-связывающий полипептид представляет собой ДНК-связывающий полипептид, специфичный к последовательности, в том смысле, что ДНК-связывающий полипептид связывается с последовательностью-мишенью с большей частотой, чем связывается с рандомизированной фоновой последовательностью. В некоторых вариантах ДНК-связывающий полипептид представляет собой мегануклеазу, гибридный белок со структурой «цинковые пальцы» или TALEN или получен из них. В некоторых вариантах гибридный белок содержит РНК-управляемый ДНК-связывающий полипептид и полипептид деаминазы. В некоторых вариантах РНК-управляемый ДНК-связывающий полипептид представляет собой РНК-управляемую нуклеазу, такую как полипептидный домен Cas9, который связывается с направляющей РНК (также называемой гРНК (гидовой РНК)), которая, в свою очередь, связывает последовательность-мишень нуклеиновой кислоты, при гибридизации цепей.

Описанные в данном контексте полипептиды деаминазы могут дезаминировать азотистое основание, такое как, например, аденин. Дезаминирование нуклеинового основания деаминазой может привести к точечной мутации в соответствующем остатке, что упоминается в данном

контексте как «редактирование нуклеиновой кислоты» или «редактирование основания». Таким образом, гибридные белки, содержащие полипептид РНК-управляемой нуклеазы (RGN) и деаминазу, можно использовать для направленного редактирования последовательностей нуклеиновых кислот.

5 Такие гибридные белки можно использовать для направленного редактирования ДНК *in vitro*, например, для конструирования генетически модифицированных клеток. Эти генетически модифицированные клетки могут представлять собой клетки растений или клетки животных. Такие гибридные белки также можно использовать для введения направленных мутаций, 10 например, для коррекции генетических дефектов в клетках млекопитающих *ex vivo*, например, в клетках, полученных от субъекта, которые впоследствии повторно вводят тому же или другому субъекту; и для введения направленных мутаций, например, коррекции генетических дефектов или введения дезактивирующих мутаций в ассоциированные с заболеванием гены у субъекта- 15 млекопитающего. Такие гибридные белки также можно использовать для введения направленных мутаций в растительные клетки, например, для введения оказывающих благоприятное воздействие или агрономически ценных признаков или аллелей.

 Термины «белок», «пептид» и «полипептид» используются в данном 20 контексте взаимозаменяемо и относятся к полимеру аминокислотных остатков, связанных друг с другом пептидными (амидными) связями. Термины относятся к белку, пептиду или полипептиду любого размера, структуры или функции. Как правило, длина белка, пептида или полипептида составляет по меньшей мере три аминокислоты. Белок, пептид или полипептид могут относиться к отдельному 25 белку или набору белков. Одна или несколько аминокислот в белке, пептиде или полипептиде могут являться модифицированными, например, при добавлении химической группы, такой как углеводная группа, гидроксильная группа, фосфатная группа, фарнезильная группа, изофарнезильная группа, группа жирной кислоты, линкер для конъюгации, функционализации или другой 30 модификации и т.д. Белок, пептид или полипептид также могут представлять собой одну молекулу или могут представлять собой мультимолекулярный комплекс. Белок, пептид или полипептид может представлять собой просто фрагмент природного белка или пептида. Белок, пептид или полипептид могут

являться природными, рекомбинантными или синтетическими или любой их комбинацией.

Любой из представленных в данном контексте белков можно получать любым способом, известным в данной области техники. Например, представленные в настоящем документе белки можно получить с использованием экспрессии и очистки рекомбинантных белков, что особенно является пригодным для гибридных белков, содержащих пептидный линкер.

Методы экспрессии и очистки рекомбинантных белков хорошо известны и включают методы, описанные в руководстве Green и Sambrook, *Molecular Cloning: A Laboratory Manual* (4-е изд., Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY (2012)), содержание которого включено в полном объеме в данный контекст в качестве ссылки.

II. Дезаминазы

Термин «дезаминаза» относится к ферменту, который катализирует реакцию дезаминирования. Дезаминазы по изобретению представляют собой дезаминазы азотистых оснований, и термины «дезаминаза» и «дезаминаза азотистых оснований» используются в данном контексте взаимозаменяемо. Дезаминаза может представлять собой природный фермент дезаминазу или его активный фрагмент или варианты. Дезаминаза может проявлять активность на одноцепочечных нуклеиновых кислотах, таких как одноцепочечные ДНК или одноцепочечные РНК, или на двухцепочечных нуклеиновых кислотах, таких как двухцепочечные ДНК или двухцепочечные РНК. В некоторых вариантах осуществления настоящего изобретения дезаминаза способна дезаминировать только одноцепочечную ДНК и не действует на двуцепочечную ДНК.

Раскрытые в данном контексте способы и композиции включают адениндезаминазу. В некоторых вариантах дезаминаза представляет собой дезаминазу семейства ADAT или ее вариант. При дезаминировании аденина, аденозина или дезоксиаденозина образуется инозин, который обрабатывается полимеразой как гуанин. На сегодняшний день не известны природные адениндезаминазы, которые дезаминируют аденин в составе ДНК. Было использовано несколько методов для конструирования и оптимизации действия адениндезаминазы на белки тРНК (ADAT), для придания ей активности в отношении молекул ДНК в клетках млекопитающих (см. статьи Gaudelli и др., (2017); Koblan, L.W. и др., *Nat Biotechnol* 36, 843-846 (2018); Richter, M.F. и др.,

Nat Biotechnol (2020), doi:10.1038/s41587-020-0562-8, каждая из которых в полном объеме включена в настоящий документ в качестве ссылки). В одном из таких методов используется анализ бактериальной селекции, при котором выживают только клетки, способные активировать устойчивость к антибиотикам за счет превращения А:Т в G:С.

Настоящее изобретение относится к новым полипептидам адениндезаминазы, которые были получены с использованием эволюции и оптимизации бактериальных дезаминаз. В данном контексте описаны новые адениндезаминазы, представленные как SEQ ID NO: 1-10 и 399-441. Дезаминазы по изобретению можно использовать для редактирования молекул ДНК или РНК. В некоторых вариантах дезаминазы по изобретению можно использовать для редактирования молекул одноцепочечной ДНК или одноцепочечной РНК. Описанные в данном контексте адениндезаминазы можно использовать в качестве дезаминаз в отдельности или в качестве компонентов гибридных белков. Гибридный белок, содержащий полипептид, направленный на ДНК, и полипептид адениндезаминазы, упоминается в данном контексте как «редактор основания А», «редактор оснований аденина» или «ABE» и его можно использовать для направленного редактирования последовательностей нуклеиновых кислот.

«Редакторы оснований» представляют собой гибридные белки, содержащие полипептид, направленный на ДНК, такой как RGN, и дезаминазу. Редакторы адениновых оснований (ABE) включают белок, направленный на ДНК, такой как RGN, и адениндезаминазу. Функция ABE заключается в дезаминировании аденина в инозин на молекуле ДНК-мишени (Gaudelli, N.M. и др. (2017). Инозин распознается полимеразой как гуанин и позволяет включать цитозин в комплементарную цепь ДНК напротив инозина. После цикла репликация-пост-дезаминирование в геноме происходит замена пары оснований А:Т на G:С. В некоторых вариантах описанные в данном контексте адениндезаминазы или их активные варианты или фрагменты вводят мутации с заменой А на N в молекулу ДНК, где N представляет собой C, G или T. В дополнительных вариантах они вводят мутации с заменой А на G в молекулу ДНК.

В тех вариантах, в которых дезаминаза направлена на конкретную область молекулы нуклеиновой кислоты после гибридизации с ДНК-связывающим полипептидом, можно измерить скорость мутации аденинов внутри

последовательности-мишени, с которой связывается ДНК-связывающий полипептид, или рядом с ней, с использованием любого метода, известного в данной области техники, включая полимеразную цепную реакцию (ПЦР), полиморфизм длин рестрикционных фрагментов (RFLP) или секвенирование ДНК.

Описанные в настоящем документе новые дезаминазы или их активные варианты или фрагменты, которые сохраняют дезаминазную активность, можно включить в клетку как часть гибридного ДНК-связывающего полипептида-дезаминазы и/или их можно совместно экспрессировать с гибридным ДНК-связывающим полипептидом-дезаминазой, чтобы повысить эффективность включения требуемой мутации с заменой А на G в молекулу ДНК-мишени.

Описанные в данном контексте дезаминазы характеризуются аминокислотной последовательностью любой из последовательностей SEQ ID NO: 1-10 и 399-441 или ее вариантом или фрагментом, сохраняющими активность дезаминазы. В

некоторых вариантах дезаминаза характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 81%, по меньшей мере на 82%, по меньшей мере на 83%, по меньшей мере на 84%, по меньшей мере на 85%, по меньшей мере на 86%, по меньшей мере на 87%, по меньшей мере на 88%, по меньшей мере на 89%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на

99% по сравнению с любой аминокислотной последовательностью SEQ ID NO: 1-10 и 399-441. В конкретных вариантах дезаминаза содержит аминокислотную последовательность, идентичную по меньшей мере на 80% любой из

последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441. В некоторых вариантах дезаминаза характеризуется идентичностью

аминокислотной последовательности по меньшей мере на 80% по сравнению с SEQ ID NO: 407. Например, дезаминаза содержит аминокислотную последовательность, идентичную по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 96%, по меньшей мере

приблизительно на 97%, по меньшей мере приблизительно на 98%, по меньшей мере на 99%, по меньшей мере приблизительно на 99,5% или по меньшей мере приблизительно на 99,9% по сравнению с последовательностью SEQ ID NO: 407.

В некоторых вариантах дезаминаза содержит аминокислотную

5 последовательность, которая характеризуется идентичностью по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99%, по меньшей мере на 99,5% или по меньшей мере на 99,9% по сравнению с последовательностью SEQ ID NO: 407. В некоторых вариантах дезаминаза
10 содержит аминокислотную последовательность SEQ ID NO: 407. В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая характеризуется идентичностью по меньшей мере на 80% по сравнению с последовательностью SEQ ID NO: 399. Например, дезаминаза содержит аминокислотную последовательность, которая характеризуется идентичностью
15 по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 96%, по меньшей мере приблизительно на 97%, по меньшей мере приблизительно на 98%, по меньшей мере приблизительно на 99%, по меньшей мере приблизительно на 99,5% или по меньшей мере приблизительно на 99,9% по сравнению с последовательностью SEQ ID NO: 399.
20

В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая характеризуется идентичностью по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на
25 99%, по меньшей мере на 99,5% или по меньшей мере на 99,9% по сравнению с последовательностью SEQ ID NO: 399. В некоторых вариантах дезаминаза содержит аминокислотную последовательность SEQ ID NO: 399. В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая характеризуется идентичностью по меньшей мере на 80% по сравнению с последовательностью SEQ ID NO: 405. Например, дезаминаза содержит
30 аминокислотную последовательность, которая характеризуется идентичностью по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 96%, по меньшей мере приблизительно на 97%, по меньшей мере

мере приблизительно на 98%, по меньшей мере приблизительно на 99%, по
меньшей мере приблизительно на 99,5% или по меньшей мере приблизительно
на 99,9% по сравнению с последовательностью SEQ ID NO: 405. В некоторых
вариантах дезаминаза содержит аминокислотную последовательность, которая
5 характеризуется идентичностью по меньшей мере на 80%, по меньшей мере на
90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на
97%, по меньшей мере на 98%, по меньшей мере на 99%, по меньшей мере на
99,5% или по меньшей мере на 99,9% по сравнению с последовательностью SEQ
ID NO: 405. В некоторых вариантах дезаминаза содержит аминокислотную
10 последовательность SEQ ID NO: 405.

III. Полипептиды, связывающие молекулу нуклеиновой кислоты

Некоторые объекты настоящего изобретения относятся к гибридным
белкам, которые содержат полипептид, связывающий молекулу нуклеиновой
кислоты, и полипептид дезаминазы. Хотя связывание и направленное
15 редактирование молекул РНК предусмотрено настоящим изобретением, в
некоторых вариантах полипептид гибридного белка, связывающий молекулу
нуклеиновой кислоты, представляет собой ДНК-связывающий полипептид.
Такие гибридные белки пригодны для направленного редактирования ДНК в
условиях *in vitro*, *ex vivo* или *in vivo*. Эти новые гибридные белки проявляют
20 активность в клетках млекопитающих и пригодны для направленного
редактирования молекул ДНК.

Термин «гибридный белок», используемый в данном контексте, относится к
гибридному полипептиду, который содержит белковые домены по меньшей мере
из двух разных белков. Гибридный белок может содержать более одного из
25 различных доменов, например, ДНК-связывающий домен и дезаминазу. В
некоторых вариантах гибридный белок находится в составе комплекса или
связан с нуклеиновой кислотой, например, РНК.

В некоторых вариантах раскрытые в настоящем документе гибридные
белки содержат ДНК-связывающий полипептид. Используемый в данном
30 контексте термин «ДНК-связывающий полипептид» относится к любому
полипептиду, способному связываться с ДНК. В некоторых вариантах ДНК-
связывающая полипептидная часть раскрываемых в данном контексте
гибридных белков связывается с двухцепочечной ДНК. В конкретных вариантах
ДНК-связывающий полипептид связывается с ДНК по специфичному к

последовательности механизму. Используемые в данном контексте термины «специфичный к последовательности» или «специфичный к последовательности механизм» относятся к селективному взаимодействию с конкретной последовательностью нуклеотидов.

5 Две полинуклеотидные последовательности можно считать в основном комплементарными, когда две последовательности гибридизируются друг с другом в жестких условиях. Аналогичным образом считается, что ДНК-связывающий полипептид связывается с конкретной последовательностью-мишенью специфичным к последовательности способом, если ДНК-связывающий полипептид связывается с последовательностью ДНК в жестких
10 условиях. Термины «жесткие условия» или «жесткие условия гибридизации» означают условия, при которых две полинуклеотидные последовательности (или полипептид связывается со своей конкретной последовательностью-мишенью) будут связываться друг с другом в значительно большей детектируемой степени, чем с другими последовательностями (например, по меньшей мере, в 2 раза
15 больше фона). Жесткие условия зависят от последовательности и будут отличаться в различных обстоятельствах. Как правило, жесткие условия означают среду, в которой концентрация соли составляет менее 1,5 М ионов Na, обычно приблизительно от 0,01 до 1,0 М концентрации ионов Na (или других солей) при pH от 7,0 до 8,3, а температура составляет меньшей мере 30°C для коротких последовательностей (например, от 10 до 50 нуклеотидов) и по меньшей мере 60°C для длинных последовательностей (например, более 50 нуклеотидов). Жесткие условия также можно обеспечить при добавлении дестабилизирующих агентов, таких как формамид. Типичные условия
20 пониженной жесткости включают гибридизацию в буферном растворе следующего состава: 30-35% формамида, 1 М NaCl, 1% ДСН (додecilсульфат натрия) при 37°C и промывку в растворе от 1X до 2X SSC (20X SSC = 3,0 М NaCl/0,3 М цитрата натрия) при температуре от 50 до 55°C. Типичные условия умеренной жесткости включают гибридизацию в буферном растворе
25 следующего состава: от 40 до 45% формамида, 1,0 М NaCl, 1% ДСН при 37°C и промывку в буферном растворе от 0,5X до 1X SSC при 55-60°C. Типичные условия высокой жесткости включают гибридизацию в буферном растворе следующего состава: 50% формамид, 1 М NaCl, 1% ДСН при 37°C и промывку в растворе 0,1X SSC при 60-65°C. Необязательно, буферные растворы для

промывки могут содержать от приблизительно 0,1% до приблизительно 1% ДСН. Продолжительность гибридизации обычно составляет приблизительно менее 24 ч, обычно от приблизительно 4 до приблизительно 12 ч. Продолжительность промывки должна составлять период, в который по меньшей мере является достаточным для достижения равновесия.

T_m означает температуру (при определенной ионной силе и pH), при которой 50% комплементарной последовательности-мишени гибридизуется с идеально совпадающей последовательностью. Для гибридов ДНК-ДНК величину T_m можно аппроксимировать по уравнению Майнкота и Вейля (Meinkoth and Wahl) Anal. Biochim. 138:267-284 (1984): $T_m = 81,5^{\circ}\text{C} + 16,6 (\log M) + 0,41 (\% \text{ГЦ}) - 0,61 (\% \text{форм}) - 500/L$; где M - молярность одновалентных катионов, %ГЦ - процентное содержание гуанозиновых и цитозиновых нуклеотидов в ДНК, % форм - процентное содержание формамида в растворе для гибридизации, а L - длина гибрида (в парах оснований). Как правило, жесткие условия выбирают таким образом, чтобы они были приблизительно на 5°C ниже температуры плавления (T_m) конкретной последовательности и ее комплементарной последовательности при определенной ионной силе и pH. Однако в очень жестких условиях можно использовать гибридизацию и/или промывку при температуре на 1, 2, 3 или 4°C ниже температуры плавления (T_m); в умеренно жестких условиях можно использовать гибридизацию и/или промывку при температуре на 6, 7, 8, 9 или 10°C ниже температуры плавления (T_m); в условиях низкой жесткости можно использовать гибридизацию и/или промывку при температуре на 11, 12, 13, 14, 15 или 20°C ниже температуры плавления (T_m). При использовании уравнения, композиции для гибридизации и промывки, а также требуемой T_m , специалистам в данной области техники представляется очевидным, что вариации жесткости растворов для гибридизации и/или промывки в основном описаны. Подробное руководство по гибридизации нуклеиновых кислот можно найти в руководствах Tijssen Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, New York) (1993); и Ausubel и др., ред. Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York) (1995). См. Sambrook и др., Molecular Cloning: A Laboratory Manual (2е ред., Cold Spring Harbor Laboratory Press, Plainview, New York (1989)).

В некоторых вариантах специфичный к последовательности ДНК-связывающий полипептид представляет собой РНК-управляемый ДНК-связывающий полипептид (RGDBP). Используемые в данном контексте термины «РНК-управляемый ДНК-связывающий полипептид» и «RGDBP» относятся к полипептидам, способным связываться с ДНК в ходе гибридизации ассоциированной молекулы РНК с последовательностью ДНК-мишени.

В некоторых вариантах ДНК-связывающий полипептид гибридного белка представляет собой нуклеазу, такую как специфичная к последовательности нуклеаза. Используемый в данном контексте термин «нуклеаза» относится к ферменту, который катализирует расщепление фосфодиэфирных связей между нуклеотидами в молекуле нуклеиновой кислоты. В некоторых вариантах ДНК-связывающий полипептид представляет собой эндонуклеазу, которая способна расщеплять фосфодиэфирные связи между нуклеотидами в молекуле нуклеиновой кислоты, тогда как в некоторых вариантах ДНК-связывающий полипептид представляет собой экзонуклеазу, которая способна расщеплять нуклеотиды либо на 5' конце либо 3' конце молекулы нуклеиновой кислоты. В некоторых вариантах специфичная к последовательности нуклеаза выбрана из группы, состоящей из мегануклеазы, нуклеазы цинковых пальцев, гибридного белка TAL-эффекторного ДНК-связывающего домена-нуклеазы (TALEN) и РНК-управляемой нуклеазы (RGN) или их вариантов, при этом нуклеазная активность является сниженной или ингибированной.

Используемый в данном контексте термин «мегануклеаза» или «хоуминг-эндонуклеаза» относится к эндонуклеазам, которые связываются с сайтом распознавания в двухцепочечной ДНК длиной от 12 до 40 п.о.

Неограничивающими примерами мегануклеаз являются мегануклеазы, которые принадлежат к семейству LAGLIDADG и содержат консервативный аминокислотный мотив LAGLIDADG (SEQ ID NO: 49). Термин «мегануклеаза» может относиться к димерной или одноцепочечной мегануклеазе.

Используемый в данном контексте термин «нуклеаза цинковых пальцев» или «ZFN» относится к химерному белку, содержащему ДНК-связывающий домен цинковых пальцев и нуклеазный домен.

Используемый в данном контексте термин «гибридный белок TAL-эффекторный ДНК-связывающий домен-нуклеаза» или «TALEN» относится к

химерному белку, содержащему эффекторный ДНК-связывающий домен TAL и нуклеазный домен.

Используемый в данном контексте термин «РНК-управляемая нуклеаза» или «RGN» относится к РНК-управляемому ДНК-связывающему полипептиду, который обладает нуклеазной активностью. RGN рассматриваются как «РНК-управляемые», поскольку направляющие РНК образуют комплекс с РНК-управляемыми нуклеазами, направляя РНК-управляемую нуклеазу на связывание с последовательностью-мишенью, а в некоторых вариантах вводят одноцепочечный или двухцепочечный разрыв в последовательности-мишени.

RGN может представлять собой CasX, CasY, C2c1, C2c2, C2c3, GeoCas9, SpCas9, SaCas9, Nme2Cas9, CjCas9, Cas12a (ранее известную как Cpf1), Cas12b, Cas12g, Cas12h, Cas12i, aLbCas12a, AsCas12a, CasMINI, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с кольцевой перестановкой, Argonaute (Ago), SmacCas9, или Spy-macCas9, домен Spy-macCas9, или RGN с аминокислотной последовательностью, любой из указанных SEQ ID NO: 41, 60, 366 или 368. В некоторых вариантах, как указано ниже, нуклеазы RGN, описанные в данном контексте, представляют собой нуклеазы RGN.

Согласно настоящему изобретению белок RGN, который был мутирован, для превращения в нуклеазо-неактивный или «мертвый», такой как, например, dCas9, можно назвать РНК-управляемым ДНК-связывающим полипептидом или нуклеазо-неактивным RGN или нуклеазо-мертвым RGN. Кроме того, можно определить пригодные нуклеазо-неактивные домены Cas9 других известных РНК-управляемых нуклеаз (RGN) (например, нуклеазо-неактивный вариант RGN APG08290.1, раскрытый в публикации патента США № 2019/0367949), содержание которого в полном объеме включено в настоящий документ в качестве ссылки).

В некоторых вариантах гибридный белок содержит RGN, гибридизированную с описанной в данном контексте дезаминазой. В вариантах осуществления гибридных белков, описанных выше, дезаминаза выбрана из дезаминаз, содержащих аминокислотную последовательность, по меньшей мере на 80% идентичную любой из последовательностей SEQ ID NO: 1-10 и 399-441. В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая по меньшей мере на 80% идентична

последовательности SEQ ID NO: 407. В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 399. В некоторых вариантах дезаминаза содержит аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 405. В некоторых вариантах RGN выбрана из CasX, CasY, C2cl, C2c2, C2c3, GeoCas9, SpCas9, SaCas9, Nme2Cas9, CjCas9, Casl2a (ранее известную как Cpfl), Cas12b, Cas12g, Cas12h, Cas12i, aLbCas12a, AsCas12a, CasMINI, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с кольцевой перестановкой, Argonaute (Ago), SmacCas9, Spy-macCas9, домена Spy-macCas9, или RGN с аминокислотной последовательностью, любой из указанных SEQ ID NO: 41, 60, 366 или 368. В конкретных вариантах гибридный белок содержит нуклеазу Cas9, гибридизированную с дезаминазой, содержащей аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 407. В некоторых вариантах гибридный белок содержит нуклеазу Cas9, гибридизированную с дезаминазой, содержащей аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 399. В конкретных вариантах гибридный белок содержит нуклеазу Cas9, гибридизированную с дезаминазой, содержащей аминокислотную последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 405. Нуклеаза Cas9 может представлять собой любую нуклеазу Cas9, раскрытую в публикации патента РСТ № WO 2020181195, содержание которого в полном объеме включено в настоящий документ в качестве ссылки.

Термин «полипептид RGN» включает полипептиды RGN, которые расщепляют только одну цепь нуклеотидной последовательности-мишени, которая в данном контексте упоминается как нуклеаза. Такие RGN включают один функциональный нуклеазный домен. Нуклеазы RGN могут являться природными нуклеазами или белками RGN, которые естественным образом расщепляют обе цепи двухцепочечной молекулы нуклеиновой кислоты, которые включали мутации в одном или нескольких доменах нуклеазы, так что нуклеазная активность этих мутантных доменов снижена или устранена, и таким образом они превращаются в нуклеазы. В некоторых вариантах RGN нуклеазы гибридного белка содержат мутацию (например, мутацию D10A), которая придает RGN

способность расщеплять только цепь-мишень, не отредактированную основаниями (цепь, которая содержит РАМ и представляет собой спаренную с гидовой РНК (гРНК)) дуплекса нуклеиновой кислоты. Эта мутация D10A изменяет первый остаток аспарагиновой кислоты в расщепленном нуклеазном домене RuvC RGN. В настоящей заявке раскрыто несколько вариантов никазы D10A или гомологичных вариантов никазы описанных RGN (см. пример 4). Варианты nAPG07433.1 и nAPG08290.1 (обозначенные как SEQ ID NO: 42 и 61 соответственно) представляют собой варианты никазы APG07433.1 и APG08290.1, которые представлены как SEQ ID NO: 41 и 60 соответственно, и описаны в заявке WO 2019/236566 (в полном объеме включена в настоящий документ в качестве ссылки). Варианты nAPG00969 (указан как SEQ ID NO: 52) и nAPG09748 (указан как SEQ ID NO: 54) представляют собой варианты никазы APG00969 и APG09748, соответственно, которые описаны в заявке WO 2020/139783 (полностью включена в настоящий документ в качестве ссылки). Варианты nAPG06646 (обозначен как SEQ ID NO: 53) и nAPG09882 (обозначен как SEQ ID NO: 55) представляют собой варианты никазы APG06646 и APG09882, соответственно, которые описаны в публикации PCT WO 2021/030344 (включенной в полном объеме в качестве ссылки). Варианты nAPG03850, nAPG07553, nAPG055886 и nAPG01604 представлены как SEQ ID NO: 56-59, соответственно, и представляют собой варианты никазы APG03850, APG07553, APG055886 и APG01604, которые описаны в находящейся на рассмотрении заявке PCT № PCT/US2021/028843 (в полном объеме включена в настоящий документ в качестве ссылки). Различные никазы RGN, их варианты и их последовательности раскрыты в публикации патента PCT № WO2020181195, содержание которого в полном объеме включено в настоящее описание в качестве ссылки. Одним типичным пригодным нуклеазо-неактивным Cas9 является мутант D10A/H840A Cas9 (см., например, Qi и др., Cell. 152(5): 1173-1183 (2013), содержание которого в полном объеме включено в настоящее описание в качестве ссылки).

В некоторых вариантах никаза RGN гибридного белка содержит мутацию (например, мутацию H840A), которая придает RGN способность расщеплять только нецелевую цепь с отредактированным основанием (цепь, которая не содержит РАМ и является не спаренной с гРНК) дуплекса нуклеиновой кислоты. Мутация H840A изменяет первый гистидин нуклеазного домена HNH. Никаза

RGN, содержащая мутацию H840A или эквивалентную мутацию, включает инактивированный домен HNH. Никаза RGN с мутацией H840A расщепляет нецелевую цепь. Никаза, содержащая мутацию D10A или эквивалентную мутацию, включает инактивированный домен нуклеазы RuvC и расщепляет цепь-мишень. Никазы D10A не способны расщеплять нецелевую цепь ДНК, то есть цепь, в которой требуется редактирование оснований.

Другие дополнительные типичные пригодные неактивные нуклеазные домены Cas9 включают, но не ограничиваясь только этим, мутантные домены D10A/D839A/H840A и D10A/D839A/H840A/N863A (см., например, статью Mali и др., Nature Biotechnology, 31(9): 833-838 (2013), содержание которой в полном объеме включено в настоящий документ в качестве ссылки). Дополнительные пригодные белки RGN, мутированные в никазы, представляются очевидными специалистам в данной области техники на основании этого описания и знаний в данной области (например, RGN, раскрытые в публикациях РСТ №№ WO 2019/236566, WO2020181195, которые в данном контексте в полном объеме включены в качестве ссылки) и включены в объем настоящего изобретения. В предпочтительных вариантах RGN, обладающая никазной активностью в отношении цепи-мишени, разрывает цепь-мишень, в то время как комплементарная цепь, нецелевая цепь модифицируется дезаминазой. Механизм репарации клеточной ДНК может восстанавливать разорванную цепь-мишень, с использованием модифицированной нецелевой цепи в качестве матрицы, тем самым в ДНК вводится мутация.

В некоторых вариантах никаза RGN, сохраняющая никазную активность, содержит аминокислотную последовательность, которая идентична по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90% по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или по меньшей мере на 99,5% по сравнению с последовательностью SEQ ID NO: 42 или любой из последовательностей SEQ ID NO: 52-59, 61, 397 и 398.

Для получения никаза или RGN без нуклеаз можно использовать любой способ, известный в данной области техники для введения мутаций в аминокислотную последовательность, такой как мутагенез, опосредованный ПЦР, и сайт-направленный мутагенез. См., например, публикацию патента США

№ 2014/0068797 и патент № 9790490, каждый из которых в полном объеме включен в настоящий документ в качестве ссылки. РНК-управляемые нуклеазы (RGN) позволяют целенаправленно манипулировать одним сайтом в геноме и их можно использовать при получении генов-мишеней в терапевтических и исследовательских целях. У различных организмов, включая млекопитающих, РНК-управляемые нуклеазы используются для конструирования генома с использованием стимуляции либо негомологичного соединения концов, либо гомологичной рекомбинации. RGN включают белки CRISPR-Cas, которые представляют собой РНК-управляемые нуклеазы, направленные на последовательность-мишень с помощью гидовой РНК (гРНК), как часть системы РНК-управляемых нуклеаз с регулярно расположенными интервалами коротких палиндромных повторов (CRISPR), или их активные варианты или фрагменты.

Кроме того, в настоящем документе представлены полипептиды RGN (и молекулы нуклеиновых кислот, кодирующие полипептиды RGN), которые содержат аминокислотную последовательность, указанную как SEQ ID NO: 41 или 60, но лишены аминокислотных остатков 590-597 последовательности SEQ ID NO: 41 или 60, или их активный вариант или его фрагмент. В некоторых вариантах полипептид RGN содержит аминокислотную последовательность, представленную как SEQ ID NO: 366, 368, 397 или 398, или его активный вариант или фрагмент.

Некоторые объекты настоящего изобретения относятся к гибридным белкам, которые содержат РНК-управляемый ДНК-связывающий полипептид и полипептид дезаминазы, в конкретном варианте полипептид адениндезаминазы. В некоторых вариантах РНК-управляемый ДНК-связывающий полипептид представляет собой РНК-управляемую нуклеазу. В дополнительных вариантах РНК-управляемая нуклеаза представляет собой природный белок CRISPR-Cas или его активный вариант или фрагмент. Системы CRISPR-Cas относятся к системам класса 1 или класса 2. Системы класса 2 содержат одну эффекторную нуклеазу и включают типы II, V и VI. Системы классов 1 и 2 подразделяются на типы (типы I, II, III, IV, V, VI), а некоторые типы далее подразделяются на подтипы (например, тип II-A, тип II-B, тип II-C, Тип В-A, Тип В-B).

В некоторых вариантах белок CRISPR-Cas представляет собой природный белок CRISPR-Cas типа II или его активный вариант или фрагмент.

Используемый в данном документе термин «белок CRISPR-Cas типа II»,

«эффекторный белок CRISPR-Cas типа II» или «Cas9» относится к
эффекторному белку CRISPR-Cas, для которого требуется наличие
трансактивирующей РНК (tracrRNA) и содержит два нуклеазных домена
(например, RuvC и HNH), каждый из которых отвечает за расщепление одной
5 цепи двухцепочечной молекулы ДНК. В некоторых вариантах настоящего
изобретения предлагается гибридный белок, содержащий раскрытую в данном
контексте дезаминазу, гибридизованную с нуклеотидной последовательностью Streptococcus pyogenes Cas9
(SpCas9), или нуклеотидной последовательности которых представлены в виде
SEQ ID NO: 555 и 556, соответственно, и описаны в патентах США 10000772 и
10 8697359, каждый из которых в полном объеме включен в настоящее описание в
качестве ссылки. В некоторых вариантах в настоящем изобретении предлагается
гибридный белок, содержащий раскрытую в данном контексте дезаминазу,
гибридизованную с нуклеотидной последовательностью Streptococcus thermophilus Cas9 (StCas9), или нуклеотидной
StCas9, последовательности которых представлены как SEQ ID NO: 557 и 558,
15 соответственно, и раскрыты в патенте США № 10113167, который в полном
объеме включен в настоящий документ в качестве ссылки. В некоторых
вариантах в настоящем изобретении предлагается гибридный белок, содержащий
раскрытую в данном контексте дезаминазу, гибридизованную с нуклеотидной
Streptococcus aureus Cas9 (SaCas9), или с нуклеотидной SaCas9, последовательности
20 которых представлены как SEQ ID NO: 559 и 560, соответственно, и раскрыты в
патенте США № 9752132, который в полном объеме включен в настоящий
документ в качестве ссылки.

В некоторых вариантах белок CRISPR-Cas представляет собой природный
белок CRISPR-Cas типа V или его активный вариант или фрагмент.
25 Используемый в данном контексте термин «белок CRISPR-Cas типа V»,
«эффекторный белок CRISPR-Cas типа V» или «Cas12» относится к
эффекторному белку CRISPR-Cas, который расщепляет двухцепочечную ДНК и
содержит один нуклеазный домен RuvC или расщепленный нуклеазный домен
RuvC и лишен домена HNH (см. статьи Zetsche и др., Cell (2015)
30 doi:10.1016/j.cell.2015.09.038; Shmakov и др., Nat Rev Microbiol (2017)
,doi:10.1038/nrmicro.2016.184; Yan и др., Science (2018)
doi:10.1126/science.aav7271; Harrington и др., Science (2018)
doi:10.1126/science.aav4294). Следует отметить, что Cas12a также обозначается
как Cpf1 и не требует tracrRNA, хотя другие белки CRISPR-Cas типа V, такие

как Cas12b, требуют наличия tracrRNA. Большинство эффекторов типа V также могут быть направлены на оцДНК (одноцепочечную ДНК), часто без требования PAM (см. статьи Zetsche и др. (2015); Yan и др. (2018); Harrington и др. (2018)).

Термин «белок CRISPR-Cas типа V» включает уникальные RGN, содержащие расщепленные домены нуклеазы RuvC, такие как домены, раскрытые в предварительных заявках №№ 62/955014, поданной 30 декабря 2019 г., и 63/058169, поданной 29 июля 2020 г., и в международной заявке № PCT/US2020/067138, поданной 28 декабря 2020 г., содержание каждой из которых в полном объеме включено в настоящий документ в качестве ссылки. В некоторых вариантах настоящее изобретение относится к гибриднему белку, содержащему раскрытую в данном контексте дезаминазу, гибридизованную с нуклеотидной последовательностью Francisella novicida Cas12a (FnCas12a), последовательность которой представлена в виде SEQ ID NO: 561 и раскрыта в патенте США No. № 9790490, который в полном объеме включен в настоящий документ в качестве ссылки, или с любым из инактивирующих нуклеазу мутантов FnCas12a, раскрытых в патенте США No. № 9790490.

В некоторых вариантах белок CRISPR-Cas представляет собой природный белок CRISPR-Cas типа VI или его активный вариант или фрагмент. Используемый в данном контексте термин «белок CRISPR-Cas типа VI», «эффекторный белок CRISPR-Cas типа VI» или «Cas13» относится к эффекторному белку CRISPR-Cas, для которого не требуется tracrRNA и содержит два домена HEPN, которые расщепляют РНК.

Термин «гидовая РНК (гРНК)» относится к нуклеотидной последовательности, которая в достаточной степени комплементарна с нуклеотидной последовательностью-мишенью для гибридизации с последовательностью-мишенью и прямого специфичного для последовательности связывания RGN, ассоциированной с нуклеотидной последовательностью-мишенью. Для нуклеаз RGN CRISPR-Cas соответствующая гидовая РНК представляет собой одну или более молекул РНК (как правило, одну или две), которые могут связываться с RGN и направлять RGN для связывания с конкретной нуклеотидной последовательностью-мишенью, и в тех случаях, когда RGN обладает нуклеазной активностью, а также расщепляет нуклеотидную последовательность-мишень.

Гидовая РНК содержит РНК CRISPR (crРНК) и, в некоторых вариантах, транскрибирующую РНК CRISPR (tracrРНК).

РНК CRISPR содержит спейсерную последовательность и повторяющуюся последовательность CRISPR. «Спейсерная последовательность» представляет собой нуклеотидную последовательность, которая напрямую гибридизуется с исследуемой нуклеотидной последовательностью-мишенью. Спейсерную последовательность конструируют таким образом, чтобы она была полностью или частично комплементарной исследуемой последовательности-мишени. В различных вариантах спейсерная последовательность содержит от

5

10

15

20

25

30

приблизительно 8 нуклеотидов до приблизительно 30 нуклеотидов или более. Например, длина спейсерной последовательности может составлять приблизительно 8, приблизительно 9, приблизительно 10, приблизительно 11, приблизительно 12, приблизительно 13, приблизительно 14, приблизительно 15, приблизительно 16, приблизительно 17, приблизительно 18, приблизительно 19, приблизительно 20, приблизительно 21, приблизительно 22, приблизительно 23, приблизительно 24, приблизительно 25, приблизительно 26, приблизительно 27, приблизительно 28, приблизительно 29, приблизительно 30 или более нуклеотидов. В некоторых вариантах длина спейсерной последовательности составляет 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 и более нуклеотидов. В некоторых вариантах длина спейсерной последовательности составляет от приблизительно 10 до приблизительно 26 нуклеотидов или от приблизительно 12 до приблизительно 30 нуклеотидов. В некоторых вариантах длина спейсерной последовательности составляет от 10 до 26 нуклеотидов или от 12 до 30 нуклеотидов. В конкретных вариантах осуществления длина спейсерной последовательности составляет приблизительно 30 нуклеотидов. В конкретных вариантах длина спейсерной последовательности составляет 30 нуклеотидов. В некоторых вариантах степень комплементарности между спейсерной последовательностью и ее соответствующей последовательностью-мишенью, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания составляет от 50% до 99% или более, включая, но не ограничиваясь только этим, приблизительно или более приблизительно 50%, приблизительно 60%, приблизительно 70%, приблизительно 75%, приблизительно 80%, приблизительно 81%, приблизительно 82%, приблизительно 83%,

приблизительно 84%, приблизительно 85%, приблизительно 86%,
приблизительно 87%, приблизительно 88%, приблизительно 89 %,
приблизительно 90%, приблизительно 91%, приблизительно 92%,
приблизительно 93%, приблизительно 94%, приблизительно 95%,
5 приблизительно 96%, приблизительно 97%, приблизительно 98%,
приблизительно 99% или более. В конкретных вариантах степень
комплементарности между спейсерной последовательностью и соответствующей
ей последовательностью-мишенью, при оптимальном выравнивании с
использованием пригодного алгоритма выравнивания составляет 50%, 60%, 70%,
10 75%, 80%, 81%, 82%, 83 %, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%,
93%, 94%, 95%, 96%, 97%, 98%, 99%, или более. В конкретных вариантах
спейсерная последовательность не образует вторичной структуры, которую
можно предсказать с использованием любого пригодного алгоритма
сворачивания полинуклеотидов, известного в данной области техники, включая,
15 но не ограничиваясь только этим, mFold (см., например, статьи Zuker и Stiegler
Nucleic Acids Res. 9:133-148 (1981) и PНК-fold (см., например, Gruber и др., Cell
106(1):23-24 (2008)).

Повторяющаяся последовательность ПНК CRISPR содержит
нуклеотидную последовательность, которая образует структуру либо сама по
20 себе, либо вместе с гибридизированной tracrПНК, которая распознается
молекулой RGN. В различных вариантах повторяющаяся последовательность
ПНК CRISPR содержит от приблизительно 8 нуклеотидов до приблизительно 30
нуклеотидов или более. В конкретных вариантах повторяющаяся
последовательность CRISPR ПНК содержит от 8 до 30 нуклеотидов или более.
25 Например, длина повторяющейся последовательности CRISPR может составлять
приблизительно 8, приблизительно 9, приблизительно 10, приблизительно 11,
приблизительно 12, приблизительно 13, приблизительно 14, приблизительно 15,
приблизительно 16, приблизительно 17, приблизительно 18, приблизительно 19,
приблизительно 20, приблизительно 21, приблизительно 22, приблизительно 23,
30 приблизительно 24, приблизительно 25, приблизительно 26, приблизительно 27,
приблизительно 28, приблизительно 29, приблизительно 30 или более
нуклеотидов. В конкретных вариантах длина повторяющейся
последовательности CRISPR составляет 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,
20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 и более нуклеотидов. В конкретных

вариантах степень комплементарности между повторяющейся последовательностью CRISPR и соответствующей ей последовательностью tracrPHK, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания, составляет от 50% до 99% или более, включая, но без

5 ограничения перечисленным, приблизительно 50%, приблизительно 60%, приблизительно 70%, приблизительно 75%, приблизительно 80%, приблизительно 81%, приблизительно 82%, приблизительно 83%, приблизительно 84%, приблизительно 85%, приблизительно 86%, приблизительно 87%, приблизительно 88%, приблизительно 89%,

10 приблизительно 90%, приблизительно 91%, приблизительно 92%, приблизительно 93%, приблизительно 94%, приблизительно 95%, приблизительно 96%, приблизительно 97%, приблизительно 98%, приблизительно 99%, или более. В конкретных вариантах степень комплементарности между повторяющейся последовательностью CRISPR и

15 соответствующей ей последовательностью tracrPHK при оптимальном выравнивании с использованием пригодного 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 гидовая РНК дополнительно содержит молекулу

20 tracrPHK. Трансактивирующая молекула РНК CRISPR или tracrRNA содержит нуклеотидную последовательность, включающую область, которая характеризуется достаточной комплементарностью для гибридизации с повторяющейся последовательностью CRISPR сгРНК, которая упоминается в данном контексте как область антиповтора. В некоторых вариантах молекула

25 tracrPHK дополнительно содержит область со вторичной структурой (например, тяж-петля) или образует вторичную структуру при гибридизации с соответствующей сгРНК. В конкретных вариантах область tracrPHK, которая является полностью или частично комплементарной повторяющейся последовательности CRISPR, находится на 5'-конце молекулы, а 3'-конец

30 tracrPHK содержит вторичную структуру. Эта область вторичной структуры обычно включает несколько структур «шпилек», в том числе некусную (связующую) шпильку, которая находится рядом с последовательностью антиповтор. В большинстве случаев на 3'-конце tracrPHK расположены концевые шпильки, которые могут различаться по структуре и количеству, но в

большинстве случаев содержат ГЦ-обогащенную Rho-независимую шпильку терминатора транскрипции, за которой следует цепочка оснований U на 3'-конце. См., например, статьи Briner и др., *Molecular Cell* 56:333-339 (2014), Briner и Barrangou, *Cold Spring Harb Protoc* (2016); doi: 10.1101/pdb.top090902 и публикацию заявки США № 2017/0275648, каждая из которых в полном объеме включена в настоящий документ в качестве ссылки.

В различных вариантах область антиповтор в составе *tracrRNA*, которая является полностью или частично комплементарной повторяющейся последовательности *CRISPR*, содержит от приблизительно 6 нуклеотидов до приблизительно 30 нуклеотидов или более. Например, область спаривания оснований между последовательностью антиповторов *tracrRNA* и повторяющейся последовательностью *CRISPR* может составлять приблизительно 6, приблизительно 7, приблизительно 8, приблизительно 9, приблизительно 10, приблизительно 11, приблизительно 12, приблизительно 13, приблизительно 14, приблизительно 15, приблизительно 16, приблизительно 17, приблизительно 18, приблизительно 19, приблизительно 20, приблизительно 21, приблизительно 22, приблизительно 23, приблизительно 24, приблизительно 25, приблизительно 26, приблизительно 27, приблизительно 28, приблизительно 29, приблизительно 30 или более нуклеотидов. В конкретных вариантах область спаривания оснований между последовательностью антиповторов *tracrRNA* и повторяющейся последовательностью *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 и более нуклеотидов. В конкретных вариантах длина области антиповтор *tracrRNA*, которая является полностью или частично комплементарной повторяющейся последовательности *CRISPR*, составляет приблизительно 10 нуклеотидов. В конкретных вариантах длина области антиповтор *tracrRNA*, является полностью или частично комплементарной повторяющейся последовательности *CRISPR*, составляет 10 нуклеотидов. В некоторых вариантах степень комплементарности между повторяющейся последовательностью *CRISPR* и соответствующей последовательностью антиповтор *tracrRNA*, при оптимальном выравнивании с использованием пригодного алгоритма выравнивания, составляет от 50% до 99% или более, включая, но не ограничиваясь только этим, приблизительно или более приблизительно 50%, приблизительно 60%, приблизительно 70%, приблизительно 75%, приблизительно 80%, приблизительно 81%,

приблизительно 82%, приблизительно 83%, приблизительно 84%,
приблизительно 85%, приблизительно 86%, приблизительно 87%,
приблизительно 88%, приблизительно 89%, приблизительно 90%,
приблизительно 91%, приблизительно 92%, приблизительно 93%,
5 приблизительно 94%, приблизительно 95%, приблизительно 96%,
приблизительно 97%, приблизительно 98%, примерно 99% или более. В
конкретных вариантах степень комплементарности между повторяющейся
последовательностью CRISPR и соответствующей ей последовательностью
антиповтор tracrPНК, при оптимальном выравнивании с использованием
10 пригодного 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% и более.

В различных вариантах полноразмерная tracrPНК содержит от
приблизительно 60 нуклеотидов до более приблизительно 210 нуклеотидов. В
15 конкретных вариантах полноразмерная tracrPНК содержит от 60 нуклеотидов до
более 210 нуклеотидов. Например, длина tracrPНК может составлять
приблизительно 60, приблизительно 65, приблизительно 70, приблизительно 75,
приблизительно 80, приблизительно 85, приблизительно 90, приблизительно 95,
приблизительно 100, приблизительно 105, приблизительно 110, приблизительно
20 115, приблизительно 120, приблизительно 125, приблизительно 130,
приблизительно 135, приблизительно 140, приблизительно 150, приблизительно
160, приблизительно 170, приблизительно 180, приблизительно 190,
приблизительно 200, приблизительно 210 или более нуклеотидов. В конкретных
вариантах длина tracrPНК составляет 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110,
25 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 190, 200, 210 и более
нуклеотидов. В конкретных вариантах длина tracrPНК составляет от
приблизительно 100 до приблизительно 210 нуклеотидов, включая
приблизительно 95, приблизительно 96, приблизительно 97, приблизительно 98,
приблизительно 99, приблизительно 100, приблизительно 105, приблизительно
30 106, приблизительно 107, приблизительно 108, приблизительно 109, и
приблизительно 100 нуклеотидов. В конкретных вариантах длина tracrPНК
составляет от 100 до 110 нуклеотидов, включая 95, 96, 97, 98, 99, 100, 105, 106,
107, 108, 109 и 110 нуклеотидов.

Гидовые (направляющие) РНК образуют комплекс с РНК-управляемым ДНК-связывающим полипептидом или РНК-управляемой нуклеазой, для направления РНК-управляемой нуклеазы на связывание с последовательностью-мишенью. Если направляющая РНК образует комплекс с RGN, связанная RGN вводит одноцепочечный или двухцепочечный разрыв в последовательности-мишени. После расщепления последовательности-мишени разрыв можно репарировать таким образом, чтобы осуществить модификацию последовательности ДНК в составе последовательности-мишени в ходе репарации. В настоящем документе предлагаются способы применения мутантных вариантов РНК-управляемых нуклеаз, которые являются либо неактивными в отношении нуклеаз, либо в отношении нуклеаз, которые присоединены к деаминазам для модификации последовательности-мишени в ДНК клеток-хозяина. Мутантные варианты РНК-управляемых нуклеаз, в которых нуклеазная активность инактивирована или значительно снижена, можно называть РНК-управляемыми ДНК-связывающими полипептидами, поскольку полипептиды способны связываться с последовательностью-мишенью, но не обязательно расщеплять ее. РНК-управляемые нуклеазы, способные расщеплять только одну цепь двухцепочечной молекулы нуклеиновой кислоты, называются в данном контексте нуклеазами.

Нуклеотидная последовательность-мишень связывается с ДНК-связывающим полипептидом, управляемым РНК, и гибридизуется с гидовой (направляющей) РНК, связанной с RGNBP. Затем последовательность-мишень можно впоследствии расщеплять, если RGNBP обладает нуклеазной активностью (т.е. является RGN), которая включает активность нуклеаз.

Гидовая (направляющая) РНК может представлять собой единую направляющую РНК или систему двух направляющих РНК. РНК с одной направляющей включает sgРНК и, необязательно, tracrРНК на одной молекуле РНК, тогда как система РНК с двумя направляющими включает sgРНК и tracrРНК, которые присутствуют на двух различных молекулах РНК, гибридизованных друг с другом через по меньшей мере часть повторяющейся последовательности CRISPR в составе sgРНК и по меньшей мере часть tracrРНК, которая может являться полностью или частично комплементарной последовательности повторов CRISPR sgРНК. В некоторых из этих вариантов, где направляющая РНК представляет собой одну направляющую РНК, sgРНК и

необязательно tracrРНК разделены линкерной нуклеотидной последовательностью.

Как правило, линкерная нуклеотидная последовательность представляет собой последовательность, которая не включает комплементарные основания, чтобы исключить образование вторичной структуры внутри или включающей нуклеотиды линкерной нуклеотидной последовательности. В некоторых вариантах длина линкерной нуклеотидной последовательности между сгРНК и tracrРНК 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, длиной не менее 12 и более нуклеотидов. В конкретных вариантах длина линкерной нуклеотидной последовательности между сгРНК и tracrРНК составляет 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 или более нуклеотидов. В конкретных вариантах длина линкерной нуклеотидной последовательности единой направляющей РНК составляет по меньшей мере 4 нуклеотида. В конкретных вариантах длина линкерной нуклеотидной последовательности единой направляющей РНК составляет 4 нуклеотида.

В некоторых вариантах направляющую РНК можно вводить в клетку-мишень, органеллу или эмбрион в виде молекулы РНК. Направляющую РНК можно транскрибировать *in vitro* или получить в ходе химического синтеза. В некоторых вариантах нуклеотидную последовательность, кодирующую направляющую РНК, вводят в клетку, органеллу или эмбрион. В некоторых вариантах нуклеотидная последовательность, кодирующая направляющую РНК, функционально связана с промотором (например, промотором РНК-полимеразы III). Промотор может представлять собой природный промотор или гетерологичный промотор по отношению к нуклеотидной последовательности, кодирующей направляющую РНК.

В различных вариантах направляющую РНК можно вводить в клетку-мишень, органеллу или эмбрион в виде рибонуклеопротеинового комплекса, как описано в настоящем документе, где направляющая РНК связана с полипептидом РНК-управляемой нуклеазы.

Направляющая РНК направляет ассоциированную РНК-управляемую нуклеазу к конкретной исследуемой нуклеотидной последовательности-мишени при гибридизации направляющей РНК с нуклеотидной последовательностью -

мишенью. Нуклеотидная последовательность-мишень может содержать ДНК, РНК или их комбинацию и может представлять собой одноцепочечную или двухцепочечную молекулу. Нуклеотидная последовательность-мишень может представлять собой геномную ДНК (т.е. хромосомальную ДНК), плазмидную ДНК или молекулу РНК (например, матричную РНК, рибосомальную РНК, транспортную РНК, микроРНК, малую интерферирующую РНК). Нуклеотидная последовательность-мишень может являться связанной (и в некоторых вариантах расщепленной) РНК-управляемым ДНК-связывающим полипептидом *in vitro* или в клетке. Хромосомальная последовательность, на которую направлен RGDBP, может представлять собой ядерную, пластидную или митохондриальную хромосомальную последовательность. В некоторых вариантах нуклеотидная последовательность-мишень является уникальной в геноме-мишени.

В некоторых вариантах нуклеотидная последовательность-мишень примыкает к мотиву, соседнему с протоспейсером (РАМ). РАМ обычно находится в интервале от приблизительно 1 до приблизительно 10 нуклеотидов от нуклеотидной последовательности-мишени, включая приблизительно 1, приблизительно 2, приблизительно 3, приблизительно 4, приблизительно 5, приблизительно 6, приблизительно 7, приблизительно 8, приблизительно 9 или приблизительно 10 нуклеотидов от нуклеотидной последовательности-мишени. В конкретных вариантах РАМ находится в интервале от 1 до 10 нуклеотидов от нуклеотидной последовательности-мишени, включая 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 нуклеотидов от нуклеотидной последовательности-мишени. РАМ можно расположить на 5'- или 3'-конце последовательности-мишени. В некоторых вариантах РАМ находится на 3'-конце последовательности-мишени. Как правило, РАМ представляет собой консенсусную последовательность примерно из 2-6 нуклеотидов, но в конкретных вариантах осуществления ее длина составляет 1, 2, 3, 4, 5, 6, 7, 8, 9 или более нуклеотидов.

РАМ ограничивает участки последовательности, которые можно направить на данную RGDBP или RGN, поскольку ее РАМ должен примыкать к нуклеотидной последовательности-мишени. После распознавания соответствующей последовательности РАМ, RGN может расщеплять нуклеотидную последовательность-мишень в конкретном сайте расщепления. Используемый в данном контексте сайт расщепления состоит из двух конкретных нуклеотидов в нуклеотидной последовательности-мишени, между

которыми нуклеотидная последовательность 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 РАМ в любом направлении 5' или 3'. Поскольку RGN могут расщеплять нуклеотидную последовательность-мишень, что приводит к смещению концов, то в некоторых вариантах сайт расщепления определяют по расстоянию двух нуклеотидов от РАМ на положительной (+) цепи полинуклеотида и расстоянию между двумя нуклеотидами от РАМ на отрицательной (-) цепи полинуклеотида.

RGDBP и RGN можно использовать для доставки гибридных полипептидов, полинуклеотидов или низкомолекулярных фрагментов полезной нагрузки в определенный участок генома.

В тех вариантах, в которых ДНК-связывающий полипептид содержит мегануклеазу, последовательность-мишень может содержать пару инвертированных «полусайтов» из 9 пар оснований (п.о.), которые разделены четырьмя парами оснований. В случае одноцепочечной мегануклеазы N-концевой домен белка контактирует с первым полусайтом, а С-концевой домен белка контактирует со вторым полусайтом. Расщепление мегануклеазой приводит к формированию 3'-тяжей из четырех п.о. В тех вариантах, где ДНК-связывающий полипептид содержит компактный TALEN, последовательность распознавания содержит первую последовательность CNNNGN, которая распознается доменом I-TevI, за которой следует неспецифический спейсер длиной 4-16 п.о., за которым следует вторая последовательность длиной 16–22 п.о., которая распознается TAL-эффекторным доменом (эта последовательность обычно включает 5'-Т-основание). В тех вариантах, в которых ДНК-связывающий полипептид содержит структуру цинковый палец, ДНК-связывающие домены обычно распознают распознаваемую последовательность длиной 18 п.о., содержащую пару «полусайтов» из девяти п.о., разделенных 2-10 парами оснований, и расщепление нуклеазой создает тупой конец или 5'-тяж переменной длины (часто четыре п.о.).

IV. Гибридные белки

В некоторых вариантах ДНК-связывающий полипептид (например, неактивный в отношении нуклеазы или никаза RGN) функционально связан с дезаминазой по изобретению. В некоторых вариантах ДНК-связывающий полипептид (например, нуклеазо-неактивная RGN или никаза RGN),

гибридизированный с дезаминазой по изобретению, можно направить на конкретное место расположения молекулы нуклеиновой кислоты (т.е. молекулу нуклеиновой кислоты-мишени), которая в некоторых вариантах представляет собой конкретный геномный локус, для изменения экспрессии требуемой последовательности. В некоторых вариантах связывание гибридного белка с последовательностью-мишенью приводит к дезаминированию азотистого основания, что приводит к превращению одного азотистого основания в другое. В некоторых вариантах связывание этого гибридного белка с последовательностью-мишенью приводит к дезаминированию азотистого основания, соседнего с последовательностью-мишенью. Нуклеотидное основание, смежное с последовательностью-мишенью, которое является дезаминированным и мутированным с использованием описанных в настоящем документе композиций и способов, может представлять собой 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 или 100 п.о. от 5' или 3'-конца последовательности-мишени (связанной гидовой гРНК) в молекуле нуклеиновой кислоты-мишени. Некоторые объекты настоящего изобретения относятся к гибридным белкам, содержащим (i) ДНК-связывающий полипептид (например, нуклеазо-неактивный или нуклеазо-неактивный полипептид RGN); (ii) полипептид дезаминазы; и необязательно (iii) вторую дезаминазу. Вторая дезаминаза может представлять собой ту же самую дезаминазу, что и первая, или может являться другой дезаминазой. В некоторых вариантах как первая, так и вторая дезаминаза представляют собой адениндезаминазы по изобретению.

