(19) Евразийское патентное ведомство

(21) 201991809 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2020.02.05

(22) Дата подачи заявки
2018.01.30

(51) Int. Cl. ***C12N 15/82*** (2006.01)
C12N 15/10 (2006.01)
C12N 15/11 (2006.01)
C12N 15/90 (2006.01)
A01H 5/00 (2018.01)

(54) СЦЕПЛЕНИЕ МАТРИЦЫ ДЛЯ РЕПАРАЦИИ С ЭНДОНУКЛЕАЗАМИ ДЛЯ ГЕНОМНОЙ ИНЖЕНЕРИИ

(31) **62/451,859**
(32) **2017.01.30**
(33) **US**
(86) **PCT/EP2018/052313**
(87) **WO 2018/138385 2018.08.02**

(71) Заявитель:
КВС ЗААТ СЕ & Ко. КГаА (DE)

(72) Изобретатель:
Лабс Матиас (US)

(74) Представитель:
Зуйков С.А. (RU)

(57) Настоящее изобретение относится к искусственным молекулярным комплексам, содержащим по меньшей мере одну сайт-специфичную нуклеазу и непосредственно взаимодействующий с ней по меньшей мере один стыковочный домен матрицы для репарации, причем указанный стыковочный домен матрицы для репарации взаимодействует по меньшей мере с одной репарационной матричной последовательностью нуклеиновой кислоты. Искусственный комплекс может дополнительно содержать по меньшей мере один взаимодействующий домен. Искусственные молекулярные комплексы выполнены с возможностью опосредования репарации последовательности ДНК-мишени в прокариотическом или эукариотическом организме с высокой точностью целенаправленным образом, и, таким образом, они могут быть использованы для геномной инженерии в прокариотической или эукариотической клетке или организме или для редактирования вирусного генома. Также предложены способы модификации по меньшей мере одной последовательности ДНК-мишени в прокариотической или эукариотической клетке или в вирусном геноме, например, для развития признака или для лечения заболевания. Кроме того, предложен способ изготовления растения, растительной клетки, растительного материала или производного, или его потомства, содержащий или отредактированный по меньшей мере одним искусственным молекулярным комплексом.

201991809 A1

201991809 A1

Сцепление матрицы для репарации с эндонуклеазами для геномной инженерии

Область техники, к которой относится изобретение

Настоящее изобретение относится к искусственным молекулярным комплексам, содержащим, по меньшей мере, одну сайт-специфичную нуклеазу и непосредственно взаимодействующий с ней, по меньшей мере, один стыковочный домен матрицы для репарации, причем, указанный стыковочный домен матрицы для репарации взаимодействует с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты. Искусственный комплекс может дополнительно содержать, по меньшей мере, один взаимодействующий домен. Искусственные молекулярные комплексы выполнены с возможностью опосредования репарации последовательности ДНК-мишени в прокариотическом или эукариотическом, или вирусном организме, или в геноме с высокой точностью целенаправленным образом и, таким образом, могут быть использованы для геномной инженерии в прокариотической или эукариотической клетке, или организме, или в геномной инженерии с прокариотическим, эукариотическим или вирусным геномом *in vivo* или *in vitro*. Также предложены способы модификации, по меньшей мере, одной последовательности ДНК-мишени в прокариотической или эукариотической клетке, или в вирусном геноме, например, для развития признака или для лечения заболевания. Кроме того, предложен способ изготовления растения, растительной клетки, растительного материала или производного, или его потомства, содержащий или отредактированный, по меньшей мере, одним искусственным молекулярным комплексом. Таким образом, предложен искусственный молекулярный комплекс, подходящий для любой сайт-специфичной нуклеазы, который направляет матрицу для репарации в непосредственную физическую близость с последовательностью ДНК-мишени, подлежащей модификации, чтобы обеспечить