В настоящем изобретении предлагаются гибридные белки различных конфигураций. В некоторых вариантах полипептид дезаминазы является гибридизованным с N-концом ДНК-связывающего полипептида (например, полипептида RGN). В некоторых вариантах полипептид дезаминазы является гибридизованным с C-концом ДНК-связывающего полипептида (например, полипептида RGN).

В некоторых вариантах дезаминаза и ДНК-связывающий полипептид (например, РНК-управляемый ДНК-связывающий полипептид) образуют гибридизованную структуру пептидный линкер. Линкер между дезаминазой и ДНК-связывающим полипептидом (например, РНК-управляемым, ДНК-связывающим полипептидом) может определять окно редактирования

гибридного белка, тем самым повышая специфичность деаминазы и уменьшая нецелевые мутации. Можно использовать линкеры различной длины и гибкости, от очень гибких линкеров формы (GGGGS)_n и (G)_n до более жестких линкеров формы (EAAAK)_n и (XP)_n для достижения оптимальной длины и жесткости и для оптимизации активности деаминазы, предназначенной для применения в конкретных вариантах. Термин «линкер», используемый в данном документе, относится к химической группе или молекуле, связывающей две молекулы или два фрагмента, например, связывающий домен и домен расщепления нуклеазы. В некоторых вариантах линкер соединяет РНК-управляемую нуклеазу и деаминазу. В некоторых вариантах линкер соединяет мертвую или неактивную RGN и деаминазу. В дополнительных вариантах линкер соединяет две деаминазы. Как правило, линкер расположен между двумя группами, молекулами или другими фрагментами или окружен ими, и соединен с каждым из них ковалентной связью, таким образом, соединяя их. В некоторых вариантах линкер представляет собой аминокислоту или множество аминокислот (например, пептид или белок). В некоторых вариантах линкер представляет собой органическую молекулу, группу, полимер или химический фрагмент. В некоторых вариантах длина линкера составляет 3-100 аминокислот, например, 3, 4, 5, 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, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150 или 150-200 аминокислот. Также рассматриваются более длинные или более короткие линкеры. В некоторых вариантах предпочтительным является более короткий линкер для уменьшения общего размера или длины гибридного белка или его кодирующей последовательности.

В некоторых вариантах линкер содержит мотив (GGGGS)_n, (G)_n и (EAAAK)_n или (XP)_n или комбинацию любого из них, где *n* независимо равен целому числу от 1 до 30, В некоторых вариантах *n* независимо равен 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 или 30, или, если присутствует более одного линкера или более одного линкерного мотива, любой их комбинации. Дополнительные пригодные линкерные мотивы и конфигурации линкеров представляются очевидными специалистам в данной области. В некоторых вариантах пригодные линкерные мотивы и конфигурации включают мотивы и конфигурации, описанные в статье Chen и др. Adv Drug Deliv Rev. 65(10):1357-1369 (2013), содержание которой в

полном объеме включено в настоящий документ в качестве ссылки.

Специалистам в данной области техники представляются очевидными дополнительные пригодные линкерные последовательности. В некоторых вариантах линкерная последовательность содержит аминокислотную

5 последовательность, представленную как SEQ ID NO: 45 или 442.

В некоторых вариантах общая структура типичных гибридных белков, представленных в настоящем документе, включает структуру: [NH₂]-[дезаминаза]-[DBP]-[COOH]; [NH₂]-[DBP]-[дезаминаза]-[COOH]; [NH₂]-[DBP]-[дезаминаза]-[дезаминаза]-[COOH]; [NH₂]-[дезаминаза]-[DBP]-[дезаминаза]-[COOH] или [NH₂]-[дезаминаза]-[дезаминаза]-[DBP]-[COOH], где DBP
10 представляет собой ДНК-связывающий полипептид, NH₂ представляет собой N-конец гибридного белка, а COOH представляет собой C-конец гибридного белка. В некоторых вариантах гибридный белок содержит более двух полипептидов дезаминазы.

15 В некоторых вариантах общая структура типичных гибридных белков, представленных в настоящем документе, включает структуру: [NH₂]-[дезаминаза]-[RGN]-[COOH]; [NH₂]-[RGN]-[дезаминаза]-[COOH]; [NH₂]-[RGN]-[дезаминаза]-[дезаминаза]-[COOH]; [NH₂]-[дезаминаза]-[RGN]-[дезаминаза]-[COOH] или [NH₂]-[дезаминаза]-[дезаминаза]-[RGN]-[COOH], где NH₂
20 представляет собой N-конец гибридного белка, а COOH представляет собой C-конец гибридного белка. В некоторых вариантах гибридный белок содержит более двух полипептидов дезаминазы.

В некоторых вариантах гибридный белок характеризуется структурой: [NH₂]-[дезаминаза]-[нуклеазо-неактивная RGN]-[COOH]; [NH₂]-[дезаминаза]-[дезаминаза]-[нуклеазо-неактивная RGN]-[COOH]; [NH₂]-[нуклеазо-неактивная RGN]-[дезаминаза]-[COOH]; [NH₂]-[дезаминаза]-[нуклеазо-неактивная RGN]-[дезаминаза]-[COOH] или [NH₂]-[нуклеазо-неактивная RGN]-[дезаминаза]-[дезаминаза]-[COOH]. Следует понимать, что «нуклеазо-неактивная RGN»
25 представляет собой любую RGN, включающую любой белок CRISPR-Cas, в который была введена мутация для превращения в нуклеазо-неактивную форму. В некоторых вариантах гибридный белок содержит более двух полипептидов дезаминазы.
30

В некоторых вариантах гибридный белок характеризуется структурой: [NH₂]-[дезаминаза]-[RGN-никаза]-[COOH]; [NH₂]-[дезаминаза]-[дезаминаза]-

[RGN никаза]-[COOH]; [NH₂]-[RGN никаза]-[дезаминаза]-[COOH]; [NH₂]-[дезаминаза]-[RGN никаза]-[дезаминаза]-[COOH]; или [NH₂]-[RGN никаза]-[дезаминаза]-[дезаминаза]-[COOH]. Следует понимать, что «никаза RGN» представляет собой любую RGN, включающую любой белок CRISPR-Cas, в
5 который была введена мутация для превращения в активную форму в качестве
никазы.

В некоторых вариантах кавычки «-», использованные в приведенной выше
общей структуре, указывают на наличие необязательной линкерной
последовательности. В некоторых вариантах гибридные белки, предложенные в
10 настоящем документе, не содержат линкерную последовательность. В некоторых
вариантах присутствует по меньшей мере одна из необязательных линкерных
последовательностей.

Другие типичные признаки, которые могут присутствовать, представляют
собой последовательности локализации, такие как последовательности ядерной
15 локализации, последовательности цитоплазматической локализации, экспортные
последовательности, такие как последовательности ядерного экспорта, или
другие последовательности локализации, а также маркерные
последовательности, которые можно использовать для солюбилизации, очистки
или детектирования гибридных белков. Пригодные сигнальные
20 последовательности локализации и маркерные последовательности белков,
которые представлены в настоящем документе, включают, но не ограничиваясь
только этим, маркеры белка-носителя биотинкарбоксилазы (BCCP), мус-
маркеры, кальмодулиновые маркеры, FLAG-маркеры, гемагглютининовые (HA)-
маркеры, полигистидиновые маркеры, также называемые гистидиновыми
25 маркерами или His-маркерами, маркеры белка, связывающего мальтозу (MBP),
nus-маркеры, маркеры глутатион-S-трансферазы (GST), маркеры зеленого
флуоресцентного белка (GFP), тиоредоксин-маркеры, S-маркеры, маркеры
Softags (например, Softag 1, Softag 3), стрептеги, маркеры биотинлигазы,
маркеры FlAsH, маркеры V5 и SBP-маркеры. Дополнительные пригодные
30 последовательности представляются очевидными специалистам в данной
области.

В некоторых вариантах раскрытые в данном контексте гибридные белки
содержат по меньшей мере один проникающий в клетку домен, который
ускоряет захват гибридного белка клеткой. Проникающие в клетку домены

известны в данной области техники и обычно содержат тяжи положительно заряженных аминокислотных остатков (т. е. поликатионные проникающие в клетку домены), чередующиеся полярные аминокислотные остатки и неполярные аминокислотные остатки (т. е. амфипатические проникающие в клетку домены), или гидрофобные аминокислотные остатки (т.е. гидрофобные проникающие в клетку домены) (см., например, статью Milletti F. Drug Discov Today 17:850-860 (2012)). Неограничивающим примером проникающего в клетку домена является трансактивирующий транскрипционный фактор (ТАТ) из вируса иммунодефицита человека 1.

В некоторых вариантах дезаминазы или гибридные белки, предложенные в настоящем документе, дополнительно содержат последовательность ядерной локализации (NLS). Сигнал ядерной локализации, сигнал плазмидной локализации, сигнал митохондриальной локализации, сигнал локализации с двойной направленностью и/или проникающий в клетку домен могут располагаться на amino-конце (N-конце), карбоксильном конце (С-конце), или внутри гибридного белка.

В некоторых вариантах NLS является гибридным с N-концом гибридного белка или дезаминазы. В некоторых вариантах NLS является гибридным с С-концом гибридного белка или дезаминазы. В некоторых вариантах NLS является гибридным с N-концом дезаминазы гибридного белка. В некоторых вариантах NLS является гибридным с С-концом дезаминазы гибридного белка. В некоторых вариантах NLS является гибридным с N-концом ДНК-связывающего полипептида (например, полипептида RGN) гибридного белка. В некоторых вариантах NLS является гибридным с С-концом ДНК-связывающего полипептида (например, полипептида RGN) гибридного белка. В некоторых вариантах NLS является гибридным с N-концом полипептида дезаминазы гибридного белка. В некоторых вариантах NLS является гибридным с С-концом полипептида дезаминазы гибридного белка. В некоторых вариантах NLS является гибридным с гибридным белком, который присоединен через один или несколько линкеров. В некоторых вариантах NLS является гибридным с гибридным белком без линкера. В некоторых вариантах NLS содержит аминокислотную последовательность любой из последовательностей NLS, представленных или упомянутых в настоящем документе. В некоторых

вариантах NLS содержит аминокислотную последовательность, представленную в виде SEQ ID NO: 43 или SEQ ID NO: 46. В некоторых вариантах гибридный белок или дезаминаза содержит SEQ ID NO: 43 на их N-конце и SEQ ID NO: 46 на их C-конце.

5 В некоторых вариантах гибридные белки, представленные в настоящем документе, содержат полноразмерную последовательность дезаминазы, например, любую из последовательностей SEQ ID NO: 1-10 и 399-441. Однако в некоторых вариантах гибридные белки, представленные в настоящем документе, содержат не полноразмерную последовательность дезаминазы, а только ее
10 фрагмент. Например, в некоторых вариантах гибридный белок, представленный в настоящем документе, дополнительно содержит ДНК-связывающий полипептидный (например, РНК-управляемый, ДНК-связывающий) домен и домен дезаминазы.

В некоторых вариантах гибридный белок по настоящему изобретению
15 содержит ДНК-связывающий полипептид (например, RGN) и дезаминазу, при этом дезаминаза содержит аминокислотную последовательность, которая идентична по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на
20 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или на 100% по сравнению с любой из последовательностей SEQ ID NO: 1-10 и 399-441. Примеры таких гибридных белков описаны в данном контексте в разделе «Примеры».

25 В некоторых вариантах гибридный белок содержит один полипептид дезаминазы. В некоторых вариантах гибридный белок содержит по меньшей мере два полипептида дезаминазы, функционально связанные либо напрямую, либо через пептидный линкер. В некоторых вариантах гибридный белок содержит один полипептид дезаминазы, а второй полипептид дезаминазы
30 коэкспрессируется с гибридным белком.

В настоящем документе также предлагается рибонуклеопротеиновый комплекс, содержащий гибридный белок, содержащий дезаминазу и RGNBP, и направляющую РНК, либо в виде одноцепочечной направляющей (гидовой),

либо в виде двухцепочечной направляющей (гидовой) РНК (также в общем случае называемой гРНК).

V. Нуклеотиды, кодирующие дезаминазы, гибридные белки и/или гРНК

В настоящем изобретении предлагаются полинуклеотиды (SEQ ID NO: 11-20 и 443-485), кодирующие описанные в данном контексте полипептиды дезаминазы. В настоящем изобретении дополнительно предлагаются полинуклеотиды, кодирующие гибридные белки, которые содержат дезаминазу и ДНК-связывающий полипептид, например, мегануклеазу, гибридный белок, включающий структуру цинковые пальцы или TALEN. В настоящем изобретении также предлагаются полинуклеотиды, кодирующие гибридные белки, которые содержат домен дезаминазы и РНК-управляемый ДНК-связывающий полипептид. Такие РНК-управляемые ДНК-связывающие полипептиды могут представлять собой RGN или вариант RGN. Вариант белка может представлять собой нуклеазо-неактивным или представлять собой никазу. RGN может представлять собой белок CRISPR-Cas или его активный вариант или фрагмент. Последовательности SEQ ID NO: 41 и 42 представляют собой неограничивающие примеры варианта RGN и никазы, соответственно. Примеры нуклеаз CRISPR-Cas хорошо известны в данной области техники, и подобные соответствующие мутации могут создавать мутантные варианты, которые также являются никазами или неактивны в отношении нуклеаз.

В варианте осуществления изобретения предлагается полинуклеотид, кодирующий гибридный белок, который содержит RGDBP и дезаминазу, описанные в настоящем документе (SEQ ID NO: 1-10 и 399-441, или их вариант). В некоторых вариантах второй полинуклеотид кодирует направляющую РНК, требуемую для RGDBP для направления на исследуемую нуклеотидную последовательность. В некоторых вариантах направляющая РНК и гибридный белок кодируются одним и тем же полинуклеотидом.

Использование термина «полинуклеотид» не предназначено для ограничения настоящего изобретения полинуклеотидами, содержащими ДНК, хотя такие полинуклеотиды ДНК рассматриваются. Специалистам в данной области представляется очевидным, что полинуклеотиды могут включать рибонуклеотиды (РНК) и комбинации рибонуклеотидов и дезоксирибонуклеотидов. Такие дезоксирибонуклеотиды и рибонуклеотиды включают как природные молекулы, так и синтетические аналоги.

Полинуклеотиды, раскрытые в настоящем документе, также охватывают все формы последовательностей, включая, но не ограничиваясь только этим, одноцепочечные формы, двухцепочечные формы, структуры типа «тяж-петля», кольцевые формы (например, включая кольцевую РНК) и т.п.

5 В варианте осуществления изобретения предлагается молекула нуклеиновой кислоты, содержащая последовательность, которая идентична по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по
10 меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или на 100% по сравнению с любой из SEQ ID NO: 11-20 и 443- 485, где молекула нуклеиновой кислоты кодирует дезаминазу, обладающую адениндезаминазной активностью. Молекула нуклеиновой кислоты может дополнительно содержать гетерологичный
15 промотор или терминатор. Молекула нуклеиновой кислоты может кодировать гибридный белок, в котором кодируемая дезаминаза функционально связана с ДНК-связывающим полипептидом, и необязательно со второй дезаминазой. В некоторых вариантах молекула нуклеиновой кислоты кодирует гибридный белок, где кодируемая дезаминаза функционально связана с RGN и,
20 необязательно, со второй дезаминазой.

В некоторых вариантах молекулы нуклеиновой кислоты, содержащие полинуклеотид, кодирующий дезаминазу по изобретению, являются оптимизированными по кодам для экспрессии в исследуемом организме. «Кодон-оптимизированная» кодирующая последовательность представляет
25 собой полинуклеотидную кодирующую последовательность, с частотой использования кодонов, сконструированную для имитации частоты предпочтительного использования кодонов или условий транскрипции конкретной клетки-хозяина. Экспрессия в конкретной клетке-хозяина или организме повышается в результате изменения одного или нескольких кодонов
30 на уровне нуклеиновой кислоты, так что транслируемая аминокислотная последовательность не изменяется. Молекулы нуклеиновых кислот могут являться оптимизированными по кодам полностью или частично. Таблицы кодонов и другие ссылки, в которых представлена информация о предпочтениях для широкого круга организмов, доступны в данной области (см., например,

статью Campbell и Gowri Plant Physiol. 92:1-11 (1990) для обсуждения использования предпочтительных для растений кодонов). В данной области техники доступны способы синтеза предпочтительных для растений генов. См., например, патенты США № 5380831 и 5436391, и статью Murray и др., Nucleic Acids Res. 17:477-498 (1989), которые включены в данный контекст в качестве ссылки.

В некоторых вариантах полинуклеотиды, кодирующие дезаминазы, гибридные белки и/или гРНК, описанные в настоящем документе, представлены в экспрессионных кассетах для экспрессии *in vitro* или экспрессии в исследуемой клетке, органелле, эмбрионе или организме. Кассета может включать 5'- и 3'-регуляторные последовательности, функционально связанные с полинуклеотидом, кодирующим дезаминазу, и/или гибридный белок, содержащий дезаминазу, РНК-управляемый ДНК-связывающий полипептид и, необязательно, вторую дезаминазу, и/или гРНК, представленные в настоящем документе, которые обеспечивает экспрессию полинуклеотида. Кассета может дополнительно содержать по меньшей мере один дополнительный ген или генетический элемент, котрансформируемый в организм. В тех участках, в которых включены дополнительные гены или элементы, компоненты являются функционально связанными. Термин «функционально связанный» предназначен для обозначения функциональной связи между двумя или более элементами. Например, функциональная связь между промотором и исследуемым кодирующим участком (например, участком, кодирующим дезаминазу, РНК-управляемый ДНК-связывающий полипептид и/или гРНК) представляет собой функциональную связь, которая обеспечивает экспрессию кодирующего участка. Функционально связанные элементы могут являться смежными или несмежными. При использовании для обозначения соединения двух областей, кодирующих белок, термин «функционально связанные» означает, что кодирующие области находятся в одной и той же рамке считывания. В некоторых вариантах дополнительный ген(ы) или элемент(ы) представлены на кассетах множественной экспрессии. Например, нуклеотидная последовательность, кодирующая раскрытую в данном контексте дезаминазу, либо отдельно, либо как компонент гибридного белка, может присутствовать на одной экспрессионной кассете, тогда как нуклеотидная последовательность, кодирующая гРНК, может находиться на отдельной экспрессионной кассете. В

другом примере может присутствовать нуклеотидная последовательность, кодирующая только описанную в данном контексте дезаминазу, на первой экспрессионной кассете, вторая экспрессионная кассета, кодирующая гибридный белок, содержащий дезаминазу, и нуклеотидная последовательность, кодирующая гРНК, на третьей экспрессионной кассете. Такая экспрессионная кассета снабжена множеством сайтов рестрикции и/или сайтов рекомбинации для вставки полинуклеотидов, предназначенных для транскрипционной регуляции регуляторных областей. Могут также присутствовать экспрессионные кассеты, которые содержат селективный маркерный ген.

Экспрессионная кассета может включать в 5'-3' направлении транскрипции область инициации транскрипции (и, в некоторых вариантах, трансляции) (т.е. промотор), кодирующий дезаминазу полинуклеотид по изобретению и транскрипционную (и в некоторых вариантах осуществления трансляционную) область терминирования (т.е. область терминирования), которая является функциональной в исследуемом организме. Промоторы по изобретению способны направлять или стимулировать экспрессию кодирующей последовательности в клетке-хозяине. Регуляторные области (например, промоторы, области регуляции транскрипции и области терминирования трансляции) могут представлять собой эндогенные или гетерологичные области по отношению к клетке-хозяину или друг к другу. Используемый в данном контексте термин «гетерологичный» по отношению к последовательности представляет собой последовательность, которая происходит от чужеродного вида или, если она принадлежит к тому же виду, в значительной степени изменена по сравнению с ее нативной формой по составу и/или геномному локусу в результате преднамеренного вмешательства человека. Как используется в данном документе, химерный ген содержит кодирующую последовательность, функционально связанную с областью инициации транскрипции, гетерологичной кодирующей последовательности.

Пригодные области терминирования находятся в Ti-плазмиде *A. tumefaciens*, такие как области терминирования октопинсинтазы и нопалинсинтазы. См. также статьи Guerineau и др., *Mol. Gen. Genet.* 262:141-144 (1991); Proudfoot, *Cell* 64:671-674 (1991); Sanfacon и др., *Genes Dev.* 5:141-149 (1991); Mogen и др., *Plant Cell* 2:1261-1272 (1990); Munroe и др., *Gene* 91:151-158 (1990); Ballas и др.,

Nucleic Acids Res. 17:7891-7903 (1989) и Joshi и др., Nucleic Acids Res. 15:9627-9639 (1987).

Дополнительные регуляторные сигналы включают, но не ограничиваясь только ими, сайты инициации транскрипции, операторы, активаторы, энхансеры, другие регуляторные элементы, сайты связывания рибосом, кодон инициации, сигналы терминации и т.п. См., например, патенты США 5039523 и 4853331; EPO 0480762A2; статьи Sambrook и др., *Molecular Cloning: A Laboratory Manual*, ред. Maniatis и др. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1992), N.Y.), далее в тексте “Sambrook 11”; Davis и др., ред. *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory Press), Cold Spring Harbor, N.Y. (1980), и публикации, цитированные в указанных документах.

При приготовлении экспрессионной кассеты можно использовать различные фрагменты ДНК, чтобы сконструировать последовательности ДНК в надлежащей ориентации и, при необходимости, в надлежащей рамке считывания. С этой целью для соединения фрагментов ДНК можно использовать адаптеры или линкеры, или можно использовать другие методы конструирования для обеспечения удобных сайтов рестрикции, удаления излишней ДНК, удаления сайтов рестрикции и т.п. Для этой цели можно использовать мутагенез *in vitro*, репарацию праймеров, рестрикцию, отжиг, повторные замены, например, транзиции и трансверсии.

На практике при осуществлении изобретения можно использовать ряд промоторов. Промоторы можно выбрать с учетом требуемого результата. Нуклеиновые кислоты можно комбинировать с конститутивными, индуцибельными, специфичными для стадии роста, специфичными для типа клеток, ткане-предпочтительными, ткане-специфичными или другими промоторами для экспрессии в исследуемом организме. См., например, промоторы, изложенные в заявке WO 99/43838 и в патентах США №№: 8575425, 7790846, 8147856, 8586832, 772369, 7534939, 6072050, 5659026, 5608149, 5608144, 5604121, 5569597, 5466785, 5399680, 5268463, 5608142 и 6177611, которые включены в настоящий документ в качестве ссылок.

Для экспрессии в растениях, конститутивные промоторы также включают промотор CaMV 35S (Odell и др., *Nature* 313:810-812 (1985)); актин из риса (McElroy и др., *Plant Cell* 2:163-171 (1990); убиквитин (Christensen и др., *Plant Mol. Biol.* 12:619-632 (1989) и Christensen и др., *Plant Mol. Biol.* 18:675-689

(1992)); pEMU (Last и др., *Theor. Appl. Genet.* 81:581-588 (1991)); и MAS (Velten и др., *EMBO J.* 3:2723-2730 (1984)).

Примерами индуцибельных промоторов являются промотор Adh1, который индуцируется гипоксией или холодовым стрессом, промотор Hsp70, который индуцируется тепловым стрессом, промотор PPDК и промотор пепкарбоксилазы, которые оба индуцируются светом. Можно также использовать химически индуцибельные промоторы, такие как промотор In2-2, который индуцируется антидотом (патент США № 5364780), промотор Axig1, который индуцируется ауксином и является тапетум-специфичным, но также проявляет активность в каллюсе (PCT US01/22169), стероид-чувствительные промоторы (см., например, ERE-промотор, который индуцируется эстрогеном, и глюкокортикоид-индуцибельный промотор, которые описаны в статьях Schena и др., *Proc. Natl. Acad. Sci. USA* 88:10421-10425 (1991) и McNellis и др. *Plant J.* 14(2):247-257 (1998)), а также промоторы, индуцируемые тетрациклином, и промоторы, репрессируемые тетрациклином (см., например, статью Gatz и др., *Mol. Gen. Genet.* 227:229-237 (1991) и патенты США No. 5814618 и 5789156), которые включены в настоящее описание в качестве ссылок.

В некоторых вариантах ткане-специфичные или ткане-предпочтительные промоторы используют для направленной экспрессии экспрессионного конструкта в конкретной ткани. В некоторых вариантах ткане-специфичные или ткане-предпочтительные промоторы проявляют активность в растительной ткани. Примеры промоторов, находящихся под контролем развития у растений, включают промоторы, которые предпочтительно иницируют транскрипцию в определенных тканях, таких как листья, корни, плоды, семена или цветы. «Ткане-специфический» промотор представляет собой промотор, который иницирует транскрипцию только в определенных тканях. В отличие от конститутивной экспрессии генов, ткане-специфичная экспрессия происходит в результате нескольких взаимодействующих уровней генной регуляции. Как таковые, промоторы из гомологичных или близкородственных видов растений могут являться предпочтительными при использовании, предназначенном для достижения эффективной и надежной экспрессии трансгенов в конкретных тканях. В некоторых вариантах экспрессия включает ткане-предпочтительный промотор. «Ткане-предпочтительный» промотор представляет собой промотор,

который предпочтительно инициирует транскрипцию, но не обязательно полностью или исключительно в определенных тканях.

В некоторых вариантах молекулы нуклеиновой кислоты, кодирующие дезаминазу, описанную в настоящем документе, содержат промотор,

5 специфичный для типа клеток. «Специфичный для типа клеток» промотор представляет собой промотор, который в первую очередь управляет экспрессией в определенных типах клеток в одном или нескольких органах. Некоторые типичные растительные клетки, в которых специфичные для типа клеток и функционирующие в растениях промоторы могут в первую очередь проявлять
10 активность, включают, например, клетки BETL, васкулярные клетки в корнях, листьях, клетки стебля и ствольные клетки. Молекулы нуклеиновых кислот также могут включать промоторы предпочтительного типа клеток. Промотор «предпочтительного типа клеток» представляет собой промотор, который главным образом в основном управляет экспрессией, но не обязательно
15 полностью или исключительно в определенных типах клеток в одном или нескольких органах. Некоторые примеры клеток растений, в которых могут предпочтительно проявлять активность промоторы предпочтительного типа клеток, функционирующие в растениях, включают, например, клетки BETL, васкулярные клетки в корнях, листьях, клетки стебля и ствольные клетки.

20 В некоторых вариантах последовательности нуклеиновых кислот, кодирующие дезаминазы, гибридные белки и/или гРНК, функционально связаны с последовательностью промотора, которая распознается фаговой РНК-полимеразой, например, для синтеза мРНК *in vitro*. В таких вариантах транскрибируемую *in vitro* РНК можно очищать для использования в описанных
25 в данном контексте способах. Например, промоторная последовательность может представлять собой промоторную последовательность T7, T3 или SP6 или вариацию промоторной последовательности T7, T3 или SP6. В таких вариантах экспрессированный белок и/или РНК можно очистить для использования в описанных в данном контексте способах модификации генома.

30 В некоторых вариантах полинуклеотид, кодирующий дезаминазу, гибридный белок и/или гРНК, связаны с сигналом полиаденилирования (например, сигналом полиА SV40 и другими сигналами, функционирующими в растениях) и/или по меньшей мере с одной последовательностью терминации транскрипции. В некоторых вариантах последовательность, кодирующая

дезаминазу или гибридный белок, связана с последовательностью(ями), кодирующей(ими) по меньшей мере один сигнал ядерной локализации, по меньшей мере один проникающий в клетку домен и/или по меньшей мере один сигнальный пептид, способный вызывать миграцию белков в определенные субклеточные участки, как описано в данном контексте.

В некоторых вариантах полинуклеотид, кодирующий дезаминазу, гибридный белок и/или гРНК, присутствует в векторе или нескольких векторах. «Вектор» относится к полинуклеотидной композиции для переноса, доставки или введения нуклеиновой кислоты в клетку-хозяина. Пригодные векторы включают плазмидные векторы, фагемиды, космиды, искусственные/мини-хромосомы, транспозоны и вирусные векторы (например, лентивирусные векторы, аденоассоциированные вирусные векторы, бакуловирусный вектор). В некоторых вариантах вектор содержит дополнительные последовательности контроля экспрессии (например, энхансерные последовательности, последовательности Козака, последовательности полиаденилирования, последовательности терминации транскрипции), селективируемые маркерные последовательности (например, гены устойчивости к антибиотикам), точки начала репликации и т.п. Дополнительную информацию можно найти в статьях “Current Protocols in Molecular Biology” Ausubel и др., John Wiley & Sons, New York (2003) или “Molecular Cloning: A Laboratory Manual” Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 3-е изд., (2001).

В некоторых вариантах вектор содержит селективируемый маркерный ген для селекции трансформированных клеток. Селективируемые маркерные гены используют для селекции трансформированных клеток или тканей. Маркерные гены включают гены, кодирующие устойчивость к антибиотикам, такие как гены, кодирующие неомицин-фосфотрансферазу II (NEO) и гигромицин-фосфотрансферазу (HPT), а также гены, придающие устойчивость к гербицидным соединениям, таким как глюфосинат аммония, бромоксинил, имидазолиноны и 2,4-дихлорфеноксиацетат (2,4-D).

В некоторых вариантах экспрессионная кассета или вектор, содержащий последовательность, кодирующую гибридный белок, содержащий РНК-управляемый ДНК-связывающий полипептид, такой как RGN, дополнительно содержит последовательность, кодирующую гРНК. В некоторых вариантах последовательность(и), кодирующая(ие) гРНК, функционально связана по

меньшей мере с одной последовательностью контроля транскрипции для экспрессии гРНК в исследуемом организме или клетке-хозяина. Например, полинуклеотид, кодирующий гРНК, можно функционально связать с последовательностью промотора, которая распознается РНК-полимеразой III (Pol III). Типичные пригодные промоторы Pol III включают, но не ограничиваясь только ими, промоторы U6, U3, H1 и 7SL РНК млекопитающих и промоторы U6 и U3 риса.

Как указано в данном контексте, экспрессионные конструкторы, содержащие нуклеотидные последовательности, кодирующие дезаминазы, гибридные белки и/или гРНК, можно использовать для трансформации исследуемых организмов. Способы трансформации включают введение нуклеотидного конструктора в исследуемый организм. Термин «введение» означает введение нуклеотидного конструктора в клетку-хозяина таким образом, чтобы конструктор получил доступ внутрь клетки-хозяина. Способы по изобретению не требуют конкретного способа введения нуклеотидного конструктора в организм-хозяина, а требуется только обеспечить доступ нуклеотидного конструктора внутрь по меньшей мере одной клетки организма-хозяина. Клетка-хозяина может представлять собой эукариотическую или прокариотическую клетку. В конкретных вариантах эукариотическая клетка-хозяин представляет собой клетку растения, клетку млекопитающего или клетку насекомого. Способы введения нуклеотидных конструкторов в растения и другие клетки-хозяина известны в данной области техники, включая, но не ограничиваясь только ими, способы стабильной трансформации, способы транзиторной трансформации и способы, опосредованные вирусом.

Способы позволяют получить трансформированный организм, такой как растение, включая целые растения, а также органы растения (например, листья, стебли, корни и т.д.), семена, клетки растений, рассеиваемые части организма, зародыши и их потомство. Клетки растений могут представлять собой дифференцированные или недифференцированные клетки (например, каллус, клетки суспензионной культуры, протопласты, клетки листа, клетки корня, клетки флоэмы, пыльца).

«Трансгенные организмы», или «трансформированные организмы», или «стабильно трансформированные» организмы, или клетки, или ткани относятся к организмам, которые включают или интегрируют полинуклеотид, кодирующий

дезаминазу по изобретению. Известно, что другие экзогенные или эндогенные последовательности нуклеиновых кислот или фрагменты ДНК также можно включать в клетку-хозяина. Трансформация, опосредованная агробактериями и биолистическим методом, остается двумя преимущественно используемыми подходами для трансформации растительных клеток. Однако трансформацию клетки-хозяина можно осуществлять при инфицировании, трансфекции, микроинъекции, электропорации, микропроецировании, биолистической бомбардировки или бомбардировки частицами, электропорации, с использованием силикагеля/углеродных волокон, при воздействии ультразвуком, при воздействии ПЭГ, соосаждении фосфатом кальция, с использованием метода поликатионов-ДМСО, DEAE-декстрана и вируса, липосом и т.п. Опосредованное вирусом введение полинуклеотида, кодирующего дезаминазу, гибридный белок и/или гРНК, включает ретровирусное, лентивирусное, аденовирусное и аденоассоциированное вирус-опосредованное введение и экспрессию, а также использование каулимовирусов (например, вируса мозаики цветной капусты), геминивирусов (например, вирус золотисто-желтой мозаики фасоли или вирус полосатости кукурузы) и РНК-вирусов растений (например, вирус табачной мозаики).

Протоколы трансформации, а также протоколы введения полипептидных или полинуклеотидных последовательностей в растения можно варьировать в зависимости от типа клетки-хозяина (например, клетки однодольного или двудольного растения), предназначенной для трансформации. Способы трансформации известны в данной области и включают методы, изложенные в патентах No: 8575425, 7692068, 8802934, 7541517; каждый из которых включен в настоящее описание в качестве ссылки. См. также статьи Rakoczy-Trojanowska, M., *Cell Mol Biol Lett.* 7:849-858 (2002); Jones и др., *Plant Methods* 1:5 (2005); Rivera и др., *Physics of Life Reviews* 9:308-345 (2012); Bartlett и др., *Plant Methods* 4:1-12 (2008); Bates G.W., *Methods in Molecular Biology* 111:359-366 (1999); Binns и Thomashow, *Annual Reviews in Microbiology* 42:575-606 (1988); Christou P. *The Plant Journal* 2:275-281 (1992); Christou P. *Euphytica* 85:13-27 (1995); Tzfira и др. *TRENDS in Genetics* 20:375-383 (2004); Yao и др., *Journal of Experimental Botany* 57:3737-3746 (2006); Zupan and Zambryski, *Plant Physiology* 107:1041-1047 (1995); Jones и др., *Plant Methods* 1:5 (2005).

Трансформация может обеспечить стабильное или временное включение нуклеиновой кислоты в клетку. Предполагается, что «стабильная трансформация» означает введение нуклеотидного конструкта в клетку-хозяина, интеграцию в геном клетки-хозяина и его способность наследоваться ее потомством. Предполагается также, что «временная трансформация» означает введение полинуклеотида в клетку-хозяина и отсутствие интеграции в геном клетки-хозяина.

Способы трансформации хлоропластов известны в данной области техники, см. статьи Svab и др., *Proc. Natl. Acad. Sci. USA* 87:8526-8530 (1990); Svab и Maliga, *Proc. Natl. Acad. Sci. USA* 90:913-917 (1993); Svab и Maliga, *EMBO J.* 12:601-606 (1993). Этот метод основан на доставке ДНК, содержащей селективный маркер, с помощью генной пушки и на направлении ДНК на пластидный геном с использованием гомологичной рекомбинации. Кроме того, пластидную трансформацию можно осуществлять с использованием трансактивации молчащего, переносимого пластидами трансгена с использованием ткане-предпочтительной экспрессии кодируемой в ядре и управляемой пластидой РНК-полимеразы. О такой системе сообщалось в статье McBride и др., *Proc. Natl. Acad. Sci. USA* 91:7301-7305 (1994).

Трансформированные клетки можно выращивать в трансгенном организме, таком как растения с использованием стандартных способов. См., например, статью McCormick и др., *Plant Cell Reports* 5:81-84 (1986). Затем эти растения можно выращивать и либо опылять одним и тем же трансформированным штаммом, либо различными штаммами, и полученный гибрид содержит идентифицированный полинуклеотид дезаминазы или гибридного белка. Два или более поколения можно выращивать, чтобы обеспечить стабильное сохранение полинуклеотида дезаминазы или гибридного белка и передачу его по наследству, а семена собирают, чтобы обеспечить присутствие полинуклеотида дезаминазы или гибридного белка. Таким образом, в настоящем изобретении предлагаются трансформированные семена (также называемые «трансгенными семенами»), содержащие нуклеотидный конструкт по изобретению, например экспрессионную кассету по изобретению, стабильно включенную в их геном.

В некоторых вариантах клетки, которые были трансформированы, вводят в организм. Эти клетки могут происходить из организма, в котором клетки трансформировали *ex vivo*.

Представленные в данном контексте последовательности можно использовать для трансформации любых видов растений, включая, но не ограничиваясь только ими, однодольные и двудольные растения. Типичные исследуемые растения включают, но не ограничиваясь только ими, кукурузу (маис), сорго, пшеницу, подсолнечник, томат, крестоцветные, перец, картофель, хлопчаточный, рис, соя, сахарную свеклу, сахарный тростник, табак, ячмень и масличный рапс, капустные *Brassica* sp., люцерну, рожь, просо, сафлор, арахис, сладкий картофель, маниоку, кофе, кокос, ананас, цитрусовые деревья, какао, чай, банан, авокадо, инжир, гуаву, манго, оливки, папайю, кешью, макадамиию, миндаль, овес, овощи, декоративные растения и хвойные деревья.

Овощи включают, но не ограничиваясь только ими, помидоры, салат, зеленую фасоль, лимскую фасоль, горох и представителей рода *Cucumis*, такие как огурец, канталупу и мускусную дыню. Декоративные растения включают, но не ограничиваясь только ими, азалии, гортензии, гибискус, розы, тюльпаны, нарциссы, петунии, гвоздики, поинсеттии и хризантемы. Растения по настоящему изобретению предпочтительно представляют собой сельскохозяйственные культуры (например, кукуруза, сорго, пшеница, подсолнечник, помидоры, крестоцветные, перец, картофель, хлопчаточный, рис, соя, сахарная свекла, сахарный тростник, табак, ячмень, масличный рапс и т.д.).

Используемый в данном контексте термин «растение» включает клетки растений, протопласты растений, культуры тканей клеток растений, из которых растения можно регенерировать, каллусы растений, группы побегов растений и клетки растений, которые являются интактными в растениях или частях растений, таких как зародыши, пыльца, семязачатки, семена, листья, цветы, ветки, плоды, зерна, колосья, початки, шелуха, стебли, корни, кончики корней, пыльники и т.п. Термин зерно означает зрелые семена, выпускаемые коммерческими фирмами для целей, отличающихся от выращивания или воспроизводства вида. Потомство, варианты и мутанты регенерированных растений также включены в объем изобретения при условии, что эти части содержат введенные полинуклеотиды. Кроме того, предлагается переработанный растительный продукт или побочный продукт, который сохраняет раскрытые в данном контексте последовательности, включая, например, соевые продукты.

В некоторых вариантах полинуклеотиды, кодирующие дезаминазы, гибридные белки и/или гРНК, используют для трансформации любых

эукариотических видов, включая, не ограничиваясь только ими, животных (например, млекопитающих, насекомых, рыб, птиц и рептилий), грибы, амёбы, водоросли и дрожжи. В некоторых вариантах полинуклеотиды, кодирующие дезаминазы, гибридные белки и/или гРНК, используют для трансформации любых прокариотических видов, включая, но не ограничиваясь только ими, археи и бактерии (например, *Bacillus* spp., *Klebsiella* spp., *Streptomyces* spp., *Rhizobium* spp., *Escherichia* spp., *Pseudomonas* spp., *Salmonella* spp., *Shigella* spp., *Vibrio* spp., *Yersinia* spp., *Mycoplasma* spp., *Agrobacterium* spp. и *Lactobacillus* spp.).

В некоторых вариантах для введения нуклеиновых кислот в клетки млекопитающих или ткани-мишени используют стандартные вирусные и невирусные способы переноса генов. Такие способы можно использовать для введения нуклеиновых кислот, кодирующих дезаминазу или гибридный белок по изобретению и необязательно гРНК, в клетки в культуре или в организм-хозяина. Невирусные векторные системы доставки включают ДНК-плазмиды, РНК (например, транскрипт вектора, описанного в настоящем документе), депротенинизированную (голую) нуклеиновую кислоту и нуклеиновую кислоту в составе комплекса с носителем доставки, таким как липосома. Вирусные векторные системы доставки включают ДНК- и РНК-вирусы, которые после доставки в клетку содержат либо эписомальные, либо интегрированные геномы. Неограничивающие примеры включают векторы, использующие каулимовирусы (например, вирус мозаики цветной капусты), геминивирусы (например, вирус золотисто-желтой мозаики фасоли или вирус кукурузного стебля) и РНК-вирусы растений (например, вирус табачной мозаики). Обзор процедур генной терапии см. в статьях Anderson, *Science* 256: 808- 813 (1992); Nabel & Feigner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada и др., in *Current Topics in Microbiology and Immunology*, Doerfler and Bohm (ред.) (1995) и Yu и др., *Gene Therapy* 1:13-26 (1994).

Методы невирусной доставки нуклеиновых кислот включают липофекцию, трансформацию, опосредованную *Agrobacterium*, нуклеофекцию,

микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, поликатион или липид: конъюгаты нуклеиновых кислот, голую ДНК, искусственные вирионы и усиленный агентом захват ДНК. Липофекция описана, например, в патентах США No. № 5049386, 4946787 и 4897355, а реагенты для

5 липофекции выпускаются коммерческими фирмами (например, Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые можно использовать для эффективной липофекции полинуклеотидов с распознаванием рецепторов, включают липиды, описанные в заявках Feigner, WO 91/17424, WO 91/16024. Доставку можно осуществлять в клетки (например, введение *in vitro* или *ex vivo*)

10 либо в ткани-мишени (например, введение *in vivo*). Приготовление комплексов липид/нуклеиновая кислота известно специалистам в данной области техники (см., например, статьи Crystal, *Science* 270:404-410 (1995); Blaese и др., *Cancer Gene Ther.* 2:291- 297 (1995); Behr и др., *Bioconjugate Chem.* 5:382-389 (1994); Remy и др., *Bioconjugate Chem.* 5:647-654 (1994); Gao и др., *Gene Therapy* 2:710-

15 722 (1995); Ahmad и др., *Cancer Res.* 52:4817-4820 (1992); патенты США No. 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028 и 4946787).

Преимущество использования систем на основе вирусов РНК или ДНК для доставки нуклеиновых кислот заключается в применении усовершенствованных

20 процессов для направления вируса на определенные клетки в организме и доставки вирусной полезной нагрузки в ядро. Вирусные векторы можно вводить напрямую пациентам (*in vivo*) или их можно использовать для обработки клеток *in vitro*, а модифицированные клетки можно необязательно вводить пациентам (*ex vivo*). Типичные вирусные системы могут включать ретровирусные,

25 лентивирусные, аденовирусные, аденоассоциированные векторы и векторы вируса простого герпеса для переноса генов. Для интеграции в геном хозяина можно использовать методы переноса генов ретровируса, лентивируса и аденоассоциированного вируса, что в большинстве случаев приводит к длительной экспрессии встроенного трансгена. Кроме того, высокая

30 эффективность трансдукции наблюдалась во многих различных типах клеток и тканях-мишенях.

Тропизм ретровируса можно изменить за счет включения чужеродных белков оболочки, расширяя потенциальную популяцию клеток-мишеней. Лентивирусные векторы представляют собой ретровирусные векторы, которые

способны трансдуцировать или инфицировать неделящиеся клетки и обычно продуцируют высокие вирусные титры. Таким образом, выбор системы переноса ретровирусного гена будет зависеть от ткани-мишени. Ретровирусные векторы состоят из цис-действующих длинных концевых повторов с упаковывающей способностью до 6-10 т.п.о. чужеродной последовательности. Минимальные цис-действующие LTR являются достаточными для репликации и упаковки векторов, которые затем используют для интеграции терапевтического гена в клетку-мишень для обеспечения постоянной экспрессии трансгена. Широко используемые ретровирусные векторы включают векторы на основе вируса лейкоза мышей (MuLV), вируса лейкоза обезьян-гibbonов (GaLV), вируса иммунодефицита обезьян (SIV), вируса иммунодефицита человека (ВИЧ) и их комбинаций, см., например, статьи Buchscher и др., *J. Virol.* 66:2731-2739 (1992); Johann и др., *J. Virol.* 66:1635-1640 (1992); Sommerfelt и др., *Virol.* 176:58-59 (1990); Wilson и др., *J. Virol.* 63:2374-2378 (1989); Miller и др., *J. Virol.* 65:2220-2224 (1991) и заявку PCT/US94/05700.

При использовании в тех случаях, когда предпочтительной является временная экспрессия, можно использовать системы на основе аденовирусов. Векторы на основе аденовируса характеризуются очень высокой эффективностью трансдукции во многих типах клеток и не требуют клеточного деления. При использовании таких векторов были получены высокие титры и уровни экспрессии. Этот вектор можно получать в больших количествах в относительно простой системе. Векторы аденоассоциированного вируса ("AAV") также можно использовать для трансдукции клеток нуклеиновыми кислотами-мишенями, например, при продуцировании нуклеиновых кислот и пептидов *in vitro*, а также для процедур генной терапии *in vivo* и *ex vivo* (см., например, статьи West и др., *Virology* 160:38-47 (1987); патент США No. 4797368; заявку WO 93/24641; статьи Katin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Конструкты рекомбинантных векторов AAV описаны в ряде публикаций, включая патент США No. 5173414; статьи Tratschin и др., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, и др., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984) и Samulski и др., *J. Virol.* 63:3822-3828 (1989). Пакующие клетки обычно используют для формирования вирусных частиц, способных инфицировать клетку-хозяина.

Такие клетки включают клетки 293, которые пакут аденовирус, и клетки ψ J2 или клетки RA317, которые пакут ретровирус.

Вirusные векторы, используемые в генной терапии, обычно конструируют при получении клеточной линии, которая пакует 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 заменены на экспрессионную кассету для полинуклеотидов, подлежащих экспрессии.

Отсутствующие вирусные функции обычно поставляются в виде трансактивной клеточной линии клеток-упаковщиков. 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 клеточную линию, которая содержит хелперную плазмиду, кодирующую другие гены AAV, а именно гер и сар, но лишенные последовательностей ITR.

Линию клеток также можно инфицировать аденовирусом в качестве хелпера. Вирус-хелпер способствует репликации вектора AAV и экспрессии генов AAV из плазмиды-хелпера. Хелперная плазида не упакована в значительных количествах из-за отсутствия последовательностей ITR.

Загрязнение аденовирусом можно уменьшить, например, с помощью термической обработки, к которой аденовирус более чувствителен, чем AAV. Дополнительные способы доставки нуклеиновых кислот в клетки известны специалистам в данной области техники. См., например, патент США 20030087817, включенный в настоящее описание в качестве ссылки.

В некоторых вариантах клетку-хозяина временно или не временно трансфектируют одним или более векторов, описанных в настоящем документе. В некоторых вариантах клетку трансфектируют естественным образом, то есть аналогичным способом, наблюдаемому у субъекта. В некоторых вариантах трансфектированную клетку отбирают из организма субъекта.

В некоторых вариантах клетка, которую трансфектируют, представляет собой эукариотическую клетку. В некоторых вариантах эукариотическая клетка представляет собой клетку животного (например, млекопитающих, насекомых, рыб, птиц и рептилий). В некоторых вариантах клетка, которую трансфектируют, представляет собой клетку человека. В некоторых вариантах

клетка, которую трансфектируют, представляет собой клетку кроветворной системы, такую как иммунная клетка (т.е. клетка врожденного или приобретенного иммунитета), включая, не ограничиваясь только ими, В клетку, Т клетку, природную клетку-киллер (NK), плюрипотентную стволовую клетку, индуцированную плюрипотентную (недифференцированную) стволовую клетку, Т клетку с химерным антигенным рецептором (CAR-T), моноцит, макрофаг и дендритную клетку.

В некоторых вариантах клетку получают из клеток, взятых у субъекта, таких как клеточная линия. В некоторых вариантах клетка или клеточная линия являются прокариотическими. В некоторых вариантах клетка или клеточная линия являются эукариотическими. В дополнительных вариантах клетка или клеточная линия происходят от видов насекомых, птиц, растений или грибов. В некоторых вариантах клетка или клеточная линия могут принадлежать млекопитающему, такому как, например, человек, обезьяна, мышь, КРС, свинья, коза, хомяк, крыса, кошка или собака. В данной области техники известно большое разнообразие клеточных линий для тканевых культур. Типичные клеточные линии включают, но не ограничиваясь только ими, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLaS3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEK_n, HEK_a, MiaPaCell, Panel, PC-3, TFI, CTLL-2, CIR, Rat6, CVI, RPTE, A1O, T24, 182, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, клетки 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, эпителиальные клетки почек обезьяны BS-C-1, эмбриональные клетки фибробластов мыши BALB/3T3, 3T3 Swiss, 3T3-L1, эмбриональные клетки фибробластов человека 132-d5; клетки фибробластов мыши 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, клетки BCP-I, 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, клетки lurkat, клетки IY, клетки K562, Ku812, KCL22, KGI, KYOI, 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, клеточные линии OPCN/OPCT, Peer, PNT-1A/ PNT 2, RenCa, RIN-5F, RMA/RMAS, клетки Saos-2, Sf-9, SkBr3, T2, T-47D, T84, клеточная линия THP1, U373, U87, U937, VCaP, клетки Vero, WM39, WT-49, X63, YAC-1, YAR и их трансгенные варианты. Клеточные линии выпускаются множеством фирм, известных специалистам в данной области техники (см., например, American Type Culture Collection (ATCC) (Manassas, Va.)).

В некоторых вариантах клетку, трансфектированную одним или более векторов, описанными в настоящем документе, используют для получения новой клеточной линии, содержащей одну или более последовательностей, полученных из вектора. В некоторых вариантах клетку, временно трансфектированную гибридным белком по изобретению и, необязательно, гРНК, или рибонуклеопротеиновым комплексом по изобретению и модифицированную за счет активности гибридного белка или рибонуклеопротеинового комплекса, используют для получения новой клеточной линии, содержащей модифицированные клетки, но лишенные какой-либо другой экзогенной последовательности. В некоторых вариантах клетки, временно или не-временно трансфектированные одним или более векторами, описанными в настоящем документе, или клеточные линии, полученные из таких клеток, используют для оценки одного или более исследуемых соединений.

В некоторых вариантах один или более векторов, описанных в настоящем документе, используют для получения трансгенного животного, отличающегося от человека или трансгенного растения. В некоторых вариантах трансгенное животное представляет собой насекомое. В дополнительных вариантах насекомое представляет собой насекомое-вредитель, такое как комар или клещ. В некоторых вариантах насекомое представляет собой вредителя растений, такого как кукурузный жук или совка травяная. В некоторых вариантах трансгенное животное представляет собой птицу, такую как курица, индейка, гусь или утка. В некоторых вариантах трансгенное животное представляет собой млекопитающее, такое как человек, мышь, крыса, хомяк, обезьяна, примат, кролик, свинья, КРС, лошадь, коза, овца, кошка или собака.

VI. Варианты и фрагменты полипептидов и полинуклеотидов

В настоящем изобретении предлагаются новые адениндезаминазы, которые проявляют активность на молекулах ДНК, аминокислотная последовательность которых указана в виде SEQ ID NO: 1-10 и 399-441, их активные варианты или
5 фрагменты и кодирующие их полинуклеотиды.

Хотя активность варианта или фрагмента можно изменить по сравнению с исследуемым полинуклеотидом или полипептидом, вариант и фрагмент должны сохранять функциональную активность исследуемого полинуклеотида или полипептида. Например, вариант или фрагмент может проявлять повышенную
10 активность, пониженную активность, может характеризоваться другим спектром активности или любым другим изменением активности по сравнению с исследуемым полинуклеотидом или полипептидом.

Фрагменты и варианты дезаминаз по изобретению, обладающие активностью адениндезаминазы, сохраняют указанную активность, если они
15 являются частью гибридного белка, дополнительно содержащего ДНК-связывающий полипептид или его фрагмент.

Термин «фрагмент» относится к части полинуклеотидной или полипептидной последовательности по изобретению. «Фрагменты» или «биологически активные части» включают полинуклеотиды, содержащие
20 достаточное количество смежных нуклеотидов для сохранения биологической активности (т.е. дезаминазной активности на нуклеиновых кислотах).

«Фрагменты» или «биологически активные части» включают полипептиды, содержащие достаточное количество смежных аминокислотных остатков для сохранения биологической активности. Фрагменты дезаминаз, раскрытых в
25 настоящем документе, включают более короткие фрагменты по сравнению с полноразмерными последовательностями, из-за использования альтернативного расположенного ниже стартового сайта. В некоторых вариантах биологически активная часть дезаминазы представляет собой полипептид, который содержит, например, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160 или
30 более смежных аминокислотных остатков любой из последовательностей SEQ ID NO: 1-10 и 399-441 или их варианта. Такие биологически активные части можно получить рекомбинантными методами и оценивать их активность.

В основном «варианты» означают в значительной степени сходные последовательности. Для полинуклеотидов вариант включает делецию и/или

включение одного или нескольких нуклеотидов в одном или нескольких внутренних сайтах в нативном полинуклеотиде и/или замену одного или нескольких нуклеотидов в одном или нескольких сайтах в нативном полинуклеотиде. Используемый в настоящем документе термин «нативный» полинуклеотид или полипептида или его «дикий тип» содержит природную нуклеотидную последовательность или аминокислотную последовательность, соответственно. Для полинуклеотидов, консервативные варианты включают те последовательности, которые из-за вырожденности генетического кода кодируют природную аминокислотную последовательность исследуемого гена.

Природные аллельные варианты, такие как указанные выше, можно идентифицировать с использованием хорошо известных методов молекулярной биологии, таких как, например, полимеразная цепная реакция (ПЦР) и методы гибридизации, как описано ниже. Варианты полинуклеотидов также включают синтетические полинуклеотиды, такие как полученные, например, с помощью сайт-направленного мутагенеза, но которые по-прежнему кодируют исследуемый полипептид или полинуклеотид. Как правило, варианты конкретного полинуклеотида, описанные в настоящем документе, являются идентичными по меньшей мере на 40%, по меньшей мере на 45%, по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более по сравнению с конкретным полинуклеотидом по данным анализа при выравнивании последовательностей с использованием программ и параметров, как указано в данном контексте.

Варианты конкретного полинуклеотида, раскрытого в настоящем документе (т.е. эталонного полинуклеотида), также можно оценивать при сравнении идентичности последовательностей в процентах между полипептидом, кодируемым вариантом полинуклеотида, и полипептидом, кодируемым эталонным полинуклеотидом. Идентичность последовательностей в процентах между любыми двумя полипептидами можно рассчитывать с использованием программ выравнивания последовательностей и параметров, описанных в

данном контексте. Когда любую заданную пару полинуклеотидов, раскрытых в настоящем документе, оценивают при сравнении идентичности в процентах последовательности, общей для двух полипептидов, которые они кодируют, идентичность последовательности в % между двумя кодируемыми

5 полипептидами составляет по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%,
10 по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или более.

В конкретных вариантах раскрытые в настоящем документе полинуклеотиды кодируют адениндезаминазу, содержащую аминокислотную последовательность, которая идентична по меньшей мере на 40%, по меньшей мере на 45%, по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 81%, по меньшей мере на 82%, по меньшей мере на 83%, по меньшей мере на 84%, по меньшей мере на 85%, по меньшей мере на 86%, по меньшей мере на 87%, по меньшей мере на 88%, по меньшей мере на 89%, по меньшей мере на 90%, по меньшей мере на 91%, не менее 92%, не менее 93%, не менее 94%, не менее 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более по сравнению с аминокислотной последовательностью любой из последовательностей SEQ ID NO: 1-10 и 399-
20 441.
25

Биологически активный вариант адениндезаминазы по изобретению может отличаться всего на 1-15 аминокислотных остатков, всего на 1-10, например, 6-10, всего на 5, всего на 4, всего на 3, всего на 2 или всего на 1 аминокислотный остаток. В конкретных вариантах полипептиды содержат N-концевое или C-
30 концевое укорочение, которое может содержать по меньшей мере делецию 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 аминокислот или более, либо с N-, либо с C-конца полипептида. В некоторых вариантах полипептиды содержат внутреннюю делецию, которая может включать по меньшей мере делецию 1, 2, 3, 4, 5, 6, 7, 8,

9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60
аминокислот или более.

Предполагалось, что предложенные в данном контексте дезаминазы можно модифицировать и конструировать варианты белков и полинуклеотидов.

5 Разработанные в ручную изменения, можно вводить с использованием методов сайт-направленного мутагенеза. В некоторых вариантах также можно было идентифицировать нативные, еще неизвестные или еще не идентифицированные полинуклеотиды и/или полипептиды, в структурном и/или функциональном отношении родственные раскрытым в данном контексте последовательностям, 10 которые включены в объем настоящего изобретения. Консервативные аминокислотные замены можно осуществлять в неконсервативных областях, которые не изменяют функцию полипептида в качестве адениндезаминазы. В некоторых вариантах были внесены модификации, улучшающие адениндезаминазную активность дезаминазы.

15 Варианты полинуклеотидов и белков также охватывают последовательности и белки, полученные в результате мутагенных и рекомбиногенных процедур, таких как перестановка экзонов ДНК. С использованием такой процедуры, одну или более различных дезаминаз, раскрытых в настоящем документе (например, SEQ ID NO: 1-10 и 399-441), 20 использовали для конструирования новой адениндезаминазы, обладающей требуемыми свойствами. Таким образом, были созданы библиотеки рекомбинантных полинуклеотидов из популяции полинуклеотидов с родственными последовательностями, включающими области последовательностей, которые характеризуются значительной идентичностью 25 последовательностей и которые можно гомологично рекомбинировать *in vitro* или *in vivo*. Например, используя этот подход, в последовательностях дезаминазы можно осуществлять перестановку мотивов последовательностей, кодирующих исследуемый домен, представленных в настоящем документе, и других впоследствии идентифицированных генов дезаминазы, при этом 30 получали новый ген, кодирующий исследуемый белок с улучшенным свойством, таким как повышенная *K_m* в случае фермента. Стратегии для перестановки экзонов ДНК известны в данной области техники, см., например, статьи Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994); Stemmer, *Nature* 370:389-391 (1994); Cramer и др., *Nature Biotech.* 15:436-438 (1997); Moore и др.,

J. Mol. Biol. 272:336-347 (1997); Zhang и др., *Proc. Natl. Acad. Sci. USA* 94:4504-4509 (1997); Cramer и др., *Nature* 391:288-291 (1998) и патенты США No 5605793 и 5837458. Нуклеиновая кислота «после процедуры перестановки

экзонов» представляет собой нуклеиновую кислоту, полученную с помощью

5 процедуры перестановки, такой как любая процедура перестановки, изложенная в настоящем документе. Нуклеиновые кислоты «после процедуры перестановки экзонов» получают при рекомбинации (физической или виртуальной) двух или более нуклеиновых кислот (или символьных строк), например, искусственным и, необязательно, рекурсивным способом (в обратном порядке). Как правило, в

10 процессах перестановки используется один или несколько этапов скрининга для идентификации исследуемых нуклеиновых кислот; причем этот этап скрининга можно выполнять до или после любого этапа рекомбинации. В некоторых (но не во всех) вариантах перестановку требуется выполнить в ходе нескольких циклов рекомбинации перед селекцией, чтобы увеличить разнообразие пула,

15 подлежащего скринингу. Общий процесс рекомбинации и селекции необязательно повторяют рекурсивно. В зависимости от контекста, перестановка может относиться к общему процессу рекомбинации и селекции или, наоборот, может относиться просто к рекомбинационным этапам общего процесса.

Используемый в данном контексте термин «идентичность

20 последовательности» или «идентичность» в контексте двух полинуклеотидных или полипептидных последовательностей относится к остаткам в двух последовательностях, которые являются одинаковыми при выравнивании для максимального соответствия в указанном окне сравнения. Когда идентичность последовательности в процентах используется в отношении белков,

25 предполагается, что положения остатков, которые не идентичны, часто отличаются консервативными аминокислотными заменами, то есть, когда аминокислотные остатки заменены другими аминокислотными остатками с аналогичными химическими свойствами (например, зарядом или гидрофобностью) и поэтому не изменяют функциональных свойств молекулы.

30 Когда последовательности отличаются консервативными заменами, идентичность последовательности в процентах можно скорректировать в сторону увеличения, чтобы скорректировать консервативный характер замены. Известно, что последовательности, отличающиеся такими консервативными заменами, характеризуются «сходством последовательностей» или «сходством».

Методы для выполнения этой корректировки хорошо известны специалистам в данной области техники. Как правило, такая процедура включает оценку консервативной замены как частичного, а не полного несоответствия, тем самым увеличивая идентичность последовательности в процентах. Так, например, если

5 идентичной аминокислоте присваивается оценка 1, а неконсервативной замене присваивается нулевая оценка, консервативной замене присваивается оценка от нуля до 1. Оценку консервативных замен рассчитывают, например, как описано в программе PC/GENE (Intelligenetics, Маунтин-Вью, Калифорния).

Используемый в данном контексте термин «идентичность

10 последовательности в процентах» означает значение, определенное при сравнении двух оптимально выровненных последовательностей в окне сравнения, где часть полинуклеотидной последовательности в окне сравнения может содержать добавления или делеции (т.е. пробелы) по сравнению с эталонной последовательностью (которая не содержит добавлений или делеций)

15 для оптимального выравнивания двух последовательностей. Процент рассчитывают при определении числа положений, в которых идентичное основание нуклеиновой кислоты или аминокислотный остаток встречается в обеих последовательностях, при этом получают число совпадающих положений, при делении числа совпадающих положений на общее число положений в окне

20 сравнения, и при умножении результата на 100, при этом получают идентичность последовательности в процентах.

Если не указано иное, приведенные в данном контексте значения идентичности/сходства последовательностей относятся к значениям, полученным с использованием приложения версии 10 GAP с использованием

25 следующих параметров: % идентичности и % сходства для нуклеотидной последовательности с использованием штрафа за открытие гэпа (GAP Weigth), равного 50, и штрафа за удаление гэпа (Length Weigth), равного 3, и файла nwsgapdna.cmp. матрица оценок; % идентичности и % сходства для аминокислотной последовательности с использованием GAP Weigth, равного 8,

30 и Length Weigth, равного 2, и матрицы оценки BLOSUM62; или любой эквивалентной этим параметрам программы. «Эквивалентной программой» называется любая программа сравнения последовательностей, которая для любых двух рассматриваемых последовательностей создает выравнивание, в котором наблюдаются идентичные совпадения нуклеотидов или

аминокислотных остатков и идентичный процент идентичности последовательностей по сравнению с соответствующим выравниванием, созданным приложением GAP версии 10.

Две последовательности являются «оптимально выровненными», когда они
5 выровнены для оценки сходства с использованием определенной матрицы аминокислотных замен (например, BLOSUM62), штрафа за наличие гэпа и штрафа за удлинение гэпа, чтобы получить максимально возможный балл для этой пары последовательностей. Матрицы аминокислотных замен и их
10 использование для количественной оценки сходства между двумя последовательностями хорошо известны в данной области техники и описаны, например, в статье Dayhoff и др., “A model of evolutionary change in proteins.”, “Atlas of Protein Sequence and Structure,”», т. 5, Suppl. 3 (ред. MO Dayhoff), стр. 345–352 (1978), и Henikoff и др., *Proc. Natl. Acad. Sci. USA* 89:10915-10919 (1992). Матрицу BLOSUM62 часто используют в качестве матрицы замены по
15 умолчанию в протоколах выравнивания последовательностей. Штраф за существование гэпа налагается за введение одного аминокислотного гэпа в одну из выровненных последовательностей, а штраф за удлинение гэпа налагается за каждое дополнительное пустое аминокислотное положение, вставленное в уже открытый гэп. Выравнивание определяется положениями аминокислот каждой
20 последовательности, с которых начинается и заканчивается выравнивание, и, необязательно, вставкой гэпа или нескольких гэпов в одну или обе последовательности, чтобы получить максимально возможную оценку. В то время как оптимальное выравнивание и оценку можно выполнять вручную, этот процесс облегчается за счет использования реализованного компьютером
25 алгоритма выравнивания, например, BLAST 2.0 с пробелами, описанного в статье Altschul и др. *Nucleic Acids Res.* 25:3389-3402 (1997), и который опубликован на веб-сайте Национального центра биотехнологической информации (www.ncbi.nlm.nih.gov). Оптимальные выравнивания, включая
множественные выравнивания, можно подготовить с использованием, например,
30 программы PSI-BLAST, доступной на сайте www.ncbi.nlm.nih.gov и описанной в статье Altschul и др., *Nucleic Acids Res.* 25:3389-3402 (1997).

Что касается аминокислотной последовательности, которая оптимально выровнена с эталонной последовательностью, аминокислотный остаток «соответствует» положению в эталонной последовательности, с которым этот

остаток спаривается при выравнивании. «Положение» обозначается числом, которое последовательно идентифицирует каждую аминокислоту в эталонной последовательности на основе ее положения с N-конца. Из-за делеций, вставок, усечений, слияний и т. д., которые необходимо учитывать при определении оптимального выравнивания, в общем случае число аминокислотных остатков в исследуемой последовательности, определяемое простым подсчетом от N-конца, необязательно будет совпадать с номером соответствующего положения в эталонной последовательности. Например, в случае делеции в выровненной исследуемой последовательности, будет отсутствовать аминокислота, соответствующая положению в эталонной последовательности в месте делеции. При наличии вставки в выровненной эталонной последовательности эта вставка не будет соответствовать ни одному положению аминокислоты в эталонной последовательности. В случае усечения или слияния могут присутствовать участки аминокислот либо в эталонной, либо в выровненной последовательности, которые не соответствуют ни одной аминокислоте в соответствующей последовательности.

VII. Антитела

Антитела к деаминазам, гибридным белкам или рибонуклеопротеинам, включающим дезаминазы по настоящему изобретению, и дезаминазы с аминокислотной последовательностью, указанной в виде любой из SEQ ID NO: 1-10 и 399-441, или их активные варианты или фрагменты, также включены в объем настоящего изобретения. Способы получения антител хорошо известны в данной области техники (см., например, Harlow и Lane Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1988) и патент США. No. 4196265). Эти антитела можно использовать в наборах для детектирования и выделения дезаминаз или гибридных белков или рибонуклеопротеинов, содержащих описанные в данном контексте дезаминазы. Таким образом, в настоящем изобретении предлагаются наборы, содержащие антитела, которые специфически связываются с полипептидами или рибонуклеопротеинами, описанными в настоящем документе, включая, например, полипептиды, содержащие последовательность, по меньшей мере на 85% идентичную любой из SEQ ID NO: 1-10 и 399-441.

VIII. Системы и рибонуклеопротеиновые комплексы для связывания и/или модификации исследуемой последовательности-мишени и способы их получения

В настоящем изобретении предлагается система, которая направлена на последовательность нуклеиновой кислоты и модифицирует последовательность-мишень нуклеиновой кислоты. В некоторых вариантах РНК-управляемый ДНК-связывающий полипептид, такой как RGN, и гРНК отвечают за направление

5 рибонуклеопротеинового комплекса на исследуемую последовательность нуклеиновой кислоты; а полипептид дезаминазы, гибридизированный с RGDBP, отвечает за модификацию последовательности-мишени нуклеиновой кислоты при замене А на N. В некоторых вариантах в результате действия дезаминазы происходит замена А на G. Направляющая РНК гибридизуется с исследуемой

10 последовательностью-мишенью, а также образует комплекс с РНК-направляемым ДНК-связывающим полипептидом, тем самым направляя РНК-направляемый ДНК-связывающий полипептид на связывание с последовательностью-мишенью. РНК-управляемый ДНК-связывающий полипептид представляет собой один домен гибридного белка; второй домен

15 представляет собой описанную в данном контексте дезаминазу. В некоторых вариантах РНК-управляемый ДНК-связывающий полипептид представляет собой RGN, такой как Cas9. Другие примеры РНК-управляемых ДНК-связывающих полипептидов включают RGN, описанные в публикациях международных патентных заявок № WO 2019/236566 и WO 2020/139783. В некоторых вариантах

20 РНК-управляемый ДНК-связывающий полипептид представляет собой полипептид CRISPR-Cas типа II или его активный вариант или фрагмент. В некоторых вариантах РНК-управляемый ДНК-связывающий полипептид представляет собой полипептид CRISPR-Cas типа V или его активный вариант или фрагмент. В некоторых вариантах РНК-управляемый ДНК-связывающий

25 полипептид представляет собой полипептид CRISPR-Cas типа VI. В некоторых вариантах для ДНК-связывающего домена гибридного белка не требуется направляющая РНК, такая как нуклеаза цинковых пальцев, TALEN или полипептид мегануклеазы. В некоторых вариантах нуклеазная активность ДНК-связывающего домена является частично или полностью инактивированной. В

30 дополнительных вариантах РНК-управляемый ДНК-связывающий полипептид содержит аминокислотную последовательность RGN, такую как, например, APG07433.1 (SEQ ID NO: 41), или ее активный вариант или фрагмент, такой как никаза nAPG07433.1. (SEQ ID NO: 42) или другие варианты RGN никазы, описанные в разделе Примеры (SEQ ID NO: 52-59, 61, 397 и 398).

В некоторых вариантах система для связывания и модификации исследуемой последовательности-мишени, предлагаемая в настоящем документе, представляет собой рибонуклеопротеиновый комплекс, который включает по меньшей мере одну молекулу РНК, связанную по меньшей мере с одним белком. Предлагаемые в данном контексте рибонуклеопротеиновые комплексы содержат по меньшей мере одну направляющую РНК в качестве РНК-компонента и гибридный белок, содержащий дезаминазу по изобретению и РНК-управляемый ДНК-связывающий полипептид в качестве белкового компонента. В некоторых вариантах рибонуклеопротеиновый комплекс выделяли и очищали из клетки или организма, которые были трансфектированы полинуклеотидами, кодирующими гибридный белок и направляющую РНК, и культивировали в условиях, обеспечивающих экспрессию гибридного белка и направляющей РНК.

В различных вариантах предлагаются рибонуклеопротеиновые комплексы, содержащие любой из гибридных белков, описанных в настоящем документе, и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка. Например, в настоящем документе предлагается рибонуклеопротеиновый комплекс, содержащий гибридный белок, включающий дезаминазу, содержащую аминокислотную последовательность, идентичную по меньшей мере на 80% по сравнению последовательностью SEQ ID NO: 407. В другом случае рибонуклеопротеиновый комплекс, содержащий гибридный белок, включающий дезаминазу, содержащую аминокислотную последовательность, которая идентична по меньшей мере на 80% по сравнению с последовательностью SEQ ID NO: 399. В еще одном примере предлагается рибонуклеопротеиновый комплекс, содержащий гибридный белок, включающий дезаминазу, содержащую аминокислотную последовательность, идентичную по меньшей мере на 80% по сравнению с последовательностью SEQ ID NO: 405. В некоторых вариантах рибонуклеопротеиновых комплексов, описанных выше, гибридный белок содержит RGN, выбранную из CasX, CasY, C2c1, C2c2, C2c3, GeoCas9, aSpCas9, SaCas9, Nme2Cas9, CjCas9, Cas12a (панее известную как Cpf1), Cas12b, Cas12g, Cas12h, Cas12i, aLbCas12a, AsCas12a, CasMINI, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с кольцевой перестановкой, домен Argonaute (Ago), SmacCas9, домен Spy-macCas9 или RGN, включающую аминокислотную последовательность, которая представлена в

любой из SEQ ID NO: 41, 60, 366 или 368. В некоторых вариантах
рибонуклеопротеиновый комплекс содержит никазу, характеризующуюся
аминокислотной последовательностью, идентичной по меньшей мере на 95% по
сравнению с любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398,
5 гибридизированной с дезаминазой, содержащей аминокислотную
последовательность, идентичную по меньшей мере на 80% SEQ ID NO: 407. В
некоторых вариантах рибонуклеопротеиновый комплекс содержит никазу с
аминокислотной последовательностью, идентичность которой составляет по
меньшей мере 95% по сравнению с любой из последовательностей SEQ ID NO:
10 42, 52-59, 61, 397 и 398, гибридизированную с дезаминазой, содержащей
аминокислотную последовательность, идентичную по меньшей мере на 80% по
сравнению с последовательностью SEQ ID NO: 399. В некоторых вариантах
рибонуклеопротеиновый комплекс содержит никазу, содержащую
аминокислотную последовательность, идентичную по меньшей мере на 95% по
15 сравнению с последовательностью любой из последовательностей SEQ ID NO:
42, 52- 59, 61, 397 и 398, гибридизированной с дезаминазой, содержащей
аминокислотную последовательность, по меньшей мере на 80% идентичную SEQ
ID NO: 405. В некоторых вариантах рибонуклеопротеиновый комплекс содержит
никазу Cas9, гибридизированную с дезаминазой, содержащей аминокислотную
20 последовательность, идентичную по меньшей мере на 80% по сравнению с
последовательностью SEQ ID NO: 407. В некоторых вариантах
рибонуклеопротеиновый комплекс содержит никазу Cas9, гибридизированную с
дезаминазой, содержащей аминокислотную последовательность, идентичную по
меньшей мере на 80% по сравнению с последовательностью SEQ ID NO: 399. В
25 некоторых вариантах рибонуклеопротеиновый комплекс содержит никазу Cas9,
гибридизированную с дезаминазой, содержащую аминокислотную
последовательность, идентичную по меньшей мере на 80% по сравнению с
последовательностью SEQ ID NO: 405. Никаза Cas9 может представлять собой
любую никазу Cas9, описанную в публикации патента PCT № WO 2020181195,
30 содержание которой в полном объеме включено в настоящее описание в
качестве ссылки. В различных вариантах, описанных в настоящем документе,
рибонуклеопротеиновый комплекс может также содержать гРНК, описанные в
настоящем документе.

В настоящем изобретении предлагаются способы получения дезаминазы, гибридного белка или комплекса гибридного белка с рибонуклеопротеином. Такие способы включают культивирование клетки, содержащей нуклеотидную последовательность, кодирующую дезаминазу, гибридный белок и, в некоторых вариантах, нуклеотидную последовательность, кодирующую направляющую РНК, в условиях, при которых дезаминаза или гибридный белок (и в некоторых вариантах направляющая РНК) являются экспрессированными. Затем дезаминазу, гибридный белок или гибридный рибонуклеопротеин можно выделить и очистить из лизата культивируемых клеток.

Способы очистки дезаминазы, гибридного белка или комплекса гибридного рибонуклеопротеина из лизата биологического образца известны в данной области техники (например, эксклюзионная и/или аффинная хроматография, 2D-ЭФ в ПААГ, ВЭЖХ, обращенно-фазовая хроматография, иммунопреципитация). В конкретных способах дезаминазу или гибридный белок получают рекомбинантным методом и они содержат метку очистки, которая упрощает его очистку, включая, но не ограничиваясь только ими, глутатион-S-трансферазу (GST), хитин-связывающий белок (CBP), мальтозо-связывающий белок, тиоредоксин (TRX), поли(NANP), метку для очистки тандемной аффинной хроматографией (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, биотинилированный карбоксилированный белок-носитель (BCCP) и кальмодулин. Как правило, меченые дезаминазу, гибридный белок или гибридный рибонуклеопротеиновый комплекс очищают с использованием иммунопреципитации или других подобных способов, известных в данной области техники.

«Выделенный» или «очищенный» полипептид или его биологически активная часть в основном или в значительной степени не содержат компоненты, которые обычно присутствуют или взаимодействуют с полипептидом, присутствующим в его окружающей среде. Таким образом, выделенный или очищенный полипептид практически не содержит другой клеточный материал или культуральную среду, если его получали рекомбинантными методами, или практически не содержит химических предшественников или других химических веществ, если его синтезировали химическим способом. Белок, практически не содержащий клеточный материал, включает препараты белка, содержащие менее

30%, менее 20%, менее 10%, менее 5% или менее 1% (в расчете на сухую массу) примесного белка. Если белок по изобретению или его биологически активную часть получают рекомбинантным методом, оптимальная культуральная среда содержит менее 30%, менее 20%, менее 10%, менее 5% или менее 1% (в расчете на сухую массу) химических предшественников или небелковых химических веществ, не представляющих интерес.

Конкретные способы, предлагаемые в настоящем документе для связывания и/или расщепления исследуемой последовательности-мишени, включают использование рибонуклеопротеинового комплекса. В некоторых вариантах сборку рибонуклеопротеинового комплекса осуществляют в условиях *in vitro*. Сборку рибонуклеопротеинового комплекса *in vitro* можно осуществлять с использованием любого метода, известного в данной области техники, согласно которому полипептид RGDBP или содержащий его гибридный белок контактирует с направляющей РНК в условиях, обеспечивающих связывание полипептида RGDBP или содержащего его гибридного белка с направляющей РНК. Используемые в данном контексте термины «контакт», «контактирование», «контактированный» относятся к смешиванию компонентов требуемой реакции в условиях, пригодных для проведения требуемой реакции. В некоторых вариантах описанных способов модификации молекулы ДНК-мишени, стадию контактирования проводят в условиях *in vitro*. В некоторых вариантах стадию контактирования проводят в условиях *in vivo*. В некоторых вариантах стадию контактирования осуществляют в организме субъекта (например, человека или животного, отличающегося от человека). В некоторых вариантах стадию контактирования проводят в клетке, такой как человеческая клетка или клетка животного, отличающегося от человека. Полипептид RGDBP или содержащий его гибридный белок можно выделить и очистить из биологического образца, клеточного лизата или культуральной среды, полученной в результате трансляции *in vitro* или синтезировать химическим методом. Направляющую РНК можно выделить и очистить из биологического образца, клеточного лизата или культуральной среды, после транскрипции *in vitro* или химического синтеза. Полипептид RGDBP или содержащий его гибридный белок, и направляющую РНК, можно смешивать в растворе (например, буферном солевом растворе), чтобы обеспечить сборку рибонуклеопротеинового комплекса *in vitro*.

IX. Способы модификации последовательности-мишени

В настоящем изобретении предлагаются способы модификации исследуемой молекулы нуклеиновой кислоты-мишени (например, молекулы ДНК-мишени). Способы включают доставку гибридного белка, содержащего ДНК-связывающий полипептид и по меньшей мере одну дезаминазу по изобретению, или кодирующий их полинуклеотид, в последовательность-мишень или в клетку, органеллу или эмбрион, содержащие последовательность-мишень. В некоторых вариантах способы включают доставку системы, содержащей по меньшей мере одну направляющую РНК или кодирующий ее полинуклеотид, и по меньшей мере один гибридный белок, содержащий по меньшей мере одну дезаминазу по изобретению и РНК-управляемый ДНК-связывающий полипептид или кодирующий их полинуклеотид, в последовательность-мишень или клетку, органеллу или эмбрион, содержащие последовательность-мишень. В некоторых вариантах гибридный белок содержит любую из аминокислотных последовательностей SEQ ID NO: 1-10 и 399-441 или ее активный вариант или фрагмент.

В некоторых вариантах способы включают контактирование молекулы ДНК с (а) гибридным белком, содержащим дезаминазу и РНК-управляемый ДНК-связывающий полипептид, такой как, например, домен Cas9, нуклеазо-неактивный или никазо-неактивный; и (б) гРНК, направляющий гибридный белок по п. (а) на нуклеотидную последовательность-мишень молекулы ДНК; при этом молекула ДНК контактирует с гибридным белком и гРНК в количестве, эффективном, и в условиях, пригодных для дезаминирования азотистого основания. В некоторых вариантах молекула ДНК-мишень содержит последовательность, связанную с заболеванием или нарушением, и при этом дезаминирование азотистого основания приводит к образованию последовательности, не связанной с заболеванием или нарушением. В некоторых вариантах заболевание или расстройство поражает животных. В дополнительных вариантах заболевание или нарушение поражает млекопитающих, таких как человек, КРС, лошади, собаки, кошки, козы, овцы, свиньи, обезьяны, крысы, мыши или хомяки. В некоторых вариантах последовательность-мишень ДНК находится в аллеле культурного растения, где конкретный аллель исследуемого признака приводит формированию растения с меньшей агрономической ценностью. Дезаминирование азотистого основания приводит к образованию

аллеля, который улучшает признак и увеличивает агрономическую ценность растения.

В тех вариантах осуществления, в которых способ включает доставку полинуклеотида, кодирующего направляющую РНК и/или гибридный белок, клетку или эмбрион затем можно культивировать в условиях, в которых экспрессируется направляющая РНК и/или гибридный белок. В различных вариантах способ включает контактирование последовательности-мишени с рибонуклеопротеиновым комплексом, содержащим гРНК и гибридный белок (который содержит дезаминазу по изобретению и ДНК-связывающий полипептид, управляемый РНК). В некоторых вариантах способ включает введение рибонуклеопротеинового комплекса по изобретению в клетку, органеллу или эмбрион, содержащие последовательность-мишень. Рибонуклеопротеиновый комплекс по изобретению может представлять собой выделенный и очищенный из биологического образца комплекс, продуцированный рекомбинантным методом, и затем очищенный или полученный после сборки *in vitro*, как описано в данном контексте. В тех вариантах, в которых рибонуклеопротеиновый комплекс, контактирующий с последовательностью-мишенью, или клеточной органеллой или эмбрионом, был собран *in vitro*, способ может дополнительно включать сборку комплекса *in vitro* до контактирования с последовательностью-мишенью, клеткой, органеллой или эмбрионом.

Очищенный или собранный *in vitro* рибонуклеопротеиновый комплекс по изобретению можно вводить в клетку, органеллу или эмбрион с использованием любого метода, известного в данной области техники, включая, но не ограничиваясь только этим, электропорацию. В некоторых вариантах гибридный белок, содержащий дезаминазу по изобретению и РНК-управляемый ДНК-связывающий полипептид и полинуклеотид, кодирующий или содержащий направляющую РНК (гРНК), вводят в клетку, органеллу или эмбрион с использованием любого метода, известного в данной области техники (например, электропорация).

При доставке в последовательность-мишень или клетку, органеллу или эмбрион, содержащие последовательность-мишень, или при контактировании с ними, направляющая РНК (гРНК) направляет гибридный белок на связывание с последовательностью-мишенью специфичным для последовательности образом.

Последовательность-мишень впоследствии можно модифицировать с использованием деаминазного домена гибридного белка. В некоторых вариантах связывание этого гибридного белка с последовательностью-мишенью приводит к модификации нуклеотида, соседнего с последовательностью-мишенью.

- 5 Нуклеотидные основания, соседние с последовательностью-мишенью, которую модифицируют дезаминазой, могут включать 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 или 100 пар оснований от 5'- или 3'-конца последовательности-мишени. При использовании гибридного белка, содержащего дезаминазу по изобретению и
- 10 РНК-управляемый ДНК-связывающий полипептид, можно вводить направленные мутации с заменой А на N и предпочтительно направленные мутации с заменой А на G в молекуле ДНК-мишени.

- В некоторых вариантах описанных способов модификации молекулы ДНК-мишени стадию контактирования проводят в условиях *in vitro*. В конкретных
- 15 вариантах стадию контактирования проводят в условиях *in vivo*. В некоторых вариантах стадию контактирования осуществляют в организме субъекта (например, человека или животного, отличающегося от человека). В некоторых вариантах стадию контактирования проводят в клетке, такой как человеческая клетка или клетка животного, отличающегося от человека.

- 20 Способы оценки связывания гибридного белка с последовательностью-мишенью известны в данной области техники и включают анализы иммунопреципитации хроматина, анализы с определением изменения (сдвига) электрофоретической подвижности в геле, анализы методом копреципитации ДНК, репортерные анализы, анализы захвата и детектирования на
- 25 микропланшетах. Аналогичным образом, способы оценки расщепления или модификации последовательности-мишени известны в данной области техники и включают анализы расщепления *in vitro* или *in vivo*, в которых расщепление подтверждается с помощью ПЦР, секвенирования или гель-электрофореза с включением или без включения соответствующей метки (например,
- 30 радиоизотоп, флуоресцентное вещество) в последовательность-мишень для упрощения детектирования продуктов деградации. В некоторых вариантах используют анализ реакции экспоненциальной амплификации, запускаемой разрывом (NTEXPAR) (см., например, статью Zhang и др. Chem. Sci. 7:4951-4957

(2016)). Расщепление *in vivo* можно оценить с помощью анализа Surveyor (Guschin и др., Methods Mol Biol 649:247-256 (2010)).

В некоторых вариантах способы включают использование РНК-связывающего ДНК-управляемого домена в составе комплекса гибридного белка с более чем одной направляющей РНК. Более чем одну направляющую РНК можно направлять на различные области единого гена или можно направлять на несколько генов. Такое множественное направлявание позволяет с помощью дезаминазного домена гибридного белка модифицировать нуклеиновые кислоты, тем самым вводя множественные мутации в исследуемую молекулу нуклеиновой кислоты-мишени (например, геном).

В тех вариантах, в которых способ включает использование РНК-управляемой нуклеазы (RGN), такой как никаза RGN (т.е. способной расщеплять только одну цепь двухцепочечного полинуклеотида, например, pAPG07433.1 (SEQ ID NO: 42 или SEQ ID NO: 50-57), способ может включать введение двух различных RGN или вариантов RGN, направленных на идентичные или перекрывающиеся последовательности-мишени, и расщепляющих различные цепи полинуклеотида. Например, никазу RGN, которая расщепляет только положительную (+) цепь двухцепочечного полинуклеотида, можно вводить вместе со второй никазой RGN, которая расщепляет только отрицательную (-) цепь двухцепочечного полинуклеотида. В некоторых вариантах предлагаются два различных гибридных белка, где каждый гибридный белок содержит различную RGN с различной последовательностью распознавания РАМ, и таким образом большее разнообразие нуклеотидных последовательностей можно направлять на мутацию.

Специалисту в данной области представляется очевидным, что любой из раскрытых в данном контексте способов можно использовать для направления на единую последовательность-мишень или несколько последовательностей-мишеней. Таким образом, способы включают использование гибридного белка, содержащего один ДНК-связывающий полипептид, управляемый РНК, в комбинации с несколькими отдельными направляющими РНК, которые можно направлять на множество отдельных последовательностей в пределах единого гена и/или нескольких генов. Затем дезаминазный домен гибридного белка вносит мутации в каждую из последовательностей-мишеней. В данном документе также предлагаются способы, в которых несколько отдельных

направляющих РНК вводят в комбинации с несколькими отдельными управляемыми РНК ДНК-связывающими полипептидами. Такие РНК-управляемые ДНК-связывающие полипептиды могут представлять собой несколько RGN или вариантов RGN. Эти направляющие РНК и системы направляющих РНК/гибридных белков можно направить на несколько различных последовательностей в пределах одного гена и/или нескольких генов.

В некоторых вариантах гибридный белок, содержащий РНК-управляемый ДНК-связывающий полипептид и полипептид дезаминазы по изобретению, можно использовать для конструирования мутаций в гене-мишени или в области-мишени исследуемого гена. В некоторых вариантах гибридный белок по изобретению можно использовать для мутагенеза с насыщением гена-мишени или области исследуемого гена-мишени с последующим высокопроизводительным прямым генетическим скринингом для выявления новых мутаций и/или фенотипов. В некоторых вариантах гибридный белок, описанный в настоящем документе, можно использовать для конструирования мутаций в геномном участке-мишени, который может содержать или не содержать кодирующую последовательность ДНК. Библиотеки клеточных линий, полученных с помощью направленного мутагенеза, описанного выше, также можно эффективно использовать для изучения функции гена или экспрессии гена.

Х. Полинуклеотиды-мишени

В одном объекте изобретения предлагаются способы модификации полинуклеотида-мишени в эукариотической клетке, которые можно осуществлять в условиях *in vivo*, *ex vivo* или *in vitro*. В некоторых вариантах способ включает отбор образцов клеток или популяции клеток человека или животного или растения (включая микроводоросли) и модификацию клетки или клеток. Культивирование можно осуществлять на любой стадии в условиях *ex vivo*. Клетка или клетки можно даже повторно вводить в организм человека, животного или в растение (включая микроводоросли).

Используя природную изменчивость, селекционеры растений комбинируют наиболее полезные гены для требуемых свойств, таких как урожайность, качество, однородность, выносливость и устойчивость к вредителям. Эти требуемые свойства также включают рост, предпочтительную продолжительность дня, требования к температуре, дату начала цветочного или

репродуктивного развития, содержание жирных кислот, устойчивость к насекомым, болезням, устойчивость к нематодам, устойчивость к грибкам, устойчивость к гербицидам, устойчивость к различным факторам окружающей среды, включая засуху, жару, сырость, холод, ветер и неблагоприятные почвенные условия, включая высокую засоленность. Источники этих полезных генов включают местные или инородные сорта, старые негибридные сорта, дикие родственники растений и индуцированные мутации, например, обработку растительного материала мутагенными агентами. На основе настоящего изобретения, селекционеры получают новый инструмент для индукции мутаций. Соответственно, специалист в данной области техники может использовать настоящее изобретение для индукции числа пригодных генов с большей точностью по сравнению с предшествующими мутагенными агентами, и, следовательно, ускорить и улучшить программы селекции растений.

Полинуклеотид-мишень дезаминазы или гибридного белка по изобретению может представлять собой любой полинуклеотид, эндогенный или экзогенный по отношению к эукариотической клетке. Например, полинуклеотид-мишень может являться полинуклеотидом, находящимся в ядре эукариотической клетки. В некоторых вариантах полинуклеотид-мишень представляет собой последовательность, кодирующую продукт гена (например, белок), или некодирующую последовательность (например, регуляторный полинуклеотид или избыточную (бесполезную) ДНК). В некоторых вариантах последовательность-мишень для гибридного белка по изобретению связана с РАМ (мотив, примыкающий к протоспейсеру); то есть с короткой последовательностью, распознаваемой ДНК-связывающим полипептидом, управляемым РНК. Точная последовательность и требования к длине для РАМ различаются в зависимости от используемого ДНК-связывающего полипептида, управляемого РНК, но РАМ обычно представляют собой последовательности из 2-5 пар оснований, примыкающие к протоспейсеру (то есть последовательности-мишени).

Полинуклеотид-мишень гибридного белка по изобретению может включать ряд генов и полинуклеотидов, ассоциированных с заболеванием, а также гены и полинуклеотиды, ассоциированные с сигнальным биохимическим путем. Примеры полинуклеотидов-мишеней включают последовательность, связанную с биохимическим путем передачи сигнала, например, ген или полинуклеотид,

связанный с биохимическим путем передачи сигнала. Примеры полинуклеотидов-мишеней включают ген или полинуклеотид, ассоциированный с заболеванием. «Связанный с заболеванием» ген или полинуклеотид относится к любому гену или полинуклеотиду, который приводит к образованию

5 продуктов транскрипции или трансляции на аномальном уровне или в аномальной форме в клетках, полученных из тканей, пораженных заболеванием, по сравнению с контрольными тканями или клетками, не пораженными заболеванием. Такой ген может представлять собой ген, который

10 экспрессируется на аномально высоком уровне, а также может представлять собой ген, который экспрессируется на аномально низком уровне, когда измененная экспрессия коррелирует с развитием и/или прогрессированием заболевания. Ассоциированный с заболеванием ген также относится к гену, включающему мутацию (мутации) или к генетической вариации, которая непосредственно ответственна или находится в неравновесном сцеплении с

15 геном (генами), ответственными за этиологию заболевания (например, причинную мутацию). Транскрибированные или транслированные продукты могут являться известными или неизвестными и, кроме того, могут находиться на нормальном или аномальном уровне.

Неограничивающие примеры генов, ассоциированных с заболеванием, на

20 которые можно воздействовать с помощью раскрытых в данном контексте способов и композиций, представлены в таблице 34. В некоторых вариантах гены, ассоциированные с заболеванием и на которые они направлены, представляют собой гены, раскрытые в таблице 34, включающие мутацию при замене G на A. Дополнительные примеры генов и полинуклеотидов,

25 ассоциированных с заболеванием, можно найти в Институте генетической медицины McKusick-Nathans Institute of Genetic Medicine, в Университете Джона Хопкинса Johns Hopkins University (Baltimore, Md.) и в Национальном центре биотехнологической информации Национальной медицинской библиотеки National Center for Biotechnology Information, National Library of Medicine

30 (Bethesda, Md.), на сайте World Wide Web.

В некоторых вариантах полинуклеотид-мишень содержит ген регулятора трансмембранной проводимости муковисцидоза (5).

Используемый в данном контексте термин «трансмембранный регулятор проводимости при муковисцидозе» или «CFTR» относится к регулируемому

сАМР хлоридному каналу, расположенному в апикальной мембране эпителиальных клеток, который катализирует прохождение малых ионов через мембрану. Неограничивающий объем изобретения пример гена CFTR представлен как SEQ ID NO: 51.

5 Используемый в данном контексте термин «мишень» или «мишени» по отношению к спейсерной последовательности и последовательности-мишени относится к расположению РНК-управляемой нуклеазы в последовательности-мишени на основе способности спейсерной последовательности в ассоциированной управляющей РНК в достаточной степени гибридизоваться с
10 последовательностью-мишенью.

 Предложены CRISPR РНК (crRNA) или кодирующие их молекулы нуклеиновой кислоты, где crРНК содержит спейсерную последовательность, направленную на последовательность-мишень CFTR. Также предложены направляющие РНК, содержащие такие crRNA, одна или более молекул
15 нуклеиновых кислот, кодирующих направляющую РНК, содержащую такие crRNA, векторы, содержащие одну или более молекул нуклеиновых кислот, кодирующих направляющую РНК, содержащую такие crRNA, и системы, содержащие такие crRNA. Также предложены способы применения таких crRNA или кодирующих их молекул нуклеиновых кислот, направляющих РНК,
20 содержащих такие crRNA, одной или более молекул нуклеиновых кислот, кодирующих направляющую РНК, содержащую такие crRNA, векторов, содержащих одну или более молекул нуклеиновых кислот, кодирующих направляющую РНК, содержащих такие crRNA, и систем, содержащих такие crRNA для связывания, отщепления и/или модулирования последовательности-мишени.
25 мишени.

 В некоторых вариантах направляющая CFTR-последовательность crRNA или последовательность направляющей РНК характеризуются любой указанной последовательностью SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345- 364, 562 и 563 или комплементарными им цепями. В некоторых вариантах
30 единственная направляющая РНК (sgRNA), содержащая crRNA, имеющую спейсерную последовательность, направленную на последовательность-мишень CFTR, характеризуется последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75% , на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на

99% или более по сравнению с последовательностью SEQ ID NO: 98- 115, 140–151, 186–202, 235–250, 287–304, 345–364 и 564.

В некоторых вариантах направляющая CFTR-последовательность sgRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 62-68, 80-85, 116-119, 128-131, 163, 164, 180, 181, 203-209, 219-225, 256-258, 274-276, 310-313 и 330-333 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью с SEQ ID NO: 53. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется наличием спейсерной последовательности, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60 %, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 98-104, 140-143, 197, 198, 235-241, 292-294 и 350-353, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60 %, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 53.

В некоторых вариантах направляющая CFTR-последовательность crRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 68-71, 86-89, 120-122, 132-134, 152-156, 169-173, 213-215, 229-231, 251-255, 269-273, 305-309 и 325-329 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 55. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%,

на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 104-107, 144-146, 186-190, 245-247, 287-291 и 345-349, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 55.

В некоторых вариантах направляющая CFTR-последовательность sgRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 72, 73, 90, 91, 161, 162, 178, 179, 265, 266, 283, и 284 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 52. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 108, 109, 195, 196, 301 и 302, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 52.

В некоторых вариантах направляющая CFTR-последовательность sgRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 74, 75, 92, 93, 123, 124, 135, 136, 167, 184, 216-218, 232-234, 259-261, 277-279, 314-317 и 334-337 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%,

на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 56. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной

последовательностью, направленной на последовательность-мишень CFTR,

идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с

последовательностью SEQ ID NO: 110, 111, 147, 148, 201, 248-250, 295-297 и 354-357, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 56.

В некоторых вариантах направляющая CFTR-последовательность crRNA

или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 76, 94, 210-212, 226-228, 322, 342, 562 и 563

или комплементарными им цепями, и ассоциированный полипептид RGN

характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%,

на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 42. В

некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR,

идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%,

на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с

последовательностью SEQ ID NO: 112, 242-244, 362 и 564, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью,

идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на

95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с

последовательностью SEQ ID NO: 42.

В некоторых вариантах направляющая CFTR-последовательность crRNA

или последовательность направляющей РНК характеризуется любой указанной

последовательностью SEQ ID NO: 77, 95, 125, 137, 157-160, 174-177, 323 и 343 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%,
5 на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 54. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%,
10 на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 113, 149, 191-194 и 363, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%,
15 на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 54.

В некоторых вариантах направляющая CFTR-последовательность crRNA или последовательность направляющей РНК характеризуется любой указанной
20 последовательностью SEQ ID NO: 78, 96, 126, 138, 168, 185, 267, 285, 318, 319, 338 и 339 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на
25 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 57. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на
30 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 114, 150, 202, 303, 358 и 359, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45% , на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%,

на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 57.

В некоторых вариантах направляющая CFTR-последовательность sgRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 79, 97, 127, 139, 262-264, 280-282, 324 и 344 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 58. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 115, 151, 298-300 и 364, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 58.

В некоторых вариантах направляющая CFTR-последовательность sgRNA или последовательность направляющей РНК характеризуется любой указанной последовательностью SEQ ID NO: 165, 166, 182, 183, 268, 286, 320, 321, 340 и 341 или комплементарными им цепями, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 59. В некоторых вариантах sgRNA, содержащая crRNA, характеризуется спейсерной последовательностью, направленной на последовательность-мишень CFTR, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с

последовательностью SEQ ID NO: 199, 200, 304, 360 и 361, и ассоциированный полипептид RGN характеризуется аминокислотной последовательностью, идентичной по меньшей мере на 40%, на 45%, на 50%, на 55%, на 60%, на 65%, на 70%, на 75%, на 80%, на 85%, на 90%, на 91%, на 92%, на 93%, на 94%, на 95%, на 96%, на 97%, на 98%, на 99% или более по сравнению с последовательностью SEQ ID NO: 59.

В некоторых вариантах способы включают приведение в контакт молекулы ДНК, содержащей последовательность ДНК-мишени, с гибридным белком ДНК-связывающей полипептид-деаминазой по настоящему изобретению, где молекула ДНК контактирует с эффективным количеством гибридного белка в условиях, подходящих для дезаминирования азотистого основания. В некоторых вариантах способы включают приведение в контакт молекулы ДНК, содержащей последовательность ДНК-мишени, с (а) гибридным белком RGN-деаминазой по настоящему изобретению и (б) gRNA, направляющей гибридный белок (а) на нуклеотидную последовательность-мишень цепи ДНК, при этом молекула ДНК контактирует с гибридным белком и эффективным количеством gRNA в условиях, подходящих для дезаминирования азотистого основания. В некоторых вариантах последовательность ДНК-мишени содержит последовательность, связанную с заболеванием или расстройством, и при этом дезаминирование азотистого основания приводит к образованию последовательности, не связанной с заболеванием или расстройством. В некоторых вариантах последовательность ДНК-мишени локализована в аллеле культурного растения, где конкретный аллель представляющего интерес признака обуславливает получение растения с пониженной агрономической ценностью. Дезаминирование азотистого основания в результате дает аллель, который улучшает признак и повышает агрономическую ценность растения.

В некоторых вариантах последовательность ДНК-мишени содержит точечную мутацию G>A, связанную с заболеванием или расстройством, и при этом дезаминирование мутантного основания A приводит к образованию последовательности, не связанной с заболеванием или расстройством. В некоторых вариантах дезаминирование корректирует точечную мутацию в последовательности, связанную с заболеванием или расстройством.

В некоторых вариантах последовательность, связанная с заболеванием или расстройством, кодирует белок, и в результате дезаминирования

терминирующий кодон вводится в последовательность, связанную с заболеванием или расстройством, что приводит к усечению кодируемого белка. В некоторых вариантах контактирование проводят *in vivo* у субъекта, предрасположенного к заболеванию, страдающего заболеванием или с 5 диагностированным заболеванием или расстройством. В некоторых вариантах заболевание или расстройство представляет собой заболевание, связанное с точечной мутацией или с мутацией одного основания в геноме. В некоторых вариантах заболевание представляет собой генетическое заболевание, рак, метаболическое заболевание или лизосомную болезнь накопления.

10 XI. Фармацевтические композиции и способы лечения

В данном контексте представлены способы лечения заболевания у субъекта, нуждающегося в таком лечении. Эти способы включают введение субъекту, нуждающемуся в таком лечении, эффективного количества гибридного белка, раскрытого в настоящем изобретении, или кодирующего его 15 полинуклеотида, или раскрытой в настоящем изобретении gRNA или кодирующего ее полинуклеотида, раскрытой в настоящем изобретении системы гибридного белка, раскрытого в настоящем изобретении рибонуклеопротеинового комплекса, либо клетки, модифицированной любой из этих композиций или содержащей любую из этих композиций.

20 В некоторых вариантах лечение включает редактирование генов *in vivo* при введении субъекту, нуждающемуся в этом лечении, раскрытого в настоящем изобретении гибридного белка, gRNA или раскрытой в настоящем изобретении системы гибридного белка или кодирующего их полинуклеотида(ов). В некоторых вариантах лечение включает редактирование генов *ex vivo*, при 25 котором клетки генетически модифицируют *ex vivo* с помощью раскрытого в настоящем изобретении гибридного белка, gRNA или раскрытой в настоящем изобретении системы гибридного белка или кодирующего этот белок полинуклеотида(ов), с последующим введением модифицированных клеток субъекту. В некоторых вариантах генетически модифицированные клетки 30 происходят от субъекта, которому затем вводят модифицированные клетки, и трансплантированные клетки названы в данном контексте аутологичными. В некоторых вариантах генетически модифицированные клетки происходят от другого субъекта (т. е. донора) того же вида, что и субъект, которому вводят модифицированные клетки (т. е. реципиент), и трансплантированные клетки

названы в данном контексте аллогенными. В некоторых примерах, описанных в данном контексте, клетки можно размножить в культуре перед введением нуждающемуся в этом субъекту.

Например, в некоторых вариантах предложен способ, который включает
5 введение субъекту, имеющему, например, такое заболевание как генетический дефект, связанный с геном CFTR, эффективного количества
рибонуклеопротеинового комплекса, включающего гибридный белок с
деаминазой, характеризующейся аминокислотной последовательностью,
которая по меньшей мере на 80% идентична последовательности,
10 представленной в любой из SEQ ID NO: 399 и 405-407. В вариантах
осуществления изобретения, описанных в данном контексте, введение
рибонуклеопротеинового комплекса корректирует точечную мутацию или
вводит деактивирующую мутацию в ассоциированный с заболеванием ген CFTR.
Другие заболевания, которые можно лечить с использованием коррекции
15 точечной мутации или введением деактивирующей мутации в ассоциированный
с заболеванием ген, известны специалистам в данной области техники, и в этом
отношении не ограничивают объем изобретения.

В некоторых вариантах заболевание, которое необходимо лечить с
помощью раскрытых в данном контексте композиций, представляет собой
20 заболевание, которое можно лечить с помощью иммунотерапии, такой как
воздействие Т-клетками с химерным антигенным рецептором (CAR). Такие
заболевания включают рак, но не ограничиваются только им.

В некоторых вариантах дезаминирование нуклеинового основания-мишени
приводит к коррекции генетического дефекта, например, к коррекции гена
25 CFTR, или к коррекции точечной мутации, которая приводит к потере функции
генного продукта. В некоторых вариантах генетический дефект связан с
заболеванием или расстройством, например, с лизосомной болезнью накопления
или метаболическим заболеванием, таким как, например, диабет I типа. Таким
образом, в некоторых вариантах заболевание, подлежащее лечению с помощью
30 раскрытых в данном контексте композиций, обусловлено последовательностью
(т.е. последовательность является причинной заболевания или расстройства или
причиной симптомов, связанных с заболеванием или расстройством), которую
подвергают мутации с целью лечения заболевания или расстройства или
уменьшения симптомов, связанных с заболеванием или расстройством.

В некоторых вариантах заболевание, которое необходимо лечить с помощью раскрытых в настоящем документе композиций, связано со случайной мутацией. Используемый в данном контексте термин «случайная мутация» относится к конкретному нуклеотиду, нуклеотидам или последовательности нуклеотидов в геноме, которые способствуют тяжести или наличию заболевания или расстройства у субъекта. Коррекция случайной мутации приводит к улучшению по меньшей мере одного симптома, возникающего в результате заболевания или расстройства. В некоторых вариантах коррекция случайной мутации приводит к улучшению по меньшей мере одного симптома, возникающего в результате заболевания или расстройства. В некоторых вариантах случайная мутация локализована рядом с сайтом РАМ, распознаваемым RGDBP (например, RGN), гибридным с описанной в данном контексте дезаминазой. Случайную мутацию можно скорректировать с помощью гибридного полипептида, содержащего RGDBP (например, RGN) и описанной в данном контексте дезаминазой. Не ограничивающие объем изобретения примеры заболеваний, связанных со случайной мутацией, включают муковисцидоз, синдром Гурлера, атаксию Фридрейха, болезнь Хантингтона и серповидно-клеточную анемию. Дополнительные не ограничивающие объем изобретения примеры генов и мутаций, связанных с заболеванием, доступны в базах Института генетической медицины МакКьюсика-Натанса Университета Джона Хопкинса (Балтимор, Мэриленд), Национального центра биотехнологической информации Национальной медицинской библиотеки (Бетезда, Мэриленд) и во всемирной компьютерной сети.

В некоторых вариантах способы, представленные в данном контексте, используются для введения деактивирующей точечной мутации в ген или аллель, который кодирует генный продукт, связанный с заболеванием или расстройством. Например, в некоторых вариантах в данном контексте предусмотрены способы, в которых используется гибридный белок для введения деактивирующей точечной мутации в онкоген (например, при лечении пролиферативного заболевания). В некоторых вариантах деактивирующая мутация может генерировать преждевременный стоп-кодон в кодирующей последовательности, что приводит к экспрессии усеченного гена-продукта, например усеченного белка, лишённого функции полноразмерного белка. В некоторых вариантах цель способов, представленных в данном контексте,

заключается в восстановлении функции дисфункционального гена в результате редактирования генома. Гибридные белки, представленные в данном контексте, можно протестировать для использования в терапии человека на основе редактирования генов *in vitro*, например, с использованием коррекции мутации, связанной с заболеванием, в культуре клеток человека. Специалистам в данной области техники будет понятно, что гибридные белки, представленные в данном контексте, например, гибридные белки, содержащие РНК-управляемый, ДНК-связывающий полипептид и полипептид деаминазы, можно использовать для коррекции любой точечной мутации G>A. Дезаминирование мутантного основания А в G приводит к коррекции мутации.

Используемые в данном контексте термины «лечение», или «воздействие», или «временное улучшение», или «смягчение» взаимозаменяемы. Эти термины относятся к подходу к получению полезных или желаемых результатов, включая, но не ограничиваясь ими, терапевтический эффект и/или профилактический эффект. Под терапевтическим эффектом подразумевается любое терапевтическое улучшение или воздействие на одно или более заболеваний, состояний или симптомов при лечении. Для достижения профилактического эффекта композиции можно вводить субъекту, подвергнутому риску развития конкретного заболевания, состояния или симптома, или субъекту, сообщаящему об одном или более физиологических симптомах заболевания, даже если заболевание, состояние или симптом, возможно, еще не проявились. В некоторых вариантах лечение можно проводить после того, как развились один или более симптомов и/или после постановки диагноза заболевания. В конкретных вариантах лечение можно проводить в отсутствие симптомов, например, для предотвращения или задержки возникновения симптома или задержки появления или прогрессирования заболевания. Например, лечение можно назначать восприимчивому человеку до появления симптомов (например, в виду анамнеза симптомов и/или в виду генетических или других факторов восприимчивости). Лечение также можно продолжать после того, как симптомы устранены, например, для их профилактики или задержки рецидива.

Термин «эффективное количество» или «терапевтически эффективное количество» относится к количеству агента, достаточному для достижения полезных или желаемых результатов. Терапевтически эффективное количество

может варьироваться в зависимости от одного или более следующих факторов: субъекта и болезненного состояния, подлежащего лечению, веса и возраста субъекта, тяжести болезненного состояния, способа введения и т.п., которые может легко определить специалист в данной области техники. Конкретную дозу можно варьировать в зависимости от одного или более следующих факторов: конкретного выбранного агента, схемы дозировки, которой следует придерживаться, того, вводят ли его в комбинации с другими соединениями, продолжительности введения и системы доставки, в которой он содержится.

Термин «введение» относится к введению активного ингредиента субъекту с помощью метода или пути, который приводит к, по меньшей мере, частичной локализации введенного активного ингредиента в желаемом месте, таком как место повреждения или восстановления, причем достигается желаемый эффект(ы). В некоторых вариантах изобретение относится к способам, включающим доставку любого из выделенных полипептидов, молекул нуклеиновых кислот, гибридных белков, рибонуклеопротеиновых комплексов, векторов, фармацевтических композиций и/или gRNA, описанных в данном контексте. В некоторых вариантах изобретения дополнительно предлагаются клетки, полученные такими способами, и организмы (такие как животные или растения), содержащие такие клетки или полученные из них. В некоторых вариантах в клетку доставляют молекулы дезаминазы, гибридного белка и/или нуклеиновой кислоты, как описано в данном контексте, в комбинации с направляющей последовательностью (и необязательно в виде комплекса с ней).

В некоторых вариантах введение включает введение с помощью вирусной доставки. Вирусные векторы, содержащие нуклеиновую кислоту, кодирующую гибридные белки, комплексы рибонуклеопротеинов или векторы, описанные в данном контексте, можно вводить непосредственно пациентам (т.е. *in vivo*) или их можно использовать для лечения клеток *in vitro*, а модифицированные клетки необязательно можно вводить пациентам (т. е. *ex vivo*). Подходящие вирусные системы могут включать, но не ограничиваясь только ими, ретровирусные, лентивирусные, аденовирусные, аденоассоциированные векторы и векторы вируса простого герпеса для переноса генов. Интеграция в геном хозяина возможна с помощью методов переноса генов ретровируса, лентивируса и аденоассоциированного вируса, что часто приводит к длительной экспрессии встроенного трансгена. Лентивирусные векторы представляют собой

ретровирусные векторы, которые способны трансдуцировать или инфицировать неделящиеся клетки и обычно продуцируют высокие вирусные титры. В приложениях, где предпочтительна временная экспрессия, могут использоваться системы на основе аденовирусов. Векторы на основе аденовируса способны к трансдукции с очень высокой эффективностью во многих типах клеток и не требуют клеточного деления.

В некоторых вариантах введение включает введение с помощью электропорации. В некоторых вариантах введение включает введение с помощью доставки наночастицами. В некоторых вариантах введение включает введение с помощью липосомной доставки. Любой эффективный способ введения можно использовать для введения эффективного количества фармацевтической композиции, описанной в данном контексте.

В некоторых вариантах введение включает введение с помощью другого способа невирусной доставки нуклеиновых кислот. Примеры невирусных способов доставки, но не ограничиваясь только ими, включают RNP-комплексы, липофекцию, нуклеофекцию, микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, поликатионы или конъюгаты липидов с нуклеиновыми кислотами, голую ДНК, искусственные вирионы и усиленное агентом поглощение ДНК. Липофекция описана, например, в патентах США № 5049386, 4946787 и 4897355, а реагенты для липофекции поставляются коммерческими фирмами (например, Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, пригодные для эффективной липофекции полинуклеотидов с распознаванием рецепторов, включают таковые от фирмы Feigner, и описаны во международных патентах WO 1991/17424, WO 1991/16024. Доставку можно осуществлять в клетки (например, введением *in vitro* или *ex vivo*) или ткани-мишени (например, введение *in vivo*).

Используемый в данном контексте термин «субъект» относится к любому индивидууму, для которого желательна диагностика, лечение или терапия. В некоторых вариантах субъект представляет собой животное. В некоторых вариантах субъект представляет собой млекопитающее. В некоторых вариантах субъектом является человек.

Эффективность лечения может определить опытный клиницист. Однако лечение считается «эффективным лечением», если какой-либо один или все признаки или симптомы заболевания или расстройства изменяются

благоприятным образом (например, уменьшаются по меньшей мере на 10%), или другие клинически приемлемые симптомы или маркеры болезни улучшаются или купируются. Эффективность также можно оценить по отсутствию ухудшения состояния пациента по оценке необходимости в госпитализации или потребности в медицинских вмешательствах (например, когда прогрессирование заболевания остановлено или, по меньшей мере, замедлено). Способы измерения этих показателей известны специалистам в данной области техники. Лечение включает: (1) подавление заболевания, например остановку или замедление прогрессирования симптомов или (2) облегчение заболевания, например, регресс симптомов и (3) предотвращение или снижение вероятности развития симптомов.

Предложены фармацевтические композиции, содержащие описанные в данном контексте полипептиды RGN или кодирующие их полинуклеотиды, раскрытые в данном контексте gRNA или кодирующие их полинуклеотиды, раскрытые в данном контексте деаминазы или кодирующие их полинуклеотиды, раскрытые в данном контексте гибридные белки, раскрытые в данном контексте системы (например, содержащие гибридный белок), раскрытый в данном контексте рибонуклеопротеиновый комплекс или клетки, содержащие любой из полипептидов RGN или полинуклеотидов, кодирующих RGN, gRNA или полинуклеотиды, кодирующие gRNA, полинуклеотиды, кодирующие гибридный белок или системы, и фармацевтически приемлемый носитель.

Используемый в данном контексте термин «фармацевтически приемлемый носитель» относится к материалу, который не вызывает значительного раздражения в организме и не подавляет активность и свойства активного ингредиента (например, дезаминазы или гибридного белка, или молекулы нуклеиновой кислоты, кодирующей его). Носители должны характеризоваться достаточно высокой чистотой и достаточно низкой токсичностью, чтобы быть пригодными для введения субъекту, проходящему лечение. Носитель может являться инертным или обладать фармацевтическими преимуществами. В некоторых вариантах фармацевтически приемлемый носитель содержит один или более совместимых твердых или жидких наполнителей, разбавителей или инкапсулирующих веществ, пригодных для введения человеку или другому позвоночному животному. В некоторых вариантах фармацевтическая композиция содержит фармацевтически приемлемый носитель, не

встречающийся в природе. В некоторых вариантах фармацевтически приемлемый носитель и активный ингредиент не встречаются вместе в природе и, таким образом, являются гетерологичными.

Фармацевтические композиции, используемые в раскрытых в данном контексте способах, можно составлять с пригодными носителями, 5 эксципиентами и другими агентами, которые обеспечивают подходящий перенос, доставку, переносимость и т.п. Специалистам в данной области техники известно множество подходящих составов. См., например, Remington, The Science and Practice of Pharmacy (21-е изд., 2005 г.). Неограничивающие примеры 10 включают стерильный разбавитель, такой как вода для инъекций, физиологический раствор, нелетучие масла, полиэтиленгликоли, глицерин, пропиленгликоль или другие синтетические растворители, антибактериальные агенты, такие как бензиловый спирт или метилпарабены, антиоксиданты, такие как аскорбиновая кислота или бисульфит натрия, хелатирующие агенты, такие как этилендиаминтетрауксусная кислота, буферные растворы, такие как ацетаты, 15 цитраты или фосфаты, и агенты для регулирования тоничности, такие как хлорид натрия или декстроза. Конкретными носителями для внутривенного введения являются физиологический раствор или фосфатно-солевой буфер (PBS). Фармацевтические композиции для перорального или парентерального применения можно приготовить в виде дозированных форм в стандартной дозе, 20 соответствующей дозе активных ингредиентов. Такие лекарственные формы в стандартной дозе включают, например, таблетки, пилюли, капсулы, инъекции (ампулы), суппозитории и т.д. Эти композиции также могут содержать адъюванты, включая консерванты, смачивающие агенты, эмульгаторы и диспергирующие агенты. Предотвращение действия микроорганизмов можно 25 обеспечить с помощью различных антибактериальных и противогрибковых средств, например парабенов, хлорбутанола, фенола, сорбиновой кислоты и т.п. Также может быть желательным включение изотонических агентов, например сахаров, хлорида натрия и т.п. Длительное всасывание инъекционной 30 лекарственной формы можно вызвать с помощью агентов, замедляющих всасывание, например моностеарата алюминия и желатина.

В некоторых вариантах субъекту вводят клетки, содержащие или модифицированные раскрытыми в данном контексте RGN, гРНК, деаминазами, гибридными белками, системами (включая содержащие гибридные белки) или

кодирующими их полинуклеотидами, клетки вводят в виде суспензии с фармацевтическим приемлемым носителем. Специалисту в данной области техники будет понятно, что фармацевтически приемлемый носитель, используемый в клеточной композиции, не должен включать буферные растворы, соединения, агенты криоконсервации, консерванты или другие агенты в количествах, которые существенно влияют на жизнеспособность клеток, подлежащих доставке субъекту. Состав, содержащий клетки, может включать, например, осмотические буферные растворы, которые позволяют поддерживать целостность клеточной мембраны, и, необязательно, питательные вещества для поддержания жизнеспособности клеток или для усиления приживания при введении. Такие составы и суспензии известны специалистам в данной области техники и/или их можно адаптировать для использования с клетками, описанными в данном контексте, с помощью обычных экспериментов.

Клеточную композицию также можно эмульгировать или представить в виде липосомальной композиции при условии, что процедура эмульгирования не оказывает неблагоприятного воздействия на жизнеспособность клеток. Клетки и любой другой активный ингредиент можно смешивать с эксципиентами, которые являются фармацевтически приемлемыми и совместимыми с активным ингредиентом, и в количествах, подходящих для использования в описанных в данном контексте терапевтических методах.

Дополнительные агенты, включенные в клеточную композицию, могут включать фармацевтически приемлемые соли ее компонентов. Фармацевтически приемлемые соли включают кислотно-аддитивные соли (образованные из свободных аминогрупп полипептида), которые образуются с неорганическими кислотами, такими как, например, соляная или фосфорная кислоты, или такими органическими кислотами, как уксусная, винная, миндальная и т.п. Соли, образованные со свободными карбоксильными группами, также можно получить из неорганических оснований, таких как, например, гидроксиды натрия, калия, аммония, кальция или железа, и таких органических оснований, как изопропиламин, триметиламин, 2-этиламиноэтанол, гистидин, прокаин и подобных.

Пригодные способы введения фармацевтической композиции, описанной в данном контексте, включают, но не ограничиваются только ими: местное, подкожное, чрескожное, внутрикожное, внутривенное, внутрисуставное,

внутрибрюшинное, внутрипузырное, чрезслизистое, десневое, внутризубное, внутрикохлеарное, транстимпанальное, внутриорганный, эпидуральный, интратекальный, внутримышечный, внутривенный, внутрисосудистый, внутрикостный, периокулярный, внутриопухолевый, внутримозговой и интрацеребровентрикулярный введение.

В некоторых вариантах фармацевтическую композицию, описанную в данном контексте, вводят локально в пораженный участок (например, в легкое). В некоторых вариантах фармацевтическую композицию, описанную в данном контексте, вводят субъекту с помощью инъекции, ингаляции (например, аэрозоля), с помощью катетера, с помощью суппозитория или с помощью имплантата, причем имплантат является пористым, непористым или желеобразным материалом, включая мембрану, такую как сиаластическая мембрана или волокно. В некоторых вариантах фармацевтическая композиция составлена для доставки субъекту, например, для редактирования генов.

В некоторых вариантах фармацевтическую композицию составляют в соответствии с обычными процедурами в виде композиции, адаптированной для внутривенного или подкожного введения субъекту, например человеку. В некоторых вариантах фармацевтические композиции для введения с использованием инъекции представляют собой растворы в стерильном изотоническом водном буферном растворе. При необходимости фармацевтический препарат может также включать солюбилизующий агент и местный анестетик, такой как лигнокаин, для облегчения боли в месте инъекции. Как правило, ингредиенты поставляются либо по отдельности, либо в смеси, например, в виде сухого лиофилизированного порошка или безводного концентрата в герметично закрытой таре, такой как ампула или саше, с указанием количества активного вещества. Если фармацевтический препарат следует вводить с использованием инфузии, его можно вводить с помощью флакона для инфузий, содержащего стерильную воду фармацевтического качества или физиологический раствор. Если фармацевтическую композицию вводят с использованием инъекции, может быть предоставлена ампула со стерильной водой для инъекций или физиологическим раствором, чтобы можно было смешать ингредиенты перед введением.

В некоторых вариантах фармацевтическая композиция может содержаться внутри липидной частицы или везикулы, такой как липосома или микрокристалл, которые также пригодны для парентерального введения.

Хотя приведенные здесь описания фармацевтических композиций в основном относятся к фармацевтическим композициям, пригодным для введения людям, специалисту в данной области техники должно быть понятно, что такие композиции обычно подходят для введения животным или организмам всех видов.

Изменение причинных мутаций с помощью редактирования основания

Примером генетически наследуемого заболевания, которое можно корректировать с помощью подхода, основанного на использовании гибридного белка RGN-деаминазы по настоящему изобретению, является муковисцидоз. Муковисцидоз (CF) представляет собой аутосомно-рецессивное заболевание, вызываемое мутациями в гене трансмембранного регулятора муковисцидоза (CFTR) (обозначен как SEQ ID NO: 51). CFTR кодирует регулируемый сАМР хлоридный канал, расположенный в апикальной мембране эпителиальных клеток, который катализирует прохождение небольших ионов через мембрану. Нарушение регуляции этого механизма вызывает нарушение солевого и жидкостного гомеостаза, что приводит к полиорганным дисфункциям и, в конечном итоге, к смерти от дыхательной недостаточности.

Было найдено, что почти 2000 мутаций в гене CFTR вызывают муковисцидоз. Мутации CFTR подразделяются на шесть классов на основе функционального дефекта синтеза белка CFTR, переноса, функции или стабильности, хотя признано, что многие мутанты CFTR имеют множественные дефекты. Мутации класса I приводят к серьезному дефекту продуцирования белка. В первую очередь это бессмысленные мутации или мутации со сдвигом рамки считывания, которые вводят кодон преждевременной терминации (РТС), что приводит к деградации нестабильной матричной РНК (мРНК) по пути распада мРНК (NMD). Бессмысленные мутации (нонсенс-мутации), вызванные изменениями отдельных нуклеотидов, составляют основное подмножество мутаций класса I (Marangi, M. и Pistritto, G, Front Pharmacol 9, 396, (2018), doi:10.3389/fphar.2018.00396, Pranke, I. и др., Front Pharmacol 10, 121, (2019), doi:10.3389/fphar.2019.00121, обе статьи включены в данное описание в качестве ссылки). Лечение пациентов с кистозным фиброзом класса I может быть

затруднено, поскольку не вырабатывается функциональный белок CFTR. Примечательно, что значительную часть этих бессмысленных мутаций потенциально можно устранить с помощью редакторов оснований от A до G (Geurts, M.H. и др., Cell Stem Cell 26, 503-510 e507, (2020),

doi:10.1016/j.stem.2020.01.019, статья включена в качестве ссылки в данном контексте).

Гертс и др. были первой группой, которая выполнила точное редактирование оснований в культивируемых эпителиальных клетках легких с мутациями класса I от пациентов с муковисцидозом, используя гибридный белок, содержащий адениндезаминазу, функционально связанную с RGN, а именно варианты SpyCas9 или xSpyCas9. SpyCas9 распознает 5'-nGG-3' PAM, тогда как вариант xSpyCas9 распознает редуцированный 5'-nG-3'. Авторы заявляют, что основным ограничением технологии редактирования оснований является требование PAM для используемого белка Cas. Они обнаружили, что большинство нонсенс-мутаций, идентифицированных в гене *CFTR*, не входят в требуемое окно таргетинга для гибридного белка, содержащего RGN SpyCas9. PAM представляет собой короткий мотив, обычно состоящий из одного-четырех нуклеотидов, в последовательности ДНК-мишени, распознаваемой RGN. Последовательность PAM присуща каждому белку RGN, так что RGN может получить доступ только к геномному пространству вокруг подходящего PAM. Кроме того, окно редактирования базы для редакторов базы ограничено, часто только частью нуклеотидов в целевой последовательности. Если интересующий нуклеотид находится слишком близко к PAM, RGN блокирует доступ к нуклеотиду. Если нуклеотид находится слишком далеко от PAM, дезаминаза, связанная с RGN, не может достичь нуклеотида. Кроме того, количество одноцепочечной ДНК, подвергаемой воздействию R-петли, ограничивает доступность дезаминаз. Настоящее изобретение включает гибридные белки RGN-дезаминазы, в которых RGN распознает PAM, который является проксимальным по отношению к мутации класса I гена *CFTR*, и дезаминаза способна успешно модифицировать целевую причинную мутацию.

Другое ограничение известных в данной области техники гибридных белков RGN-дезаминазы состоит в том, что векторная конструкция, кодирующая гибридный белок, слишком велика для способов доставки *in vivo*. Доставка этих гибридных белков с помощью AAV не подходит для гибридных белков на

основе SpyCas9, поскольку их размер превышает предел для эффективной упаковки AAV. Компонент RGN гибридных белков, описанных в данном контексте, имеет меньший размер и, следовательно, является жизнеспособным кандидатом для стратегии доставки вектора AAV. Настоящее изобретение также раскрывает направляющие РНК, которые специфичны для RGN, описанных в данном контексте, и которые направляют гибридные белки по настоящему изобретению к сайтам-мишеням нонсенс-мутаций в гене *CFTR*, которые ранее были недоступны. Настоящее изобретение также относится к способам использования указанных гибридных белков для целенаправленного редактирования оснований посредством доставки вектора AAV *in vivo*.

В идеальном случае кодирующая последовательность гибридного белка RGN-дезаминазы по настоящему изобретению и соответствующая направляющая РНК для направления гибридного белка на ген *CFTR* можно упаковать в один вектор AAV. Общепринятый предельный размер векторов AAV составляет 4,7 т.п.н., хотя могут рассматриваться и большие размеры за счет снижения эффективности упаковки. Никазы RGN в таблице 28 имеют длину кодирующей последовательности примерно 3,15-3,45 т.п.н. Чтобы убедиться, что экспрессионные кассеты как для гибридного белка, так и для его соответствующей направляющей РНК могут вписаться в вектор AAV, в данном контексте описаны новые активные делеционные варианты RGN. В дополнение к укорочению аминокислотной последовательности и, следовательно, кодирующей последовательности RGN гибридного белка также можно укоротить пептидный линкер, который связывает RGN и дезаминазу. Наконец, генетические элементы, такие как промоторы, энхансеры и/или терминаторы, также можно сконструировать с помощью анализа делеций для определения минимального размера, необходимого для того, чтобы каждый из них был функциональным.

Некоторые варианты осуществления настоящего изобретения обеспечивают способы редактирования нуклеиновой кислоты с использованием дезаминаз или комплексов RGN, описанных в данном контексте, для достижения замены азотистых оснований, например, пары оснований А:Т на пару оснований G:C. В некоторых вариантах способ представляет собой способ редактирования азотистого основания нуклеиновой кислоты (например, пары оснований двухцепочечной последовательности ДНК). В некоторых вариантах дезаминазы

или комплексы RGN, описанные в данном контексте, используют для введения точечной мутации в нуклеиновую кислоту с использованием дезаминирования и вырезания целевого азотистого основания «А». В некоторых вариантах дезаминирование и удаление целевого азотистого основания приводит к

5 коррекции генетического дефекта, например, к коррекции точечной мутации в гене *CFTR*. В некоторых вариантах генетический дефект связан с заболеванием, нарушением или состоянием, например, муковисцидозом. Например, в некоторых вариантах в предложенных в данном контексте способах используются комплексы RGN для редактирования оснований, содержащих
10 гибридный белок с дезаминазой, аминокислотная последовательность которого по меньшей мере на 80% идентична последовательности, указанной в любой из SEQ ID NO: 399 и 405-407, для исправления гена, связанного с генетическим дефектом, например, для исправления точечной мутации в гене *CFTR* (например, при лечении пролиферативного заболевания). В конкретных вариантах
15 осуществления изобретения целевая последовательность в гене *CFTR* представляет собой 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 или 563.

В некоторых вариантах цель способов, представленных в данном контексте, заключается в восстановлении функции дисфункционального гена с помощью
20 редактирования генома. Пригодность представленные в данном контексте белков-редакторов оснований в качестве компонентов для терапии человека можно оценивать на основе редактирования генов *in vitro*, например, с использованием коррекции мутации, связанной с заболеванием, в культуре клеток человека. Специалистам в данной области техники должно быть понятно,
25 что гибридные белки и/или комплексы RGN, представленные в данном контексте, содержащие белок, связывающий нуклеиновую кислоту (например, pCas9), и домен модификации азотистых оснований (например, деаминазу с аминокислотной последовательностью, представленной в SEQ ID №: 407, 399 или 405) можно использовать для исправления любой отдельной точки Т на G
30 или изменения пары Т:А на G:С.

В некоторых вариантах в данном контексте представлены способы лечения субъекта, у которого диагностировано заболевание, связанное с точечной мутацией или вызванное ею (например, мутацией в гене *CFTR*), которую можно скорректировать с помощью гибридного белка или комплексов RGN, описанных

в данном контексте. Например, в некоторых вариантах предложен способ, который включает введение субъекту, страдающему таким заболеванием, как, например, муковисцидоз, эффективного количества гибридного белка или комплекса RGN, раскрытых в данном контексте, которые корректируют точечную мутацию или вводят деактивирующую мутацию в ген, ассоциированный с заболеванием. В некоторых вариантах предложен способ, который включает введение субъекту, страдающему таким заболеванием, как рак, связанный с точечной мутацией, как описано выше, эффективного количества гибридного белка, комплекса RGN или фармацевтической композиции, описанной в данном контексте, которая корректирует точечную мутацию или вводит деактивирующую мутацию в ген, ассоциированный с заболеванием. В конкретных вариантах осуществления изобретения предусмотрены способы лечения муковисцидоза наряду со способами уменьшения по меньшей мере одного симптома муковисцидоза с использованием введения эффективного количества фармацевтической композиции, описанной в данном контексте. Эффективное количество фармацевтической композиции для лечения или уменьшения симптома муковисцидоза может уменьшить симптом (т.е. лечить) муковисцидоза приблизительно на 5%, на 10%, на 15%, на 20%, на 25%, на 30%, на 40%, на 50%, на 60%, на 70%, на 80%, на 90%, на 95% или более, или приблизительно на 10-20%, 15-25%, 20-40%, 30-50%, 40-60%, 50-70%, 60-80%, 70-90%, 80-95% или 90-95% по сравнению с контрольным пациентом. В конкретных вариантах контрольным пациентом может быть тот же пациент до введения эффективного количества фармацевтической композиции, раскрытой в данном контексте. Симптомы муковисцидоза могут включать, но не ограничиваются только ими: чихание, постоянный кашель с выделением слизи или мокроты, одышку, особенно при физических нагрузках, рецидивирующие легочные инфекции, заложенность носа, заложенность носовых пазух, жирный зловонный стул, запоры, тошнота, вздутие живота, потеря аппетита, среди прочего. Способы идентификации и оценки симптомов муковисцидоза известны в данной области техники.

В некоторых вариантах описанных способов модификации молекулы ДНК-мишени стадию контактирования проводят *in vitro*. В конкретных вариантах стадию контактирования проводят *in vivo*. В некоторых вариантах этап

контактирования осуществляют у субъекта (например, у человека или животного, отличного от человека). В некоторых вариантах стадию контактирования проводят в клетке, такой как человеческая клетка или клетка животного, отличного от человека.

5 ХII. Клетки, содержащие полинуклеотидную генетическую модификацию

В данном контексте представлены клетки и организмы, содержащие представляющую интерес молекулу нуклеиновой кислоты-мишени, которую модифицировали с использованием процесса, опосредованного гибридным белком, необязательно с gRNA, как описано в данном контексте. В некоторых вариантах гибридный белок содержит полипептид дезаминазы, содержащий аминокислотную последовательность любой из SEQ ID NO: 1-10 и 399-441, или его активный вариант или фрагмент. В некоторых вариантах гибридный белок содержит адениндезаминазу, содержащую аминокислотную последовательность, имеющую по меньшей мере 50%, по меньшей мере 55%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичности любой из SEQ ID NO: 1-10 и 399 -441. В некоторых вариантах гибридный белок содержит дезаминазу и ДНК-связывающий полипептид (например, РНК-направляемый ДНК-связывающий полипептид). В других вариантах гибридный белок содержит дезаминазу и RGN или их вариант, такой как, например, APG07433.1 (SEQ ID NO: 41) или его вариант никазы nAPG07433.1 (SEQ ID NO: 42). В некоторых вариантах гибридный белок содержит дезаминазу и Cas9 или их вариант, такой как, например, dCas9 или никаза Cas9. В некоторых вариантах гибридный белок содержит неактивный в отношении нуклеаз или «никазный» вариант полипептида CRISPR-Cas типа II. В некоторых вариантах гибридный белок содержит неактивный в отношении нуклеаз или «никазный» вариант полипептида CRISPR-Cas типа V. В некоторых вариантах гибридный белок содержит неактивный в отношении нуклеаз или «никазный» вариант полипептида CRISPR-Cas типа VI.

Модифицированные клетки могут являться эукариотическими (например, клетки млекопитающих, растений, насекомых, птиц) или прокариотическими. Также предусмотрены органеллы и эмбрионы, содержащие по меньшей мере

одну нуклеотидную последовательность, которую модифицировали в процессе с использованием гибридного белка, как описано в данном контексте. Генетически модифицированные клетки, организмы, органеллы и эмбрионы могут являться гетерозиготными или гомозиготными по модифицированной нуклеотидной последовательности. Мутация(и), введенная(ые) деаминазным доменом гибридного белка, может приводить к измененной экспрессии (повышающая или отрицательная регуляция), инактивации или экспрессии измененного белкового продукта или интегрированной последовательности. В тех случаях, когда мутация (мутации) приводит либо к инактивации гена, либо к экспрессии нефункционального белкового продукта, генетически модифицированная клетка, организм, органелла или эмбрион называют «нокаутными». Нокаутный фенотип может являться результатом делеционной мутации (т. е. делеции 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).

В некоторых вариантах мутация(и), введенная(ые) деаминазным доменом гибридного белка, приводит к получению вариантного белкового продукта. Продукт экспрессированного варианта белка может характеризоваться заменой по меньшей мере одной аминокислоты и/или добавлением или делецией по меньшей мере одной аминокислоты. Вариантный белковый продукт может проявлять модифицированные характеристики или активность по сравнению с белком дикого типа, включая, но не ограничиваясь только этим, измененную ферментативную активность или субстратную специфичность.

В некоторых вариантах мутация(и), введенная(ые) деаминазным доменом гибридного белка, приводит к измененному характеру экспрессии белка. В качестве неограничивающего примера, мутация(и) в регуляторных областях, контролирующих экспрессию белкового продукта, может приводить к сверхэкспрессии или подавлению белкового продукта или к измененному тканевому или временному профилю экспрессии.

Клетки, которые были промодифицированы, можно вживлять в организм, такой как растение, в соответствии с известными способами. См., например, McCormick и др. *Plant Cell Reports* 5:81-84 (1986). Затем эти растения можно вырастить и либо опылить одним и тем же модифицированным штаммом, либо разными штаммами, и полученный гибрид будет характеризоваться

генетической модификацией. В настоящем изобретении предложены генетически модифицированные семена. Потомство, варианты и мутанты регенерированных растений также входят в объем настоящего изобретения при условии, что эти части содержат генетическую модификацию. Кроме того, предложен переработанный растительный продукт или побочный продукт, который сохраняет генетическую модификацию, включая, например, соевый шрот.

Представленные в данном контексте способы можно использовать для модификации любых видов растений, включая, но не ограничиваясь только ими, однодольные и двудольные растения. Примеры представляющих интерес растений включают, но не ограничиваются только ими, кукурузу (маис), сорго, пшеницу, подсолнечник, томат, крестоцветные, перец, картофель, хлопок, рис, сою, сахарную свеклу, сахарный тростник, табак, ячмень и масличный рапс, капусту, люцерну, рожь, просо, сафлор, арахис, сладкий картофель, маниоку, кофе, кокос, ананас, цитрусовые деревья, какао, чай, банан, авокадо, инжир, гуаву, манго, оливки, папайю, кешью, макадамию, миндаль, овес, овощи, декоративные растения и хвойные деревья.

Овощи включают, но не ограничиваясь только ими, помидоры, салат, зеленую фасоль, лимскую фасоль, горох и представителей рода *Cucumis*, таких как огурец, дыня и мускусная дыня. Декоративные растения включают, но не ограничиваясь только ими, азалии, гортензии, гибискусы, розы, тюльпаны, нарциссы, петунии, гвоздики, пуансеттии и хризантемы. Предпочтительно, растения по настоящему изобретению представляют собой сельскохозяйственные культуры (например, кукуруза, сорго, пшеница, подсолнечник, помидоры, крестоцветные, перец, картофель, хлопок, рис, соя, сахарная свекла, сахарный тростник, табак, ячмень, масличный рапс и т. д.).

Предложенные здесь способы также можно использовать для генетической модификации любых прокариотических видов, включая, помимо прочего, 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.).

Предложенные в данном контексте способы можно использовать для генетической модификации любых видов эукариот или их клеток, включая, но не ограничиваясь только ими, животных (например, млекопитающих, насекомых,

рыб, птиц и рептилий), грибы, амёбы, водоросли и дрожжи. В некоторых вариантах клетки, модифицированные раскрытыми в данном контексте способами, включают клетки гемопоэтического происхождения, такие как иммунные клетки (т.е. клетки врожденной или адаптивной иммунной системы), включая, но не ограничиваясь только ими, В-клетки, Т-клетки, естественные киллеры (NK), плюрипотентные стволовые клетки, индуцированные плюрипотентные стволовые клетки, Т-клетки с химерным антигенным рецептором (CAR-T), моноциты, макрофаги и дендритные клетки.

Модифицированные клетки можно вводить в организм. Такие клетки могут происходить из одного и того же организма (например, человека) в случае аутологичных клеточных трансплантатов, когда клетки модифицируют *ex vivo*. В некоторых вариантах клетки происходят из другого организма того же вида (например, другого человека) в случае аллогенных клеточных трансплантатов.

XIII. Наборы

В некоторых объектах по настоящему изобретению предложены наборы, содержащие дезаминазу по настоящему изобретению. В некоторых вариантах настоящего изобретения предложены наборы, включающие гибридный белок, содержащий дезаминазу по настоящему изобретению и ДНК-связывающий полипептид (например, РНК-направляемый ДНК-связывающий полипептид, такой как полипептид RGN, например, неактивный в отношении домена нуклеазы Cas9) и, необязательно, линкер, расположенный между ДНК-связывающим полипептидным доменом и дезаминазой. Кроме того, в некоторых вариантах набор содержит подходящие реагенты, буферные растворы и/или инструкции по применению гибридного белка, например, для редактирования ДНК или РНК *in vitro* или *in vivo*. В некоторых вариантах набор содержит инструкции относительно конструирования и использования подходящих gRNA для целенаправленного редактирования последовательности нуклеиновой кислоты.

В некоторых вариантах фармацевтическая композиция может являться фармацевтическим набором, включающим (а) контейнер, содержащий композицию по настоящему изобретению в лиофилизованной форме, и (б) второй контейнер, содержащий фармацевтически приемлемый разбавитель (например, стерильную воду) для инъекций. Фармацевтически приемлемый разбавитель можно использовать для восстановления или разбавления

лиофилизованного соединения по настоящему изобретению. Необязательно с таким контейнером (контейнерами) может содержаться уведомление в форме, предписанной государственным органом, регулирующим производство, использование или продажу фармацевтических или биологических продуктов, которое отражает одобрение агентства по производству, использованию или

5 продаже для введения человеку.

Артикли «а» и «an» используются в данном контексте для обозначения одного или более чем одного (т. е. по меньшей мере одного) грамматического объекта при артикле. Например, «полипептид» означает один или более

10 полипептидов.

Все публикации и патентные заявки, упомянутые в описании, соответствуют уровню специалистов в данной области техники, к которой относится данное изобретение. Все публикации и патентные заявки включены в данном контексте в качестве ссылки в той же степени, как если бы каждая

15 отдельная публикация или патентная заявка были конкретно и индивидуально указаны при включении в настоящий контекст в качестве ссылки.

Хотя предыдущее изобретение было описано довольно подробно с использованием иллюстрации и примера в целях ясности понимания, должно быть очевидно, что в пределах объема прилагаемой формулы изобретения

20 можно осуществлять некоторые изменения и модификации.

Неограничивающие варианты по настоящему изобретению включают:

1. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441, где

25 указанный полипептид характеризуется дезаминазной активностью.

2. Выделенный полипептид по варианту 1, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

3. Выделенный полипептид по варианту 1, содержащий аминокислотную последовательность, на 100% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

30

4. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, при этом дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

5 5. Молекула нуклеиновой кислоты по варианту 4, где дезаминаза кодируется нуклеотидной последовательностью, которая по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485.

10 6. Молекула нуклеиновой кислоты по варианту 4, где дезаминаза кодируется нуклеотидной последовательностью, которая по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450. и 452-485.

15 7. Молекула нуклеиновой кислоты по варианту 4, где дезаминаза кодируется нуклеотидной последовательностью, которая имеет 100%-ую идентичность с любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485.

20 8. Молекула нуклеиновой кислоты по варианту 4, отличающаяся тем, что полипептид дезаминазы содержит аминокислотную последовательность, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406, и 408-441.

25 9. Молекула нуклеиновой кислоты по варианту 4, где полипептид дезаминазы имеет аминокислотную последовательность, имеющую 100% идентичность с любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408. -441.

10. Молекула нуклеиновой кислоты по любому из вариантов 4-9, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

30 11. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из вариантов 1-3 или молекулу нуклеиновой кислоты по любому из вариантов 4-10.

12. Фармацевтическая композиция по варианту 11, в которой фармацевтически приемлемый носитель гетерологичен указанному полипептиду или указанной молекуле нуклеиновой кислоты.

5 13. Фармацевтическая композиция по вариантам 11 или 12, в которой фармацевтически приемлемый носитель не является природным.

14. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, последовательность которого по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

10 15. Гибридный белок по варианту 14, отличающийся тем, что указанная дезаминаза по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

15 16. Гибридный белок по варианту 14, где указанная дезаминаза на 100% идентична любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

17. Гибридный белок по любому из вариантов 14-16, где дезаминаза представляет собой адениндезаминазу.

20 18. Гибридный белок по любому из вариантов 14-17, где ДНК-связывающий полипептид представляет собой мегануклеазу, цинк-пальцевый гибридный белок или TALEN.

19. Гибридный белок по любому из вариантов 14-17, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

25 20. Гибридный белок по варианту 19, отличающийся тем, что РНК-управляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

21. Гибридный белок по варианту 20, где RGN представляет собой полипептид CRISPR-Cas типа II.

30 22. Гибридный белок по варианту 20, где RGN представляет собой полипептид CRISPR-Cas типа V.

23. Гибридный белок по любому из вариантов 20-22, где RGN представляет собой RGN-никазу.

24. Гибридный белок по варианту 20, отличающийся тем, что RGN имеет аминокислотную последовательность, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

5 25. Гибридный белок по варианту 20, отличающийся тем, что RGN имеет аминокислотную последовательность любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

26. Гибридный белок по варианту 23, где RGN-никаза представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

10 27. Гибридный белок по любому из вариантов 14-26, отличающийся тем, что гибридный белок дополнительно содержит по меньшей мере один сигнал ядерной локализации (NLS).

28. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, при этом дезаминаза кодируется нуклеотидной
15 последовательностью, которая:

а) характеризуется по меньшей мере 80%-ой идентичностью с любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485 или

20 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

29. Молекула нуклеиновой кислоты по варианту 28, где указанная нуклеотидная последовательность по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485.

25 30. Молекула нуклеиновой кислоты по варианту 28, где указанная нуклеотидная последовательность по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485.

30 31. Молекула нуклеиновой кислоты по варианту 28, отличающаяся тем, что указанная нуклеотидная последовательность на 100% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485.

32. Молекула нуклеиновой кислоты по варианту 28, где указанная нуклеотидная последовательность кодирует аминокислотную последовательность, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

33. Молекула нуклеиновой кислоты по варианту 28, где указанная нуклеотидная последовательность кодирует аминокислотную последовательность, на 100% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

5 34. Молекула нуклеиновой кислоты по любому из вариантов 28-33, где дезаминаза представляет собой адениндезаминазу.

35. Молекула нуклеиновой кислоты по любому из вариантов 28-34, где ДНК-связывающий полипептид представляет собой мегануклеазу, цинк-пальцевый гибридный белок или TALEN.

10 36. Молекула нуклеиновой кислоты по любому из вариантов 28-34, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

37. Молекула нуклеиновой кислоты по варианту 36, где РНК-направляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

15 38. Молекула нуклеиновой кислоты по варианту 37, где RGN представляет собой полипептид CRISPR-Cas типа II.

39. Молекула нуклеиновой кислоты по варианту 37, где RGN представляет собой полипептид CRISPR-Cas типа V.

20 40. Молекула нуклеиновой кислоты по любому из вариантов 37-39, где RGN представляет собой RGN-никазу.

41. Молекула нуклеиновой кислоты по варианту 37, отличающаяся тем, что RGN характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

25 42. Молекула нуклеиновой кислоты по варианту 37, где RGN представляет собой последовательность SEQ ID NO: 41, 60, 366 или 368.

43. Молекула нуклеиновой кислоты по варианту 40, где RGN-никаза представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 44. Молекула нуклеиновой кислоты по любому из вариантов 28-43, где полинуклеотид, кодирующий гибридный белок, функционально связан по своему 5'-концу с гетерологичным промотором.

45. Молекула нуклеиновой кислоты по любому из вариантов 28-44, где полинуклеотид, кодирующий гибридный белок, функционально связан по своему 3'-концу с гетерологичным терминатором.

5 46. Молекула нуклеиновой кислоты по любому из вариантов 28-45, где гибридный белок содержит один или более сигналов ядерной локализации.

47. Молекула нуклеиновой кислоты по любому из вариантов 28-46, где гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

10 48. Молекула нуклеиновой кислоты по любому из вариантов 28-46, где гибридный белок оптимизирован по кодонам для экспрессии в прокариотической клетке.

49. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 28-48.

15 50. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 28-48, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (гРНК), способную гибридизироваться с последовательностью-мишенью.

51. Вектор по варианту 50, где gRNA представляет собой одиночную направляющую РНК.

20 52. Вектор по варианту 50, где gRNA представляет собой двойную направляющую РНК.

53. Клетка, содержащая гибридный белок по любому из вариантов 14-27.

54. Клетка, содержащая гибридный белок по любому из вариантов 14-27, где клетка дополнительно содержит направляющую РНК.

25 55. Клетка, содержащая молекулу нуклеиновой кислоты по любому из вариантов 28-48.

56. Клетка, содержащая вектор по настоящему изобретению по любому из вариантов 49-52.

30 57. Клетка по любому из вариантов 53-56, отличающаяся тем, что клетка представляет собой прокариотическую клетку.

58. Клетка по любому из вариантов 53-56, где клетка представляет собой эукариотическую клетку.

59. Клетка по варианту 58, где эукариотическая клетка представляет собой клетку млекопитающего.

60. Клетка по варианту 59, где клетка млекопитающего представляет собой клетку человека.

61. Клетка по варианту 60, где клетка человека представляет собой иммунную клетку.

5 62. Клетка по варианту 61, где иммунная клетка представляет собой стволовую клетку.

63. Клетка по варианту 62, где стволовая клетка представляет собой индуцированную плюрипотентную стволовую клетку.

10 64. Клетка по варианту 58, где эукариотическая клетка представляет собой клетку насекомого или птицы.

65. Клетка по варианту 58, где эукариотическая клетка представляет собой грибковую клетку.

66. Клетка по варианту 58, где эукариотическая клетка представляет собой растительную клетку.

15 67. Растение, содержащее клетку по варианту 66.

68. Семя, содержащее клетку по варианту 66.

20 69. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из вариантов 14-27, молекулу нуклеиновой кислоты по любому из вариантов 28-48, вектор по любому из вариантов 49-52 или клетку по любому из вариантов 59-63.

70. Способ получения гибридного белка, включающий культивирование клетки по любому из вариантов 53-66 в условиях, в которых экспрессируется гибридный белок.

25 71. Способ получения гибридного белка, включающий введение в клетку молекулы нуклеиновой кислоты по любому из вариантов 28-48 или вектора по любому из вариантов 49-52 и культивирование клетки в условиях, в которых экспрессируется гибридный белок.

72. Способ по вариантам 70 или 71, дополнительно включающий очистку указанного гибридного белка.

30 73. Способ получения гибридного рибонуклеопротеинового комплекса RGN, включающий введение в клетку молекулы нуклеиновой кислоты по любому из вариантов 37-43 и молекулы нуклеиновой кислоты, содержащей кассету экспрессии, кодирующую направляющую РНК, или вектор по любому из вариантов 50-52, и культивирование клетки в условиях, в которых гибридный

белок и gRNA экспрессируются и образуют гибридный рибонуклеопротеиновый комплекс RGN.

74. Способ по варианту 73, дополнительно включающий очистку указанного гибридного рибонуклеопротеинового комплекса RGN.

5 75. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 10 407, 405, 399, 1-10, 400-404, 406 и 408-441, или нуклеотидную последовательность, кодирующую указанный гибридный белок, и

б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более 15 направляющих РНК (гРНК), и где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

76. Система по варианту 75, отличающаяся тем, что указанная деаминаза 20 характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408- 441.

77. Система по варианту 75, где указанная деаминаза характеризуется аминокислотной последовательностью, на 100% идентичной любой 25 последовательности из SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

78. Система по любому из вариантов 75-77, где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующих одну или более направляющих РНК, и указанных нуклеотидных последовательностей, кодирующих гибридный белок, функционально связана с промотором, 30 гетерологичным указанной нуклеотидной последовательности.

79. Система по любому из вариантов 75-78, где последовательность ДНК-мишени представляет собой последовательность ДНК-мишени эукариот.

80. Система по любому из вариантов 75-79, где целевая последовательность ДНК расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается RGN.

5 81. Система по любому из вариантов 75-80, где молекула ДНК-мишени находится внутри клетки.

82. Система по варианту 81, где клетка представляет собой эукариотическую клетку.

83. Система по варианту 82, где эукариотическая клетка представляет собой растительную клетку.

10 84. Система по варианту 82, где эукариотическая клетка представляет собой клетку млекопитающего.

85. Система по варианту 84, где клетка млекопитающего представляет собой клетку человека.

15 86. Система по варианту 85, в которой человеческая клетка представляет собой иммунную клетку.

87. Система по варианту 86, где иммунная клетка представляет собой стволовую клетку.

88. Система по варианту 87, где стволовая клетка представляет собой индуцированную плюрипотентную стволовую клетку.

20 89. Система по варианту 82, где эукариотическая клетка представляет собой клетку насекомого.

90. Система по варианту 81, где клетка является прокариотической клеткой.

91. Система по любому из вариантов 75-90, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II.

25 92. Система по любому из вариантов 75-90, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа V.

93. Система по любому из вариантов 75-90, отличающаяся тем, что RGN гибридного белка имеет аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности из последовательностей SEQ ID NO: 41, 60, 366 или 368.

30 94. Система по любому из вариантов 75-90, отличающаяся тем, что RGN гибридного белка характеризуется любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 и 368.

95. Система по любому из вариантов 75-90, где RGN гибридного белка представляет собой RGN-никазу.

96. Система по варианту 95, где никаза RGN представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

5 97. Система по любому из вариантов 75-96, где гибридный белок содержит один или более сигналов ядерной локализации.

98. Система по любому из вариантов 75-97, отличающаяся тем, что гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

10 99. Система по любому из вариантов 75-98, где нуклеотидные последовательности, кодирующие одну или более направляющих РНК, и нуклеотидная последовательность, кодирующая гибридный белок, расположены на одном векторе.

15 100. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и систему по любому из вариантов 75-99.

101. Способ модификации молекулы ДНК-мишени, содержащей последовательность ДНК-мишени, причем указанный способ включает доставку системы по любому из вариантов 75-99 в указанную молекулу ДНК-мишени или в клетку, содержащую молекулу ДНК-мишени.

20 102. Способ по варианту 101, где указанная модифицированная молекула ДНК-мишени содержит мутацию A>N по меньшей мере одного нуклеотида в молекуле ДНК-мишени, где N представляет собой C, G или T.

25 103. Способ по варианту 102, где указанная модифицированная молекула ДНК-мишени содержит мутацию A>G по меньшей мере одного нуклеотида в молекуле ДНК-мишени.

104. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза *in vitro* с использованием объединения:

30 i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК и

ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная

последовательность дезаминазы по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441 в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

5 б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с комплексом рибонуклеотидов RGN-дезаминазы, собранным *in vitro*,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

105. Способ по варианту 104, где указанная дезаминаза характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408- 441.

106. Способ по варианту 104, где указанная дезаминаза характеризуется аминокислотной последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

107. Способ по любому из вариантов 104-106, отличающийся тем, что указанная модифицированная молекула ДНК-мишени содержит мутацию A>N по меньшей мере одного нуклеотида в молекуле ДНК-мишени, где N представляет собой C, G или T.

108. Способ по варианту 107, где указанная модифицированная молекула ДНК-мишени содержит мутацию A>G по меньшей мере одного нуклеотида в молекуле ДНК-мишени.

109. Способ по любому из вариантов 104-108, отличающийся тем, что RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II.

110. Способ по любому из вариантов 104-108, отличающийся тем, что RGN гибридного белка представляет собой полипептид CRISPR-Cas типа V.

111. Способ по любому из вариантов 104-108, отличающийся тем, что RGN гибридного белка характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 41, 60, 366 или 368.

112. Способ по любому из вариантов 104-108, отличающийся тем, что RGN гибридного белка характеризуется любой аминокислотной последовательностью любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

5 113. Способ по любому из вариантов 104-108, где RGN гибридного белка представляет собой RGN-никазу.

114. Способ по варианту 113, отличающийся тем, что никаза RGN представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

10 115. Способ по любому из вариантов 104-114, где гибридный белок содержит один или более сигналов ядерной локализации.

116. Способ по любому из вариантов 104-115, где гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

15 117. Способ по любому из вариантов 104-116, где указанная последовательность ДНК-мишени представляет собой последовательность ДНК-мишени эукариот.

118. Способ по любому из вариантов 104-117, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (РАМ).

20 119. Способ по любому из вариантов 104-118, где молекула ДНК-мишени находится внутри клетки.

120. Способ по варианту 119, где клетка представляет собой эукариотическую клетку.

121. Способ по варианту 120, где эукариотическая клетка представляет собой растительную клетку.

25 122. Способ по варианту 120, где эукариотическая клетка представляет собой клетку млекопитающего.

123. Способ по варианту 122, где клетка млекопитающего представляет собой клетку человека.

30 124. Способ по варианту 123, где человеческая клетка представляет собой иммунную клетку.

125. Способ по варианту 124, где иммунная клетка представляет собой стволовую клетку.

126. Способ по варианту 125, где стволовая клетка представляет собой индуцированную плюрипотентную стволовую клетку.

127. Способ по варианту 120, где эукариотическая клетка представляет собой клетку насекомого.

128. Способ по варианту 119, где клетка является прокариотической клеткой.

5 129. Способ по любому из вариантов 119-128, дополнительно включающий отбор клетки, содержащей указанную модифицированную молекулу ДНК.

130. Клетка, содержащая модифицированную последовательность ДНК-мишени в соответствии со способом по варианту 129.

10 131. Клетка по варианту 130, где клетка представляет собой эукариотическую клетку.

132. Клетка по варианту 131, где эукариотическая клетка представляет собой растительную клетку.

133. Растение, содержащее клетку по варианту 132.

134. Семя, содержащее клетку по варианту 132.

15 135. Клетка по варианту 131, где эукариотическая клетка представляет собой клетку млекопитающего.

136. Клетка по варианту 135, где клетка млекопитающего представляет собой клетку человека.

20 137. Клетка по варианту 136, где клетка человека представляет собой иммунную клетку.

138. Клетка по варианту 137, где иммунная клетка представляет собой стволовую клетку.

139. Клетка по варианту 138, где стволовая клетка представляет собой индуцированную плюрипотентную стволовую клетку.

25 140. Клетка по варианту 131, где эукариотическая клетка представляет собой клетку насекомого.

141. Клетка по варианту 130, где клетка является прокариотической клеткой.

30 142. Фармацевтическая композиция, содержащая клетку по любому из вариантов 135-139 и фармацевтически приемлемый носитель.

143. Способ получения генетически модифицированной клетки с коррекцией причинной мутации наследственного заболевания, включающий введение в клетку:

а) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность дезаминазы характеризуется по меньшей мере 90%-ой идентичностью с любой из последовательностей SEQ ID NO: 407, 405, 399, 1- 10, 400-404, 406 и 408-441, или полинуклеотида, кодирующего указанный гибридный белок, где указанный полинуклеотид, кодирующий гибридный белок, функционально связан с промотором для обеспечения экспрессии гибридного белка в клетке, и

б) одну или более направляющих РНК (gRNA), способных гибридизоваться с целевой последовательностью ДНК, или полинуклеотид, кодирующий указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, функционально связан с промотором для обеспечения экспрессии gRNA в клетке,

при этом гибридный белок и gRNA направлены на геномное местоположение причинной мутации и модифицируют геномную последовательность с целью удаления причинной мутации.

144. Способ по варианту 143, где указанная дезаминаза характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408- 441.

145. Способ по варианту 143, где указанная деаминаза характеризуется аминокислотной последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

146. Способ по любому из вариантов 143-145, где указанный RGN гибридного белка представляет собой RGN-никазу.

147. Способ по варианту 146, отличающийся тем, что RGN-никаза представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

148. Способ по любому из вариантов 143-147, где модификация генома включает введение мутации A>G по меньшей мере одного нуклеотида в последовательность ДНК-мишени.

149. Способ по любому из вариантов 143-148, где клетка представляет собой клетку животного.

150. Способ по варианту 149, где клетка животного представляет собой клетку млекопитающего.

151. Способ по варианту 150, где клетку получают от собаки, кошки, мыши, крысы, кролика, лошади, овцы, козы, коровы, свиньи или человека.

152. Способ по любому из вариантов 143-151, отличающийся тем, что коррекция причинной мутации включает коррекцию нонсенс-мутации.

5 153. Способ по варианту 149, где генетически наследуемое заболевание представляет собой заболевание, указанное в таблице 34.

154. Способ по варианту 149, где генетически наследуемое заболевание представляет собой муковисцидоз.

10 155. Способ по варианту 154, где гРНК дополнительно содержит спейсерную последовательность, нацеленную на любую из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563, или их комплементарные последовательности.

15 156. Способ по варианту 155, где гРНК содержит любую из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

157. CRISPR РНК (crRNA) или молекула нуклеиновой кислоты, кодирующая ее, где указанная CRISPR РНК содержит спейсерную последовательность, которая нацелена на целевую последовательность ДНК в гене трансмембранного регулятора проводимости при муковисцидозе (CFTR),
20 где указанная целевая последовательность имеет любую из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364, 562 и 563 или их комплементарных последовательностей.

158. Направляющая РНК, содержащая crRNA по варианту 157.

25 159. Направляющая РНК по варианту 158, где указанная направляющая РНК представляет собой двойную направляющую РНК.

160. Направляющая РНК по варианту 158, где указанная направляющая РНК представляет собой одиночную направляющую РНК (sgRNA).

30 161. Направляющая РНК по варианту 160, где указанная sgRNA характеризуется по меньшей мере на 90% идентичностью с любой из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

162. Направляющая РНК по варианту 160, где указанная sgRNA характеризуется по меньшей мере на 95% идентичностью с любой из

последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

163. Направляющая РНК по варианту 160, где указанная sgRNA характеризуется последовательностью, представленную в виде любой из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

164. Вектор, содержащий одну или более молекул нуклеиновой кислоты, кодирующих указанную направляющую РНК по любому из вариантов 158-163.

165. Система для связывания целевой последовательности молекулы ДНК-мишени, содержащая:

а) одну или более направляющих РНК, способных гибридизоваться с указанной целевой последовательностью ДНК, или одним или более полинуклеотидами, содержащими одну или более нуклеотидных последовательностей, кодирующих одну или более направляющих РНК (gRNA) и

б) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и адениндезаминазу, или полинуклеотид, содержащий нуклеотидную последовательность, кодирующую гибридный белок,

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

где одна или более направляющих РНК способны образовывать комплекс с полипептидом RGN с целью направить указанный полипептид RGN на связывание с указанной целевой последовательностью ДНК-мишени и

где по меньшей мере одна направляющая РНК содержит CRISPR РНК (crРНК), содержащую спейсерную последовательность, нацеленную на целевую последовательность ДНК в гене трансмембранного регулятора проводимости при муковисцидозе (CFTR), где указанная целевая последовательность имеет последовательность, указанную в виде любой из последовательностей SEQ ID NOs: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364, 562 и 563 или их комплементарных последовательностей.

166. Система по варианту 165, где по меньшей мере одна из указанных нуклеотидных последовательностей кодирует одну или более направляющих РНК, и указанная нуклеотидная последовательность, кодирующая гибридный

белок, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности.

167. Система для связывания целевой последовательности молекулы ДНК-мишени, содержащая:

- 5 а) одну или более направляющих РНК, способных гибридизоваться с указанной целевой последовательностью ДНК, или одним или более полинуклеотидами, содержащими одну или более нуклеотидных последовательностей, кодирующих одну или более направляющих РНК (gRNA) и
- 10 б) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и адениндезаминазу,
- где одна или более направляющих РНК способны гибридизоваться с последовательностью ДНК-мишени,
- где одна или более направляющих РНК способны образовывать комплекс с
- 15 полипептидом RGN, с целью направить указанный полипептид RGN на связывание с указанной последовательностью ДНК-мишени, и
- где по меньшей мере одна направляющая РНК содержит CRISPR РНК (сгРНК), содержащую спейсерную последовательность, направленную на последовательность ДНК-мишени в гене трансмембранного регулятора
- 20 проводимости при муковисцидозе (CFTR), где указанная целевая последовательность имеет любую из последовательностей SEQ ID NOs: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364, 562 и 563 или их комплементарных последовательностей.

168. Система по варианту 167, в которой по меньшей мере одна из

25 указанных нуклеотидных последовательностей, кодирующих одну или более направляющих РНК, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности.

169. Система по любому из вариантов 165-168, отличающаяся тем, что аминокислотная последовательность дезаминазы по меньшей мере на 90%

30 идентична любой из последовательностей SEQ ID NO: 1-10 и 399-441.

170. Система по любому из вариантов 165-168, отличающаяся тем, что аминокислотная последовательность дезаминазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 1-10 и 399-441.

171. Система по любому из вариантов 165-168, отличающаяся тем, что дезаминаза характеризуется аминокислотной последовательностью, указанной в любой из последовательностей SEQ ID NO: 1-10 и 399-441.

5 172. Система по любому из вариантов 165-171, где указанный полипептид RGN и указанная одна или более направляющих РНК в природе не образуют комплексов друг с другом.

173. Система по любому из вариантов 165-172, отличающаяся тем, что:

10 а) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 62-68, 80-85, 116-119, 128-131, 163, 164, 180, 181, 203-209, 219-225, 256-258, 274-276, 310-313 и 330-333 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичную последовательности с SEQ ID NO: 53,

15 б) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 68-71, 86-89, 120-122, 132-134, 152-156, 169-173, 213-215, 229-231, 251-255, 269-273, 305-309 и 325-329 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 55,

20 в) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 72, 73, 90, 91, 161, 162, 178, 179, 265, 266, 283 и 284 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 52,

25 г) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 74, 75, 92, 93, 123, 124, 135, 136, 167, 184, 216-218, 232-234, 259-261, 277-279, 314-317 и 334-337 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 56,

30 д) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 76, 94, 210-212, 226-228, 322, 342, 562 и 563 или комплементарными им последовательностями, и где указанный полипептид

RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 42,

е) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 77, 95, 125, 137, 157-160, 174-177, 323 и 343 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 54,

ж) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 78, 96, 126, 138, 168, 185, 267, 285, 318, 319, 338 и 339, или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 57,

з) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 79, 97, 127, 139, 262-264, 280-282, 324 и 344, или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 58, и

и) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 165, 166, 182, 183, 268, 286, 320, 321, 340 и 341 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности с SEQ ID NO: 59.

174. Система по любому из вариантов 165-172, отличающаяся тем, что:

а) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 62-68, 80-85, 116-119, 128-131, 163, 164, 180, 181, 203-209, 219-225, 256-258, 274-276, 310-313 и 330-333 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 53,

б) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 68-71, 86-89, 120-122, 132-134, 152-156, 169-173, 213-215, 229-231, 251-255, 269-273, 305-309 и 325-329 или комплементарными им последовательностями, и где указанный полипептид RGN

характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 55,

5 в) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 72, 73, 90, 91, 161, 162, 178, 179, 265, 266, 283 и 284, или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 52,

10 г) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 74, 75, 92, 93, 123, 124, 135, 136, 167, 184, 216-218, 232-234, 259-261, 277-279, 314-317 и 334-337 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 56,

15 д) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 76, 94, 210-212, 226-228, 322, 342, 562 и 563 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 42,

20 е) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 77, 95, 125, 137, 157-160, 174-177, 323 и 343 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 54,

25 ж) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 78, 96, 126, 138, 168, 185, 267, 285, 318, 319, 338 и 339, или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 57,

30 з) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 79, 97, 127, 139, 262-264, 280-282, 324 и 344, или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности с SEQ ID NO: 58, и

и) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 165, 166, 182, 183, 268, 286, 320, 321, 340 и 341 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95%

5 идентичной последовательности с SEQ ID NO: 59.

175. Система по любому из вариантов 165-172, где:

а) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 62-68, 80-85, 116-119, 128-131, 163, 164, 180, 181, 203-209, 219-225, 256-258, 274-276, 310-313 и 330-333 или

10 комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 53,

б) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 68-71, 86-89, 120-122, 132-134, 152-156, 169-
15 173, 213-215, 229-231, 251-255, 269-273, 305-309 и 325-329 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичную последовательности с SEQ ID NO: 55,

в) указанная последовательность ДНК-мишени характеризуется любой из
20 последовательностей SEQ ID NO: 72, 73, 90, 91, 161, 162, 178, 179, 265, 266, 283 и 284 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 52,

г) указанная последовательность ДНК-мишени характеризуется любой из
25 последовательностей SEQ ID NO: 74, 75, 92, 93, 123, 124, 135, 136, 167, 184, 216-218, 232-234, 259-261, 277-279, 314-317 и 334-337 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 56,

д) указанная последовательность ДНК-мишени характеризуется любой из
30 последовательностей SEQ ID NO: 76, 94, 210-212, 226-228, 322, 342, 562 и 563 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 42,

е) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 77, 95, 125, 137, 157-160, 174-177, 323 и 343 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной

5 последовательности SEQ ID NO: 54,

ж) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 78, 96, 126, 138, 168, 185, 267, 285, 318, 319, 338 и 339 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной

10 последовательности SEQ ID NO: 57,

з) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 79, 97, 127, 139, 262-264, 280-282, 324 и 344 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной

15 последовательности SEQ ID NO: 58, и

и) указанная последовательность ДНК-мишени характеризуется любой из последовательностей SEQ ID NO: 165, 166, 182, 183, 268, 286, 320, 321, 340 и 341 или комплементарными им последовательностями, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной

20 последовательности SEQ ID NO: 59.

176. Система по любому из вариантов 165-175, где по меньшей мере одна направляющая РНК представляет собой двойную направляющую РНК.

177. Система по любому из вариантов 165-175, где по меньшей мере одна направляющая РНК представляет собой одиночную направляющую РНК

25 (sgRNA).

178. Система по варианту 177, в которой:

а) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 98-104, 140-143, 197, 198, 235-241, 292-294 и 350-353, и где указанный полипептид RGN характеризуется

30 последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 53,

б) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 104-107, 144-146, 186-190, 245-247, 287-291 и 345-349, и где указанный полипептид RGN характеризуется

последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 55,

5 в) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 108, 109, 195, 196, 301 и 302, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 52,

10 г) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 110, 111, 147, 148, 201, 248-250, 295-297 и 354-357, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 56,

д) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 112, 242-244, 362 и 564, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 42,

15 е) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 113, 149, 191-194 и 363, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 54,

20 ж) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 114, 150, 202, 303, 358 и 359, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: HET: 57,

25 з) указанная sgRNA по меньшей мере на 90% идентична любой последовательности SEQ ID NO: 115, 151, 298-300 и 364, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 58, и

и) указанная sgRNA по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 199, 200, 304, 360 и 361, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 59.

179. Система по варианту 177, в которой:

а) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 98-104, 140-143, 197, 198, 235-241, 292-294 и 350-353, и где указанный полипептид RGN характеризуется

последовательностью по меньшей мере на 95% идентичной последовательности SEQ ID NO: 53,

б) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 104-107, 144-146, 186-190, 245-247, 287-291 и 345-349, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 55,

в) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 108, 109, 195, 196, 301 и 302, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере 95% идентичной последовательности SEQ ID NO: HET: 52,

г) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 110, 111, 147, 148, 201, 248-250, 295-297 и 354-357, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 56,

д) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 112, 242-244, 362 и 564, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 42,

е) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 113, 149, 191-194 и 363, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 54,

ж) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 114, 150, 202, 303, 358 и 359, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: HET: 57,

з) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 115, 151, 298-300 и 364, и где указанный полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 58, и

и) указанная sgRNA по меньшей мере на 95% идентична любой из последовательностей с SEQ ID NO: 199, 200, 304, 360 и 361, и где указанный

полипептид RGN характеризуется последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 59.

180. Система по варианту 177, в которой:

- а) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 98-104, 140-143, 197, 198, 235-241, 292-294 и 350-353, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 53,
- б) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 104-107, 144-146, 186-190, 245-247, 287-291 и 345-349, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 55,
- в) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 108, 109, 195, 196, 301 и 302, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 52,
- г) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 110, 111, 147, 148, 201, 248-250, 295-297 и 354-357, и где указанный полипептид RGN характеризуется последовательностью, на 100 % идентичной последовательности SEQ ID NO: 56,
- д) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 112, 242-244, 362 и 564, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 42,
- е) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 113, 149, 191-194 и 363, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 54,
- ж) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 114, 150, 202, 303, 358 и 359, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 57,

з) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 115, 151, 298-300 и 364, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 58, и

5 и) указанная sgRNA характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 199, 200, 304, 360 и 361, и где указанный полипептид RGN характеризуется последовательностью, на 100% идентичной последовательности SEQ ID NO: 59.

10 181. Клетка, содержащая молекулу crRNA или нуклеиновой кислоты по варианту 157, направляющую РНК по любому из вариантов 158-163, вектор по варианту 164 или систему по любому из вариантов 165-180.

15 182. Фармацевтическая композиция, содержащая crRNA или молекулу нуклеиновой кислоты по варианту 157, направляющую РНК по любому из вариантов 158-163, вектор по варианту 164, клетку по варианту 181 или систему по любому из вариантов 165- 180 и фармацевтически приемлемый носитель.

183. Композиция, содержащая:

а) гибридный белок, содержащий ДНК-связывающий полипептид и адениндезаминазу, или молекулу нуклеиновой кислоты, кодирующую гибридный белок, и

20 б) вторую адениндезаминазу, характеризующуюся последовательностью, по меньшей мере на 90% идентичной любой последовательности SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441 или молекулу нуклеиновой кислоты, кодирующую дезаминазу.

25 184. Композиция по варианту 183, где указанная вторая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

30 185. Композиция по варианту 183, где указанная вторая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

186. Композиция по любому из вариантов 183-185, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

187. Композиция по любому из вариантов 183-186, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

5 188. Композиция по любому из вариантов 183-186, где первая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408- 441.

10 189. Композиция по любому из вариантов 183-188, где ДНК-связывающий полипептид представляет собой мегануклеазу, цинк-пальцевый гибридный белок или TALEN.

190. Композиция по любому из вариантов 183-189, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

15 191. Композиция по варианту 190, где РНК-направляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

192. Композиция по варианту 191, где RGN представляет собой RGN-никазу.

20 193. Вектор, содержащий молекулу нуклеиновой кислоты, кодирующую гибридный белок, и молекулу нуклеиновой кислоты, кодирующую вторую дезаминазу, где указанный гибридный белок содержит ДНК-связывающий полипептид и первую адениндезаминазу, и где указанная вторая адениндезаминаза содержит последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

25 194. Вектор по варианту 193, где указанная вторая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 30 408-441.

195. Вектор по варианту 193, где указанная вторая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

196. Вектор по любому из вариантов 193-195, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

5 197. Вектор по любому из вариантов 193-195, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

10 198. Вектор по любому варианту 193-195, где первая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

199. Вектор по любому из вариантов 193-198, где ДНК-связывающий полипептид представляет собой мегануклеазу, гибридный цинк-пальцевый белок или TALEN.

15 200. Вектор по любому из вариантов 193-198, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

20 201. Вектор по варианту 200, где РНК-направляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

202. Вектор по варианту 201, где RGN представляет собой RGN-никазу.

203. Клетка, содержащая вектор по любому из вариантов 193-202.

204. Клетка, содержащая:

25 а) гибридный белок, содержащий ДНК-связывающий полипептид и первую адениндезаминазу, или молекула нуклеиновой кислоты, кодирующая гибридный белок, и

б) вторую адениндезаминазу, характеризующуюся последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441, или молекула нуклеиновой кислоты, кодирующая вторую адениндезаминазу.

30 205. Клетка по варианту 204, где указанная вторая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

206. Клетка по варианту 204, где указанная вторая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

5 207. Клетка по любому из вариантов 204-206, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

10 208. Клетка по любому из вариантов 204-206, где первая адениндезаминаза характеризуется последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

209. Клетка по любому из вариантов 204-206, где первая адениндезаминаза характеризуется последовательностью, на 100% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408- 441.

15 210. Клетка по любому из вариантов 204-209, где ДНК-связывающий полипептид представляет собой мегануклеазу, гибридный цинк-пальцевый белок или TALEN.

20 211. Клетка по любому из вариантов 204-209, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

212. Клетка по варианту 211, где РНК-направляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

213. Клетка по варианту 212, где RGN представляет собой RGN-никазу.

25 214. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и композицию по любому из вариантов 183-192, вектор по любому из вариантов 193-202 или клетку по любому из вариантов 203-213.

30 215. Способ лечения заболевания, включающий введение субъекту, нуждающемуся в таком введении, эффективного количества фармацевтической композиции по любому из вариантов 69, 100, 142 и 214.

216. Способ по варианту 215, где указанное заболевание связано с причинной мутацией, и указанное эффективное количество указанной фармацевтической композиции корректирует указанную причинную мутацию.

217. Применение гибридного белка по любому из вариантов 14-27, молекула нуклеиновой кислоты по любому из вариантов 28-48, вектор по любому из вариантов 49-52 и 193-202, клетка по любому из вариантов 59-63, 135-139 и 203-213, система по любому из вариантов 75-99 или композиция по
5 любому из вариантов 183-192 для лечения заболевания у субъекта.

218. Применение по варианту 217, где указанное заболевание связано с причинной мутацией, и указанное лечение включает коррекцию указанной причинной мутации.

219. Применение гибридного белка по любому из вариантов 14-27, молекулы нуклеиновой кислоты по любому из вариантов 28-48, вектора по
10 любому из вариантов 49-52 и 193-202, клетки по любому из вариантов 59-63, 135-139 и 203-213, системы по любому из вариантов 75-99 или композиции по любому из вариантов 183-192 для изготовления лекарственного средства, пригодного для лечения заболевания.

15 220. Применение по варианту 219, где указанное заболевание связано с причинной мутацией, и эффективное количество указанного лекарственного средства корректирует указанную причинную мутацию.

221. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид РНК-направляемой нуклеазы (RGN), где указанный
20 полинуклеотид содержит нуклеотидную последовательность, кодирующую полипептид RGN, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 41 или 60, но не содержащую аминокислотные остатки с 590 по 597 в последовательности SEQ ID NO: 41 или 60,

25 где указанный полипептид RGN способен связываться с последовательностью ДНК-мишени в РНК-направляемой последовательности специфическим образом при связывании с направляющей РНК (gRNA), способной гибридизоваться с указанной последовательностью ДНК-мишени.

222. Молекула нуклеиновой кислоты по варианту 221, где указанный
30 полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

223. Молекула нуклеиновой кислоты по вариантам 221 или 222, где указанный полипептид RGN содержит аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 366 или 368.

224. Молекула нуклеиновой кислоты по вариантам 221 или 222, где указанный полипептид RGN содержит аминокислотную последовательность SEQ ID NO: 366 или 368.

5 225. Молекула нуклеиновой кислоты по любому из вариантов 221-223, где указанный полипептид RGN представляет собой мертвую нуклеазу или функционирует как никаза.

226. Молекула нуклеиновой кислоты по варианту 225, где указанная никаза имеет аминокислотную последовательность SEQ ID NO: 397 или 398.

10 227. Молекула нуклеиновой кислоты по любому из вариантов 221-226, где полипептид RGN функционально слит с полипептидом, редактирующим азотистое основание.

228. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 221-227.

15 229. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 41 или 60, но не содержащий аминокислотных остатков с 590 по 597 в последовательности SEQ ID NO: 41 или 60, где указанный полипептид представляет собой РНК- направляемую нуклеазу.

20 230. Выделенный полипептид по варианту 229, где указанный полипептид RGN содержит аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 366 или 368.

231. Выделенный полипептид по варианту 230, где указанный полипептид RGN содержит аминокислотную последовательность SEQ ID NO: 366 или 368.

25 232. Выделенный полипептид по вариантам 229 или 230, где указанный полипептид RGN является мертвой нуклеазой или функционирует как никаза.

233. Выделенный полипептид по варианту 232, где указанная никаза характеризуется аминокислотной последовательностью SEQ ID NO: 397 или 398.

30 234. Выделенный полипептид по любому из вариантов 229-233, где полипептид RGN функционально слит с полипептидом, редактирующим азотистое основание.

235. Клетка, содержащая молекулу нуклеиновой кислоты по любому из вариантов 221-227, вектор по варианту 228 или полипептид по любому из вариантов 229-234.

236. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 407, где указанный полипептид обладает деаминазной активностью.

5 237. Выделенный полипептид по варианту 236, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 407, где указанный полипептид обладает деаминазной активностью.

238. Выделенный полипептид по варианту 236, где полипептид содержит аминокислотную последовательность SEQ ID NO: 407.

10 239. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 451, или

15 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407.

240. Молекула нуклеиновой кислоты по варианту 239, где дезаминаза кодируется нуклеотидной последовательностью, которая по меньшей мере на 90% идентична последовательности SEQ ID NO: 451.

20 241. Молекула нуклеиновой кислоты по варианту 239, где дезаминаза кодируется нуклеотидной последовательностью, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 451.

242. Молекула нуклеиновой кислоты по варианту 239, где дезаминаза кодируется нуклеотидной последовательностью, которая по меньшей мере на 25 100% идентична последовательности SEQ ID NO: 451.

243. Молекула нуклеиновой кислоты по вариантам 239-242, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

30 244. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из вариантов 236-238 или молекулу нуклеиновой кислоты по любому из вариантов 239-242.

245. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 407.

246. Гибридный белок по варианту 245, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующуюся последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 407.

5 247. Гибридный белок по варианту 245, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующуюся последовательностью, на 100% идентичной последовательности SEQ ID NO: 407.

248. Гибридный белок по любому из вариантов 245-247, где ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

10 249. Гибридный белок по варианту 248, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

250. Гибридный белок по любому из вариантов 248-249, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, 15 xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

20 251. Гибридный белок по любому из вариантов 248-250, где полипептид RGN представляет собой нуклеазу.

252. Гибридный белок по варианту 251, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

25 253. Гибридный белок по варианту 251, где аминокислотная последовательность нуклеазы на 100% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 254. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 451, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 407.

255. Молекула нуклеиновой кислоты по варианту 254, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 451.

5 256. Молекула нуклеиновой кислоты по варианту 254, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 451.

257. Молекула нуклеиновой кислоты по варианту 254, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 100% идентичной последовательности SEQ ID NO: 451.

10 258. Молекула нуклеиновой кислоты по любому из вариантов 254-257, где ДНК-связывающий полипептид представляет собой полипептид RGN.

259. Молекула нуклеиновой кислоты по варианту 258, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

15 260. Молекула нуклеиновой кислоты по любому из вариантов 258-259, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или
20 полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

261. Молекула нуклеиновой кислоты по любому из вариантов 258-260, где полипептид RGN представляет собой нуклеазу.

25 262. Молекула нуклеиновой кислоты по варианту 261, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

263. Молекула нуклеиновой кислоты по варианту 262, где аминокислотная последовательность нуклеазы на 100% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 264. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 254-263.

265. Вектор по варианту 264, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизоваться с последовательностью-мишенью.

266. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из вариантов 245-253 и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

267. Клетка, содержащая гибридный белок по любому из вариантов 245-253, молекулу нуклеиновой кислоты по любому из вариантов 254-263, вектор по любому из вариантов 264-265 или комплекс RNP по варианту 266.

268. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 407, или нуклеотидная последовательность, кодирующая указанный гибридный белок, и

б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (gRNA), и

где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

269. Система по варианту 268, где указанная деаминаза характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 407.

270. Система по варианту 268, где указанная деаминаза характеризуется аминокислотной последовательностью, на 100% идентичной последовательности SEQ ID NO: 407.

271. Система по любому из вариантов 268-270, где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующей одну или более направляющих РНК, и указанная нуклеотидная последовательность, кодирующая гибридный белок, функционально связаны с промотором, гетерологичным указанной нуклеотидной последовательности.

272. Система по любому из вариантов 268-271, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается полипептидом RGN.

273. Система по любому из вариантов 268-272, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563 или комплементарных им последовательностей.

274. Система по любому из вариантов 268-273, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

275. Система по любому из вариантов 268-274, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

276. Система по любому из вариантов 272-275, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

277. Система по варианту 276, где полипептид RGN представляет собой никазу.

278. Система по варианту 277, где аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

279. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из вариантов 245-253, молекулу нуклеиновой кислоты по любому из вариантов 254-263, вектор по любому из вариантов 264-265, комплекс RNP по варианту 266, клетку по варианту 267 или систему по любому из вариантов 268-278.

280. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза с использованием объединения:

i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и

ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная

5 последовательность дезаминазы характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 407,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

10 б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и
15 происходит модификация молекулы ДНК-мишени.

281. Способ по варианту 280, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305- 344, 562 и 563 или их комплементарных им последовательностей.

20 282. Способ по любому из вариантов 280-281, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

283. Способ по любому из вариантов 280-283, где способ осуществляется in vitro, in vivo или ex vivo.

284. Способ лечения субъекта, имеющего или подверженного риску развития заболевания, нарушения или состояния, включающий:

30 введение субъекту гибридного белка по любому из вариантов 245-253, молекулы нуклеиновой кислоты по любому из вариантов 254-263, вектора по любому из вариантов 264-265, комплекса RNP по варианту 266, клетки по варианту 267, системы по любому из вариантов 268-28 или фармацевтической композиции по варианту 279.

285. Способ по варианту 284, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из

группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

286. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 405, где указанный полипептид обладает деаминазной активностью.

287. Выделенный полипептид по варианту 286, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 405, где указанный полипептид обладает деаминазной активностью.

288. Выделенный полипептид по варианту 286, где полипептид содержит аминокислотную последовательность SEQ ID NO: 407.

289. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 449 или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 405.

290. Молекула нуклеиновой кислоты по варианту 289, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 449.

291. Молекула нуклеиновой кислоты по варианту 289, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 449.

292. Молекула нуклеиновой кислоты по варианту 289, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 100% идентичной последовательности SEQ ID NO: 449.

293. Молекула нуклеиновой кислоты по вариантам 289-292, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

294. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из вариантов 286-288 или молекулу нуклеиновой кислоты по любому из вариантов 289-293.

295. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 405.

5 296. Гибридный белок по варианту 295, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующуюся последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 405.

297. Гибридный белок по варианту 295, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, на 100% идентичной последовательности SEQ ID NO: 405.

10 298. Гибридный белок по любому из вариантов 295-297, где ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

299. Гибридный белок по варианту 298, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

15 300. Гибридный белок по любому из вариантов 298-299, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или
20 полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

301. Гибридный белок по любому из вариантов 298-300, где полипептид RGN представляет собой нуклеазу.

25 302. Гибридный белок по варианту 301, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

303. Гибридный белок по варианту 301, где аминокислотная последовательность нуклеазы на 100% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 304. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 449, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 405.

5 305. Молекула нуклеиновой кислоты по варианту 304, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 449.

10 306. Молекула нуклеиновой кислоты по варианту 304, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 449.

307. Молекула нуклеиновой кислоты по варианту 304, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 100% идентичной последовательности SEQ ID NO: 449.

15 308. Молекула нуклеиновой кислоты по любому из вариантов 304-307, где ДНК-связывающий полипептид представляет собой полипептид RGN.

309. Молекула нуклеиновой кислоты по варианту 308, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 310. Молекула нуклеиновой кислоты по любому из вариантов 308-309, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из
25 последовательностей SEQ ID NO: 41, 60, 366 или 368.

311. Молекула нуклеиновой кислоты по любому из вариантов 308-310, где полипептид RGN представляет собой никазу.

30 312. Молекула нуклеиновой кислоты по варианту 311, где аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

313. Молекула нуклеиновой кислоты по варианту 312, отличающаяся тем, что аминокислотная последовательность никазы на 100% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

314. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 304-313.

315. Вектор по варианту 314, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизоваться с последовательностью-мишенью.

316. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из вариантов 295-303 и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

317. Клетка, содержащая гибридный белок по любому из вариантов 295-303, молекулу нуклеиновой кислоты по любому из вариантов 304-313, вектор по любому из вариантов 314-315 или комплекс RNP по варианту 316.

318. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 405, или нуклеотидная последовательность, кодирующая указанный гибридный белок, и

б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (gRNA), и

где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишенью и модифицировать молекулу ДНК-мишени.

319. Система по варианту 318, где указанная деаминаза характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 405.

320. Система по варианту 318, где указанная деаминаза характеризуется аминокислотной последовательностью, на 100% идентичной последовательности SEQ ID NO: 405.

321. Система по любому из вариантов 318-320, где по меньшей мере одна из указанной нуклеотидной последовательности, кодирующей одну или более направляющих РНК, и указанная нуклеотидная последовательность,

кодирующая гибридный белок, функционально связаны с промотором, гетерологичным указанной нуклеотидной последовательности.

5 322. Система по любому из вариантов 318-321, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается полипептидом RGN.

323. Система по любому из вариантов 318-322, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251. -286, 305-344, 562 и 563 или их дополнения.

10 324. Система по любому из вариантов 318-323, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

15 325. Система по любому из вариантов 318-324, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 326. Система по любому из вариантов 322-325, отличающаяся тем, что полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

25 327. Система по варианту 326, где полипептид RGN представляет собой никазу.

328. Система по варианту 327, отличающаяся тем, что аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 329. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из вариантов 295-303, молекулу нуклеиновой кислоты по любому из пунктов 304-313, вектор по любому из вариантов 314-315, комплекс RNP по варианту 316, клетку по варианту 317 или систему по любому из вариантов 318-328.

330. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза с использованием объединения:

5 i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и

ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная последовательность дезаминазы по меньшей мере на 90% идентична

10 последовательности с SEQ ID NO: 405,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

15 где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

20 331. Способ по варианту 330, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305- 344, 562 и 563 или комплементарных им последовательностей.

25 332. Способ по любому из вариантов 330-331, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

333. Способ по любому из вариантов 330-332, где способ осуществляется *in vitro*, *in vivo* или *ex vivo*.

30 334. Способ лечения субъекта, страдающего или подверженного риску развития заболевания, нарушения или состояния, включающий:

введение субъекту гибридного белка по любому из вариантов 295-303, молекулы нуклеиновой кислоты по любому из вариантов 304-313, вектора по любому из вариантов 314-315, комплекса RNP по варианту 316, клетки по

варианту 317, системы по любому из вариантов 318-328 или фармацевтической композиции по варианту 329.

335. Способ по варианту 334, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

336. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 399, где указанный полипептид обладает деаминазной активностью.

337. Выделенный полипептид по варианту 336, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 399, где указанный полипептид обладает деаминазной активностью.

338. Выделенный полипептид по варианту 336, где полипептид содержит аминокислотную последовательность SEQ ID NO: 399.

339. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 443, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 399.

340. Молекула нуклеиновой кислоты по варианту 339, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 443.

341. Молекула нуклеиновой кислоты по варианту 339, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 443.

342. Молекула нуклеиновой кислоты по варианту 339, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 100% идентичной последовательности SEQ ID NO: 443.

343. Молекула нуклеиновой кислоты по вариантам 339-342, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

344. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из вариантов 336-338 или молекулу нуклеиновой кислоты по любому из вариантов 339-342.

5 345. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 399.

346. Гибридный белок по варианту 345, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 399.

10 347. Гибридный белок по варианту 345, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, на 100% идентичной последовательности SEQ ID NO: 399.

15 348. Гибридный белок по любому из вариантов 345-347, где ДНК-связывающий полипептид представляет собой полипептид РНК-управляемой нуклеазы (RGN).

349. Гибридный белок по варианту 348, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 350. Гибридный белок по любому из вариантов 348-349, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью, из последовательностей SEQ ID NO: 41, 60, 366 или 368.

25 351. Гибридный белок по любому из вариантов 348-350, где полипептид RGN представляет собой нуклеазу.

352. Гибридный белок по варианту 351, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 353. Гибридный белок по варианту 351, отличающийся тем, что аминокислотная последовательность нуклеазы на 100% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

354. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и

дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 443 или

5 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 399.

355. Молекула нуклеиновой кислоты по варианту 354, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 443.

10 356. Молекула нуклеиновой кислоты по варианту 354, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 443.

357. Молекула нуклеиновой кислоты по варианту 354, где дезаминаза кодируется нуклеотидной последовательностью, по меньшей мере на 100% идентичной последовательности SEQ ID NO: 443.

358. Молекула нуклеиновой кислоты по любому из вариантов 354-357, где ДНК-связывающий полипептид представляет собой полипептид RGN.

359. Молекула нуклеиновой кислоты по варианту 358, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 360. Молекула нуклеиновой кислоты по любому из вариантов 358-359, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

361. Молекула нуклеиновой кислоты по любому из вариантов 358-360, где полипептид RGN представляет собой никазу.

30 362. Молекула нуклеиновой кислоты по варианту 361, где аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей с SEQ ID NO: 42, 52-59, 61, 397 и 398.

363. Молекула нуклеиновой кислоты по варианту 362, где аминокислотная последовательность никазы на 100% идентична любой из последовательностей с SEQ ID NO: 42, 52-59, 61, 397 и 398.

5 364. Вектор, содержащий молекулу нуклеиновой кислоты по любому из вариантов 354-363.

365. Вектор по варианту 364, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизоваться с последовательностью-мишенью.

10 366. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из вариантов 345-353, и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

367. Клетка, содержащая гибридный белок по любому из вариантов 345-353, молекулу нуклеиновой кислоты по любому из вариантов 354-363, вектор по любому из вариантов 364-365 или комплекс RNP по варианту 366.

15 368. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности с SEQ ID NO: 399, или
20 нуклеотидная последовательность, кодирующая указанный гибридный белок, и

б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (gRNA), и

25 где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

369. Система по варианту 368, где указанная деаминаза характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 399.

370. Система по варианту 368, где указанная деаминаза имеет аминокислотную последовательность, на 100% идентичную последовательности SEQ ID NO: 399.

371. Система по любому из вариантов 368-370, где по меньшей мере одна из указанной нуклеотидной последовательности, кодирующей одну или более направляющих РНК, и указанной нуклеотидной последовательности, кодирующей гибридный белок, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности.

372. Система по любому из вариантов 368-371, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (РАМ), который распознается полипептидом RGN.

373. Система по любому из вариантов 368-372, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251. -286, 305-344, 562 и 563 или комплементарных им последовательностей.

374. Система по любому из вариантов 368-373, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

375. Система по любому из вариантов 368-374, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

376. Система по любому из вариантов 372-375 где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

377. Система по варианту 376, где полипептид RGN представляет собой нуклеотид.

378. Система по варианту 377, где аминокислотная последовательность нуклеотидов по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

379. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из вариантов 345-353,

молекулу нуклеиновой кислоты по любому из вариантов 354-363, вектор по любому из вариантов 364-365, комплекс RNP по варианту 366, клетку по варианту 367 или систему по любому из вариантов 368-378.

380. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза с использованием объединения:

i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и

ii) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная последовательность дезаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 399,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

381. Способ по варианту 380, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305- 344, 562 и 563 или комплементарных им последовательностей.

382. Способ по любому из вариантов 380-381, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

383. Способ по любому из вариантов 380-382, где способ осуществляется *in vitro*, *in vivo* или *ex vivo*.

384. Способ лечения субъекта, имеющего или подверженного риску развития заболевания, нарушения или состояния, включающий:

введение субъекту гибридного белка по любому из вариантов 345-353, молекулы нуклеиновой кислоты по любому из вариантов 354-363, вектора по любому из вариантов 364-365, комплекса RNP по варианту 366, клетки по варианту 367, системы по любому из вариантов 368-378 или фармацевтической композиции по варианту 379.

385. Способ по варианту 384, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

386. Способ лечения или купирования по меньшей мере одного симптома муковисцидоза, включающий введение субъекту, нуждающемуся в таком лечении, эффективного количества:

а) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность дезаминазы характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441, или полинуклеотида, кодирующего указанный гибридный белок, где указанный полинуклеотид, кодирующий гибридный белок, функционально связан с промотором для обеспечения экспрессии гибридного белка в клетке, и

б) одной или более направляющих РНК (gRNA), способных гибридизоваться с последовательностью ДНК-мишени, или полинуклеотида, кодирующего указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, функционально связан с промотором для обеспечения экспрессии gRNA в клетке,

где гибридный белок и gRNA направлены на геномное местоположение причинной мутации и модифицируют геномную последовательность для удаления причинной мутации.

387. Способ по варианту 386, где gRNA содержит спейсерную последовательность, направленную на любую из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562, и 563 или комплементарные им последовательности.

388. Способ по вариантам 386 или 387, где gRNA содержит любую из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

5 389. Способ по любому из вариантов 386-388, где указанный RGN характеризуется аминокислотной последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

10 390. Способ по любому из вариантов 386-389, где указанный RGN характеризуется аминокислотной последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

Следующие примеры служат исключительно в качестве иллюстрации и не ограничивают объем настоящего изобретения.

Экспериментальная часть

Пример 1

15 Демонстрация редактирования нуклеотидных оснований в клетках млекопитающих

Деаминазы, показанные ниже в таблице 1, получали на основе встречающихся в природе дезаминаз, которые затем подвергали мутациям и отбирали по активности адениндезаминазы в прокариотических клетках.

20 Таблица 1

Последовательности дезаминазы

Дезаминаза	SEQ ID NO.
APG09982	1
APG03724	2
APG09949	3
APG08196	4
APG06333	5
APG06489	6
APG08449	7
APG05174	8
APG09102	9
APG05723	10

Чтобы определить, способны ли дезаминазы, представленные в таблице 1, осуществлять редактирование адениновых оснований в клетках млекопитающих, каждую дезаминазу функционально гибридизовали с RGN-никазой для получения гибридного белка. Идентифицировали остатки, предположительно дезактивирующие домен RuvC в RGN APG07433.1 (указан как SEQ ID NO: 41, описан во международном патенте WO 2019/236566, включенном в данном контексте в качестве ссылки), и RGN модифицировали с получением варианта никазы (nAPG07433.1, SEQ ID NO: 42). Никазный вариант RGN упоминается в данном контексте как «nRGN». Следует понимать, что любой вариант никазы RGN можно использовать для получения гибридного белка по настоящему изобретению.

Кодоны нуклеотидных последовательностей дезаминазы и nRGN, оптимизированные для экспрессии у млекопитающих, синтезировали в виде гибридных белков с N-концевой меткой ядерной локализации и клонировали в экспрессирующей плазмиде pTwist CMV (от фирмы Twist Biosciences). Каждый гибридный белок содержит, начиная с N-конца, SV40 NLS (SEQ ID NO: 43), функционально связанный на C-конце с меткой 3X FLAG Tag (SEQ ID NO: 44), функционально связанный на C-конце с дезаминазой, функционально связанный на C-конце с пептидным линкером (SEQ ID NO: 45), функционально связанный на C-конце с nRGN (например, с nAPG07433.1, представляющим собой SEQ ID NO: 42), наконец, функционально связанным на C-конце с нуклеоплазмином NLS (SEQ ID NO: 45). Все гибридные белки содержали по меньшей мере один NLS и метку 3X FLAG, как описано выше.

Также получены экспрессионные плазмиды, содержащие кассету экспрессии, кодирующую sgRNA, экспрессируемую промотором U6 человека (SEQ ID NO: 50). Геномные последовательности-мишени человека и последовательности sgRNA для направления гибридных белков к геномным мишеням указаны в таблице 2.

Таблица 2

Последовательности направляющих РНК

sgRNA ID	Последовательность-мишень	sgRNA-последовательность	Прямой праймер для амплификации	Обратный праймер для амплификации
SGN000930	21	26	31	32
SGN000186	22	27	33	34

sgRNA ID	Последовательность-мишень	sgRNA-последовательность	Прямой праймер для амплификации	Обратный праймер для амплификации
SGN000194	23	28	35	36
SGN000143	24	29	37	38
SGN000139	25	30	39	40

500 нг плазмиды, содержащей кассету экспрессии, содержащую кодирующую последовательность для гибридного белка для каждой дезаминазы, указанной в таблице 1, и 500 нг плазмиды, содержащей кассету экспрессии, кодирующую sgRNA, указанную в таблице 2, совместно трансфицировали в клетки HEK293FT при 75-90% конfluентности в 24-луночных планшетах с использованием реагента Lipofectamine 2000 (от фирмы Life Technologies). Затем клетки инкубировали при 37°C в течение 72 ч. После инкубирования геномную ДНК экстрагировали с использованием набора NucleoSpin 96 Tissue (от фирмы Macherey-Nagel) в соответствии с протоколом производителя. Геномную область, фланкирующую целевой геномный сайт, амплифицировали с помощью ПЦР с использованием праймеров, указанных в таблице 2, а продукты очищали с помощью системы ZR-96 DNA Clean and Concentrator (от фирмы Zymo Research) в соответствии с протоколом производителя. Очищенные ПЦР-продукты подвергали процедуре Next Generation Sequencing (секвенированию следующего поколения) с помощью системы Illumina MiSeq. Как правило, на ампликон генерируются 100000 считываний на 250 пар оснований при считывании спаренных концевых фрагментов (2 x 100000 считываний). Считывания анализировали с помощью приложения CRISPResso (Pinello и др., Nature Biotech, 34:695-697 (2016)) для расчета скорости редактирования. Выходные данные уравнивали с целью анализа INDEL (введение или делеция нуклеотидных оснований в геном – прим. пер.) или введения специфических адениновых мутаций. В таблицах с 3 по 7 показано редактирование аденинового основания для каждого гибридного белка, содержащего nAPG07433.1 и дезаминазу, указанную в таблице 1, и направляющую РНК, указанную в таблице 2. Указан дезаминазный компонент каждого гибридного белка. Указана скорость редактирования аденинов внутри целевой последовательности или вблизи нее. «A5» указывает, например, на аденин в положении 5 целевой последовательности. Положение каждого нуклеотида в последовательности-

мишени определяли с помощью нумерации первого нуклеотида в последовательности-мишени, ближайшего к РАМ, как положение 1, и номер положения увеличивался в 3'-направлении от последовательности РАМ.

5 Таблицы также показывают, на какой нуклеотид был заменен аденин и с какой скоростью. Например, в таблице 3 показано, что для гибридного белка APG09982-пAPG07433.1 аденин в положении 13 мутировал в гуанин со скоростью 1,2%.

Таблица 3

Скорость редактирования А>N при направляющем действии SGN000139

Дезаминаза		A5	A12	A13	A20	A22
APG09982	C	0	0	0	0.3	0
	G	0	0.5	1.2	0	0
	T	0	0	0	0	0
APG03724	C	0	0	0	0.3	0
	G	0	0.7	0.7	0.1	0
	T	0	0	0	0	0
APG09949	C	0	0	0	0.3	0.1
	G	0.1	0.6	0.7	0	0
	T	0	0	0	0	0
APG08196	C	0	0	0	0.6	0.1
	G	0	0.6	0.6	0	0
	T	0	0	0	0	0
APG06333	C	0	0	0	0.2	0
	G	0	0.5	1	0	0
	T	0	0	0	0	0
APG06489	C	0	0	0	0.2	0
	G	0	0.6	0.4	0	0
	T	0	0	0	0	0
APG08449	C	0	0	0	0.3	0.1
	G	0	0.8	0.8	0	0
	T	0	0	0	0	0
APG05174	C	0	0	0	0.6	0.1
	G	0	0.6	0.7	0	0
	T	0	0	0	0	0
APG09102	C	0	0	0	0.1	0
	G	0	0.6	0.6	0	0
	T	0	0	0	0	0
APG05723	C	0	0	0	0.1	0
	G	0	0.4	0.5	0.1	0
	T	0	0	0	0	0

Все гибридные белки показали детектируемую конверсию A>G в положениях A12 и A13. APG09982 и APG06333 показали не менее 1% редактирования в позиции A13.

5 Таблица 4
Скорость редактирования A>N при направляющем действии SGN000143

Дезаминаза		A1	A4	A6	A9	A11	A14	A19	A30
APG09982	C	0	0	0	0	0	0	0	0
	G	0	0	0	0.1	4.5	1.7	0	0
	T	0	0	0	0	0	0	0	0
APG03724	C	0	0	0	0	0	0	0	0
	G	0	0	0.1	0.1	1.3	1.1	0	0
	T	0	0	0	0	0	0	0	0
APG09949	C	0	0	0	0	0	0	0	0.1
	G	0	0	0	0.1	0.8	0.7	0	0
	T	0	0	0	0	0	0	0	0
APG08196	C	0	0	0	0	0	0	0	0
	G	0	0	0	0.4	0.7	0.5	0.1	0
	T	0	0	0	0	0	0	0	0
APG06333	C	0	0	0	0	0	0	0	0
	G	0	0	0	0	1.3	0.8	0.1	0
	T	0	0	0	0	0	0	0	0
APG06489	C	0	0	0	0	0	0	0	0
	G	0	0	0.1	0.6	1.8	0.8	0.1	0
	T	0	0	0	0	0	0	0	0
APG08449	C	0	0	0	0	0	0	0	0.1
	G	0	0	0	0	2.4	1.2	0	0
	T	0	0	0	0	0	0	0	0
APG05174	C	0	0	0	0	0	0	0	0
	G	0	0	0	0	1.5	0.7	0	0
	T	0	0	0	0	0	0	0	0
APG09102	C	0	0	0	0	0	0	0	0
	G	0	0	0	0	2.6	1.6	0	0
	T	0	0	0	0	0	0	0	0
APG05723	C	0	0	0	0	0.1	0	0	0
	G	0	0	0.1	0.1	1.1	0.5	0	0
	T	0	0	0	0	0	0	0	0

Все гибридные белки продемонстрировали конверсию A>G в положениях A11 и A14. APG09982 показал 4,5%-ую конверсию A11 в G и 1,7%-ую конверсию A14 в G.

Таблица 5

5 Скорость редактирования A>N при направляющем действии SGN000186

Дезаминаза		A9	A16	A18	A22	A25	A28	A30
APG09982	C	0	0	0	0	0	0	0
	G	1.7	4.5	2	0	0	0	0
	T	0	0	0	0	0	0	0
APG03724	C	0	0	0	0	0.1	0	0
	G	0.7	4.1	1.4	0	0	0	0
	T	0	0	0	0	0	0	0
APG09949	C	0	0	0.1	0	0.1	0	0
	G	0.6	3.4	1.1	0	0	0	0
	T	0	0	0	0	0.1	0	0
APG08196	C	0	0	0.1	0	0.1	0	0
	G	1	3.3	1.4	0	0	0.1	0
	T	0	0	0	0	0.1	0	0
APG06333	C	0	0	0	0	0	0	0
	G	1.4	4.2	1.9	0	0	0	0
	T	0	0	0	0	0	0	0
APG06489	C	0	0	0	0	0	0	0
	G	1.7	2.5	1.4	0	0	0	0.1
	T	0	0	0	0	0	0	0
APG08449	C	0	0	0.1	0	0.1	0	0
	G	1.5	5.3	1.6	0	0	0	0
	T	0	0	0	0	0.1	0	0
APG05174	C	0	0	0.1	0	0	0	0
	G	0.9	3.2	1	0	0	0.1	0
	T	0	0	0	0	0.1	0	0
APG09102	C	0	0	0	0	0	0	0
	G	2.3	6.2	2.1	0	0	0	0
	T	0	0	0	0	0	0	0
APG05723	C	0	0	0	0	0	0	0
	G	1.1	1.9	1.2	0	0	0	0
	T	0	0	0	0	0	0	0

Все гибридные белки продемонстрировали редактирование оснований более чем на 1% в нескольких местах в мишени SGN000186. APG09102 показал конверсию A>G 6,2% в положении A16, он также показал более 2% базового

редактирования на позициях A9 и A18. Для всех протестированных гибридных белков положение A16 подвергалось редактированию в наибольшей степени.

Таблица 6

Скорость редактирования A>N при направляющем действии SGN000194

Дезаминаза		A6	A10	A13	A15	A21	A23	A26	A27
APG09982	C	0	0	0	0	0	0	0	0
	G	0	0.3	0.6	1.5	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG03724	C	0	0	0	0	0	0	0	0
	G	0	0.1	0.3	1	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG09949	C	0	0	0	0	0	0	0	0
	G	0	0.2	0.3	1.6	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG08196	C	0	0	0	0	0	0	0	0
	G	0.1	0.4	0.1	0.9	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG06333	C	0	0	0	0	0	0	0	0
	G	0	0.2	0.3	1	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG06489	C	0	0	0	0	0	0	0	0
	G	0	0.4	0.2	1.1	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG08449	C	0	0	0	0	0	0	0	0
	G	0.1	0.3	0.4	1.8	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG05174	C	0	0	0	0	0	0	0	0
	G	0.1	0.1	0.3	0.9	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG09102	C	0	0	0	0	0	0	0	0
	G	0	0.2	0.7	1.6	0	0	0	0
	T	0	0	0	0	0	0	0	0
APG05723	C	0	0	0	0	0	0	0	0
	G	0	0	0.1	0.9	0	0	0	0
	T	0	0	0	0	0	0	0	0

5

С SGN00194 все гибридные белки показали 0,9–1,8% редактирования A>G в положении A15. В позициях A21, A23, A26 и A27 заметного редактирования не наблюдали.

Таблица 7

Скорость редактирования A>N при направляющем действии SGN000930

Дизами- наза		A2	A4	A5	A8	A9	A10	A14	A15	A16	A20	A21	A23	A24	A26	A27	A29	A30
APG09982	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0.3	0.7	0.1	0.2	0.5	0.2	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG03724	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0.1	0.1	0.4	0.5	0.2	0.2	0.3	0.1	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG09949	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0.1	0.5	0.3	0.3	0.4	0	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG08196	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0.1	0	0.2	0.7	0.3	0.2	0.4	0.1	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG06333	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0.1	0.1	0	0.3	0.4	0.3	0.9	0.2	0.1	0.1	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG06489	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0.3	0.1	0.2	0.8	0.3	0.4	0.6	0	0.1	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG08449	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0.1	0.1	0.3	0.6	0.4	0.2	0.4	0.1	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG05174	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0.1	0.2	0.8	0.3	0.4	0.2	0.2	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG09102	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	0.9	0.1	0.1	0.6	0.3	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
APG05723	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0.1	0.1	1.2	0.6	0.2	0.5	0	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

5 A14 являлась наиболее редактируемой позицией в SGN000930 со всеми протестированными гибридными белками. Скорость редактирования варьировалась от 0,3% до 1,2% для конверсий при замене A>G.

Пример 2

Флуоресцентный анализ для направленного редактирования аденинового азотистого основания

Вектор, экспрессирующий усиленный зеленый флуоресцентный белок (EGFP), содержащий мутацию W58X, вызывающую образование преждевременного стоп-кодона (GFP-STOP, SEQ ID NO: 47), был сконструирован таким образом, что кодон W58 можно обратить из стоп-кодона (TGA) в остаток триптофана дикого типа (TGG) с использованием адениндезаминазы для изменения А в третьем положении на G. Успешное преобразование А в G приводит к экспрессии EGFP, которую можно определить количественно. Также был получен второй вектор, способный экспрессировать направляющую РНК, которая направляет гибридный белок деаминаза-RGN на область в районе мутации W58X (SEQ ID NO: 48).

Этот репортерный вектор GFP-STOP вместе с векторами, способными экспрессировать гибридный белок деаминазу-nRGN и соответствующую направляющую РНК, трансфицировали в клетки HEK293T с использованием либо липофекции, либо электропорации. С целью липофекции клетки высевали в количестве 1×10^5 клеток/лунку в 24-луночные планшеты за день до трансфекции в питательную среду (DMEM + 10% эмбриональная телячья сыворотка + 1% пенициллин/стрептомицин). По 500 нг репортерного вектора GFP-STOP, вектора экспрессии дезаминазы-RGN и вектора экспрессии направляющей РНК трансфицировали с использованием реагента Lipofectamine® 3000 (от фирмы Thermo Fisher Scientific) в соответствии с инструкциями производителя. С целью электропорации клетки подвергали электропорации с использованием системы трансфекции Neon® (от фирмы Thermo Fisher Scientific) в соответствии с инструкциями производителя.

В дополнение к транзientной трансфекции флуоресцентного репортера GFP-STOP была создана стабильная клеточная линия, несущая хромосомно-интегрированную кассету GFP-STOP. После создания стабильной линии для трансфекции клетки высевали в количестве 1×10^5 клеток/лунку в 24-луночные планшеты за день до трансфекции в питательную среду (DMEM + 10% эмбриональная телячья сыворотка + 1% пенициллин/стрептомицин). 500 нг каждого экспрессионного вектора дезаминазы-nRGN и вектора экспрессии направляющей РНК трансфицировали с использованием реагента Lipofectamine®

3000 (от фирмы Thermo Fisher Scientific) в соответствии с инструкциями производителя. С целью электропорации клетки подвергали электропорации с использованием системы трансфекции Neon® (от фирмы Thermo Fisher Scientific) в соответствии с инструкциями производителя.

5 Через 24-48 ч после липофекции или электропорации определяли экспрессию GFP с использованием микроскопического исследования клеток на наличие GFP⁺ клеток. После визуального осмотра можно определить долю клеток GFP⁺ по сравнению с клетками GFP⁻. Флуоресценцию наблюдали в клетках млекопитающих, экспрессирующих каждый из гибридных белков
10 деаминаза-nRGN, что указывает на то, что гибридный белок был успешно направлен на мутацию GFP-STOP и отредактировал мутацию для восстановления флуоресценции белка GFP.

После микроскопического анализа клетки лизировали в буферном растворе RIPA и полученный лизат анализировали на флуоресцентном планшет-ридере
15 для определения интенсивности флуоресценции GFP (таблица 8). Специалисту в данной области техники будет понятно, что клетки можно анализировать с помощью проточной цитометрии или сортировки клеток с активированной флуоресценцией для определения точного соотношения клеток GFP⁺ и GFP⁻.

Таблица 8

20 Результаты анализа GFP-STOP

Деаминаза гибридного белка	Обнаруженные GFP ⁺ клетки
APG09982	++
APG03724	++
APG09949	++
APG08196	++
APG06333	+++
APG06489	++
APG08449	++
APG05174	+++
APG09102	++
APG05723	++

N.D = не обнаружено; + = обнаружено несколько клеток GFP⁺; ++ = обнаружено несколько клеток GFP⁺; +++ = обнаружено много клеток GFP⁺

Пример 3

Демонстрация редактирования азотистого основания в клетках
млекопитающих

5 Деаминазы, ниже показанные в таблице 9, получали на основе
встречающихся в природе дезаминаз, которые затем подвергали мутации и
отобирали по активности адениндезаминазы в прокариотических клетках.

Таблица 9

Последовательности дезаминазы

Дезаминаза	SEQ ID NO.	Дезаминаза	SEQ ID NO.
LPG50140	399	LPG50162	421
LPG50141	400	LPG50163	422
LPG50142	401	LPG50164	423
LPG50143	402	LPG50165	424
LPG50144	403	LPG50166	425
LPG50145	404	LPG50167	426
LPG50146	405	LPG50168	427
LPG50147	406	LPG50169	428
LPG50148	407	LPG50170	429
LPG50149	408	LPG50171	430
LPG50150	409	LPG50172	431
LPG50151	410	LPG50173	432
LPG50152	411	LPG50174	433
LPG50153	412	LPG50175	434
LPG50154	413	LPG50176	435
LPG50155	414	LPG50177	436
LPG50156	415	LPG50178	437
LPG50157	416	LPG50179	438
LPG50158	417	LPG50180	439
LPG50159	418	LPG50181	440
LPG50160	419	LPG50182	441
LPG50161	420		

10 Чтобы определить, способны ли дезаминазы, указанные в таблице 9,
осуществлять редактирование адениновых оснований в клетках млекопитающих,
каждую дезаминазу функционально гибридизовали с RGN-никазой для
получения гибридного белка. Идентифицировали остатки, предположительно
15 дезактивирующие домен RuvC в RGN APG07433.1 (указан как SEQ ID NO: 41,
описан во международном патенте WO 2019/236566, включенном в данный

контекст в качестве ссылки), а RGN модифицировали в никазный вариант (nAPG07433.1, SEQ ID NO: 42). Никазный вариант RGN упоминается в данном контексте как «nRGN». Следует понимать, что любой никазный вариант RGN можно использовать для получения гибридного белка по настоящему изобретению.

Нуклеотидные последовательности дезаминазы и nRGN, оптимизированные кодоном для экспрессии у млекопитающих, синтезировали в виде гибридных белков с N-концевой меткой ядерной локализации и клонировали в экспрессирующую плазмиду pTwist CMV (от фирмы Twist Biosciences). Каждый гибридный белок содержит, начиная с amino-конца, NLS SV40 (SEQ ID NO: 43), функционально связанный на C-конце с меткой 3X FLAG (SEQ ID NO: 44), функционально связанной на C-конце с дезаминазой, функционально связанной на C-конце с пептидным линкером (SEQ ID NO: 442), функционально связанным на C-конце с nRGN (например, nAPG07433.1, представляющей собой последовательность SEQ ID NO: 42), наконец, функционально связанной на C-конце с нуклеоплазмином NLS (SEQ ID NO: 46). Нуклеотидные последовательности nAPG07433.1 и пептидного линкера, оптимизированные для экспрессии у млекопитающих, представлены как SEQ ID NO: 486 и 487 соответственно. В таблице 10 указаны полученные гибридные белки, для которых определена активность. Все гибридные белки содержат по меньшей мере один NLS и метку 3X FLAG, как описано выше.

Таблица 10

Последовательности гибридных белков с N-концевым SV40 NLS, меткой 3X FLAG и C-концевым нуклеоплазмином NLS

Гибридный белок	SEQ ID
LPG50140-nAPG07433.1	488
LPG50141-nAPG07433.1	489
LPG50142-nAPG07433.1	490
LPG50143-nAPG07433.1	491
LPG50144-nAPG07433.1	492
LPG50145-nAPG07433.1	493
LPG50146-nAPG07433.1	494
LPG50147-nAPG07433.1	495
LPG50148-nAPG07433.1	496
LPG50149-nAPG07433.1	497
LPG50150-nAPG07433.1	498

Гибридный белок	SEQ ID
LPG50151-nAPG07433.1	499
LPG50152-nAPG07433.1	500
LPG50153-nAPG07433.1	501
LPG50154-nAPG07433.1	502
LPG50155-nAPG07433.1	503
LPG50156-nAPG07433.1	504
LPG50157-nAPG07433.1	505
LPG50158-nAPG07433.1	506
LPG50159-nAPG07433.1	507
LPG50160-nAPG07433.1	508
LPG50161-nAPG07433.1	509
LPG50162-nAPG07433.1	510
LPG50163-nAPG07433.1	511
LPG50164-nAPG07433.1	512
LPG50165-nAPG07433.1	513
LPG50166-nAPG07433.1	514
LPG50167-nAPG07433.1	515
LPG50168-nAPG07433.1	516
LPG50169-nAPG07433.1	517
LPG50170-nAPG07433.1	518
LPG50171-nAPG07433.1	519
LPG50172-nAPG07433.1	520
LPG50173-nAPG07433.1	521
LPG50174-nAPG07433.1	522
LPG50175-nAPG07433.1	523
LPG50176-nAPG07433.1	524
LPG50177-nAPG07433.1	525
LPG50178-nAPG07433.1	526
LPG50179-nAPG07433.1	527
LPG50180-nAPG07433.1	528
LPG50181-nAPG07433.1	529
LPG50182-nAPG07433.1	530

Также получали экспрессионные плазмиды, содержащие кассету экспрессии, кодирующую sgRNA. Геномные последовательности-мишени человека и последовательности sgRNA для направления гибридных белков к геномным мишеням указаны в таблице 11.

Таблица 11

Последовательности направляющей РНК

sgRNA ID	Последовательность-мишень	Последовательность sgRNA	Прямой праймер для амплификации	Обратный праймер для амплификации
SGN000139	537	531	543	549
SGN000143	538	532	544	550
SGN000186	539	533	545	551
SGN000194	540	534	546	552
SGN000930	541	535	547	553
SGN001681	542	536	548	554

500 нг плазмиды, содержащей экспрессионную кассету, содержащую
5 кодирующую последовательность гибридного белка, указанную в таблице 10, и
500 нг плазмиды, содержащей экспрессионную кассету, кодирующую sgRNA,
указанную в таблице 11, совместно трансфицировали в клетки НЕК293FT при
75-90% конфлюэнтности в 24-луночных планшетах с использованием реагента
Lipofectamine 2000 (от фирмы Life Technologies). Затем клетки инкубировали
10 при 37°C в течение 72 ч. После инкубации геномную ДНК экстрагировали с
использованием набора NucleoSpin 96 Tissue (от фирмы Macherey-Nagel) в
соответствии с протоколом производителя. Геномную область, примыкающую к
целевому геномному участку, амплифицировали с помощью ПЦР с
использованием праймеров, указанных в таблице 11, а продукты очищали с
15 помощью набора ZR-96 DNA Clean and Concentrator (от фирмы Zymo Research) в
соответствии с протоколом производителя. Очищенные ПЦР-продукты
подвергали секвенированию следующего поколения на системе Illumina MiSeq.
Как правило, на ампликон генерируется 100000 считываний спаренных
концевых фрагментов на 250 пар оснований (2 x 100000 считываний).
20 Считывания анализировали с использованием программы CRISPResso (Pinello и
др., Nature Biotech, 34:695-697 (2016)) для расчета скорости редактирования.
Выходные уравнивали с целью анализа образования INDEL или введения
специфических адениновых мутаций.

В таблице 12 указаны все редактирования адениновых оснований для
25 каждого гибрида адениндезаминазы, указанных в таблице 10 и направляющей
РНК, указанной в таблице 12. В Таблицах 13-27 указан профиль конкретных
нуклеотидных мутаций для выбранных типичных образцов. Указана скорость

редактирования аденинов внутри последовательности-мишени или вблизи нее.

«A5» указывает, например, на аденин в положении 5 последовательности-мишени. Положение каждого нуклеотида в последовательности-мишени определяли, нумеруя первый нуклеотид в последовательности-мишени,

- 5 ближайший к РАМ (который находится на 3'-конце мишени для APG07433.1), как положение 1, а номер положения увеличивался в направлении 5'-конца от последовательности РАМ. В таблице также указано, на какой нуклеотид был заменен аденин и с какой скоростью. Например, в таблице 13 указано, что для гибридного белка LPG50148-nAPG07433.1 аденин в положении 13 мутировал в
- 10 гуанин со скоростью 9,7%.

Таблица 12

Оценка скоростей редактирования азотистых оснований для каждой адениндезаминазы

Дезаминаза	SGN	Мутированные считывания, %	Дезаминаза	SGN	Мутированные считывания, %
LPG50140	SGN001681	30.01%	LPG50161	SGN000930	0%
LPG50140	SGN000139	6.91%	LPG50162	SGN001681	21.73%
LPG50140	SGN000143	16.09%	LPG50162	SGN000139	2%
LPG50140	SGN000186	18.76%	LPG50162	SGN000143	5%
LPG50140	SGN000194	9.77%	LPG50162	SGN000186	14%
LPG50140	SGN000930	3.51%	LPG50162	SGN000194	6%
LPG50141	SGN001681	21.37%	LPG50162	SGN000930	5%
LPG50141	SGN000139	2.43%	LPG50163	SGN001681	12.80%
LPG50141	SGN000143	6.93%	LPG50163	SGN000139	0%
LPG50141	SGN000186	9.79%	LPG50163	SGN000143	2%
LPG50141	SGN000194	4.45%	LPG50163	SGN000186	10%
LPG50141	SGN000930	5.29%	LPG50163	SGN000194	4%
LPG50142	SGN001681	34.19%	LPG50163	SGN000930	3%
LPG50142	SGN000139	3.10%	LPG50164	SGN001681	4.28%
LPG50142	SGN000143	8.67%	LPG50164	SGN000139	0%
LPG50142	SGN000186	14.12%	LPG50164	SGN000143	3.36%
LPG50142	SGN000194	10.04%	LPG50164	SGN000186	7.38%
LPG50142	SGN000930	6.78%	LPG50164	SGN000194	2.73%
LPG50143	SGN001681	20.62%	LPG50164	SGN000930	1.47%
LPG50143	SGN000139	1.99%	LPG50165	SGN001681	25.66%
LPG50143	SGN000143	6.09%	LPG50165	SGN000139	2%
LPG50143	SGN000186	10.58%	LPG50165	SGN000143	5.11%
LPG50143	SGN000194	5.60%	LPG50165	SGN000186	9.88%
LPG50143	SGN000930	3.98%	LPG50165	SGN000194	3.97%

Дезаминаза	SGN	Мутированные считывания, %	Дезаминаза	SGN	Мутированные считывания, %
LPG50144	SGN001681	28.26%	LPG50165	SGN000930	3.18%
LPG50144	SGN000139	3.55%	LPG50166	SGN000139	2%
LPG50144	SGN000143	5.77%	LPG50166	SGN000143	4%
LPG50144	SGN000186	12.22%	LPG50166	SGN000186	8%
LPG50144	SGN000194	6.40%	LPG50166	SGN000194	2%
LPG50144	SGN000930	5.81%	LPG50166	SGN000930	4%
LPG50145	SGN001681	29.23%	LPG50167	SGN001681	20.56%
LPG50145	SGN000139	2.53%	LPG50167	SGN000139	2%
LPG50145	SGN000143	3.75%	LPG50167	SGN000143	4%
LPG50145	SGN000186	9.93%	LPG50167	SGN000186	8%
LPG50145	SGN000194	3.98%	LPG50167	SGN000194	5%
LPG50145	SGN000930	3.84%	LPG50167	SGN000930	4%
LPG50146	SGN001681	32.53%	LPG50168	SGN001681	13.81%
LPG50146	SGN000139	5.95%	LPG50168	SGN000139	2%
LPG50146	SGN000143	11.30%	LPG50168	SGN000143	3%
LPG50146	SGN000186	17.78%	LPG50168	SGN000186	7%
LPG50146	SGN000194	7.38%	LPG50168	SGN000194	2%
LPG50146	SGN000930	7.13%	LPG50168	SGN000930	3%
LPG50147	SGN001681	49.10%	LPG50169	SGN001681	25.73%
LPG50147	SGN000139	3.26%	LPG50169	SGN000139	4%
LPG50147	SGN000143	8.59%	LPG50169	SGN000143	8%
LPG50147	SGN000186	12.61%	LPG50169	SGN000186	13%
LPG50147	SGN000194	8.80%	LPG50169	SGN000194	9%
LPG50147	SGN000930	4.96%	LPG50169	SGN000930	8%
LPG50148	SGN001681	49.39%	LPG50170	SGN001681	12.87%
LPG50148	SGN000139	10.80%	LPG50170	SGN000139	1.50%
LPG50148	SGN000143	12.49%	LPG50170	SGN000143	3.14%
LPG50148	SGN000186	32.65%	LPG50170	SGN000186	12.16%
LPG50148	SGN000194	16.60%	LPG50170	SGN000194	2.76%
LPG50148	SGN000930	7.61%	LPG50170	SGN000930	4.10%
LPG50149	SGN001681	27.62%	LPG50171	SGN001681	27.16%
LPG50149	SGN000139	2.83%	LPG50171	SGN000139	1.75%
LPG50149	SGN000143	9.33%	LPG50171	SGN000143	6.14%
LPG50149	SGN000186	22.12%	LPG50171	SGN000186	12.65%
LPG50149	SGN000194	7.94%	LPG50171	SGN000194	5.60%
LPG50149	SGN000930	7.06%	LPG50171	SGN000930	4.55%
LPG50150	SGN001681	28.46%	LPG50172	SGN001681	1.78%
LPG50150	SGN000139	3.06%	LPG50172	SGN000139	0%
LPG50150	SGN000143	6.00%	LPG50172	SGN000143	0%
LPG50150	SGN000186	23.67%	LPG50172	SGN000186	0%

Дезаминаза	SGN	Мутированные считывания, %	Дезаминаза	SGN	Мутированные считывания, %
LPG50150	SGN000194	9.47%	LPG50172	SGN000194	0%
LPG50150	SGN000930	5.41%	LPG50172	SGN000930	0%
LPG50151	SGN001681	3.01%	LPG50173	SGN001681	12.64%
LPG50151	SGN000139	0%	LPG50173	SGN000139	1.00%
LPG50151	SGN000143	1.53%	LPG50173	SGN000143	3.23%
LPG50151	SGN000186	7.76%	LPG50173	SGN000186	7.88%
LPG50151	SGN000194	1.43%	LPG50173	SGN000194	2.66%
LPG50151	SGN000930	0%	LPG50173	SGN000930	1.77%
LPG50152	SGN001681	26.06%	LPG50174	SGN001681	14.11%
LPG50152	SGN000139	2%	LPG50174	SGN000139	0%
LPG50152	SGN000143	3%	LPG50174	SGN000143	3%
LPG50152	SGN000186	18%	LPG50174	SGN000186	8%
LPG50152	SGN000194	3%	LPG50174	SGN000194	2%
LPG50152	SGN000930	6%	LPG50174	SGN000930	3%
LPG50153	SGN001681	1.12%	LPG50175	SGN001681	22.29%
LPG50153	SGN000139	0%	LPG50175	SGN000139	4%
LPG50153	SGN000143	0%	LPG50175	SGN000143	9%
LPG50153	SGN000186	0%	LPG50175	SGN000186	14%
LPG50153	SGN000194	1%	LPG50175	SGN000194	13%
LPG50153	SGN000930	0%	LPG50175	SGN000930	5%
LPG50154	SGN001681	2.26%	LPG50176	SGN001681	9.52%
LPG50154	SGN000139	0%	LPG50176	SGN000139	0%
LPG50154	SGN000143	0%	LPG50176	SGN000143	2%
LPG50154	SGN000186	0%	LPG50176	SGN000186	7%
LPG50154	SGN000194	1%	LPG50176	SGN000194	2%
LPG50154	SGN000930	0%	LPG50176	SGN000930	0%
LPG50155	SGN001681	14.91%	LPG50177	SGN001681	7.98%
LPG50155	SGN000139	2%	LPG50177	SGN000139	2%
LPG50155	SGN000143	4%	LPG50177	SGN000143	4%
LPG50155	SGN000186	17%	LPG50177	SGN000186	11%
LPG50155	SGN000194	7%	LPG50177	SGN000194	3%
LPG50155	SGN000930	5%	LPG50177	SGN000930	9%
LPG50156	SGN001681	11.19%	LPG50178	SGN000139	2.00%
LPG50156	SGN000139	3.79%	LPG50178	SGN000143	6.19%
LPG50156	SGN000143	6.44%	LPG50178	SGN000186	12.94%
LPG50156	SGN000186	12.69%	LPG50178	SGN000194	5.51%
LPG50156	SGN000194	6.87%	LPG50178	SGN000930	3.95%
LPG50156	SGN000930	4.10%	LPG50179	SGN001681	23.35%
LPG50157	SGN001681	20.66%	LPG50179	SGN000139	2.00%
LPG50157	SGN000139	3.37%	LPG50179	SGN000143	5.08%

Дезаминаза	SGN	Мутированные считывания, %	Дезаминаза	SGN	Мутированные считывания, %
LPG50157	SGN000143	6.91%	LPG50179	SGN000186	12.50%
LPG50157	SGN000186	12.15%	LPG50179	SGN000194	4.49%
LPG50157	SGN000194	9.98%	LPG50179	SGN000930	4.62%
LPG50157	SGN000930	5.55%	LPG50180	SGN001681	1.80%
LPG50158	SGN001681	1.56%	LPG50180	SGN000139	0%
LPG50158	SGN000139	0%	LPG50180	SGN000143	0%
LPG50158	SGN000143	1.15%	LPG50180	SGN000186	0%
LPG50158	SGN000186	4.91%	LPG50180	SGN000194	0%
LPG50158	SGN000194	1.73%	LPG50180	SGN000930	0%
LPG50158	SGN000930	0%	LPG50181	SGN001681	7.93%
LPG50159	SGN001681	5.85%	LPG50181	SGN000139	2.88%
LPG50159	SGN000139	0%	LPG50181	SGN000143	3.78%
LPG50159	SGN000143	2.78%	LPG50181	SGN000186	12.56%
LPG50159	SGN000186	6.99%	LPG50181	SGN000194	3.39%
LPG50159	SGN000194	4.40%	LPG50181	SGN000930	1.20%
LPG50159	SGN000930	2.60%	LPG50182	SGN001681	16.49%
LPG50160	SGN001681	22.20%	LPG50182	SGN000139	1.00%
LPG50160	SGN000139	4%	LPG50182	SGN000143	5%
LPG50160	SGN000143	8%	LPG50182	SGN000186	9%
LPG50160	SGN000186	16%	LPG50182	SGN000194	6%
LPG50160	SGN000194	5%	LPG50182	SGN000930	3%
LPG50160	SGN000930	6%			
LPG50161	SGN001681	1.47%			
LPG50161	SGN000139	0%			
LPG50161	SGN000143	0%			
LPG50161	SGN000186	0%			
LPG50161	SGN000194	0%			

Таблица 13

Скорость редактирования A>N с использованием дезаминазы LPG50148 и направляющей SGN000139

		SGN000139				
		A5	A12	A13	A20	A22
LPG50148	C	0	0	0	0.1	0
	G	0	2.2	9.7	0.2	0
	T	0	0	0	0	0

- 5 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A12 и A13. LPG50148 характеризовалась более 9% редактирования в проржении A13.

Таблица 14

Скорость редактирования A>N с использованием дезаминазы LPG50148 и направляющей SGN000143

		SGN000143							
		A1	A4	A6	A9	A11	A14	A19	A30
LPG50148	C	0	0	0	0	0	0	0	0
	G	0	0	0.1	1.2	11	6.7	0.1	0
	T	0	0	0	0	0	0	0	0

5 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A11 и A14. LPG50148 показала более 11% редактирования в положении A11.

Таблица 15

Скорость редактирования A>N с использованием дезаминазы LPG50148 и направляющей SGN000186

10

		SGN000186						
		A9	A16	A18	A22	A25	A28	A30
LPG50148	C	0	0	0	0	0.4	0	0
	G	23.7	29.2	4.1	0.2	0	0	0
	T	0	0	0	0	0	0	0

LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A16 и A18. LPG50148 показала более 23% редактирования в положениях A9 и A16.

15

Таблица 16

Скорость редактирования A>N с использованием дезаминазы LPG50148 и направляющей SGN000194

		SGN000194							
		A6	A10	A13	A15	A21	A23	A26	A27
LPG50148	C	0	0	0	0	0	0	0	0
	G	0.3	5.3	13	14	0	0	0	0
	T	0	0	0	0	0	0	0	0

20 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A13 и A15. LPG50148 показала более 12% редактирования в положениях A13 и A15.

Таблица 17

Скорость редактирования A>N с использованием дезаминазы LPG50148 и направляющей SGN000930

		SGN000930																
		A 2	A 4	A 5	A 8	A 9	A 10	A 14	A 15	A 16	A 20	A 21	A 23	A 24	A 26	A 27	A 29	A 30
LPG5 0148	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0. 2	2	2. 2	1. 1	2. 2	2. 2	2. 5	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- 5 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A10, A14, A15, A16, A20 и A21. LPG50148 характеризовалась более 2% редактирования в положениях A10, A14, A16, A20 и A21.

Таблица 18

- 10 Скорость редактирования A>N с использованием дезаминазы LPG50146 и направляющей SGN000139

		SGN000139				
		A5	A12	A13	A20	A22
LPG50146	C	0	0	0	0.4	0.1
	G	0	2.1	4.1	0	0
	T	0	0	0	0	0

- 15 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A12 и A13. LPG50146 характеризовалась более 4% редактирования в положении A13.

Таблица 19

Скорость редактирования A>N с использованием дезаминазы LPG50146 и направляющей SGN000143

		SGN000143							
		A1	A4	A6	A9	A11	A14	A19	A30
LPG50146	C	0	0	0	0	0	0	0	0
	G	0	0	0	0.8	8.4	5	0	0
	T	0	0	0	0	0	0	0	0

LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A11 и A14. LPG50146 характеризовалась более 8% редактирования в позиции A11.

Таблица 20

5 Скорость редактирования A>N с использованием дезаминазы LPG50146 и
направляющей SGN000186

		SGN000186						
		A9	A16	A18	A22	A25	A28	A30
LPG50146	C	0	0	0	0	0.2	0	0
	G	7.4	13.4	3.1	0.1	0	0	0
	T	0	0	0	0	0	0	0

LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A16 и A18. LPG50146
10 характеризовалась более 13% редактирования в позиции A16.

Таблица 21

Скорость редактирования A>N с использованием дезаминазы LPG50146 и направляющей SGN000194

		SGN000194							
		A6	A10	A13	A15	A21	A23	A26	A27
LPG50146	C	0	0	0	0	0	0	0	0
	G	0	1.8	3.2	4.5	0	0	0	0
	T	0	0	0	0	0	0	0	0

15 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A13 и A15. LPG50146 характеризовалась более 3% редактирования в положениях A13 и A15.

Таблица 22

20 Скорость редактирования A>N с использованием дезаминазы LPG50146 и
направляющей SGN000930

[illegible]

LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A10, A14, A15, A16, A20 и A21. LPG50146 характеризовалась более 2% редактирования в положениях A14 и A16.

Таблица 23

5 Скорость редактирования A>N с использованием дезаминазы LPG50140 и направляющей SGN000139

		SGN000139				
		A5	A12	A13	A20	A22
LPG50140	C	0	0	0	0.4	0
	G	0	0.5	5.5	0	0
	T	0	0	0	0	0

10 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A12 и A13. LPG50140 характеризовалась более 5% редактирования в положении A13.

Таблица 24

Скорость редактирования A>N с использованием дезаминазы LPG50140 и направляющей SGN000143

		SGN000143							
		A1	A4	A6	A9	A11	A14	A19	A30
LPG50140	C	0	0	0	0	0	0	0	0
	G	0	0	0	1.2	14	5.6	0	0
	T	0	0	0	0	0	0	0	0

15 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A11 и A14. LPG50140 характеризовалась 14% редактирования в положении A11.

Таблица 25

20 Скорость редактирования A>N с использованием дезаминазы LPG50140 и направляющей SGN000186

		SGN000186						
		A9	A16	A18	A22	A25	A28	A30
LPG50140	C	0	0	0	0	0.2	0	0
	G	9.4	15	1.7	0	0	0	0
	T	0	0	0	0	0	0	0

LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A9, A16 и A18. LPG50140 характеризовалась более 9% редактирования в позициях A9 и A16.

Таблица 26

5 Скорость редактирования A>N с использованием дезаминазы LPG50140 и направляющей SGN000194

		SGN000194							
		A6	A10	A13	A15	A21	A23	A26	A27
LPG50140	C	0	0	0	0	0	0	0	0
	G	0	0	6.7	7.8	0	0	0	0
	T	0	0	0	0	0	0	0	0

10 LPG50140, LPG50146 и LPG50148 характеризовались детектируемое преобразованием A>G в положениях A13 и A15. LPG50140 характеризовалась более 6% редактирования в положениях A13 и A15.

Таблица 27

Скорость редактирования A>N с использованием дезаминазы LPG50140 и направляющей SGN000930

		SGN000930																
		A2	A4	A5	A8	A9	A10	A14	A15	A16	A20	A21	A23	A24	A26	A27	A29	A30
LPG50140	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0.4	1.4	0.6	1.1	0.4	0.5	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

15 LPG50140, LPG50146 и LPG50148 характеризовались детектируемым преобразованием A>G в положениях A10, A14, A15, A16, A20 и A21. LPG50140 показала более 1% редактирования в положениях A14 и A16.

20 В приведенной ниже таблице 28 указаны средние скорости редактирования для LPG50148-nAPG07433.1 при использовании нескольких направляющих, протестированных в клетках HEK293T с помощью липофекции двух плазмид. Редактор оснований был закодирован на одной плазмиде, а направляющая РНК была закодирована на второй плазмиде. Общую скорость замещения в мишени использовали для измерения скорости редактирования азотистых оснований.

Таблица 28

Средняя скорость редактирования для LPG50148-nAPG07433.1

Ген	SGN	Средняя скорость замещения, %	N
Ген А	SGN000139	10.8	1
Ген А	SGN000143	29.65	2
Ген В	SGN000487	34.68	2
Ген В	SGN000488	39.94	1
Ген В	SGN001061	9.18	2
Ген В	SGN001062	32.77	1
Ген В	SGN001270	8.34	3
Ген В	SGN001946	5.1	1
Ген В	SGN001947	16.43	1
Ген В	SGN001948	0.46	1
Ген В	SGN001949	1.44	1
Ген В	SGN001950	10.96	1
Ген В	SGN001951	5.38	1
Ген В	SGN001952	6.29	1
Ген В	SGN001953	5.28	1
Ген В	SGN001954	7.95	1
Ген В	SGN001955	7.83	1
Ген В	SGN001956	4.78	1
Ген В	SGN001959	1.43	1
Ген В	SGN001960	17.4	1
Ген В	SGN001961	1.46	1
Ген В	SGN001962	1.62	1
Ген В	SGN001963	11.31	1
Ген В	SGN001964	2.03	1
Ген В	SGN001965	9.3	1
Ген В	SGN001966	1.51	1
CFTR	SGN001101	17.06	1
Ген D	SGN001196	14.58	1
Ген D	SGN001199	42.05	1
Ген E	SGN001681	48.85	1
Ген F	SGN000169	55.13	2
Ген F	SGN000173	47.13	1
Ген G	SGN000412	16.58	1
Ген G	SGN000414	14.5	2
Ген G	SGN001259	24.16	1
Ген G	SGN001274	10.45	2
Ген G	SGN001275	5.25	1
Ген H	SGN000186	32.65	1

Ген	SGN	Средняя скорость замещения, %	N
Ген I	SGN000754	30.76	1
Ген I	SGN000909	21.57	2
Ген I	SGN000927	3.8	1
Ген I	SGN000928	28.77	1
Ген I	SGN000929	17.58	2
Ген I	SGN000949	26.43	1
Ген I	SGN001268	16.64	2
Ген I	SGN001269	6.42	1
Ген I	SGN001967	1.45	1
Ген I	SGN001968	5.61	1
Ген I	SGN001973	5.14	1
Ген I	SGN001975	0.16	1
Ген I	SGN001976	0.62	1
Ген I	SGN001977	0.65	1
Ген I	SGN001978	3.09	1
Ген I	SGN001981	2.34	1

LPG50148-nAPG07433.1 характеризуется редактированием во многих различных направляющих по всему геному.

5 В таблице 29 указаны скорости редактирования адениновых оснований в каждой направляющей из LPG50148-nAPG07433.1. Ниже указаны только положения аденина. Скорость превращения аденина представляет собой среднее значение нескольких экспериментов, где это целесообразно.

Таблица 29

Частота редактирования нуклеотидов А в клетках млекопитающих для 10 первых направляющих

SGN	Положение																					
	A1	A3	A4	A5	A6	A7	A8	A9	A10	A11	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A24	A25
SGN001681					13				47													
SGN000169				0.2		1.3	17			22					43		11			1.7		
SGN001199						3.5					42											
SGN000186								24						29		4.1				0.2		0.4
SGN000754				0		0		1.3				6.1			29							
SGN000143	0		0		0.4			4.4		27		17					0.3					
SGN000928	0.3			0.2	0.3		6.1										26					
SGN000487	0.2	0.2							12			25		8.7		7.6			14			
SGN001259			0						12						16						1	
SGN001062			0		0.7			0					10			5.8		13		2.4	0.1	0

LPG50148-nAPG07433.1 характеризовалась редактированием аденинового основания в положениях 6-21 в области-мишени в зависимости от используемой направляющей РНК. Скорости редактирования изменяются в зависимости от используемой направляющей РНК.

5 Пример 4

Коррекция нонсенс-мутаций муковисцидоза I класса

Пример 4.1

Идентификация RGN и направляющих РНК

Муковисцидоз обычно вызывается вредными мутациями в гене CFTR (SEQ ID NO: 51). Шесть наиболее распространенных нонсенс-мутаций представляют собой G542X, W1282X, R553X, R1162X, E60X, R785X и Q493X. Каждую из этих стоп-мутаций можно отредактировать для восстановления кодирующего кодона описанным в данном контексте гибридным белком RGN-дезаминазой. Для направления на каждую мутацию необходимо определить следующее: 1) RGN, которая содержит сайт распознавания РАМ, проксимальный к нонсенс-мутации; и 2) направляющую РНК, которая оптимально направляет гибридный белок RGN-дезаминазу на ДНК-мишень. В приведенной ниже таблице 30 указаны никазные варианты RGN, которые содержат РАМ, расположенные проксимальнее каждой из шести нонсенс-мутаций, и количество направляющих РНК, которые можно использовать для каждой RGN. В таблице 31 описаны генетические локусы для каждой направляющей РНК. Сайт узнавания РАМ для каждого генетического локуса подчеркнут. Также указаны последовательность-мишень для направляющей РНК и сама последовательность направляющей РНК.

Таблица 30

25 Никазы RGN и количество направляющих РНК для нонсенс-мутаций в CFTR

RGN-никаза	SEQ ID NO. RGN-никазы	E60X	G542X	Q493X	R1162X	R553X	W1282X
nAPG00969	52	2		2	2		
nAPG07433.1	42	1				3	1
nAPG06646	53	6	4	2	3	7	4
nAPG09748	54	1	1	4			1
nAPG09882	55	4	3	5	5	3	5
nAPG03850	56	2	2	1	3	3	4

RGN-никаза	SEQ ID NO. RGN-никазы	E60X	G542X	Q493X	R1162X	R553X	W1282X
nAPG07553	57	1	1	1	1	1	2
nAPG05586	58	1	1		3		1
nAPG01604	59			2	1		2

Таблица 31

Направляющие РНК для нонсенс-мутаций в CFTR

ID направляющей РНК	Генетический локус	Генетический локус (SEQ ID NO.)	Мишень (SEQ ID NO.)	гРНК (SEQ ID NO.)
E60X пAPG06646 Мишень 1	AATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGT	62	80	98
E60X пAPG06646 Мишень 2	ATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAA	63	81	99
E60X пAPG06646 Мишень 3	GCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAA	64	82	100
E60X пAPG06646 Мишень 4	AAGGGCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGA	65	83	101
E60X пAPG06646 Мишень 5	GAAGGGCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAG	66	84	102
E60X пAPG06646 Мишень 6	CGAAGGGCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAA	67	85	103
E60X пAPG09882 Мишень 1	GAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTGGG	68	86	104
E60X пAPG09882 Мишень 2	TGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTGG	69	87	105
E60X пAPG09882 Мишень 3	ATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTG	70	88	106
E60X пAPG09882 Мишень 4	AGGGCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAA	71	89	107
E60X пAPG00969 Мишень 1	GTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTGGGAC	72	90	108
E60X пAPG00969 Мишень 2	AGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTGGGA	73	91	109
E60X пAPG03850 Мишень 1	GGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTGGGAC	74	92	110
E60X пAPG03850 Мишень 2	AGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAG	75	93	111
E60X пAPG07433.1 Мишень 1	GAAGGGCATTAATGAGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAG	76	94	112
E60X пAPG09748 Мишень 1	GTCCCACTTTTTATTCTTTTGCAGAGAATGGGATAGATAGCTGGCTTCAAAGAAAAATCC	77	95	113
E60X пAPG07553 Мишень 1	AGTTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAG	78	96	114
E60X пAPG05586 Мишень 1	TTTAGGATTTTTCTTTGAAGCCAGCTATCTATCCCATTCTCTGCAAAAGAATAAAAAGTG	79	97	115
G542X пAPG06646 Мишень 1	CGTTGACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACTT	116	128	140
G542X пAPG06646 Мишень 2	GACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACTTGGAG	117	129	141
G542X пAPG06646 Мишень 3	CCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACTTGGAGAT	118	130	142
G542X пAPG06646 Мишень 4	CCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACTTGGAGATGTC	119	131	143
G542X пAPG09882 Мишень 1	TCTTGCTCGTTGACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTG	120	132	144
G542X пAPG09882 Мишень 2	TTGCTCGTTGACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCA	121	133	145
G542X пAPG09882 Мишень 3	CACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACTTGGAGATGTCC	122	134	146

ID направляющей РНК	Генетический локус	Генетический локус (SEQ ID NO.)	Мишень (SEQ ID NO.)	гРНК (SEQ ID NO.)
G542X пAPG03850 Мишень 1	TGACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAAC	123	135	147
G542X пAPG03850 Мишень 2	TCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACCTGGAGATGT	124	136	148
G542X пAPG09748 Мишень 1	AGAGAAAGACAATATAGTTCTTTGAGAAGGTGGAATCACACTGAGTGGAGGTCAACGAGC	125	137	149
G542X пAPG07553 Мишень 1	TCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCAAACCTGGAGATGT	126	138	150
G542X пAPG05586 Мишень 1	CGTTGACCTCCACTCAGTGTGATTCCACCTTCTCAAAGAACTATATTGTCTTTCTCTGCA	127	139	151
Q493X пAPG09882 Мишень 1	GATATTTTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCAC	152	169	186
Q493X пAPG09882 Мишень 2	ATATTTTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACT	153	170	187
Q493X пAPG09882 Мишень 3	TTTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTG	154	171	188
Q493X пAPG09882 Мишень 4	TTTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGC	155	172	189
Q493X пAPG09882 Мишень 5	TTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGCT	156	173	190
Q493X пAPG09748 Мишень 1	TAAGCACAGTGGAAGAATTTCAATTCTGTTCTTAGTTTTCTGGATTATGCCTGGCACCATT	157	174	191
Q493X пAPG09748 Мишень 2	AAGCACAGTGGAAGAATTTCAATTCTGTTCTTAGTTTTCTGGATTATGCCTGGCACCATT	158	175	192
Q493X пAPG09748 Мишень 3	ACAGTGGAAGAATTTCAATTCTGTTCTTAGTTTTCTGGATTATGCCTGGCACCATTAAAG	159	176	193
Q493X пAPG09748 Мишень 4	GGAAGAATTTCAATTCTGTTCTTAGTTTTCTGGATTATGCCTGGCACCATTAAAGAAAAT	160	177	194
Q493X пAPG00969 Мишень 1	GATATTTTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCAC	161	178	195
Q493X пAPG00969 Мишень 2	TTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGCT	162	179	196
Q493X пAPG06646 Мишень 1	TTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGCTTAA	163	180	197
Q493X пAPG06646 Мишень 2	AATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGCTTAATTT	164	181	198
Q493X пAPG01604 Мишень 1	TTCTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACT	165	182	199
Q493X пAPG01604 Мишень 2	TTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGTGC	166	183	200
Q493X пAPG03850 Мишень 1	CTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGT	167	184	201
Q493X пAPG07553 Мишень 1	CTTTAATGGTGCCAGGCATAATCCAGGAAAACCTAAGAACAGAAATGAAATTCTTCCACTGT	168	185	202
R553X пAPG06646 Мишень 1	CCAATAATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCC	203	219	235
R553X пAPG06646 Мишень 2	CAATAATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCC	204	220	236
R553X пAPG06646 Мишень 3	ATAATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCAC	205	221	237
R553X пAPG06646 Мишень 4	AATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCT	206	222	238
R553X пAPG06646 Мишень 5	TCACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCAAGAA	207	223	239
R553X пAPG06646 Мишень 6	CACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCAAGAAC	208	224	240

ID направляющей РНК	Генетический локус	Генетический локус (SEQ ID NO.)	Мишень (SEQ ID NO.)	гРНК (SEQ ID NO.)
R553X пAPG06646 Мишень 7	CCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCAAGAACTA	209	225	241
R553X пAPG07433.1 Мишень 1	CCAATAATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCC	210	226	242
R553X пAPG07433.1 Мишень 2	TCACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCAAGAA	211	227	243
R553X пAPG07433.1 Мишень 3	CCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCAAGAACTA	212	228	244
R553X пAPG09882 Мишень 1	AATAATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCA	213	229	245
R553X пAPG09882 Мишень 2	ATTAGTTATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTT	214	230	246
R553X пAPG09882 Мишень 3	TATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCA	215	231	247
R553X пAPG03850 Мишень 1	TATTACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTC	216	232	248
R553X пAPG03850 Мишень 2	TTCACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTC	217	233	249
R553X пAPG03850 Мишень 3	CACCTTGCTAAAGAAATTCTTGCTCATTGACCTCCACTCAGTGTGATTCCACCTTCTCCA	218	234	250
R1162X пAPG09882 Мишень 1	GGTTTACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAA	251	269	287
R1162X пAPG09882 Мишень 2	ACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATA	252	270	288
R1162X пAPG09882 Мишень 3	CTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAAC	253	271	289
R1162X пAPG09882 Мишень 4	TGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAACA	254	272	290
R1162X пAPG09882 Мишень 5	GTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAACAT	255	273	291
R1162X пAPG06646 Мишень 1	TTTACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAA	256	274	292
R1162X пAPG06646 Мишень 2	TACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAAT	257	275	293
R1162X пAPG06646 Мишень 3	TGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAA	258	276	294
R1162X пAPG03850 Мишень 1	TACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATA	259	277	295
R1162X пAPG03850 Мишень 2	TTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAA	260	278	296
R1162X пAPG03850 Мишень 3	TGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAA	261	279	297
R1162X пAPG05586 Мишень 1	TTACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAAT	262	280	298
R1162X пAPG05586 Мишень 2	CTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATA	263	281	299
R1162X пAPG05586 Мишень 3	TGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAACATTT	264	282	300
R1162X пAPG00969 Мишень 1	GGTTTACCTTCTGTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAA	265	283	301
R1162X пAPG00969 Мишень 2	GTTGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAACAT	266	284	302
R1162X пAPG07553 Мишень 1	TGGCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAA	267	285	303
R1162X пAPG01604 Мишень 1	GCATGTCAATGAACTTAAAGACTCAGCTCACAGATCGCATCTGAAATAAAAAATAACAACA	268	286	304

ID направляющей РНК	Генетический локус	Генетический локус (SEQ ID NO.)	Мишень (SEQ ID NO.)	гРНК (SEQ ID NO.)
W1282X пAPG09882 Мишень 1	GTGTGTCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTG	305	325	345
W1282X пAPG09882 Мишень 2	GTCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCA	306	326	346
W1282X пAPG09882 Мишень 3	CTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAA	307	327	347
W1282X пAPG09882 Мишень 4	GGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAAGG	308	328	348
W1282X пAPG09882 Мишень 5	GATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAAGGAC	309	329	349
W1282X пAPG06646 Мишень 1	TCGATGGTGTGTCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCA	310	330	350
W1282X пAPG06646 Мишень 2	TTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAA	311	331	351
W1282X пAPG06646 Мишень 3	TGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAAG	312	332	352
W1282X пAPG06646 Мишень 4	GGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAAGGA	313	333	353
W1282X пAPG03850 Мишень 1	TGTCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGG	314	334	354
W1282X пAPG03850 Мишень 2	GTCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGT	315	335	355
W1282X пAPG03850 Мишень 3	CTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGA	316	336	356
W1282X пAPG03850 Мишень 4	TGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGC	317	337	357
W1282X пAPG07553 Мишень 1	CTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGA	318	338	358
W1282X пAPG07553 Мишень 2	TGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGC	319	339	359
W1282X пAPG01604 Мишень 1	TCTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTG	320	340	360
W1282X пAPG01604 Мишень 2	CTTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGA	321	341	361
W1282X пAPG07433.1 Мишень 1	TTGGGATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAAA	322	342	362
W1282X пAPG09748 Мишень 1	GTATCACTCCAAAGGCTTTCTTCACTGTTGCAAAGTTATTGAATCCCAAGACACACCAT	323	343	363
W1282X пAPG05586 Мишень 1	GATTCAATAACTTTGCAACAGTGAAGGAAAGCCTTTGGAGTGATACCACAGGTGAGCAAA	324	344	364
F508de пAPG07433.1 SGN001101 Мишень 1	ACCAAAGATGATATTTTCTTTAATGGTGCCAGGCATAATCCAGGAAACTGAGAACAGAATGAAA	562	563	564

В таблице 28, представленной в примере 3, указаны данные редактирования для sgРНК SGN001101, направленной на CFTR.

Для анализа активности других направляющих РНК представлена направляющая РНК из таблицы 31 с соответствующим никазным вариантом каждого RGN, описанного в таблице 30, который функционально связан с деаминазой по настоящему изобретению для получения гибридного белка. Предполагалось, что неактивные в отношении нуклеазы варианты каждого RGN также можно протестировать аналогичным образом. Каждую комбинацию направляющего и гибридного белков анализировали на способность редактировать в целевой локализации в 16HBE14о-иммortalизованных клетках бронхиального эпителия. В настоящее время доступны три клеточные линии HBE, содержащие нонсенс-мутации CFTR (от фирмы Cystic Fibrosis Foundation, Легстингтон, Массачусетс). Эти клеточные линии использовали для анализа нонсенс-мишеней G542X, W1282X и R1162X и сравнения с линией 16HBE14о. Гибридный белок и направляющая РНК доставляются в клетки в виде рибонуклеопротеинов (RNP), которые нуклеофицируются в клеточную линию 16HBE14о с помощью методов культивирования и трансформации, описанных в статье Valley и др. (Valley и др., J Cyst Fibros 18, 476- 483 (2019), включенной в данный контекст в качестве ссылки). Направляющая РНК представлена в виде одной направляющей РНК или в молярном соотношении 1:1 или 1:1,2 дуплекса tracrRNA:crRNA с белками RGN. Нуклеофекцию RNP в клетки проводили на модуле Lonza 4D-Nucleofector. Затем клетки инкубировали при 37°C в течение 72 ч. В некоторых вариантах осуществления изобретения гибридный белок и gRNA доставляли в клетки в виде молекул РНК, при этом гибридный белок кодируется в мРНК.

Поскольку отсутствуют доступные клеточные линии для E60X, R553X и Q493X, такие мутации анализировали в клетках HEK293 с использованием варианта анализа восстановления GFP, описанного в примере 2, где мутантный локус, содержащий нонсенс-мутацию, клонируется во 2-ой рамке считывания GFP.

После инкубации геномную ДНК экстрагировали с использованием набора NucleoSpin 96 Tissue (от фирмы Macherey-Nagel) в соответствии с протоколом производителя. Геномную область, примыкающую к целевому геномному

участку, амплифицировали с помощью ПЦР, а продукты очищали с использованием набора ZR-96 DNA Clean and Concentrator (от фирмы Zymo Research) в соответствии с протоколом производителя. Очищенные ПЦР-продукты затем отправляли на секвенирование следующего поколения с использованием системы Illumina MiSeq. Как правило, на ампликон генерируется 100000 считываний спаренных концевых фрагментов на 250 пар оснований (2 x 100000 считываний). Считывания анализировали с использованием программы CRISPResso (Pinello и др. (2016)) для расчета скорости редактирования. Выходные исследуемые мутации с отредактированным азотистым основанием, а также для выявления нежелательного образования INDEL.

В дополнение к эффективности редактирования азотистых оснований белковый продукт гена CFTR после редактирования азотистых оснований оценивали на предмет функции. Для двух нонсенс-мутаций Glu60X и Gly542X отредактированная по азотистому основанию замена аденина на гуанин не восстанавливает последовательность дикого типа, поскольку эти мутации вызваны трансверсиями гуанина в тимин. Направляющая активность гибридного белка изменяет Glu60X на Glu60Gln и Gly452X на Gly542Arg. В то время как эти мутации позволяют получить полноразмерный белок, стабильность и функциональность белка CFTR также подтверждены.

Пример 4.2

Конструирование RGN для уменьшения размера

В идеальном случае кодирующая последовательность гибридного белка RGN-деаминазы по настоящему изобретению и соответствующая направляющая РНК для направления гибридного белка на ген CFTR упакованы в один вектор AAV. Общепринятый предельный размер векторов AAV составляет 4,7 тысяч пар оснований, хотя можно рассматривать и большие размеры за счет снижения эффективности упаковки. Никазы RGN в таблице 30 характеризуются длиной кодирующей последовательности примерно в 3,15-3,45 тысяч пар оснований. Чтобы экспрессионные кассеты как для гибридного белка, так и для его соответствующей направляющей РНК могли вписаться в вектор AAV, желательно сократить длину аминокислотной последовательности RGN и соответствующей кодирующей последовательности нуклеиновой кислоты.

При выравнивании с близкородственными гомологами уникальный участок из 8 аминокислот в положениях 590-597 идентифицировали в APG07433.1 и его близком гомологе APG08290.1 (описанном в международном патенте WO 2019/236566 и приведенном в данном контексте как SEQ ID NO: 60). Эту область, представленную как SEQ ID NO: 365 для APG07433.1 и SEQ ID NO: 367 для APG08290.1, удалили из обоих белков, в результате чего получили варианты RGN APG07433.1-del (SEQ ID NO: 366) и APG08290.1-del (SEQ ID NO: 368). Эти делеционные варианты и соответствующие им RGN дикого типа анализировали с целью определения активности редактирования в клетках HEK293T с использованием направляющих РНК, указанных в таблицах 32 и 33, в соответствии со способами, аналогичными описанными в примере 1. Скорости редактирования последовательностей-мишеней показаны ниже в таблицах 32 и 33.

Таблица 32

Скорость редактирования для вариантов делеции белка APG07433.1

Направляющая РНК	Мишень (SEQ ID NO.)	sgRNA (SEQ ID NO.)	APG07433.1	APG07433.1-del
SGN000139	369	383	11.09%	1.00%
SGN000143	370	384	2.68%	0.71%
SGN000169	371	385	13.37%	15.48%
SGN000173	372	386	13.65%	15.37%
SGN000186	373	387	14.72%	15.16%
SGN000194	374	388	11.91%	7.66%
SGN000927	376	390	9.53%	11.47%
SGN000929	378	392	6.14%	13.10%
SGN000930	379	393	7.52%	9.51%
SGN000935	381	395	11.08%	15.99%
SGN001101	382	396	6.16%	6.75%

Для мишеней SGN000169, SGN000173, SGN000186, SGN000927, SGN000930 и SGN001101 скорость редактирования белка APG07433.1 дикого типа и сконструированного варианта была одинаковой. Для мишеней SGN000139, SGN000143 и SGN000194 скорость редактирования снижается при использовании сконструированного варианта по сравнению с белком дикого типа. Для SGN000929 и SGN000935 скорость редактирования увеличилась со

сконструированным вариантом APG07433.1 по сравнению с последовательностью дикого типа.

Таблица 33

Скорость редактирования для вариантов делеции белка APG08290.1

sgRNA ID	Мишень (SEQ ID NO.)	sgRNA (SEQ ID NO.)	APG08290.1	APG08290.1- del
SGN000926	375	389	N.D.	6.47%
SGN000929	378	392	1.83%	0.61%
SGN000930	379	393	9.93%	6.47%
SGN000928	377	391	N.D.	0.13%
SGN000931	380	394	0%	0%

5

N.D. = не определено

Делеционный вариант APG08290.1 характеризовался редактированием во всех образцах, где белок APG08290.1 дикого типа также характеризовался редактированием. Самая низкая обнаруженная скорость редактирования составила 0,13% для сконструированного белка. Мишень SGN000926 харктеризовалась самым высоким уровнем редактирования: 9,17%.

10

Гибридные белки, содержащие APG07433.1-del или APG08290.1-del и дезаминазу по настоящему изобретению, получали и анализировали с целью оценки активности редактирования азотистых оснований с использованием способов, аналогичных способам из примера 1.

15

Гибридный белок содержит RGN и дезаминазу, связанные гибким пептидным линкером, таким как SEQ ID NO: 45. Длина линкера SEQ ID NO: 45 составляет 16 аминокислот, его длину можно укоротить для уменьшения размера кодирующей последовательности гибридного белка. Можно получить пептидные линкеры из менее чем 16 аминокислот и функционально связать RGN APG07433.1-del или APG08290.1-del и дезаминазу по настоящему изобретению и протестировать на активность редактирования азотистых оснований с использованием способов, аналогичных способам, описанным в примере 1. Поскольку пептидный линкер между RGN и дезаминазой может определять окно редактирования гибридного белка, тестирование альтернативных линкеров с различной длиной и жесткостью также может привести к повышению эффективности редактирования при одновременном снижении нецелевых

20

25

мутаций. Поэтому гибридные белки с самой высокой скоростью редактирования затем анализировали, используя способы, аналогичные способам, описанным в примере 4.1, для определения эффективности редактирования для каждой из последовательностей-мишеней CFTR. Комбинации гибридных белков с гРНК с наибольшей эффективностью редактирования выбирали в качестве предпочтительных направляющих для редактирования в данной локализации и использовали для конструирования вектора AAV.

Пример 4.3

Доставка AAV

Кодирующие последовательности для проверенных комбинаций гибридный белок/гРНК с самой высокой скоростью редактирования упаковываются в векторы AAV. Доставка AAV характеризуется рядом преимуществ, включая отсутствие патогенности, низкую иммуногенность, высокую скорость трансдукции и определенный способ продуцирования. Кроме того, было установлено, что введение AAV в легкие безопасно и, по меньшей мере, в некоторой степени эффективно как при однократном, так и при повторном введении (Guggino и др., Expert Opin Biol Ther 17, 1265-1273 (2017)). После того, как комбинацию гибридный белок/гРНК клонировали в вектор AAV, ее можно упаковать в несколько различных серотипов для оптимизации тканеспецифичной инфекционности. Для лечения муковисцидоза, мишенью для редактирования оснований являются клетки-предшественники апикального эпителия легких, что позволит такой коррекции сохраняться на протяжении всего обновления клеток. Для направления на респираторный эпителий использовали капсид для серотипов AAV1, AAV5 или AAV6, поскольку было установлено, что эти серотипы обладают высокой инфекционностью в клетках респираторного эпителия (Zabner и др., J Virol 74, 3852-3858 (2000)).

После получения векторов AAV их трансдуцировали в эпителиальные клетки дыхательных путей человека в культуре. Три клеточные линии HBE, содержащие мишени для nonsense-мутаций CFTR G542X, R1162X и W1282X, использовали для проверки конструкторов для коррекции этих мутаций. Линию 16HBE14o- использовали для тестирования конструкторов, исправляющих другие nonsense-мутации. Проверяли диапазон множественности заражения (MOI). В каждом случае оценивали реверсию nonsense-мутации в последовательность CFTR дикого типа. После 2–3 дней культивирования собирали геномную ДНК,

генерировали ампликоны вокруг участков-мишеней с помощью ПЦР и выполняли NGS для определения скорости редактирования в каждом локусе, аналогично методам, описанным в примере 1. Поскольку использовали эпителиальные клетки дыхательных путей, скорость введения и редактирования AAV в наибольшей степени являлись сходными с этими показателями при лечении *in vivo* при использовании системы культивируемых клеток. AAV с различными серотипами сравнивали, чтобы определить, какой серотип является оптимальным для доставки комбинации гибридный белок/гРНК в клетки дыхательных путей. Скорости редактирования, достигнутые за счет введения этих систем в AAV, сравнивали со скоростями редактирования RNP, наблюдаемыми в примере 4.2.

Поскольку клеточные линии для нонсенс-мутаций R553X, E60X и Q493X недоступны, системы гибридных белков/гРНК, направленные на эти мутации, оценивали в клетках 16HBE14o- дикого типа для анализа внедрения AAV, экспрессии редактора оснований и скоростей нецелевого редактирования в исследуемый участок. Чтобы определить скорость коррекции стоп-кодона, мутантный локус клонировали в GFP для анализа восстановления GFP, как описано в примере 4.1.

Параллельно с определением скорости редактирования с помощью NGS собирали лизаты общего белка из клеток, несущих мутации CFTR, отредактированные с помощью систем гибридного белка/гРНК, и определяли уровни полноразмерного белка CFTR с помощью вестерн-блоттинга. Чтобы проверить, образуется ли функциональный белок CFTR, проводили анализы активации форсколина с использованием методов, аналогичных тем, которые описаны в статье Devor и др. (Am J Physiol Cell Physiol 279, C461-479 (2000), включенной в настоящий контекст в качестве ссылки) и/или статье Dousmaïs и др. (J Gen Physiol 119, 545-559 (2002), включенной в настоящий контекст в качестве ссылки). В этих экспериментах отредактированные мутантные клетки CFTR обрабатывали форсколином, активатором аденилатциклазы, для повышения внутриклеточного уровня цАМФ. Затем повышенные уровни цАМФ активировали CFTR, а приток Cl^- измеряли либо генетически кодируемым желтым флуоресцентным белком на основе Cl^- -сенсора, либо низкомолекулярным флуоресцентным индикатором хлорида, таким как MQAE. В

этом анализе тестировали отредактированные клеточные линии G542X, R1162X и W1282X.

Чтобы определить скорость нецелевых мутаций, использовали биоинформационный подход, основанный на информации о затравочной области и гибком пространстве распознавания нецелевого РАМ каждой конкретной нуклеазы. Эти фрагменты информации определяли на основе биоданных для каждого белка и использовали для ранжирования вероятности нецелевой активности для каждого белка.

В дополнение к биоинформативному прогнозированию нецелевых объектов также проводили биохимическое обнаружение нецелевых объектов с помощью модифицированного протокола SITE-seq (см. статью Cameron и др., Nat Methods 14, 600-606 (2017), включенную в данный контекст в качестве ссылки). Краткое описание метода: получали геномную ДНК из эпителиальных клеток дыхательных путей человека. Затем эту ДНК обрабатывали исследуемой RGN в нескольких различных концентрациях. Любые двухцепочечные разрывы ДНК метят, селективно выделяют и амплифицировали с помощью ПЦР с адаптерными последовательностями, которые допускают NGS. Затем считывания при секвенировании сопоставляли с геномом, и «нагромождения» считываний идентифицировали в местах двухцепочечных разрывов, отмечая предполагаемые нецелевые местоположения. В последующей серии экспериментов клетки редактировали исследуемым гибридным белком RGN или RGN-дезаминазой, и эти предполагаемые сайты индивидуально секвенировали, чтобы подтвердить, действительно ли они не являются мишенями. Поскольку окружение хроматина, доступность ДНК и другие факторы могут влиять на эффективность редакторов генома в живых клетках, биохимические методы обычно переоценивают количество нецелевых объектов. Таким образом, как биоинформационные, так и биохимические методы вместе обеспечивают дополнительные способы для выявления предполагаемых нецелевых сайтов, но эти сайты должны быть проверены с помощью секвенирования ампликонов, чтобы получить точную оценку нецелевого редактирования.

Как только предполагаемые нецелевые сайты идентифицированы, секвенирование ампликона 16HBE на эпителиальных клетках дыхательных путей, отредактированных с использованием того же оптимизированного гибридного белка и направляющей(их), гарантирует, что нецелевой профиль,

установленный для этих систем, максимально точно соответствует ожидаемому профилю в легких пациента.

Чтобы определить, вызывают ли описанные в данном контексте гибридные белки изменения в клеточной РНК, необходим тщательный анализ клеточного транскриптома после редактирования. К счастью, методы секвенирования РНК для оценки нецелевых эффектов редактирования адениновых оснований стали традиционными (Grunewald и др., Nature 569, 433-437 (2017), Zhou и др., Nature 571, 275-278, обе статьи включены в настоящий контекст в качестве ссылки).

Краткое описание метода: после редактирования клеток с помощью систем гибридный белок/гРНК, определенных в примере 4.2, собирали общую клеточную мРНК и подвергали секвенированию РНК. Транскрипты отредактированных клеток сравнивали с клетками, трансфицированными только ABE, и выявляли существенные различия в последовательности РНК.

Пример 5

Направленное редактирование азотистого основания для исправления мутаций, вызывающих заболевание

Базу данных клинических вариантов получали из базы данных NCBI ClinVar, которая доступна в Интернете на веб-сайте NCBI ClinVar. Случаи патогенного однонуклеотидного полиморфизма (SNP) идентифицировали с помощью этого списка. Используя информацию о геномном локусе, идентифицировали мишени CRISPR в области, перекрывающейся и окружающей каждый SNP. Выбор SNP, которые можно скорректировать с помощью редактирования азотистого основания в сочетании с RGN, например, RGN, указанный в таблице 30, или его вариант для направления на причинную мутацию («Cas1 Mut.»), указан в таблице 34. В таблице 34 ниже указано только одно обозначение каждой болезни. Обозначение «RS#» соответствует регистрационному номеру RS в базе данных SNP на веб-сайте NCBI. Обозначение «AlleleID» соответствует идентификационному номеру причинной аллели. Столбец «Наименование» содержит идентификатор генетического локуса, название гена, местоположение мутации в гене и изменение, полученное в результате мутации.

Таблица 34

Мишени при заболевании для редактирования азотистого основания

RS#	ID аллеля	Наименование	Обозначение гена
36053993	20333	NM_001128425.1(MUTYH):c.1187G>A (p.Gly396Asp)	MUTYH
41293455	32714	NM_007294.3(BRCA1):c.4327C>T (p.Arg1443Ter)	BRCA1
62625308	32710	NM_007294.3(BRCA1):c.3607C>T (p.Arg1203Ter)	BRCA1
41293465	70268	NM_007294.3(BRCA1):c.5503C>T (p.Arg1835Ter)	BRCA1
80357123	70147	NM_007294.3(BRCA1):c.5251C>T (p.Arg1751Ter)	BRCA1
137929307	171217	NM_000527.4(LDLR):c.1775G>A (p.Gly592Glu)	LDLR
80356898	45982	NM_007294.3(BRCA1):c.1687C>T (p.Gln563Ter)	BRCA1
28936415	22745	NM_000303.2(PMM2):c.422G>A (p.Arg141His)	PMM2
11555217	34125	NM_001360.2(DHCR7):c.452G>A (p.Trp151Ter)	DHCR7
55770810	70063	NM_007294.3(BRCA1):c.5095C>T (p.Arg1699Trp)	BRCA1
28934906	26850	NM_004992.3(MECP2):c.473C>T (p.Thr158Met)	MECP2
28929474	33006	NM_001127701.1(SERPINA1):c.1096G>A (p.Glu366Lys)	SERPINA1
371898076	52045	NM_000257.4(MYH7):c.1988G>A (p.Arg663His)	MYH7
5030858	15616	NM_000277.3(PAH):c.1222C>T (p.Arg408Trp)	PAH
80356945	69207	NM_007294.3(BRCA1):c.2338C>T (p.Gln780Ter)	BRCA1
1800553	22927	NM_000350.2(ABCA4):c.5882G>A (p.Gly1961Glu)	ABCA4
80356962	70247	NM_007294.3(BRCA1):c.5444G>A (p.Trp1815Ter)	BRCA1
104894396	32041	NM_004004.6(GJB2):c.71G>A (p.Trp24Ter)	GJB2
113994095	28535	NM_002693.2(POLG):c.1399G>A (p.Ala467Thr)	POLG
61749721	26868	NM_004992.3(MECP2):c.763C>T (p.Arg255Ter)	MECP2
137852700	23943	NM_000310.3(PPT1):c.451C>T (p.Arg151Ter)	PPT1
75527207	22159	NM_000492.3(CFTR):c.1652G>A (p.Gly551Asp)	CFTR
78655421	22148	NM_000492.3(CFTR):c.350G>A (p.Arg117His)	CFTR
80356885	69888	NM_007294.3(BRCA1):c.4524G>A (p.Trp1508Ter)	BRCA1
113994098	28541	NM_002693.2(POLG):c.2542G>A (p.Gly848Ser)	POLG
61750240	26854	NM_004992.3(MECP2):c.808C>T (p.Arg270Ter)	MECP2
61751362	26858	NM_001110792.1(MECP2):c.916C>T (p.Arg306Ter)	MECP2
80357260	69792	NM_007294.3(BRCA1):c.4183C>T (p.Gln1395Ter)	BRCA1
80359071	67203	NM_000059.3(BRCA2):c.8243G>A (p.Gly2748Asp)	BRCA2
62625307	69596	NM_007294.3(BRCA1):c.3598C>T (p.Gln1200Ter)	BRCA1
76992529	28465	NM_000371.3(TTR):c.424G>A (p.Val142Ile)	TTR
77010898	22168	NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)	CFTR
80359003	67069	NM_000059.3(BRCA2):c.7757G>A (p.Trp2586Ter)	BRCA2
61750420	22555	NM_000466.2(PEX1):c.2528G>A (p.Gly843Asp)	PEX1
80357284	46214	NM_007294.3(BRCA1):c.5346G>A (p.Trp1782Ter)	BRCA1
200411226	174776	NM_000256.3(MYBPC3):c.1484G>A (p.Arg495Gln)	MYBPC3
5030857	98638	NM_000277.3(PAH):c.1208C>T (p.Ala403Val)	PAH
28935468	26863	NM_004992.3(MECP2):c.916C>T (p.Arg306Cys)	MECP2
62642937	15667	NM_000277.3(PAH):c.1139C>T (p.Thr380Met)	PAH

RS#	ID аллеля	Наименование	Обозначение гена
80356989	69812	NM_007294.3(BRCA1):c.4222C>T (p.Gln1408Ter)	BRCA1
28942080	18735	NM_000527.4(LDLR):c.1567G>A (p.Val523Met)	LDLR
121908039	18778	NM_000527.4(LDLR):c.551G>A (p.Cys184Tyr)	LDLR
267607213	18780	NM_000527.4(LDLR):c.131G>A (p.Trp44Ter)	LDLR
3218716	52071	NM_000257.3(MYH7):c.2389G>A (p.Ala797Thr)	MYH7
104895097	17588	NM_000243.2(MEFV):c.2282G>A (p.Arg761His)	MEFV
397516074	51962	NM_000256.3(MYBPC3):c.772G>A (p.Glu258Lys)	MYBPC3
119455955	17682	NM_000391.3(TPP1):c.622C>T (p.Arg208Ter)	TPP1
75184679	16301	NM_024570.3(RNASEH2B):c.529G>A (p.Ala177Thr)	RNASEH2B
80338901	26909	NM_000137.2(FAH):c.1062+5G>A	FAH
119450941	17501	NM_000026.3(ADSL):c.1277G>A (p.Arg426His)	ADSL
121965019	26947	NM_000203.4(IDUA):c.1205G>A (p.Trp402Ter)	IDUA
141659620	21858	NM_003119.3(SPG7):c.1045G>A (p.Gly349Ser)	SPG7
41276738	15335	NM_000552.4(VWF):c.2561G>A (p.Arg854Gln)	VWF
80338940	32068	NM_004004.5(GJB2):c.-23+1G>A	GJB2
80357292	46268	NM_007294.3(BRCA1):c.962G>A (p.Trp321Ter)	BRCA1
121913627	29130	NM_000257.3(MYH7):c.1816G>A (p.Val606Met)	MYH7
137854601	24416	NM_198056.2(SCN5A):c.5350G>A (p.Glu1784Lys)	SCN5A
80338933	17521	NM_024577.3(SH3TC2):c.2860C>T (p.Arg954Ter)	SH3TC2
80338948	32048	NM_004004.5(GJB2):c.427C>T (p.Arg143Trp)	GJB2
80356903	69645	NM_007294.3(BRCA1):c.3718C>T (p.Gln1240Ter)	BRCA1
80356969	70213	NM_007294.3(BRCA1):c.5353C>T (p.Gln1785Ter)	BRCA1
80357010	45971	NM_007294.3(BRCA1):c.1480C>T (p.Gln494Ter)	BRCA1
116987552	17337	NM_005609.3(PYGM):c.148C>T (p.Arg50Ter)	PYGM
121913625	29128	NM_000257.4(MYH7):c.1357C>T (p.Arg453Cys)	MYH7
387907267	45725	NM_000256.3(MYBPC3):c.2827C>T (p.Arg943Ter)	MYBPC3
28934897	26968	NM_000431.3(MVK):c.1129G>A (p.Val377Ile)	MVK
76713772	22151	NM_000492.3(CFTR):c.1585-1G>A	CFTR
137852959	19587	NM_153638.3(PANK2):c.1561G>A (p.Gly521Arg)	PANK2
199682486	101428	NM_013339.4(ALG6):c.257+5G>A	ALG6
397507389	46666	NM_000059.3(BRCA2):c.7618-1G>A	BRCA2
769370816	228176	NM_000527.4(LDLR):c.1618G>A (p.Ala540Thr)	LDLR
36211715	29159	NM_000257.4(MYH7):c.2609G>A (p.Arg870His)	MYH7
76434661	53916	NM_004004.5(GJB2):c.416G>A (p.Ser139Asn)	GJB2
104894368	29104	NM_000432.3(MYL2):c.64G>A (p.Glu22Lys)	MYL2
104894635	20146	NM_000199.3(SGSH):c.734G>A (p.Arg245His)	SGSH
121913628	29131	NM_000257.3(MYH7):c.2770G>A (p.Glu924Lys)	MYH7
193922390	45304	NM_000257.4(MYH7):c.5135G>A (p.Arg1712Gln)	MYH7
397515757	51454	NM_000138.4(FBN1):c.1468+5G>A	FBN1
11549407	30441	NM_000518.5(HBB):c.118C>T (p.Gln40Ter)	HBB
61751374	22933	NM_000350.2(ABCA4):c.3113C>T (p.Ala1038Val)	ABCA4
121434420	21793	NM_004572.3(PKP2):c.235C>T (p.Arg79Ter)	PKP2

RS#	ID аллеля	Наименование	Обозначение гена
137853007	20631	NM_007194.4(CHEK2):c.433C>T (p.Arg145Trp)	CHEK2
1137887	18083	NM_000051.3(ATM):c.2250G>A (p.Lys750=)	ATM
28934872	27436	NM_000548.3(TSC2):c.1832G>A (p.Arg611Gln)	TSC2
80224560	47062	NM_000492.3(CFTR):c.2657+5G>A	CFTR
80359004	46672	NM_000059.3(BRCA2):c.7758G>A (p.Trp2586Ter)	BRCA2
121434274	18627	NM_000016.5(ACADM):c.799G>A (p.Gly267Arg)	ACADM
121908529	38436	NM_000030.2(AGXT):c.508G>A (p.Gly170Arg)	AGXT
121918007	28709	NM_000478.4(ALPL):c.571G>A (p.Glu191Lys)	ALPL
121918243	16464	NM_015506.2(MMACHC):c.482G>A (p.Arg161Gln)	MMACHC
397518423	94255	NM_005026.4(PIK3CD):c.3061G>A (p.Glu1021Lys)	PIK3CD
587781629	150997	NM_000059.3(BRCA2):c.1909+1G>A	BRCA2
765696008	228162	NM_000527.4(LDLR):c.1187-10G>A	LDLR
3218713	29127	NM_000257.3(MYH7):c.746G>A (p.Arg249Gln)	MYH7
5030855	15646	NM_000277.3(PAH):c.1066-11G>A	PAH
55851803	69067	NM_007294.3(BRCA1):c.191G>A (p.Cys64Tyr)	BRCA1
62508698	15619	NM_000277.1(PAH):c.838G>A (p.Glu280Lys)	PAH
62516152	108520	NM_000277.3(PAH):c.688G>A (p.Val230Ile)	PAH
62644499	15656	NM_000277.3(PAH):c.1243G>A (p.Asp415Asn)	PAH
80338815	18090	NM_000487.5(ARSA):c.465+1G>A	ARSA
121908987	21885	NM_016203.3(PRKAG2):c.905G>A (p.Arg302Gln)	PRKAG2
121964962	15156	NM_000071.2(CBS):c.919G>A (p.Gly307Ser)	CBS
5030851	15628	NM_000277.3(PAH):c.842C>T (p.Pro281Leu)	PAH
63750871	24273	NM_000535.6(PMS2):c.400C>T (p.Arg134Ter)	PMS2
80338853	21822	NM_001360.2(DHCR7):c.278C>T (p.Thr93Met)	DHCR7
80356893	68976	NM_007294.3(BRCA1):c.1612C>T (p.Gln538Ter)	BRCA1
80357131	46031	NM_007294.3(BRCA1):c.2563C>T (p.Gln855Ter)	BRCA1
80357223	69350	NM_007294.3(BRCA1):c.2800C>T (p.Gln934Ter)	BRCA1
80357318	46112	NM_007294.3(BRCA1):c.3937C>T (p.Gln1313Ter)	BRCA1
104886457	27086	NM_000136.2(FANCC):c.1642C>T (p.Arg548Ter)	FANCC
137852944	19147	NM_138694.3(PKHD1):c.107C>T (p.Thr36Met)	PKHD1
180177083	132139	NM_024675.3(PALB2):c.196C>T (p.Gln66Ter)	PALB2
180177110	152117	NM_024675.3(PALB2):c.2257C>T (p.Arg753Ter)	PALB2
199475575	108459	NM_000277.3(PAH):c.526C>T (p.Arg176Ter)	PAH
387906843	39241	NM_002878.3(RAD51D):c.556C>T (p.Arg186Ter)	RAD51D
529008617	152318	NM_001128425.1(MUTYH):c.1214C>T (p.Pro405Leu)	MUTYH
587780021	133177	NM_000465.3(BARD1):c.1690C>T (p.Gln564Ter)	BARD1
34637584	16979	NM_198578.3(LRRK2):c.6055G>A (p.Gly2019Ser)	LRRK2
78802634	22233	NM_000492.3(CFTR):c.3266G>A (p.Trp1089Ter)	CFTR
80358809	66611	NM_000059.3(BRCA2):c.581G>A (p.Trp194Ter)	BRCA2
80359011	46678	NM_000059.3(BRCA2):c.7857G>A (p.Trp2619Ter)	BRCA2
104894503	27495	NM_001018005.1(TPM1):c.523G>A (p.Asp175Asn)	TPM1
121908641	21368	NM_000050.4(ASS1):c.1168G>A (p.Gly390Arg)	ASS1

RS#	ID аллеля	Наименование	Обозначение гена
121918593	28009	NM_000540.2(RYR1):c.7300G>A (p.Gly2434Arg)	RYR1
140108514	100191	NM_003494.3(DYSF):c.2643+1G>A	DYSF
145138923	98271	NM_000048.3(ASL):c.35G>A (p.Arg12Gln)	ASL
150726175	45795	NM_022787.3(NMNAT1):c.769G>A (p.Glu257Lys)	NMNAT1
267607578	45138	NM_170707.3(LMNA):c.1412G>A (p.Arg471His)	LMNA
376607329	48992	NM_002834.4(PTPN11):c.794G>A (p.Arg265Gln)	PTPN11
587776934	48407	NM_005027.3(PIK3R2):c.1117G>A (p.Gly373Arg)	PIK3R2
62508588	15630	NM_000277.1(PAH):c.728G>A (p.Arg243Gln)	PAH
62637014	20604	NM_014336.4(AIPL1):c.834G>A (p.Trp278Ter)	AIPL1
80356860	46194	NM_007294.3(BRCA1):c.5117G>A (p.Gly1706Glu)	BRCA1
80357268	70265	NM_007294.3(BRCA1):c.5497G>A (p.Val1833Met)	BRCA1
80357418	70077	NM_007294.3(BRCA1):c.5136G>A (p.Trp1712Ter)	BRCA1
80358145	46229	NM_007294.3(BRCA1):c.5467+1G>A	BRCA1
121918166	15994	NM_000275.2(OCA2):c.1327G>A (p.Val443Ile)	OCA2
140342925	150591	NM_001128425.1(MUTYH):c.734G>A (p.Arg245His)	MUTYH
148660051	195093	NM_206933.2(USH2A):c.10073G>A (p.Cys3358Tyr)	USH2A
193922672	45341	NM_004572.3(PKP2):c.1613G>A (p.Trp538Ter)	PKP2
267607144	20039	NM_021625.4(TRPV4):c.806G>A (p.Arg269His)	TRPV4
397516083	51977	NM_000256.3(MYBPC3):c.927-9G>A	MYBPC3
397516357	52565	NM_000363.4(TNNI3):c.557G>A (p.Arg186Gln)	TNNI3
587782958	165560	NM_000256.3(MYBPC3):c.3190+5G>A	MYBPC3
28934907	26853	NM_004992.3(MECP2):c.316C>T (p.Arg106Trp)	MECP2
28934908	26862	NM_004992.3(MECP2):c.419C>T (p.Ala140Val)	MECP2
28940893	18091	NM_000487.5(ARSA):c.1283C>T (p.Pro428Leu)	ARSA
63751422	96795	NM_000535.5(PMS2):c.1927C>T (p.Gln643Ter)	PMS2
74315366	27817	NM_003000.2(SDHB):c.268C>T (p.Arg90Ter)	SDHB
80338856	34127	NM_001360.2(DHCR7):c.724C>T (p.Arg242Cys)	DHCR7
80357038	69707	NM_007294.3(BRCA1):c.3895C>T (p.Gln1299Ter)	BRCA1
80357136	69535	NM_007294.3(BRCA1):c.3403C>T (p.Gln1135Ter)	BRCA1
80357208	69682	NM_007294.3(BRCA1):c.3817C>T (p.Gln1273Ter)	BRCA1
80357234	69166	NM_007294.3(BRCA1):c.220C>T (p.Gln74Ter)	BRCA1
80357262	69729	NM_007294.3(BRCA1):c.3967C>T (p.Gln1323Ter)	BRCA1
80357305	69822	NM_007294.3(BRCA1):c.4258C>T (p.Gln1420Ter)	BRCA1
80357350	69232	NM_007294.3(BRCA1):c.241C>T (p.Gln81Ter)	BRCA1
104894636	20147	NM_000199.3(SGSH):c.220C>T (p.Arg74Cys)	SGSH
111401431	44742	NM_000138.4(FBN1):c.4588C>T (p.Arg1530Cys)	FBN1
121918624	27928	NM_006920.5(SCN1A):c.664C>T (p.Arg222Ter)	SCN1A
137852981	19794	NM_014795.3(ZEB2):c.2083C>T (p.Arg695Ter)	ZEB2
137854476	31491	NM_000138.4(FBN1):c.1585C>T (p.Arg529Ter)	FBN1
137854480	31500	NM_000138.4(FBN1):c.718C>T (p.Arg240Cys)	FBN1
180177100	133574	NM_024675.3(PALB2):c.1240C>T (p.Arg414Ter)	PALB2
193922109	44392	NM_000053.3(ATP7B):c.3955C>T (p.Arg1319Ter)	ATP7B

RS#	ID аллеля	Наименование	Обозначение гена
200640585	96857	NM_000535.6(PMS2):c.943C>T (p.Arg315Ter)	PMS2
201431517	48426	NM_139242.3(MTFMT):c.626C>T (p.Ser209Leu)	MTFMT
397516037	51905	NM_000256.3(MYBPC3):c.3697C>T (p.Gln1233Ter)	MYBPC3
587780104	133350	NM_002878.3(RAD51D):c.694C>T (p.Arg232Ter)	RAD51D
765123255	181726	NM_001128425.1(MUTYH):c.325C>T (p.Arg109Trp)	MUTYH
63751657	95331	NM_000249.3(MLH1):c.1731G>A (p.Ser577=)	MLH1
75549581	22162	NM_000492.3(CFTR):c.1675G>A (p.Ala559Thr)	CFTR
80338851	16303	NM_194318.3(B3GLCT):c.660+1G>A	B3GLCT
80358544	46368	NM_000059.3(BRCA2):c.2979G>A (p.Trp993Ter)	BRCA2
111033178	52388	NM_000260.3(MYO7A):c.3719G>A (p.Arg1240Gln)	MYO7A
121908188	19535	NM_020451.2(SELENON):c.943G>A (p.Gly315Ser)	SELENON
139770721	180483	NM_000051.3(ATM):c.6095G>A (p.Arg2032Lys)	ATM
199476315	40542	NM_001018005.1(TPM1):c.574G>A (p.Glu192Lys)	TPM1
267607004	15310	NM_001134363.2(RBM20):c.1907G>A (p.Arg636His)	RBM20
267608122	94980	NM_000179.2(MSH6):c.4001G>A (p.Arg1334Gln)	MSH6
377349459	150947	NM_000051.3(ATM):c.7913G>A (p.Trp2638Ter)	ATM
387906303	18745	NM_000527.4(LDLR):c.670G>A (p.Asp224Asn)	LDLR
587779227	94719	NM_000179.2(MSH6):c.2057G>A (p.Gly686Asp)	MSH6
587780290	134019	NM_000070.2(CAPN3):c.2243G>A (p.Arg748Gln)	CAPN3
727504317	49251	NM_002755.3(MAP2K1):c.199G>A (p.Asp67Asn)	MAP2K1
5030869	25402	NM_000402.4(G6PD):c.1093G>A (p.Ala365Thr)	G6PD
9332964	18390	NM_000348.3(SRD5A2):c.680G>A (p.Arg227Gln)	SRD5A2
36211723	45266	NM_000256.3(MYBPC3):c.2308G>A (p.Asp770Asn)	MYBPC3
72549410	78547	NM_000335.4(SCN5A):c.1231G>A (p.Val411Met)	SCN5A
80357498	45948	NM_007294.3(BRCA1):c.116G>A (p.Cys39Tyr)	BRCA1
80358079	70118	NM_007294.3(BRCA1):c.5194-12G>A	BRCA1
121434529	33201	NM_000262.2(NAGA):c.973G>A (p.Glu325Lys)	NAGA
121908627	21067	NM_005476.5(GNE):c.2086G>A (p.Val696Met)	GNE
387906592	38552	NM_001613.2(ACTA2):c.536G>A (p.Arg179His)	ACTA2
397515907	51711	NM_000256.3(MYBPC3):c.1505G>A (p.Arg502Gln)	MYBPC3
397516089	51992	NM_000257.4(MYH7):c.1106G>A (p.Arg369Gln)	MYH7
397516248	52239	NM_000257.4(MYH7):c.5401G>A (p.Glu1801Lys)	MYH7
397516349	52554	NM_000363.4(TNNI3):c.434G>A (p.Arg145Gln)	TNNI3
5030846	15627	NM_000277.3(PAH):c.727C>T (p.Arg243Ter)	PAH
28941784	18134	NM_052845.3(MMAB):c.556C>T (p.Arg186Trp)	MMAB
34126013	181693	NM_001128425.1(MUTYH):c.721C>T (p.Arg241Trp)	MUTYH
62541771	21074	NM_001128227.2(GNE):c.1985C>T (p.Ala662Val)	GNE
62625303	68931	NM_007294.3(BRCA1):c.1471C>T (p.Gln491Ter)	BRCA1
74315379	27453	NM_001001430.2(TNNT2):c.421C>T (p.Arg141Trp)	TNNT2
76687508	108539	NM_000277.3(PAH):c.721C>T (p.Arg241Cys)	PAH
80338794	20654	NM_012434.4(SLC17A5):c.115C>T (p.Arg39Cys)	SLC17A5
80356866	69689	NM_007294.3(BRCA1):c.3841C>T (p.Gln1281Ter)	BRCA1

RS#	ID аллеля	Наименование	Обозначение гена
80357134	69569	NM_007294.3(BRCA1):c.34C>T (p.Gln12Ter)	BRCA1
80357229	69904	NM_007294.3(BRCA1):c.4609C>T (p.Gln1537Ter)	BRCA1
111033260	19972	NM_033056.3(PCDH15):c.733C>T (p.Arg245Ter)	PCDH15
121909398	17403	NM_201548.4(CERKL):c.769C>T (p.Arg257Ter)	CERKL
121913637	29143	NM_000257.4(MYH7):c.2155C>T (p.Arg719Trp)	MYH7
200495564	50200	NM_001128425.1(MUTYH):c.733C>T (p.Arg245Cys)	MUTYH
267607203	20760	NM_194456.1(KRIT1):c.1363C>T (p.Gln455Ter)	KRIT1
587776527	132239	NM_024675.3(PALB2):c.3256C>T (p.Arg1086Ter)	PALB2
587777219	125784	NM_172107.3(KCNQ2):c.794C>T (p.Ala265Val)	KCNQ2
587778617	96774	NM_000535.5(PMS2):c.1261C>T (p.Arg421Ter)	PMS2
587783057	166274	NM_001128425.1(MUTYH):c.1171C>T (p.Gln391Ter)	MUTYH
730880099	178699	NM_000138.4(FBN1):c.1633C>T (p.Arg545Cys)	FBN1
2309689	33868	NM_000018.3(ACADVL):c.1322G>A (p.Gly441Asp)	ACADVL
28933093	29543	NM_170707.3(LMNA):c.481G>A (p.Glu161Lys)	LMNA
28937873	20571	NM_014249.3(NR2E3):c.932G>A (p.Arg311Gln)	NR2E3
59332535	77828	NM_170707.3(LMNA):c.746G>A (p.Arg249Gln)	LMNA
62645748	48213	NM_201253.2(CRB1):c.2843G>A (p.Cys948Tyr)	CRB1
63750828	96748	NM_000251.2(MSH2):c.998G>A (p.Cys333Tyr)	MSH2
80358456	65843	NM_000059.3(BRCA2):c.1689G>A (p.Trp563Ter)	BRCA2
80359101	67273	NM_000059.3(BRCA2):c.8489G>A (p.Trp2830Ter)	BRCA2
80359148	131733	NM_000059.3(BRCA2):c.8969G>A (p.Trp2990Ter)	BRCA2
80359149	67384	NM_000059.3(BRCA2):c.8970G>A (p.Trp2990Ter)	BRCA2
80359211	46791	NM_000059.3(BRCA2):c.9380G>A (p.Trp3127Ter)	BRCA2
111033565	26915	NM_002769.4(PRSS1):c.365G>A (p.Arg122His)	PRSS1
113994205	19482	NM_004937.2(CTNS):c.414G>A (p.Trp138Ter)	CTNS
116840778	23322	NM_033337.2(CAV3):c.80G>A (p.Arg27Gln)	CAV3;SSUH2
118192158	76835	NM_000540.2(RYR1):c.14818G>A (p.Ala4940Thr)	RYR1
121434278	18633	NM_000016.5(ACADM):c.583G>A (p.Gly195Arg)	ACADM
121434346	17058	NM_001003841.2(SLC6A19):c.517G>A (p.Asp173Asn)	SLC6A19
121908011	18814	NM_000372.4(TYR):c.1147G>A (p.Asp383Asn)	TYR
121908638	21365	NM_000050.4(ASS1):c.539G>A (p.Ser180Asn)	ASS1
121912938	32219	NM_001848.2(COL6A1):c.850G>A (p.Gly284Arg)	COL6A1
137853096	22694	NM_000414.3(HSD17B4):c.46G>A (p.Gly16Ser)	HSD17B4
151344631	45847	NM_000218.2(KCNQ1):c.613G>A (p.Val205Met)	KCNQ1
192838388	98283	NM_000050.4(ASS1):c.787G>A (p.Val263Met)	ASS1
267607768	95759	NM_000249.3(MLH1):c.588+5G>A	MLH1
376107921	213634	NM_000070.2(CAPN3):c.1319G>A (p.Arg440Gln)	CAPN3
397507981	67234	NM_000059.3(BRCA2):c.8364G>A (p.Trp2788Ter)	BRCA2
398124321	101692	NM_017780.3(CHD7):c.5405-7G>A	CHD7
730882246	181441	NM_194279.3(ISCA2):c.229G>A (p.Gly77Ser)	ISCA2
778906552	195186	NM_000016.5(ACADM):c.443G>A (p.Arg148Lys)	ACADM
139428292	39421	NM_005105.4(RBM8A):c.-21G>A	RBM8A

RS#	ID аллеля	Наименование	Обозначение гена
28934891	15165	NM_000071.2(CBS):c.1330G>A (p.Asp444Asn)	CBS
28937316	24408	NM_198056.2(SCN5A):c.4931G>A (p.Arg1644His)	SCN5A
33930165	30165	NM_000518.4(HBB):c.19G>A (p.Glu7Lys)	HBB
35004220	30493	NM_000518.5(HBB):c.93-21G>A	HBB
45546039	48043	NM_198056.2(SCN5A):c.665G>A (p.Arg222Gln)	SCN5A
61751402	105177	NM_000350.2(ABCA4):c.4469G>A (p.Cys1490Tyr)	ABCA4
72549387	22776	NM_000104.3(CYP1B1):c.171G>A (p.Trp57Ter)	CYP1B1
75822236	19350	NM_000157.3(GBA):c.1604G>A (p.Arg535His)	GBA
79389353	20821	NM_014270.4(SLC7A9):c.544G>A (p.Ala182Thr)	SLC7A9
80338862	34124	NM_001360.2(DHCR7):c.1228G>A (p.Gly410Ser)	DHCR7
80338892	27366	NM_199292.2(TH):c.698G>A (p.Arg233His)	TH
80356935	68777	NM_007294.3(BRCA1):c.1059G>A (p.Trp353Ter)	BRCA1
80357468	68802	NM_007294.3(BRCA1):c.1116G>A (p.Trp372Ter)	BRCA1
104894365	27628	NM_004985.4(KRAS):c.40G>A (p.Val14Ile)	KRAS
104894639	20153	NM_000199.3(SGSH):c.1339G>A (p.Glu447Lys)	SGSH
111033364	17396	NM_206933.2(USH2A):c.11864G>A (p.Trp3955Ter)	USH2A
119103251	17338	NM_005609.3(PYGM):c.613G>A (p.Gly205Ser)	PYGM
119455954	17681	NM_000391.3(TPP1):c.1094G>A (p.Cys365Tyr)	TPP1
121913638	29144	NM_000257.4(MYH7):c.2146G>A (p.Gly716Arg)	MYH7
137854478	31496	NM_000138.4(FBN1):c.3217G>A (p.Glu1073Lys)	FBN1
143353451	179937	NM_001128425.1(MUTYH):c.545G>A (p.Arg182His)	MUTYH
151045328	20182	NM_005709.3(USH1C):c.216G>A (p.Val72=)	USH1C
151344623	24127	NM_001287174.1(ABCC8):c.3992-9G>A	ABCC8
193922204	44739	NM_000138.4(FBN1):c.4460-8G>A	FBN1
193922219	51564	NM_000138.4(FBN1):c.5788+5G>A	FBN1
193922680	33370	NM_005159.4(ACTC1):c.301G>A (p.Glu101Lys)	ACTC1
267608172	96804	NM_000535.5(PMS2):c.2174+1G>A	PMS2
397516202	52163	NM_000257.3(MYH7):c.4135G>A (p.Ala1379Thr)	MYH7
397516209	52176	NM_000257.4(MYH7):c.428G>A (p.Arg143Gln)	MYH7
397517159	49176	NM_005633.3(SOS1):c.2536G>A (p.Glu846Lys)	SOS1
587776576	18532	NM_024426.5(WT1):c.1447+5G>A	WT1
727503246	175600	NM_000257.4(MYH7):c.4066G>A (p.Glu1356Lys)	MYH7
730881687	181107	NM_007194.4(CHEK2):c.793-1G>A	CHEK2
748170941	181727	NM_001128425.1(MUTYH):c.309G>A (p.Trp103Ter)	MUTYH
140583	260073	NM_000138.4(FBN1):c.2581C>T (p.Arg861Ter)	FBN1
2754158	175617	NM_000257.3(MYH7):c.2572C>T (p.Arg858Cys)	MYH7
28931570	33013	NM_001127701.1(SERPINA1):c.187C>T (p.Arg63Cys)	SERPINA1
34424986	22089	NM_004562.2(PRKN):c.823C>T (p.Arg275Trp)	PRKN
61750130	22943	NM_000350.2(ABCA4):c.4139C>T (p.Pro1380Leu)	ABCA4
61750200	22937	NM_000350.2(ABCA4):c.634C>T (p.Arg212Cys)	ABCA4
63750451	24281	NM_000535.5(PMS2):c.1882C>T (p.Arg628Ter)	PMS2
72653706	21598	NM_001171.5(ABCC6):c.3421C>T (p.Arg1141Ter)	ABCC6

RS#	ID аллеля	Наименование	Обозначение гена
74503222	108557	NM_000277.3(PAH):c.745C>T (p.Leu249Phe)	PAH
76296470	15620	NM_000277.3(PAH):c.331C>T (p.Arg111Ter)	PAH
80338860	21826	NM_001360.2(DHCR7):c.1054C>T (p.Arg352Trp)	DHCR7
80356682	29578	NM_000228.2(LAMB3):c.1903C>T (p.Arg635Ter)	LAMB3
80356771	19334	NM_001005741.2(GBA):c.1504C>T (p.Arg502Cys)	GBA
80356904	68978	NM_007294.3(BRCA1):c.1621C>T (p.Gln541Ter)	BRCA1
80356932	69850	NM_007294.3(BRCA1):c.4372C>T (p.Gln1458Ter)	BRCA1
80356947	70087	NM_007294.3(BRCA1):c.514C>T (p.Gln172Ter)	BRCA1
80356992	69906	NM_007294.3(BRCA1):c.4612C>T (p.Gln1538Ter)	BRCA1
80357133	70034	NM_007294.3(BRCA1):c.505C>T (p.Gln169Ter)	BRCA1
80357215	68781	NM_007294.3(BRCA1):c.1066C>T (p.Gln356Ter)	BRCA1
104894419	22712	NM_002312.3(LIG4):c.2440C>T (p.Arg814Ter)	LIG4
113871094	44746	NM_000138.4(FBN1):c.4786C>T (p.Arg1596Ter)	FBN1
118203682	58105	NM_000368.4(TSC1):c.2356C>T (p.Arg786Ter)	TSC1
121908177	19611	NM_031885.3(BBS2):c.823C>T (p.Arg275Ter)	BBS2
121908715	16998	NM_000022.2(ADA):c.986C>T (p.Ala329Val)	ADA
121909122	22411	NM_001083962.1(TCF4):c.1153C>T (p.Arg385Ter)	TCF4
121917901	16740	NM_000124.3(ERCC6):c.2203C>T (p.Arg735Ter)	ERCC6
121964964	15158	NM_000071.2(CBS):c.341C>T (p.Ala114Val)	CBS
137852924	18422	NM_147127.4(EVC2):c.1195C>T (p.Arg399Ter)	EVC2
137854466	31478	NM_000138.4(FBN1):c.8326C>T (p.Arg2776Ter)	FBN1
137854467	31479	NM_000138.4(FBN1):c.364C>T (p.Arg122Cys)	FBN1
137854604	24422	NM_000335.4(SCN5A):c.5126C>T (p.Ser1709Leu)	SCN5A
150518260	51200	NM_000232.4(SGCB):c.341C>T (p.Ser114Phe)	SGCB
200432447	133521	NM_007194.4(CHEK2):c.1555C>T (p.Arg519Ter)	CHEK2
201587138	176561	NM_144612.6(LOXHD1):c.4480C>T (p.Arg1494Ter)	LOXHD1
367543286	70502	NM_002609.3(PDGFRB):c.1681C>T (p.Arg561Cys)	PDGFRB
372827156	54183	NM_004572.3(PKP2):c.1237C>T (p.Arg413Ter)	PKP2
374950566	181683	NM_001128425.1(MUTYH):c.884C>T (p.Pro295Leu)	MUTYH
397514558	48266	NM_000138.4(FBN1):c.2920C>T (p.Arg974Cys)	FBN1
397515992	51839	NM_000256.3(MYBPC3):c.2905C>T (p.Gln969Ter)	MYBPC3
397516456	52796	NM_000364.3(TNNT2):c.304C>T (p.Arg102Trp)	TNNT2
587780082	133292	NM_001128425.1(MUTYH):c.1012C>T (p.Gln338Ter)	MUTYH
587782705	152480	NM_000546.5(TP53):c.455C>T (p.Pro152Leu)	TP53
727503974	177432	NM_172107.3(KCNQ2):c.821C>T (p.Thr274Met)	KCNQ2
730881864	180279	NM_002485.4(NBN):c.2140C>T (p.Arg714Ter)	NBN
767215758	188057	NM_002485.4(NBN):c.1030C>T (p.Gln344Ter)	NBN
45517259	27442	NM_000548.3(TSC2):c.2714G>A (p.Arg905Gln)	TSC2
61195471	57234	NM_170707.3(LMNA):c.607G>A (p.Glu203Lys)	LMNA
61753185	18815	NM_000372.4(TYR):c.230G>A (p.Arg77Gln)	TYR
63749869	28021	NM_000540.2(RYR1):c.14582G>A (p.Arg4861His)	RYR1
63749939	32145	NM_000249.3(MLH1):c.200G>A (p.Gly67Glu)	MLH1

RS#	ID аллеля	Наименование	Обозначение гена
63750119	150580	NM_000179.2(MSH6):c.3725G>A (p.Arg1242His)	MSH6
72554308	26053	NM_000531.5(OTC):c.119G>A (p.Arg40His)	OTC
79891110	32671	NM_000719.6(CACNA1C):c.1216G>A (p.Gly406Arg)	CACNA1C
80338707	22758	NM_000303.2(PMM2):c.691G>A (p.Val231Met)	PMM2
80338802	32652	NM_000070.2(CAPN3):c.2306G>A (p.Arg769Gln)	CAPN3
80356700	32571	NM_000083.2(CLCN1):c.689G>A (p.Gly230Glu)	CLCN1
80359803	67339	NM_000059.3(BRCA2):c.8754G>A (p.Glu2918=)	BRCA2
81002809	67078	NM_000059.3(BRCA2):c.7805+1G>A	BRCA2
104886142	35796	NM_000495.4(COL4A5):c.1871G>A (p.Gly624Asp)	COL4A5
104894423	17048	NM_000231.2(SGCG):c.787G>A (p.Glu263Lys)	SGCG
104894525	22747	NM_000303.2(PMM2):c.385G>A (p.Val129Met)	PMM2
113994049	20984	NM_003907.3(EIF2B5):c.338G>A (p.Arg113His)	EIF2B5
121434372	17127	NM_000159.3(GCDH):c.1198G>A (p.Val400Met)	GCDH
121908099	19299	NM_000784.3(CYP27A1):c.1214G>A (p.Arg405Gln)	CYP27A1
121908192	23730	NM_005262.2(GFER):c.581G>A (p.Arg194His)	GFER
121908753	22237	NM_000492.3(CFTR):c.1055G>A (p.Arg352Gln)	CFTR
121918013	28716	NM_000478.4(ALPL):c.346G>A (p.Ala116Thr)	ALPL
139729994	68418	NM_000492.3(CFTR):c.3468G>A (p.Leu1156=)	CFTR
142637046	98272	NM_000048.3(ASL):c.446+1G>A	ASL
142761835	177782	NM_002225.3(IVD):c.367G>A (p.Gly123Arg)	IVD
146015592	46845	NM_000060.4(BTD):c.470G>A (p.Arg157His)	BTD
150877497	226470	NM_003494.3(DYSF):c.3113G>A (p.Arg1038Gln)	DYSF
199472815	67686	NM_000218.2(KCNQ1):c.1781G>A (p.Arg594Gln)	KCNQ1
199474738	79199	NM_001042492.2(NF1):c.1885G>A (p.Gly629Arg)	NF1
199476112	24747	NC_012920.1:m.11778G>A	MT-ND4
199476317	40544	NM_001018005.1(TPM1):c.688G>A (p.Asp230Asn)	TPM1
201540674	51186	RTEL1:c.2402G>A (p.Arg801His)	RTEL1
267606640	16147	NM_000642.2(AGL):c.3980G>A (p.Trp1327Ter)	AGL
386834233	76679	NM_183050.3(BCKDHB):c.832G>A (p.Gly278Ser)	BCKDHB
397515355	19301	NM_000784.3(CYP27A1):c.1263+1G>A	CYP27A1
397515404	48194	NM_020822.2(KCNT1):c.1421G>A (p.Arg474His)	KCNT1
398123787	100221	NM_003494.3(DYSF):c.4253G>A (p.Gly1418Asp)	DYSF
398124641	44139	NM_024531.4(SLC52A2):c.916G>A (p.Gly306Arg)	SLC52A2
587776783	132342	NM_000321.2(RB1):c.1215+1G>A	RB1
587776889	39757	NM_015506.2(MMACHC):c.609G>A (p.Trp203Ter)	MMACHC
587777721	165903	NM_014191.3(SCN8A):c.4850G>A (p.Arg1617Gln)	SCN8A
587779818	132798	NM_000051.3(ATM):c.170G>A (p.Trp57Ter)	ATM
587780537	136457	NM_004360.4(CDH1):c.715G>A (p.Gly239Arg)	CDH1
587783050	166264	NM_004360.5(CDH1):c.1137G>A (p.Thr379=)	CDH1
751995154	200340	NM_000018.4(ACADVL):c.1376G>A (p.Arg459Gln)	ACADVL
781404312	186796	NM_000051.3(ATM):c.3G>A (p.Met1Ile)	ATM
786202112	184694	NM_001042492.2(NF1):c.5609G>A (p.Arg1870Gln)	NF1

RS#	ID аллеля	Наименование	Обозначение гена
794727152	191718	NM_021007.2(SCN2A):c.2558G>A (p.Arg853Gln)	SCN2A
796051858	18086	NM_000051.3(ATM):c.496+5G>A	ATM
796052505	201880	NM_000816.3(GABRG2):c.316G>A (p.Ala106Thr)	GABRG2
863223408	210238	NM_000020.2(ACVRL1):c.1451G>A (p.Arg484Gln)	ACVRL1
863225082	188114	NM_006245.3(PPP2R5D):c.592G>A (p.Glu198Lys)	PPP2R5D
875989911	228151	NM_000527.4(LDLR):c.938G>A (p.Cys313Tyr)	LDLR
5030852	15638	NM_000277.3(PAH):c.842+1G>A	PAH
5030859	15651	NM_000277.3(PAH):c.1223G>A (p.Arg408Gln)	PAH
28930068	32662	NM_000069.2(CACNA1S):c.3716G>A (p.Arg1239His)	CACNA1S
56264519	55267	NM_024022.2(TMPRSS3):c.1276G>A (p.Ala426Thr)	TMPRSS3
61750641	105317	NM_000350.2(ABCA4):c.6089G>A (p.Arg2030Gln)	ABCA4
61751276	104715	NM_000329.2(RPE65):c.11+5G>A	RPE65
62507336	108472	NM_000277.3(PAH):c.561G>A (p.Trp187Ter)	PAH
62508613	108291	NM_000277.2(PAH):c.1199+17G>A	PAH
72645357	32351	NM_000088.3(COL1A1):c.994G>A (p.Gly332Arg)	COL1A1
80338777	32664	NM_000069.2(CACNA1S):c.1583G>A (p.Arg528His)	CACNA1S
80356908	68776	NM_007294.3(BRCA1):c.1058G>A (p.Trp353Ter)	BRCA1
80357093	69031	NM_007294.3(BRCA1):c.182G>A (p.Cys61Tyr)	BRCA1
80357219	70211	NM_007294.3(BRCA1):c.5345G>A (p.Trp1782Ter)	BRCA1
104886460	99352	NM_001005741.2(GBA):c.115+1G>A	GBA
104894129	27501	NM_003289.3(TPM2):c.349G>A (p.Glu117Lys)	TPM2
104894401	32056	NM_004004.5(GJB2):c.428G>A (p.Arg143Gln)	GJB2
104895085	17592	NM_000243.2(MEFV):c.1958G>A (p.Arg653His)	MEFV
111033299	53902	NM_004004.5(GJB2):c.283G>A (p.Val95Met)	GJB2
113994139	33347	NM_139276.2(STAT3):c.1909G>A (p.Val637Met)	STAT3
120074135	18010	NM_000271.4(NPC1):c.2848G>A (p.Val950Met)	NPC1
121909334	23512	NM_007126.4(VCP):c.572G>A (p.Arg191Gln)	VCP
121918491	28307	NM_000141.4(FGFR2):c.1032G>A (p.Ala344=)	FGFR2
137852314	25406	NM_000402.4(G6PD):c.577G>A (p.Gly193Ser)	G6PD
137852327	25425	NM_000402.4(G6PD):c.961G>A (p.Val321Met)	G6PD
137853285	166061	NM_000053.3(ATP7B):c.2128G>A (p.Gly710Ser)	ATP7B
138213197	133488	NM_006361.5(HOXB13):c.251G>A (p.Gly84Glu)	HOXB13
148311934	44907	NM_000162.5(GCK):c.676G>A (p.Val226Met)	GCK
199473684	25807	NM_000169.2(GLA):c.639+919G>A	GLA
200482683	131950	NM_014625.3(NPHS2):c.868G>A (p.Val290Met)	NPHS2
371418985	232124	NM_007194.4(CHEK2):c.1232G>A (p.Trp411Ter)	CHEK2
387907281	45778	NM_152296.4(ATP1A3):c.2443G>A (p.Glu815Lys)	ATP1A3
397509284	70248	NM_007294.3(BRCA1):c.5445G>A (p.Trp1815Ter)	BRCA1
397514495	152034	NM_000546.5(TP53):c.542G>A (p.Arg181His)	TP53
397514581	48359	NM_172107.3(KCNQ2):c.638G>A (p.Arg213Gln)	KCNQ2
397516101	52008	NM_000257.4(MYH7):c.1358G>A (p.Arg453His)	MYH7
397516264	52270	NM_000257.3(MYH7):c.715G>A (p.Asp239Asn)	MYH7

RS#	ID аллеля	Наименование	Обозначение гена
398122822	48057	NM_001111.5(ADAR):c.3019G>A (p.Gly1007Arg)	ADAR
587777446	141325	NM_022168.4(IFIH1):c.2336G>A (p.Arg779His)	IFIH1
587782962	165566	NM_000257.4(MYH7):c.3158G>A (p.Arg1053Gln)	MYH7
606231435	170985	NM_152296.4(ATP1A3):c.2267G>A (p.Arg756His)	ATP1A3
727504247	172354	NM_001001430.2(TNNT2):c.860G>A (p.Trp287Ter)	TNNT2
730881833	179933	NM_001128425.1(MUTYH):c.857G>A (p.Gly286Glu)	MUTYH
762307622	232266	NM_001128425.1(MUTYH):c.467G>A (p.Trp156Ter)	MUTYH
777759523	17038	NM_199242.2(UNC13D):c.1389+1G>A	UNC13D
794728625	197538	NM_130799.2(MEN1):c.784-9G>A	MEN1
1060499814	389282	NM_024675.3(PALB2):c.108+1G>A	PALB2
25403	51465	NM_000138.4(FBN1):c.184C>T (p.Arg62Cys)	FBN1
28931591	32539	NM_000744.6(CHRNA4):c.851C>T (p.Ser284Leu)	CHRNA4
28942108	18015	NM_000271.4(NPC1):c.2932C>T (p.Arg978Cys)	NPC1
61750152	105192	NM_000350.2(ABCA4):c.4577C>T (p.Thr1526Met)	ABCA4
61750654	105349	NM_000350.2(ABCA4):c.6445C>T (p.Arg2149Ter)	ABCA4
61751404	105219	NM_000350.2(ABCA4):c.4918C>T (p.Arg1640Trp)	ABCA4
61751408	22921	NM_000350.2(ABCA4):c.6079C>T (p.Leu2027Phe)	ABCA4
63751466	24276	NM_000535.5(PMS2):c.2404C>T (p.Arg802Ter)	PMS2
72552255	44374	NM_000053.3(ATP7B):c.2930C>T (p.Thr977Met)	ATP7B
74315369	27822	NM_003000.2(SDHB):c.79C>T (p.Arg27Ter)	SDHB
80338680	16726	NM_000528.3(MAN2B1):c.2248C>T (p.Arg750Trp)	MAN2B1
80356952	68980	NM_007294.3(BRCA1):c.1630C>T (p.Gln544Ter)	BRCA1
80357011	69802	NM_007294.3(BRCA1):c.4186C>T (p.Gln1396Ter)	BRCA1
80357296	69580	NM_007294.3(BRCA1):c.3544C>T (p.Gln1182Ter)	BRCA1
80357367	70140	NM_007294.3(BRCA1):c.5239C>T (p.Gln1747Ter)	BRCA1
80357377	69340	NM_007294.3(BRCA1):c.2761C>T (p.Gln921Ter)	BRCA1
80357471	69016	NM_007294.3(BRCA1):c.178C>T (p.Gln60Ter)	BRCA1
80357497	69389	NM_007294.3(BRCA1):c.2923C>T (p.Gln975Ter)	BRCA1
104893950	18137	NM_005670.3(EPM2A):c.721C>T (p.Arg241Ter)	EPM2A
104894787	26252	NM_004006.2(DMD):c.10108C>T (p.Arg3370Ter)	DMD
111231312	51536	NM_000138.4(FBN1):c.4615C>T (p.Arg1539Ter)	FBN1
112645512	178700	NM_000138.4(FBN1):c.1285C>T (p.Arg429Ter)	FBN1
113001196	51577	NM_000138.4(FBN1):c.6658C>T (p.Arg2220Ter)	FBN1
113249837	51552	NM_000138.4(FBN1):c.5368C>T (p.Arg1790Ter)	FBN1
113812345	51455	NM_000138.4(FBN1):c.1546C>T (p.Arg516Ter)	FBN1
116100695	16552	NM_000298.5(PKLR):c.1456C>T (p.Arg486Trp)	PKLR
118203631	58047	NM_000368.4(TSC1):c.2074C>T (p.Arg692Ter)	TSC1
118203963	16148	NM_025137.3(SPG11):c.6100C>T (p.Arg2034Ter)	SPG11
118204437	15739	NM_000512.4(GALNS):c.1156C>T (p.Arg386Cys)	GALNS
121434526	33315	NM_001613.3(ACTA2):c.445C>T (p.Arg149Cys)	ACTA2
121908547	20943	NM_000334.4(SCN4A):c.3938C>T (p.Thr1313Met)	SCN4A
121912504	29459	NM_000238.3(KCNH2):c.1682C>T (p.Ala561Val)	KCNH2

RS#	ID аллеля	Наименование	Обозначение гена
121913120	31271	NM_000143.3(FH):c.301C>T (p.Arg101Ter)	FH
121913122	31274	NM_000143.3(FH):c.1027C>T (p.Arg343Ter)	FH
121917783	27083	NM_000136.2(FANCC):c.553C>T (p.Arg185Ter)	FANCC
121918775	79496	NM_006920.4(SCN1A):c.2803C>T (p.Arg935Cys)	SCN1A
121964972	15170	NM_000071.2(CBS):c.1058C>T (p.Thr353Met)	CBS
128627256	26327	NM_004006.2(DMD):c.8713C>T (p.Arg2905Ter)	DMD
137854613	24413	NM_198056.2(SCN5A):c.4867C>T (p.Arg1623Ter)	SCN5A
137886232	39244	NM_002878.3(RAD51D):c.757C>T (p.Arg253Ter)	RAD51D
138996609	181608	NM_003000.2(SDHB):c.688C>T (p.Arg230Cys)	SDHB
144500145	202960	NM_002693.2(POLG):c.2554C>T (p.Arg852Cys)	POLG
180177111	132156	NM_024675.3(PALB2):c.2323C>T (p.Gln775Ter)	PALB2
185492864	99918	NM_001918.3(DBT):c.901C>T (p.Arg301Cys)	DBT
193922185	44706	NM_000138.4(FBN1):c.1948C>T (p.Arg650Cys)	FBN1
199472944	38732	NM_000238.3(KCNH2):c.1841C>T (p.Ala614Val)	KCNH2
199472990	78275	NM_000238.3(KCNH2):c.2254C>T (p.Arg752Trp)	KCNH2
199473161	78626	NM_198056.2(SCN5A):c.2440C>T (p.Arg814Trp)	SCN5A
199473524	78188	NM_000238.3(KCNH2):c.1838C>T (p.Thr613Met)	KCNH2
273898674	69115	NM_007294.3(BRCA1):c.2059C>T (p.Gln687Ter)	BRCA1
368796923	151096	NM_032043.2(BRIP1):c.1240C>T (p.Gln414Ter)	BRIP1
376128990	215031	NM_052845.3(MMAB):c.571C>T (p.Arg191Trp)	MMAB
397509283	70244	NM_007294.3(BRCA1):c.5431C>T (p.Gln1811Ter)	BRCA1
397515812	51535	NM_000138.4(FBN1):c.4567C>T (p.Arg1523Ter)	FBN1
397516005	51860	NM_000256.3(MYBPC3):c.3181C>T (p.Gln1061Ter)	MYBPC3
397516042	51914	NM_000256.3(MYBPC3):c.3811C>T (p.Arg1271Ter)	MYBPC3
397516127	52044	NM_000257.3(MYH7):c.1987C>T (p.Arg663Cys)	MYH7
397516201	52162	NM_000257.4(MYH7):c.4130C>T (p.Thr1377Met)	MYH7
397516435	52758	NM_000546.5(TP53):c.586C>T (p.Arg196Ter)	TP53
397517689	56466	NM_001267550.2(TTN):c.71602C>T (p.Arg23868Ter)	TTN
398123585	99539	NM_001165963.1(SCN1A):c.1837C>T (p.Arg613Ter)	SCN1A
549794342	360820	NM_001271208.1(NEB):c.24094C>T (p.Arg8032Ter)	NEB
574660186	178478	NM_001267550.2(TTN):c.67495C>T (p.Arg22499Ter)	TTN
575822089	227149	NM_001163435.2(TBCK):c.376C>T (p.Arg126Ter)	TBCK
587778618	138806	NM_000535.7(PMS2):c.1687C>T (p.Arg563Ter)	PMS2
587779343	96837	NM_000535.5(PMS2):c.697C>T (p.Gln233Ter)	PMS2
587780088	133302	NM_001128425.1(MUTYH):c.55C>T (p.Arg19Ter)	MUTYH
587781269	150486	NM_007194.4(CHEK2):c.283C>T (p.Arg95Ter)	CHEK2
587781756	151166	NM_002878.3(RAD51D):c.451C>T (p.Gln151Ter)	RAD51D
672601370	171771	NM_001244008.1(KIF1A):c.946C>T (p.Arg316Trp)	KIF1A
727505006	176130	NM_000138.4(FBN1):c.3373C>T (p.Arg1125Ter)	FBN1
794728165	197808	NM_000138.4(FBN1):c.1090C>T (p.Arg364Ter)	FBN1
794728228	197690	NM_000138.4(FBN1):c.4621C>T (p.Arg1541Ter)	FBN1
794728283	197585	NM_000138.4(FBN1):c.8038C>T (p.Arg2680Cys)	FBN1

RS#	ID аллеля	Наименование	Обозначение гена
879255678	247653	NM_144997.5(FLCN):c.1429C>T (p.Arg477Ter)	FLCN
886041116	263863	NM_015339.4(ADNP):c.2188C>T (p.Arg730Ter)	ADNP
1553547838	512805	NM_001172509.1(SATB2):c.1375C>T (p.Arg459Ter)	SATB2
45507199	59122	NM_000548.3(TSC2):c.5228G>A (p.Arg1743Gln)	TSC2
60458016	29564	NM_170707.3(LMNA):c.1072G>A (p.Glu358Lys)	LMNA
61672878	29534	NM_170707.3(LMNA):c.1130G>A (p.Arg377His)	LMNA
61750173	24396	NM_000180.3(GUCY2D):c.2513G>A (p.Arg838His)	GUCY2D
61753180	18833	NM_000372.4(TYR):c.140G>A (p.Gly47Asp)	TYR
61754375	18835	NM_000372.4(TYR):c.896G>A (p.Arg299His)	TYR
62636275	20778	NM_201253.2(CRB1):c.3307G>A (p.Gly1103Arg)	CRB1
63750453	95615	NM_000249.3(MLH1):c.304G>A (p.Glu102Lys)	MLH1
63750604	95363	NM_000249.3(MLH1):c.1790G>A (p.Trp597Ter)	MLH1
63751632	95404	NM_000249.3(MLH1):c.1896G>A (p.Glu632=)	MLH1
74315205	19565	NM_006005.3(WFS1):c.2590G>A (p.Glu864Lys)	WFS1
74503330	22256	NM_000492.3(CFTR):c.3752G>A (p.Ser1251Asn)	CFTR
80282562	57854	NM_000492.3(CFTR):c.532G>A (p.Gly178Arg)	CFTR
80356702	32581	NM_000083.2(CLCN1):c.950G>A (p.Arg317Gln)	CLCN1
80358543	131539	NM_000059.3(BRCA2):c.2978G>A (p.Trp993Ter)	BRCA2
80358810	46556	NM_000059.3(BRCA2):c.582G>A (p.Trp194Ter)	BRCA2
80358997	67062	NM_000059.3(BRCA2):c.7721G>A (p.Trp2574Ter)	BRCA2
80359205	67482	NM_000059.3(BRCA2):c.9317G>A (p.Trp3106Ter)	BRCA2
81002873	67120	NM_000059.3(BRCA2):c.7976+1G>A	BRCA2
104894317	18840	NM_000372.4(TYR):c.1336G>A (p.Gly446Ser)	TYR
104894590	16599	NM_000263.3(NAGLU):c.2021G>A (p.Arg674His)	NAGLU
111033270	19955	NM_022124.5(CDH23):c.5237G>A (p.Arg1746Gln)	CDH23
111436401	226974	NM_000540.2(RYR1):c.10347+1G>A	RYR1
112406105	200333	NM_000018.4(ACADVL):c.1097G>A (p.Arg366His)	ACADVL
113560320	15440	NM_017841.2(SDHAF2):c.232G>A (p.Gly78Arg)	SDHAF2
113690956	16661	NM_000018.2(ACADVL):c.1182+1G>A	ACADVL
113994171	33871	NM_000018.3(ACADVL):c.1679-6G>A	ACADVL
113994207	19490	NM_004937.2(CTNS):c.589G>A (p.Gly197Arg)	CTNS
114925667	260377	NM_024818.4(UBA5):c.1111G>A (p.Ala371Thr)	UBA5
118192122	76888	NM_000540.2(RYR1):c.7361G>A (p.Arg2454His)	RYR1
118192176	28015	NM_000540.2(RYR1):c.6502G>A (p.Val2168Met)	RYR1
118203982	16396	NM_001080.3(ALDH5A1):c.612G>A (p.Trp204Ter)	ALDH5A1
119462987	18289	NM_007171.3(POMT1):c.2005G>A (p.Ala669Thr)	POMT1
120074190	18179	NM_000218.2(KCNQ1):c.1766G>A (p.Gly589Asp)	KCNQ1
121434544	32653	NM_000070.2(CAPN3):c.1715G>A (p.Arg572Gln)	CAPN3
121434548	32661	NM_000070.2(CAPN3):c.1469G>A (p.Arg490Gln)	CAPN3;POMT1
121908153	19416	NM_001243133.1(NLRP3):c.907G>A (p.Asp303Asn)	NLRP3
121908185	19531	NM_020451.2(SELENON):c.1397G>A (p.Arg466Gln)	SELENON
121908419	20395	NM_014384.2(ACAD8):c.1129G>A (p.Gly377Ser)	ACAD8

RS#	ID аллеля	Наименование	Обозначение гена
121908759	44497	NM_000492.3(CFTR):c.1865G>A (p.Gly622Asp)	CFTR
121908889	21460	NM_003060.3(SLC22A5):c.506G>A (p.Arg169Gln)	SLC22A5
121909013	22181	NM_000492.3(CFTR):c.1651G>A (p.Gly551Ser)	CFTR
121909019	22197	NM_000492.3(CFTR):c.3197G>A (p.Arg1066His)	CFTR
121909092	22321	NM_001005360.2(DNM2):c.1102G>A (p.Glu368Lys)	DNM2
121918009	28711	NM_000478.5(ALPL):c.1001G>A (p.Gly334Asp)	ALPL
121918592	28008	NM_000540.2(RYR1):c.1021G>A (p.Gly341Arg)	RYR1
137852871	17416	NM_000709.3(BCKDHA):c.868G>A (p.Gly290Arg)	BCKDHA
141158996	22214	NM_000492.3(CFTR):c.2490+1G>A	CFTR
141554661	208401	NM_004287.4(GOSR2):c.336+1G>A	GOSR2
148032587	194820	NM_000303.2(PMM2):c.442G>A (p.Asp148Asn)	PMM2
193922503	44492	NM_000492.3(CFTR):c.1585-8G>A	CFTR
199472687	77968	NM_000218.2(KCNQ1):c.421G>A (p.Val141Met)	KCNQ1
201016593	245339	NM_000527.4(LDLR):c.11G>A (p.Trp4Ter)	LDLR
267606997	21861	NM_058216.2(RAD51C):c.773G>A (p.Arg258His)	RAD51C
267607914	96367	NM_000251.2(MSH2):c.212-1G>A	MSH2
369560930	98197	NM_000018.4(ACADVL):c.520G>A (p.Val174Met)	ACADVL
370523609	227889	NM_000016.5(ACADM):c.600-18G>A	ACADM
370950728	186993	NM_000152.3(GAA):c.655G>A (p.Gly219Arg)	GAA
374143224	187013	NM_000152.3(GAA):c.1979G>A (p.Arg660His)	GAA
397508045	67476	NM_000059.3(BRCA2):c.92G>A (p.Trp31Ter)	BRCA2
397508200	67910	NM_000492.3(CFTR):c.1393-1G>A	CFTR
397509418	75098	NM_021942.5(TRAPPC11):c.1287+5G>A	TRAPPC11
397515330	76388	NM_001098512.2(PRKG1):c.530G>A (p.Arg177Gln)	PRKG1
398122711	97208	NM_000059.3(BRCA2):c.8633-1G>A	BRCA2
398123139	98311	NM_000060.4(BTD):c.626G>A (p.Arg209His)	BTD
398123763	100162	NM_003494.3(DYSF):c.1053+1G>A	DYSF
587777057	77012	NM_020988.2(GNAO1):c.607G>A (p.Gly203Arg)	GNAO1
587777570	150453	NM_004522.2(KIF5C):c.709G>A (p.Glu237Lys)	KIF5C
587778777	76741	NM_000784.3(CYP27A1):c.1184+1G>A	CYP27A1
587779110	96248	NM_000251.2(MSH2):c.1760-1G>A	MSH2
587780639	139490	NM_000051.3(ATM):c.7788G>A (p.Glu2596=)	ATM
587781894	151348	NM_000051.3(ATM):c.9023G>A (p.Arg3008His)	ATM
587782719	152505	NM_000051.3(ATM):c.8122G>A (p.Asp2708Asn)	ATM
727503030	176785	NM_001278939.1(ELN):c.1150+1G>A	ELN
730881581	180665	NM_000059.3(BRCA2):c.8174G>A (p.Trp2725Ter)	BRCA2
730882035	180121	NM_000551.3(VHL):c.482G>A (p.Arg161Gln)	VHL
750663117	234071	NM_000051.3(ATM):c.3078-1G>A	ATM
756039188	243266	NM_000527.4(LDLR):c.12G>A (p.Trp4Ter)	LDLR
796053216	202741	NM_014191.3(SCN8A):c.4423G>A (p.Gly1475Arg)	SCN8A
876661242	231905	NM_000059.3(BRCA2):c.9381G>A (p.Trp3127Ter)	BRCA2
879254600	245669	NM_000527.4(LDLR):c.626G>A (p.Cys209Tyr)	LDLR

RS#	ID аллеля	Наименование	Обозначение гена
1057519632	362622	NM_003718.4(CDK13):c.2149G>A (p.Gly717Arg)	CDK13
10250779	15457	NM_000290.3(PGAM2):c.233G>A (p.Trp78Ter)	PGAM2
28928905	29469	NM_000238.3(KCNH2):c.1468G>A (p.Ala490Thr)	KCNH2
28931593	32066	NM_004004.5(GJB2):c.224G>A (p.Arg75Gln)	GJB2
28937318	24429	NM_198056.2(SCN5A):c.1100G>A (p.Arg367His)	SCN5A
61749397	15329	NM_000552.4(VWF):c.3946G>A (p.Val1316Met)	VWF
61751403	105220	NM_000350.2(ABCA4):c.4919G>A (p.Arg1640Gln)	ABCA4
62514907	15633	NM_000277.3(PAH):c.442-1G>A	PAH
62514956	98659	NM_000277.3(PAH):c.912+1G>A	PAH
62516146	108608	NM_000277.1(PAH):c.842+5G>A	PAH
62642939	98657	NM_000277.2(PAH):c.890G>A (p.Arg297His)	PAH
62644503	108560	NM_000277.3(PAH):c.755G>A (p.Arg252Gln)	PAH
63749856	21618	NM_001171.5(ABCC6):c.3904G>A (p.Gly1302Arg)	ABCC6
63750783	30442	NM_000518.5(HBB):c.47G>A (p.Trp16Ter)	HBB
66555264	414003	NM_000088.3(COL1A1):c.1821+1G>A	COL1A1
72645321	414022	NM_000088.3(COL1A1):c.769G>A (p.Gly257Arg)	COL1A1
74315368	27820	NM_003000.2(SDHB):c.725G>A (p.Arg242His)	SDHB
74315471	18113	NM_000487.5(ARSA):c.739G>A (p.Gly247Arg)	ARSA
78973108	19367	NM_001005741.2(GBA):c.887G>A (p.Arg296Gln)	GBA
80338735	33917	NM_000156.5(GAMT):c.327G>A (p.Lys109=)	GAMT
80338857	34128	NM_001360.2(DHCR7):c.725G>A (p.Arg242His)	DHCR7
80338864	21831	NM_001360.2(DHCR7):c.1342G>A (p.Glu448Lys)	DHCR7
80338944	32040	NM_004004.5(GJB2):c.231G>A (p.Trp77Ter)	GJB2
80356914	70276	NM_007294.3(BRCA1):c.5511G>A (p.Trp1837Ter)	BRCA1
80357212	70255	NM_007294.3(BRCA1):c.5467G>A (p.Ala1823Thr)	BRCA1
80357307	70275	NM_007294.3(BRCA1):c.5510G>A (p.Trp1837Ter)	BRCA1
80358252	18013	NM_000271.4(NPC1):c.530G>A (p.Cys177Tyr)	NPC1
104894103	19470	NM_175073.2(APTX):c.837G>A (p.Trp279Ter)	APTX
104894415	20583	NM_006783.4(GJB6):c.31G>A (p.Gly11Arg)	GJB6
104894519	21096	NM_004862.3(LITAF):c.334G>A (p.Gly112Ser)	LITAF
104894727	27461	NM_000363.4(TNNI3):c.586G>A (p.Asp196Asn)	TNNI3
104894828	25754	NM_000169.2(GLA):c.902G>A (p.Arg301Gln)	GLA
111683277	175150	NM_000256.3(MYBPC3):c.3190+1G>A	MYBPC3
111984349	258823	NM_000138.4(FBN1):c.7828G>A (p.Glu2610Lys)	FBN1
113403872	16550	NM_000298.5(PKLR):c.1529G>A (p.Arg510Gln)	PKLR
121434249	18383	NM_000348.3(SRD5A2):c.682G>A (p.Ala228Thr)	SRD5A2
121908216	23534	NM_001127221.1(CACNA1A):c.4982G>A (p.Arg1661His)	CACNA1A
121908551	20948	NM_000334.4(SCN4A):c.3877G>A (p.Val1293Ile)	SCN4A
121908552	20949	NM_000334.4(SCN4A):c.1333G>A (p.Val445Met)	SCN4A
121908557	20958	NM_000334.4(SCN4A):c.2024G>A (p.Arg675Gln)	SCN4A
121908716	16996	NM_000022.2(ADA):c.632G>A (p.Arg211His)	ADA
121908723	17007	NM_000022.3(ADA):c.646G>A (p.Gly216Arg)	ADA

RS#	ID аллеля	Наименование	Обозначение гена
121909768	21834	NM_001360.2(DHCR7):c.1055G>A (p.Arg352Gln)	DHCR7
121913039	31702	NM_001953.4(TYMP):c.622G>A (p.Val208Met)	TYMP
137853050	22116	NM_006009.3(TUBA1A):c.1265G>A (p.Arg422His)	TUBA1A
137853283	166064	NM_000053.3(ATP7B):c.2336G>A (p.Trp779Ter)	ATP7B
137854612	24434	NM_198056.2(SCN5A):c.4222G>A (p.Gly1408Arg)	SCN5A
139751448	187031	NM_000271.4(NPC1):c.1211G>A (p.Arg404Gln)	NPC1
150038620	187049	NM_004646.3(NPHS1):c.2335-1G>A	NPHS1
180177122	132185	NM_024675.3(PALB2):c.2718G>A (p.Trp906Ter)	PALB2
181087667	40103	NM_007055.3(POLR3A):c.2617-1G>A	POLR3A
193922110	44393	NM_000053.3(ATP7B):c.4058G>A (p.Trp1353Ter)	ATP7B
199473565	78528	NM_198056.2(SCN5A):c.1066G>A (p.Asp356Asn)	SCN5A
199474703	40437	NM_000258.2(MYL3):c.281G>A (p.Arg94His)	MYL3
199971687	216058	NM_052845.3(MMAB):c.291-1G>A	MMAB
201188361	40345	NM_014714.3(IFT140):c.634G>A (p.Gly212Arg)	IFT140
202160208	75126	NM_013334.3(GMPPB):c.860G>A (p.Arg287Gln)	GMPPB
281875334	38553	NM_001101.3(ACTB):c.587G>A (p.Arg196His)	ACTB
386134249	45185	NM_000244.3(MEN1):c.1277G>A (p.Cys426Tyr)	MEN1
387906623	38652	NM_000138.4(FBN1):c.5284G>A (p.Gly1762Ser)	FBN1
387906905	39430	NM_021625.4(TRPV4):c.947G>A (p.Arg316His)	TRPV4
397507479	48850	NM_004333.5(BRAF):c.1595G>A (p.Cys532Tyr)	BRAF
397514494	48018	NM_021625.4(TRPV4):c.557G>A (p.Arg186Gln)	TRPV4
397515854	51599	NM_000138.4(FBN1):c.7606G>A (p.Gly2536Arg)	FBN1
397515982	51820	NM_000256.3(MYBPC3):c.2670G>A (p.Trp890Ter)	MYBPC3
397516031	51898	NM_000256.3(MYBPC3):c.3627+1G>A	MYBPC3
397516471	52818	NM_001001430.2(TNNT2):c.518G>A (p.Arg173Gln)	TNNT2
398122853	38917	NM_004006.2(DMD):c.9G>A (p.Trp3Ter)	DMD
483352809	65656	NM_006087.3(TUBB4A):c.745G>A (p.Asp249Asn)	TUBB4A
515726205	40114	NM_001031726.3(C19orf12):c.205G>A (p.Gly69Arg)	C19orf12
564069299	200114	NM_000255.3(MMUT):c.1106G>A (p.Arg369His)	MMUT
574673404	182906	NM_002485.4(NBN):c.37+1G>A	NBN
587780345	134590	NM_000162.3(GCK):c.544G>A (p.Val182Met)	GCK
606231324	136674	NM_000257.3(MYH7):c.1573G>A (p.Glu525Lys)	MYH7
727504382	49283	NM_030662.3(MAP2K2):c.619G>A (p.Glu207Lys)	MAP2K2
730880850	29166	NM_000257.3(MYH7):c.732+1G>A	MYH7
730882175	181517	NM_002238.3(KCNH1):c.1405G>A (p.Gly469Arg)	KCNH1
751604696	425943	NM_001360.2(DHCR7):c.1337G>A (p.Arg446Gln)	DHCR7
753288303	216044	NM_000255.3(MMUT):c.1280G>A (p.Gly427Asp)	MMUT
767399782	213656	NM_006087.3(TUBB4A):c.763G>A (p.Val255Ile)	TUBB4A
794728208	197723	NM_000138.4(FBN1):c.3712G>A (p.Asp1238Asn)	FBN1
796756333	410338	NM_024422.4(DSC2):c.943-1G>A	DSC2
797044872	205316	NM_004977.2(KCNC3):c.1268G>A (p.Arg423His)	KCNC3
797045586	207083	NM_032682.5(FOXP1):c.1541G>A (p.Arg514His)	FOXP1

RS#	ID аллеля	Наименование	Обозначение гена
863223403	209408	NM_002140.4(HNRNPK):c.257G>A (p.Arg86His)	HNRNPK
876658367	232176	NM_003000.2(SDHB):c.587G>A (p.Cys196Tyr)	SDHB
1057517585	358911	NM_024675.3(PALB2):c.3G>A (p.Met1Ile)	PALB2
1555582065	431537	NM_014233.3(UBTF):c.628G>A (p.Glu210Lys)	UBTF
140630	197685	NM_000138.4(FBN1):c.4930C>T (p.Arg1644Ter)	FBN1
28940869	19031	NM_017739.3(POMGNT1):c.1324C>T (p.Arg442Cys)	POMGNT1
34451549	30497	NM_000518.5(HBB):c.316-197C>T	HBB
41556519	31832	NM_000400.3(ERCC2):c.2047C>T (p.Arg683Trp)	ERCC2
45611033	175462	NM_000257.4(MYH7):c.3133C>T (p.Arg1045Cys)	MYH7
55832599	151478	NM_000546.5(TP53):c.799C>T (p.Arg267Trp)	TP53
59616921	18036	NM_000226.3(KRT9):c.487C>T (p.Arg163Trp)	KRT9
60399023	29651	NM_000526.4(KRT14):c.373C>T (p.Arg125Cys)	KRT14
61749409	104973	NM_000350.2(ABCA4):c.1804C>T (p.Arg602Trp)	ABCA4
61749423	105003	NM_000350.2(ABCA4):c.2041C>T (p.Arg681Ter)	ABCA4
61750645	105327	NM_000350.2(ABCA4):c.6229C>T (p.Arg2077Trp)	ABCA4
61751383	22946	NM_000350.2(ABCA4):c.6088C>T (p.Arg2030Ter)	ABCA4
61752871	28154	NM_000329.2(RPE65):c.271C>T (p.Arg91Trp)	RPE65
61757582	21827	NM_001360.2(DHCR7):c.1210C>T (p.Arg404Cys)	DHCR7
61816761	31358	NM_002016.1(FLG):c.1501C>T (p.Arg501Ter)	FLG
62507344	15662	NM_000277.2(PAH):c.1066-3C>T	PAH
72559722	186816	NM_001287174.1(ABCC8):c.2509C>T (p.Arg837Ter)	ABCC8
72646846	56340	NM_001256850.1(TTN):c.56953C>T (p.Arg18985Ter)	TTN
72648250	225057	NM_001256850.1(TTN):c.88243C>T (p.Arg29415Ter)	TTN
72650700	39295	NM_001171.5(ABCC6):c.1552C>T (p.Arg518Ter)	ABCC6
72651642	271557	NM_000088.3(COL1A1):c.2089C>T (p.Arg697Ter)	COL1A1
72653170	32386	NM_000088.3(COL1A1):c.3040C>T (p.Arg1014Cys)	COL1A1
74315348	20408	NM_014625.3(NPHS2):c.871C>T (p.Arg291Trp)	NPHS2
74315391	22425	NM_172107.3(KCNQ2):c.619C>T (p.Arg207Trp)	KCNQ2
74315442	23435	NM_000100.3(CSTB):c.202C>T (p.Arg68Ter)	CSTB
74315472	18114	NM_000487.5(ARSA):c.827C>T (p.Thr276Met)	ARSA
75166491	108429	NM_000277.3(PAH):c.472C>T (p.Arg158Trp)	PAH
75949023	39947	NM_144612.6(LOXHD1):c.4714C>T (p.Arg1572Ter)	LOXHD1
78635798	16299	NM_032193.3(RNASEH2C):c.205C>T (p.Arg69Trp)	RNASEH2C
80338652	18848	NM_000081.3(LYST):c.3310C>T (p.Arg1104Ter)	LYST
80338826	29117	NM_002473.5(MYH9):c.2104C>T (p.Arg702Cys)	MYH9
80338934	17522	NM_024577.3(SH3TC2):c.3325C>T (p.Arg1109Ter)	SH3TC2
80338957	20935	NM_000334.4(SCN4A):c.2111C>T (p.Thr704Met)	SCN4A
80356680	29580	NM_000228.2(LAMB3):c.124C>T (p.Arg42Ter)	LAMB3
80356779	76552	NM_001876.3(CPT1A):c.1436C>T (p.Pro479Leu)	CPT1A
80356973	69370	NM_007294.3(BRCA1):c.2869C>T (p.Gln957Ter)	BRCA1
80356982	69227	NM_007294.3(BRCA1):c.2410C>T (p.Gln804Ter)	BRCA1
80357067	69840	NM_007294.3(BRCA1):c.4339C>T (p.Gln1447Ter)	BRCA1

RS#	ID аллеля	Наименование	Обозначение гена
80357089	69512	NM_007294.3(BRCA1):c.3331C>T (p.Gln1111Ter)	BRCA1
80357352	69958	NM_007294.3(BRCA1):c.4810C>T (p.Gln1604Ter)	BRCA1
80357485	69485	NM_007294.3(BRCA1):c.3286C>T (p.Gln1096Ter)	BRCA1
80359818	31157	NM_006516.3(SLC2A1):c.376C>T (p.Arg126Cys)	SLC2A1
80359826	201142	NM_006516.3(SLC2A1):c.988C>T (p.Arg330Ter)	SLC2A1
104894003	33314	NM_001101.4(ACTB):c.547C>T (p.Arg183Trp)	ACTB
104894261	31727	NM_130799.2(MEN1):c.1579C>T (p.Arg527Ter)	MEN1
104894267	31731	NM_130799.2(MEN1):c.1378C>T (p.Arg460Ter)	MEN1
104894364	27627	NM_004985.4(KRAS):c.173C>T (p.Thr58Ile)	KRAS
104894621	23472	NM_000304.3(PMP22):c.215C>T (p.Ser72Leu)	PMP22
104894714	19826	NM_181882.2(PRX):c.2857C>T (p.Arg953Ter)	PRX
104894797	26321	NM_004006.2(DMD):c.9568C>T (p.Arg3190Ter)	DMD
111033297	53892	NM_004004.5(GJB2):c.169C>T (p.Gln57Ter)	GJB2
111033538	17382	NM_032601.3(MCEE):c.139C>T (p.Arg47Ter)	MCEE
111687884	51571	NM_000138.4(FBN1):c.643C>T (p.Arg215Ter)	FBN1
112901682	76366	NM_001141945.2(ACTA2):c.115C>T (p.Arg39Cys)	ACTA2
114368325	38634	NM_000782.4(CYP24A1):c.1186C>T (p.Arg396Trp)	CYP24A1
118192226	34614	NM_172107.3(KCNQ2):c.1342C>T (p.Arg448Ter)	KCNQ2
118192251	34269	NM_004519.3(KCNQ3):c.988C>T (p.Arg330Cys)	KCNQ3
118203427	58245	NM_000368.4(TSC1):c.682C>T (p.Arg228Ter)	TSC1
118203434	58253	NM_000368.4(TSC1):c.733C>T (p.Arg245Ter)	TSC1
118203542	57958	NM_000368.4(TSC1):c.1525C>T (p.Arg509Ter)	TSC1
118203999	16285	NM_024675.3(PALB2):c.2962C>T (p.Gln988Ter)	PALB2
118204429	15511	NM_000035.4(ALDOB):c.178C>T (p.Arg60Ter)	ALDOB
121907916	18505	NM_000280.4(PAX6):c.607C>T (p.Arg203Ter)	PAX6
121908212	23527	NM_001127221.1(CACNA1A):c.1997C>T (p.Thr666Met)	CACNA1A
121908427	20365	NM_133647.1(SLC12A6):c.3031C>T (p.Arg1011Ter)	SLC12A6
121908489	20807	NM_003919.2(SGCE):c.289C>T (p.Arg97Ter)	SGCE
121912708	33034	NM_001182.4(ALDH7A1):c.328C>T (p.Arg110Ter)	ALDH7A1
121913344	151858	NM_000546.5(TP53):c.916C>T (p.Arg306Ter)	TP53
121917784	27085	NM_000136.2(FANCC):c.37C>T (p.Gln13Ter)	FANCC
121918167	15995	NM_000275.2(OCA2):c.2228C>T (p.Pro743Leu)	OCA2
121918244	16869	NM_001023570.3(IQCB1):c.1381C>T (p.Arg461Ter)	IQCB1
121918257	16926	NM_000255.3(MMUT):c.322C>T (p.Arg108Cys)	MMUT
122445105	26774	NM_000489.4(ATRX):c.736C>T (p.Arg246Cys)	ATRX
122445108	26781	NM_000489.4(ATRX):c.109C>T (p.Arg37Ter)	ATRX
122453121	26733	NM_004484.3(GPC3):c.1159C>T (p.Arg387Ter)	GPC3
128626235	26264	NM_004006.2(DMD):c.433C>T (p.Arg145Ter)	DMD
137852897	17803	NM_024312.4(GNPTAB):c.3565C>T (p.Arg1189Ter)	GNPTAB
137852994	19999	NM_018136.4(ASPM):c.9178C>T (p.Gln3060Ter)	ASPM
137853229	21102	NM_004260.3(RECQL4):c.2269C>T (p.Gln757Ter)	RECQL4
138049878	171163	NM_000257.4(MYH7):c.2608C>T (p.Arg870Cys)	MYH7

RS#	ID аллеля	Наименование	Обозначение гена
138119149	39897	NM_020745.3(AARS2):c.1774C>T (p.Arg592Trp)	AARS2
139675596	40180	NM_023073.3(CPLANE1):c.7477C>T (p.Arg2493Ter)	CPLANE1
140511594	39892	NM_024753.4(TTC21B):c.626C>T (p.Pro209Leu)	TTC21B
143343083	169011	NM_004004.5(GJB2):c.298C>T (p.His100Tyr)	GJB2
148865119	210450	NM_000071.2(CBS):c.146C>T (p.Pro49Leu)	CBS
180177091	132277	NM_024675.3(PALB2):c.751C>T (p.Gln251Ter)	PALB2
199422209	33004	NM_001127701.1(SERPINA1):c.1178C>T (p.Pro393Leu)	SERPINA1
199473556	78702	NM_198056.2(SCN5A):c.361C>T (p.Arg121Trp)	SCN5A
200075782	39327	NM_003560.3(PLA2G6):c.109C>T (p.Arg37Ter)	PLA2G6
200287925	151917	NM_002485.4(NBN):c.127C>T (p.Arg43Ter)	NBN
200309328	176122	NM_000138.4(FBN1):c.8080C>T (p.Arg2694Ter)	FBN1
200440128	205749	NM_012160.4(FBXL4):c.64C>T (p.Arg22Ter)	FBXL4
201632198	55279	NM_024022.2(TMPRS3):c.325C>T (p.Arg109Trp)	TMPRS3
267606919	21912	NM_004646.3(NPHS1):c.3478C>T (p.Arg1160Ter)	NPHS1
267607143	20038	NM_021625.4(TRPV4):c.943C>T (p.Arg315Trp)	TRPV4
267607258	46918	NM_002437.5(MPV17):c.293C>T (p.Pro98Leu)	MPV17
375699023	223602	NM_024675.3(PALB2):c.1042C>T (p.Gln348Ter)	PALB2
387906799	39125	NM_001244008.1(KIF1A):c.296C>T (p.Thr99Met)	KIF1A
387906904	39429	NM_021625.4(TRPV4):c.694C>T (p.Arg232Cys)	TRPV4
387907329	51081	NM_007075.3(WDR45):c.700C>T (p.Arg234Ter)	WDR45
397507215	46080	NM_007294.3(BRCA1):c.3352C>T (p.Gln1118Ter)	BRCA1
397507447	47625	NM_024312.4(GNPTAB):c.1123C>T (p.Arg375Ter)	GNPTAB
397509002	69322	NM_007294.3(BRCA1):c.2713C>T (p.Gln905Ter)	BRCA1
397509151	69806	NM_007294.3(BRCA1):c.4201C>T (p.Gln1401Ter)	BRCA1
397509330	70405	NM_007294.3(BRCA1):c.850C>T (p.Gln284Ter)	BRCA1
397514477	40113	NM_001031726.3(C19orf12):c.32C>T (p.Thr11Met)	C19orf12
397515848	51592	NM_000138.4(FBN1):c.7180C>T (p.Arg2394Ter)	FBN1
397516463	52805	NM_001001430.2(TNNT2):c.388C>T (p.Arg130Cys)	TNNT2
398123061	76995	NM_012160.4(FBXL4):c.1444C>T (p.Arg482Trp)	FBXL4
398123168	98367	NM_000143.3(FH):c.952C>T (p.His318Tyr)	FH
398123832	100328	NM_004006.2(DMD):c.10171C>T (p.Arg3391Ter)	DMD
398123929	100476	NM_004006.2(DMD):c.3151C>T (p.Arg1051Ter)	DMD
398124478	102281	NM_138694.3(PKHD1):c.2341C>T (p.Arg781Ter)	PKHD1
536907995	137626	NM_007194.4(CHEK2):c.58C>T (p.Gln20Ter)	CHEK2
587776407	153707	NM_024675.3(PALB2):c.451C>T (p.Gln151Ter)	PALB2
587776935	48413	NM_005465.4(AKT3):c.1393C>T (p.Arg465Trp)	AKT3
587780062	133253	NM_000535.5(PMS2):c.823C>T (p.Gln275Ter)	PMS2
587780226	133611	NM_032043.2(BRIP1):c.1315C>T (p.Arg439Ter)	BRIP1
587781948	151416	NM_000465.3(BARD1):c.1921C>T (p.Arg641Ter)	BARD1
587783685	168920	NM_003482.3(KMT2D):c.12592C>T (p.Arg4198Ter)	KMT2D
587784339	169779	NM_003560.3(PLA2G6):c.1903C>T (p.Arg635Ter)	PLA2G6
724159971	172085	NM_152778.2(MFSD8):c.1444C>T (p.Arg482Ter)	MFSD8

RS#	ID аллеля	Наименование	Обозначение гена
727503504	176073	NM_000363.4(TNNI3):c.508C>T (p.Arg170Trp)	TNNI3
727503513	172503	NM_001001430.2(TNNT2):c.280C>T (p.Arg94Cys)	TNNT2
727504136	177069	NM_001165963.1(SCN1A):c.3733C>T (p.Arg1245Ter)	SCN1A
730881422	179951	NM_000465.3(BARD1):c.1996C>T (p.Gln666Ter)	BARD1
730882029	180988	NM_000546.5(TP53):c.1024C>T (p.Arg342Ter)	TP53
747604569	185305	NM_032043.2(BRIP1):c.484C>T (p.Arg162Ter)	BRIP1
750621215	184806	NM_002878.3(RAD51D):c.898C>T (p.Arg300Ter)	RAD51D
753330544	195505	NM_206933.2(USH2A):c.13316C>T (p.Thr4439Ile)	USH2A
761494650	185659	NM_007194.4(CHEK2):c.85C>T (p.Gln29Ter)	CHEK2
763091520	197655	NM_000138.4(FBN1):c.6169C>T (p.Arg2057Ter)	FBN1
768933093	226933	NM_024685.4(BBS10):c.145C>T (p.Arg49Trp)	BBS10
773770609	264863	NM_177550.4(SLC13A5):c.997C>T (p.Arg333Ter)	SLC13A5
778989252	236615	NM_007194.4(CHEK2):c.1315C>T (p.Gln439Ter)	CHEK2
786202064	184902	NM_007294.3(BRCA1):c.4834C>T (p.Gln1612Ter)	BRCA1
786203821	184272	NM_024675.3(PALB2):c.940C>T (p.Gln314Ter)	PALB2
794726710	187772	NM_001165963.1(SCN1A):c.3637C>T (p.Arg1213Ter)	SCN1A
794726730	187817	NM_001165963.1(SCN1A):c.2134C>T (p.Arg712Ter)	SCN1A
794728195	197755	NM_000138.4(FBN1):c.2645C>T (p.Ala882Val)	FBN1
796051885	199890	NM_003239.4(TGFB3):c.898C>T (p.Arg300Trp)	TGFB3
797044883	205286	NM_019066.4(MAGEL2):c.1912C>T (p.Gln638Ter)	MAGEL2
869312892	226683	NM_139276.2(STAT3):c.2147C>T (p.Thr716Met)	STAT3
876658461	232175	NM_003000.2(SDHB):c.640C>T (p.Gln214Ter)	SDHB
886037684	248861	NM_177438.2(DICER1):c.2062C>T (p.R688*)	DICER1
886038001	249129	NM_007294.3(BRCA1):c.2599C>T (p.Gln867Ter)	BRCA1
886039480	260102	NM_024675.3(PALB2):c.2368C>T (p.Gln790Ter)	PALB2
886040218	261660	NM_007294.3(BRCA1):c.4225C>T (p.Gln1409Ter)	BRCA1
886041222	264422	NM_000280.4(PAX6):c.781C>T (p.Arg261Ter)	PAX6
1057521083	366251	NM_015265.3(SATB2):c.1165C>T (p.Arg389Cys)	SATB2

Пример 6

Демонстрация активности редактирования генов в растительных клетках

Активность гибридного белка RGN-деаминазы по редактированию азотистых оснований по настоящему изобретению продемонстрирована в растительных клетках с использованием протоколов, как описано в статье Li и др., (Nat. Biotech. 31:688-691 (2013)). Краткое описание эксперимента: экспрессионный вектор, содержащий кассету экспрессии, способную экспрессировать в растительных клетках гибридный белок RGN-деаминазу, функционально связанный с сигналом ядерной локализации SV40 (SEQ ID NO: 43), и вторую кассету экспрессии, кодирующую направляющую РНК,

направленную на один или более сайтов в гене PDS растения, которые фланкируют соответствующую последовательность PAM, вводили в протопласты мезофилла *Nicotiana benthamiana* с помощью PEG-опосредованной трансформации. Трансформированные протопласты инкубировали в темноте в течение вплоть до 36 ч. Геномную ДНК выделяли из протопластов с помощью набора DNeasy Plant Mini Kit (от фирмы Qiagen). Геномную область, примыкающую к сайту-мишени RGN, амплифицировали с помощью ПЦР, продукты очищали, а очищенные продукты ПЦР анализировали с использованием секвенирования нового поколения в системе Illumina MiSeq. Как правило, на ампликон генерируются 100000 считываний на 250 пар оснований при считывании спаренных концевых фрагментов (2 x 100000 считываний). Считывания анализировали с использованием программы CRISPResso (Pinello, и др., Nature Biotech, 34:695-697 (2016)) для расчета скорости редактирования. Выходные участки выравнивания оценивали с целью анализа INDEL или введения специфических адениновых мутаций.

Пример 7

Тестирование доставки мРНК

Чтобы оценить способность редакторов оснований осуществлять доставку в различных форматах, доставку мРНК тестировали с первичными Т клетками.

Очищенные CD3⁺ Т-клетки или PBMC оттаивали, активировали с использованием гранул CD3/CD28 (ThermoFisher) в течение 3 дней, затем подвергали нуклеофекции с использованием устройства Lonza 4D-Nucleofector X и полосок Nucleocuvette. Набор P3 Primary Cell использовали для доставки как мРНК, так и RNP. Клетки трансфектировали с использованием программ EO-115 и EH-115 для доставки мРНК и RNP соответственно. Клетки культивировали в среде для размножения Т-клеток CTS OpTimizer (ThermoFisher), содержащей IL-2, IL-7 и IL-15 (Miltenyi Biotec), в течение 4 дней после нуклеофекции перед сбором с использованием набора для выделения геномной ДНК Nucleospin Tissue (Machery Nagel).

Ампликоны, окружающие сайты редактирования, генерировали с помощью ПЦР с использованием праймеров, указанных в таблице 35, и подвергали секвенированию NGS с использованием платформы Illumina Nexterra и секвенирования парных концов 2x250 п.о. Предполагаемую скорость редактирования оснований определяли при расчете общей скорости замещения

для каждого образца. Среднее значение и количество образцов для каждой тестируемой направляющей показаны ниже.

Таблица 35

Средняя скорость редактирования для LPG50148-nAPG07433.1 с помощью

5 доставки мРНК

SGN	Редактирование, средн. %	N
SGN002352	7.84	2
SGN002364	29.79	2
SGN002367	0.1	2
SGN001061	0.37	1
SGN001062	71.81	1
SGN001064	3.99	1
SGN002254	8.92	2
SGN002255	5.26	2
SGN002256	8.32	2
SGN002290	2.88	2
SGN002293	9.68	2
SGN002299	27.05	2
SGN002132	29.11	2
SGN002137	7.77	2
SGN002139	6.00	2
SGN001770	1.22	2
SGN001773	0.49	2
SGN002212	29.63	2
SGN002216	2.58	2
SGN002218	36.13	2
SGN002230	14.32	2
SGN002231	33.18	2
SGN000753	6.84	2
SGN000754	26.41	1
SGN001856	0.5	2
SGN002248	9.91	2
SGN002249	40.19	2

ПЕРВОНАЧАЛЬНАЯ ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441, где указанный полипептид характеризуется дезаминазной активностью.

2. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, при этом дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

3. Молекула нуклеиновой кислоты по п. 2, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

4. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по п. 1 или молекулу нуклеиновой кислоты по п. 2 или п. 3.

5. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, последовательность которого по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

6. Гибридный белок по любому из вариантов 14-16, где дезаминаза представляет собой адениндезаминазу.

7. Гибридный белок по п. 5 или п. 6, где ДНК-связывающий полипептид представляет собой мегануклеазу, гибридный белок со структурой цинк-пальцы или TALEN.

5 8. Гибридный белок по п. 5 или п. 6, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.

10 9. Гибридный белок по 8, где РНК-управляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

10. Гибридный белок по п. 9, где RGN представляет собой полипептид CRISPR-Cas типа II.

15 11. Гибридный белок по п.9, где RGN представляет собой полипептид CRISPR-Cas типа V.

20 12. Гибридный белок по любому из пунктов 9-11, где RGN представляет собой RGN-никазу.

13. Гибридный белок по п. 9, где RGN содержит аминокислотную последовательность, по меньшей мере на 95% идентичную любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

25 14. Гибридный белок по п. 12, где RGN содержит аминокислотную последовательность любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 368.

30 15. Гибридный белок по любому из пунктов, где гибридный белок дополнительно содержит по меньшей мере один сигнал ядерной локализации (NLS).

16. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и

дезаминазу, при этом дезаминаза кодируется нуклеотидной последовательностью, которая:

а) характеризуется по меньшей мере 80%-ой идентичностью с любой из последовательностей SEQ ID NO: 451, 449, 443, 11-20, 444-448, 450 и 452-485
5 или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441.

10 17. Молекула нуклеиновой кислоты по п. 16, где дезаминаза представляет собой адениндезаминазу.

18. Молекула нуклеиновой кислоты по п. 16 или п. 17, где ДНК-связывающий полипептид представляет собой мегануклеазу, гибридный белок со
15 структурой цинковые пальцы или TALEN.

19. Молекула нуклеиновой кислоты по п. 16 или п. 17, где ДНК-связывающий полипептид представляет собой РНК-направляемый ДНК-связывающий полипептид.
20

20. Молекула нуклеиновой кислоты по п. 19, где РНК-направляемый ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

25 21. Молекула нуклеиновой кислоты по п. 20, где RGN представляет собой полипептид CRISPR-Cas типа II.

22. Молекула нуклеиновой кислоты по п. 20, где RGN представляет собой полипептид CRISPR-Cas типа V.
30

23. Молекула нуклеиновой кислоты по п. 20, где RGN представляет собой RGN-никазу.

24. Молекула нуклеиновой кислоты по п. 20, где RGN характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

5 25. Молекула нуклеиновой кислоты по п. 23, где RGN представляет собой последовательность SEQ ID NO: 42, 52-59, 61, 397 или 398.

10 26. Молекула нуклеиновой кислоты по любому из пунктов 16-25, где полинуклеотид, кодирующий гибридный белок, функционально связан по своему 5'-концу с гетерологичным промотором.

15 27. Молекула нуклеиновой кислоты по любому пунктов 16-26, где полинуклеотид, кодирующий гибридный белок, функционально связан по своему 3'-концу с гетерологичным терминатором.

28. Молекула нуклеиновой кислоты по любому из пунктов 16-27, где гибридный белок содержит один или более сигналов ядерной локализации.

20 29. Молекула нуклеиновой кислоты по любому из пунктов 16-28, где гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

25 30. Молекула нуклеиновой кислоты по любому из пунктов 16-28, где гибридный белок оптимизирован по кодонам для экспрессии в прокариотической клетке.

31. Вектор, содержащий молекулу нуклеиновой кислоты по любому из пунктов 16-30.

30 32. Вектор по п. 31, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (гРНК), способную гибридизироваться с последовательностью-мишенью.

33. Вектор по п. 32, где гРНК представляет собой единую направляющую РНК.

5 34. Вектор по п. 32, где гРНК представляет собой двойную направляющую РНК.

10 35. Клетка, содержащая гибридный белок по любому из пунктов 5-15, нуклеиновую кислоту по любому из пунктов 16-30 или вектор по любому из пунктов 31-34.

36. Клетка, содержащая гибридный белок по любому из пунктов 5-15, где клетка дополнительно содержит направляющую РНК.

15 37. Способ получения гибридного белка, включающий культивирование клетки по любому из п. 35 или п. 36 в условиях, в которых экспрессируется гибридный белок.

20 38. Способ получения гибридного белка, включающий введение в клетку молекулы нуклеиновой кислоты по любому из пунктов 16-30 или вектора по любому из пунктов 31-34 и культивирование клетки в условиях, в которых экспрессируется гибридный белок.

25 39. Способ по п. 37 или п. 38, дополнительно включающий очистку указанного гибридного белка.

30 40. Способ получения гибридного рибонуклеопротеинового комплекса RGN, включающий введение в клетку молекулы нуклеиновой кислоты по любому из пунктов 16-30 и молекулы нуклеиновой кислоты, содержащей кассету экспрессии, кодирующую направляющую РНК, или вектор по любому из пунктов 31-34, и культивирование клетки в условиях, в которых гибридный белок и гРНК экспрессируются и образуют гибридный рибонуклеопротеиновый комплекс RGN.

41. Способ по п. 40, дополнительно включающий очистку указанного гибридного рибонуклеопротеинового комплекса RGN.

42. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична любой из последовательностей SEQ ID NO: 407, 399, 405, 1-10, 400-404, 406 и 408-441, или нуклеотидную

последовательность, кодирующую указанный гибридный белок, и

б) одну или более направляющих РНК, способных гибридизироваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (гРНК), и где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

43. Система по п. 42, где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующих одну или более направляющих РНК, и указанных нуклеотидных последовательностей, кодирующих гибридный белок, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности.

44. Система по п. 42 или п. 43, где последовательность ДНК-мишени представляет собой последовательность ДНК-мишени эукариот.

45. Система по любому из пунктов 42-44, где целевая последовательность ДНК расположена рядом с мотивом, прилегающим к протоспейсеру (РАМ), который распознается RGN.

46. Система по любому из пунктов 42-46, где молекула ДНК-мишени находится внутри клетки.

47. Система по любому из пунктов 42-46, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II.

5 48. Система по любому из пунктов 42-46, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа V.

10 49. Система по любому из пунктов 42-46, где RGN гибридного белка имеет аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности из последовательностей SEQ ID NO: 41, 60, 366 или 368.

50. Система по любому из пунктов 42-46, где RGN гибридного белка представляет собой RGN никазу.

15 51. Система по п. 50, где никаза RGN характеризуется любой аминокислотной последовательностью из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 388.

20 52. Система по любому из пунктов 42-51, где гибридный белок содержит один или более сигналов ядерной локализации.

53. Система по любому из пунктов 42-52, где гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

25 54. Система по любому из пунктов 42-53, где нуклеотидные последовательности, кодирующие одну или более направляющих РНК, и нуклеотидная последовательность, кодирующая гибридный белок, расположены на одном векторе.

30 55. Фармацевтическая композиция, содержащая фармацевтический приемлемый носитель и гибридный белок по любому из пунктов 5-15, молекулу нуклеиновой кислоты по любому из пунктов 16-30, вектор по любому из пунктов 31-34, клетку по п. 35 или п. 36 или систему по любому из пунктов 42-54.

56. Способ модификации молекулы ДНК-мишени, содержащей последовательность ДНК-мишени, причем указанный способ включает доставку системы по любому из пунктов 42-54 в указанную молекулу ДНК-мишени или в клетку, содержащую молекулу ДНК-мишени.

5

57. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза с использованием объединения:

- 10 i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и
- ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная последовательность дезаминазы по меньшей мере на 90% идентична
- 15 последовательности с SEQ ID NO: 407, 399, 405, 1-10, 400-404, 406 и 408-441, в условиях, пригодных для образования рибонуклеотидного комплекса RGN-дезаминаза, и

- б) контактирование указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным *in vivo*
- 20 рибонуклеотидным комплексом RGN-дезаминаза,
- где одна или более направляющих РНК гибридизируются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

25

58. Способ по п. 56 или п. 57, где указанная модифицированная молекула ДНК-мишени содержит мутацию А>N по меньшей мере одного нуклеотида в молекуле ДНК-мишени, где N представляет собой С, G или Т.

- 30 59. Способ по п. 58, где указанная модифицированная молекула ДНК-мишени содержит мутацию А>G по меньшей мере одного нуклеотида в молекуле ДНК-мишени.

60. Способ по любому из пунктов 56-59, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II.

5 61. Способ по любому из пунктов 56-59, где RGN гибридного белка представляет собой полипептид CRISPR-Cas типа V.

10 62. Способ по любому из пунктов 56-59, где RGN гибридного белка характеризуется аминокислотной последовательностью, по меньшей мере на 95% идентичной последовательности SEQ ID NO: 41, 60, 366 или 368.

63. Способ по любому из пунктов 56-59, где RGN гибридного белка представляет собой RGN-никазу.

15 64. Способ по п. 63, где никаза RGN представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

65. Способ по любому из пунктов 56-64, где гибридный белок содержит один или более сигналов ядерной локализации.

20 66. Способ по пунктам 56-64, где гибридный белок оптимизирован по кодонам для экспрессии в эукариотической клетке.

25 67. Способ по любому из пунктов 56-66, где указанная последовательность ДНК-мишени представляет собой последовательность ДНК-мишени эукариот.

68. Способ по любому из пунктов 56-67, где указанная последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM).

30 69. Способ по любому из пунктов 56-68, где молекула ДНК-мишени находится внутри клетки.

70. Способ по п. 69, дополнительно включающий отбор клетки, содержащей указанную модифицированную молекулу ДНК.

71. Клетка, содержащая модифицированную последовательность ДНК-мишени в соответствии со способом по п. 70.

5 72. Фармацевтическая композиция, содержащая клетку по п. 71 и фармацевтически приемлемый носитель.

10 73. Способ получения генетически модифицированной клетки с коррекцией причинной мутации наследственного заболевания, включающий введение в клетку:

а) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность дезаминазы характеризуется по меньшей мере 90%-ой идентичностью с любой из последовательностей SEQ ID NO: 407, 405, 399, 1- 10, 400-404, 406 и 408-441, или полинуклеотида, кодирующего указанный гибридный белок, где указанный полинуклеотид, кодирующий гибридный белок, функционально связан с промотором для обеспечения экспрессии гибридного белка в клетке, и

б) одну или более направляющих РНК (гРНК), способных гибридизоваться с целевой последовательностью ДНК, или полинуклеотид, кодирующий указанную гРНК, где указанный полинуклеотид, кодирующий гРНК, функционально связан с промотором для обеспечения экспрессии гРНК в клетке, при этом гибридный белок и гРНК направлены на геномное местоположение причинной мутации и модифицируют геномную последовательность с целью удаления причинной мутации.

25 74. Способ по п. 73, где указанный RGN гибридного белка представляет собой RGN-никазу.

30 75. Способ по п. 74, где RGN-никаза представляет собой любую из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

76. Способ по любому из пунктов 73-75, где модификация генома включает введение мутации А>G по меньшей мере одного нуклеотида в последовательность ДНК-мишени.

77. Способ по любому из пунктов 73-76, где коррекция причинной мутации включает коррекцию нонсенс-мутации.

5 78. Способ по п. 73, где генетически наследуемое заболевание представляет собой заболевание, указанное в таблице 34.

79. Способ по п. 73 где генетически наследуемое заболевание представляет собой муковисцидоз.

10

80. Способ лечения заболевания, включающий введение субъекту, нуждающемуся в таком введении, эффективного количества фармацевтической композиции по пунктам 55 или 72.

15 81. Способ по п. 80, где указанное заболевание связано с причинной мутацией, и указанное эффективное количество указанной фармацевтической композиции корректирует указанную причинную мутацию.

20 82. Применение гибридного белка по любому из пунктов 5-15, молекула нуклеиновой кислоты по любому из пунктов 16-30, вектор по любому из пунктов 31-34, клетка по любому из пунктов 35, 36 и 71, или система по любому из пунктов 42-54 для лечения заболевания у субъекта.

25 83. Применение по п. 82, где указанное заболевание связано с причинной мутацией, и указанное лечение включает коррекцию указанной причинной мутации.

30 84. Применение гибридного белка по любому из пунктов 5-15, молекулы нуклеиновой кислоты по любому из пунктов 16-30, вектора по любому из пунктов 31-34, клетки по любому из пунктов 35, 36 и 71, или системы по любому из пунктов 42-54 для изготовления лекарственного средства, пригодного для лечения заболевания.

85. Применение по п. 84, где указанное заболевание связано с причинной мутацией, и эффективное количество указанного лекарственного средства корректирует указанную причинную мутацию.

5 86. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид РНК-направляемой нуклеазы (RGN), где указанный полинуклеотид содержит нуклеотидную последовательность, кодирующую полипептид RGN, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 41 или 60,
10 но не содержащую аминокислотные остатки с 590 по 597 в последовательности SEQ ID NO: 41 или 60,

 где указанный полипептид RGN способен связываться с последовательностью ДНК-мишени в РНК-направляемой последовательности специфическим образом при связывании с направляющей РНК (гРНК),
15 способной гибридизоваться с указанной последовательностью ДНК-мишени.

 87. Молекула нуклеиновой кислоты по п. 86, где указанный полинуклеотид, кодирующий полипептид RGN, функционально связан с промотором, гетерологичным указанному полинуклеотиду.

20 88. Молекула нуклеиновой кислоты по п. 86 или п. 87, где указанный полипептид RGN представляет собой мертвую нуклеазу или функционирует как никаза.

25 89. Молекула нуклеиновой кислоты по любому из пунктов 86-88, где полипептид RGN функционально слит с полипептидом, редактирующим азотистое основание.

 90. Вектор, содержащий молекулу нуклеиновой кислоты по любому из
30 пунктов 86-89.

 91. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 41 или 60, но не содержащий аминокислотных остатков с 590 по 597

в последовательности SEQ ID NO: 41 или 60, где указанный полипептид представляет собой РНК-направляемую нуклеазу.

5 92. Выделенный полипептид по п. 91, где указанный полипептид RGN содержит аминокислотную последовательность, по меньшей мере на 95% идентичную последовательности SEQ ID NO: 366 или 368.

10 93. Выделенный полипептид по п. 91 или п. 92, где указанный полипептид RGN является мертвой нуклеазой или функционирует как никаза.

94. Выделенный полипептид по любому из пунктов 91 или 93, где полипептид RGN функционально слит с полипептидом, редактирующим азотистое основание.

15 95. Клетка, содержащая молекулу нуклеиновой кислоты по любому из пунктов 86-89, вектор по п. 90 или полипептид по любому из пунктов 91-94.

20 96. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 407, где указанный полипептид обладает деаминазной активностью.

97. Выделенный полипептид по п. 96, где полипептид содержит аминокислотную последовательность SEQ ID NO: 407.

25 98. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 451, или

30 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 407.

99. Молекула нуклеиновой кислоты по п. 98, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

5 100. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из пунктов 96-97 или молекулу нуклеиновой кислоты по любому из пунктов 98-99.

10 101. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 407.

102. Гибридный белок по п. 101, где ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

15 103. Гибридный белок по п. 102, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 104. Гибридный белок по любому из пунктов 101-103, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID
25 NO: 41, 60, 366 или 368.

105. Гибридный белок по любому из пунктов 102-104, где полипептид RGN представляет собой нуклеазу.

30 106. Гибридный белок по п. 105, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

107. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

5 а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 451, или

 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 407.

10 108. Молекула нуклеиновой кислоты по п. 107, где ДНК-связывающий полипептид представляет собой полипептид RGN.

 109. Молекула нуклеиновой кислоты по п. 108, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

15 110. Молекула нуклеиновой кислоты по любому из пунктов 107-109, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой
20 перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

 111. Молекула нуклеиновой кислоты по любому из пунктов 108-110, где
25 полипептид RGN представляет собой нуклеазу.

 112. Молекула нуклеиновой кислоты по п. 111, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 113. Вектор, содержащий молекулу нуклеиновой кислоты по любому из пунктов 107-112.

114. Вектор по п.113, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизоваться с последовательностью-мишенью.

5 115. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из пунктов 101-106 и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

10 116. Клетка, содержащая гибридный белок по любому из пунктов 101-106, молекулу нуклеиновой кислоты по любому из пунктов 107-112, вектор по любому из пунктов 113-114 или комплекс RNP по п. 115.

117. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

15 а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 407, или нуклеотидная последовательность, кодирующая указанный гибридный белок, и

20 б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (гРНК), и

25 где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

30 118. Система по п. 117, где по меньшей мере одна из указанных нуклеотидных последовательностей, кодирующей одну или более направляющих РНК, и указанная нуклеотидная последовательность, кодирующая гибридный белок, функционально связаны с промотором, гетерологичным указанной нуклеотидной последовательности.

119. Система по любому из пунктов 117-118, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается полипептидом RGN.

5 120. Система по любому из пунктов 117-119, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563 или комплементарных им последовательностей.

10

121. Система по любому из пунктов 117-120, где последовательность гРНК содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

15

122. Система по любому из пунктов 117-121, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 123. Система по любому из пунктов 117-121, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой
25 аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

124. Система по п. 123, где полипептид RGN представляет собой никазу.

30 125. Система по п. 124, где аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

126. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из пунктов 101-106, молекулу нуклеиновой кислоты по любому из пунктов 107-112, вектор по любому из пунктов 113-114, комплекс RNP по п. 115, клетку по п. 116 или систему по любому из пунктов 117-125.

127. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза путем объединения:

i) одной или более направляющих РНК, способных гибридизироваться с целевой последовательностью ДНК, и

ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная

последовательность дезаминазы характеризуется последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 407,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и

происходит модификация молекулы ДНК-мишени.

128. Способ по п. 127, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563 или их комплементарных им последовательностей.

129. Способ по любому из пунктов 127-128, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы,

состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

5 130. Способ по любому из пунктов 127-129, где способ осуществляется *in vitro*, *in vivo* или *ex vivo*.

131. Способ лечения субъекта, имеющего или подверженного риску развития заболевания, нарушения или состояния, включающий:

10 введение субъекту гибридного белка по любому из пунктов 101-106, молекулы нуклеиновой кислоты по любому из пунктов 107-112, вектора по любому из пунктов 113-114, комплекса RNP по п. 115, клетки по п. 116, системы по любому из пунктов 117-125 или фармацевтической композиции по п. 126.

15 132. Способ по п. 131, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

20 133. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 405, где указанный полипептид обладает деаминазной активностью.

25 134. Выделенный полипептид по п. 133, где полипептид содержит аминокислотную последовательность SEQ ID NO: 405.

135. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

30 а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 449 или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 405.

136. Молекула нуклеиновой кислоты по п. 135, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

5 137. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из пунктов 133-134 или молекулу нуклеиновой кислоты по любому из пунктов 135-136.

10 138. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 405.

139. Гибридный белок по п. 138, где ДНК-связывающий полипептид представляет собой полипептид РНК-направляемой нуклеазы (RGN).

15 140. Гибридный белок по п. 139, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 141. Гибридный белок по любому из пунктов 138-140, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID
25 NO: 41, 60, 366 или 368.

142. Гибридный белок по любому из пунктов 139-141, где полипептид RGN представляет собой нуклеазу.

30 143. Гибридный белок по п. 142, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

144. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

5 а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 449, или

 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 405.

10 145. Молекула нуклеиновой кислоты по п. 144, где ДНК-связывающий полипептид представляет собой полипептид RGN.

 146. Молекула нуклеиновой кислоты по п. 145, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

15 147. Молекула нуклеиновой кислоты по любому из пунктов 144-146, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой

20 перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

 148. Молекула нуклеиновой кислоты по любому из пунктов 145-147, где

25 полипептид RGN представляет собой нуклеотид.

 149. Молекула нуклеиновой кислоты по п. 148, где аминокислотная последовательность нуклеотида по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 150. Вектор, содержащий молекулу нуклеиновой кислоты по любому из пунктов 144-149.

151. Вектор по п. 150, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизоваться с последовательностью-мишенью.

5 152. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из пунктов 138-141 и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

10 153. Клетка, содержащая гибридный белок по любому из пунктов 138-143, молекулу нуклеиновой кислоты по любому из пунктов 144-149, вектор по любому из пунктов 150-151 или комплекс RNP по п. 152.

154. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

15 а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 405, или нуклеотидная последовательность, кодирующая указанный гибридный белок, и

20 б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (gRNA), и

25 где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишенью и модифицировать молекулу ДНК-мишени.

30 155. Система по п. 154, где по меньшей мере одна из указанной нуклеотидной последовательности, кодирующей одну или более направляющих РНК, и указанная нуклеотидная последовательность, кодирующая гибридный белок, функционально связаны с промотором, гетерологичным указанной нуклеотидной последовательности.

156. Система по любому из пунктов 154-155, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается полипептидом RGN.

5 157. Система по любому из пунктов 154-156, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251. -286, 305-344, 562 и 563 или комплементарных им последовательностей.

10

158. Система по любому из пунктов 154-157, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

15

159. Система по любому из пунктов 154-158, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 160. Система по любому из пунктов 154-159, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой
25 аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

161. Система по п. 160, где полипептид RGN представляет собой никазу.

30 162. Система по п. 161, отличающаяся тем, что аминокислотная последовательность никазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

163. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из пунктов 138-143, молекулу нуклеиновой кислоты по любому из пунктов 144-149, вектор по любому из пунктов 150-151, комплекс RNP по п. 152, клетку по п. 153 или систему по любому из пунктов 154-162.

164. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза путем объединения:

i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и

ii) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная последовательность дезаминазы по меньшей мере на 90% идентична последовательности с SEQ ID NO: 405,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

165. Способ по п. 164, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563 или комплементарных им последовательностей.

166. Способ по любому из пунктов 164-165, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы,

состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

5 167. Способ по любому из пунктов 164-166, где способ осуществляется *in vitro*, *in vivo* или *ex vivo*.

168. Способ лечения субъекта, страдающего или подверженного риску развития заболевания, нарушения или состояния, включающий:

10 введение субъекту гибридного белка по любому из пунктов 138-143, молекулы нуклеиновой кислоты по любому из пунктов 144-149, вектора по любому из пунктов 150-151, комплекса RNP по п. 152, клетки по п. 153, системы по любому из пунктов 154-162 или фармацевтической композиции по п. 163.

15 169. Способ по п. 168, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

20 170. Выделенный полипептид, содержащий аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 399, где указанный полипептид обладает деаминазной активностью.

25 171. Выделенный полипептид по п. 170, где полипептид содержит аминокислотную последовательность SEQ ID NO: 399.

172. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий полипептид дезаминазы, где дезаминаза кодируется нуклеотидной последовательностью, которая:

30 а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 443, или

б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную любой из последовательностей SEQ ID NO: 399.

173. Молекула нуклеиновой кислоты по п. 172, где указанная молекула нуклеиновой кислоты дополнительно содержит гетерологичный промотор, функционально связанный с указанным полинуклеотидом.

5 174. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и полипептид по любому из пунктов 170-171 или молекулу нуклеиновой кислоты по любому из пунктов 172-173.

10 175. Гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, характеризующийся последовательностью, по меньшей мере на 90% идентичной последовательности SEQ ID NO: 399.

15 176. Гибридный белок по п. 175, где ДНК-связывающий полипептид представляет собой полипептид РНК-управляемой нуклеазы (RGN).

177. Гибридный белок по п. 176, где полипептид RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 178. Гибридный белок по любому из пунктов 175-177, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью, из последовательностей SEQ ID
25 NO: 41, 60, 366 или 368.

179. Гибридный белок по любому из пунктов 176-178, где полипептид RGN представляет собой нуклеазу.

30 180. Гибридный белок по п. 179, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

181. Молекула нуклеиновой кислоты, содержащая полинуклеотид, кодирующий гибридный белок, содержащий ДНК-связывающий полипептид и дезаминазу, где дезаминаза кодируется нуклеотидной последовательностью, которая:

5 а) по меньшей мере на 80% идентична последовательности SEQ ID NO: 443 или

 б) кодирует аминокислотную последовательность, по меньшей мере на 90% идентичную последовательности SEQ ID NO: 399.

10 182. Молекула нуклеиновой кислоты п. 181, где ДНК-связывающий полипептид представляет собой полипептид RGN.

 183. Молекула нуклеиновой кислоты по п. 182, где RGN представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

15 184. Молекула нуклеиновой кислоты по любому из пунктов 181-183, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Cas12a, Cas12b, Cas12g, Cas12h, Cas12i, Cas13b, Cas13c, Cas13d, Cas14, Csn2, xCas9, SpCas9-NG, LbCas12a, AsCas12a, Cas9-KKH, Cas9 с круговой
20 перестановкой, белок Argonaute (Ago), SmacCas9, домен Spy-macCas9 или полипептид RGN с любой аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

 185. Молекула нуклеиновой кислоты по любому из пунктов 182-184, где
25 полипептид RGN представляет собой нуклеотид.

 186. Молекула нуклеиновой кислоты по п. 185, где аминокислотная последовательность нуклеотида по меньшей мере на 95% идентична любой из последовательностей с SEQ ID NO: 42, 52-59, 61, 397 и 398.

30 187. Вектор, содержащий молекулу нуклеиновой кислоты по любому из пунктов 181-186.

188. Вектор по п. 187, дополнительно содержащий по меньшей мере одну нуклеотидную последовательность, кодирующую направляющую РНК (gRNA), способную гибридизироваться с последовательностью-мишенью.

5 189. Рибонуклеопротеиновый (RNP) комплекс, содержащий гибридный белок по любому из пунктов 175-180, и направляющую РНК, связанную с ДНК-связывающим полипептидом гибридного белка.

10 190. Клетка, содержащая гибридный белок по любому из пунктов 175-180, молекулу нуклеиновой кислоты по любому из пунктов 181-186, вектор по любому из пунктов 187-188 или комплекс RNP по п. 189.

191. Система для модификации молекулы ДНК-мишени, включающая последовательность ДНК-мишени, причем указанная система включает:

15 а) гибридный белок, содержащий полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность деаминазы по меньшей мере на 90% идентична последовательности с SEQ ID NO: 399, или нуклеотидная последовательность, кодирующая указанный гибридный белок, и

20 б) одну или более направляющих РНК, способных гибридизоваться с указанной последовательностью ДНК-мишени, или одной или более последовательностями нуклеотидов, кодирующими одну или более направляющих РНК (gRNA), и

25 где одна или более направляющих РНК способны образовывать комплекс с гибридным белком, чтобы направить указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени и модифицировать молекулу ДНК-мишени.

192. Система по п. 191, где по меньшей мере одна из указанной нуклеотидной последовательности, кодирующей одну или более направляющих РНК, и указанной нуклеотидной последовательности, кодирующей гибридный белок, функционально связана с промотором, гетерологичным указанной нуклеотидной последовательности.

30

193. Система по любому из пунктов 191-192, где последовательность ДНК-мишени расположена рядом с мотивом, прилегающим к протоспейсеру (PAM), который распознается полипептидом RGN.

5 194. Система по любому из пунктов 191-193, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251. -286, 305-344, 562 и 563 или комплементарных им последовательностей.

10

195. Система по любому из пунктов 191-194, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287- 304, 345-364 и 564.

15

196. Система по любому из пунктов 191-195, где полипептид RGN гибридного белка представляет собой полипептид CRISPR-Cas типа II или полипептид CRISPR-Cas типа V.

20 197. Система по любому из пунктов 191-196, где полипептид RGN представляет собой Cas9, CasX, CasY, Cpf1, C2c1, C2c2, C2c3, GeoCas9, CjCas9, Casl2a, Casl2b, Casl2g, Casl2h, Casl2i, Casl3b, Casl3c, Casl3d, Casl4, Csn2, xCas9, SpCas9-NG, LbCasl2a, AsCasl2a, Cas9-KKH, Cas9 с круговой перестановкой, белок Argonaute (Ago), домен SmacCas9, Spy-macCas9 или RGN с любой
25 аминокислотной последовательностью из последовательностей SEQ ID NO: 41, 60, 366 или 368.

198. Система по п. 197, где полипептид RGN представляет собой нуклеазу.

30 199. Система по п. 198, где аминокислотная последовательность нуклеазы по меньшей мере на 95% идентична любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.

200. Фармацевтическая композиция, содержащая фармацевтически приемлемый носитель и гибридный белок по любому из пунктов 175-180, молекулу нуклеиновой кислоты по любому из пунктов 181-186, вектор по любому из пунктов 187-188, комплекс RNP по п. 189, клетку по п. 190 или систему по любому из пунктов 191-199.

201. Способ модификации молекулы ДНК-мишени, содержащей целевую последовательность, включающий:

а) сборку рибонуклеотидного комплекса RGN-дезаминаза путем объединения:

i) одной или более направляющих РНК, способных гибридизоваться с целевой последовательностью ДНК, и

ii) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и по меньшей мере одну дезаминазу, где аминокислотная последовательность дезаминазы по меньшей мере на 90% идентична последовательности SEQ ID NO: 399,

в условиях, подходящих для образования рибонуклеотидного комплекса RGN-дезаминаза, и

б) приведение в контакт указанной молекулы ДНК-мишени или клетки, содержащей указанную молекулу ДНК-мишени, с собранным рибонуклеотидным комплексом RGN-дезаминаза,

где одна или более направляющих РНК гибридизуются с последовательностью ДНК-мишени, тем самым направляя указанный гибридный белок на связывание с указанной последовательностью ДНК-мишени, и происходит модификация молекулы ДНК-мишени.

202. Способ по п. 201, где последовательность ДНК-мишени содержит последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 62-97, 116-139, 152-185, 203-234, 251-286, 305-344, 562 и 563 или комплементарных им последовательностей.

203. Способ по любому из пунктов 201-202, где последовательность gRNA содержит последовательность нуклеиновой кислоты, выбранную из группы,

состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

204. Способ по любому из пунктов 201-203, где способ осуществляется *in vitro*, *in vivo* или *ex vivo*.

205. Способ лечения субъекта, имеющего или подверженного риску развития заболевания, нарушения или состояния, включающий:

10 введение субъекту гибридного белка по любому из пунктов 175-180, молекулы нуклеиновой кислоты по любому из пунктов 181-186, вектора по любому из пунктов 187-188, комплекса RNP по п. 189, клетки по п. 190, системы по любому из пунктов 191-199 или фармацевтической композиции по п. 200.

15 206. Способ по п. 205, дополнительно включающий введение любой из gRNA, содержащей последовательность нуклеиновой кислоты, выбранную из группы, состоящей из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

20 207. Способ лечения или купирования по меньшей мере одного симптома муковисцидоза, включающий введение субъекту, нуждающемуся в таком лечении, эффективного количества:

а) гибридного белка, содержащего полипептид РНК-направляемой нуклеазы (RGN) и деаминазу, где аминокислотная последовательность дезаминазы характеризуется последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 407, 405, 399, 1-10, 400-404, 406 и 408-441, или полинуклеотида, кодирующего указанный гибридный белок, где указанный полинуклеотид, кодирующий гибридный белок, функционально связан с промотором для обеспечения экспрессии гибридного белка в клетке, и

30 б) одной или более направляющих РНК (gRNA), способных гибридизоваться с последовательностью ДНК-мишени, или полинуклеотида, кодирующего указанную gRNA, где указанный полинуклеотид, кодирующий gRNA, функционально связан с промотором для обеспечения экспрессии gRNA в клетке,

где гибридный белок и gRNA направлены на геномное местоположение причинной мутации и модифицируют геномную последовательность для удаления причинной мутации.

5 208. Способ по п. 207, где gRNA содержит спейсерную последовательность, направленную на любую из последовательностей SEQ ID NO: 62–97, 116–139, 152–185, 203–234, 251–286, 305–344, 562, и 563 или комплементарные им последовательности.

10 209. Способ по пунктам 207 или 208, где gRNA содержит любую из последовательностей SEQ ID NO: 98-115, 140-151, 186-202, 235-250, 287-304, 345-364 и 564.

15 210. Способ по любому из пунктов 207-209, где указанный RGN характеризуется аминокислотной последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 41, 60, 366 и 368.

20 211. Способ по любому из пунктов 207-209, где указанный RGN характеризуется аминокислотной последовательностью, по меньшей мере на 90% идентичной любой из последовательностей SEQ ID NO: 42, 52-59, 61, 397 и 398.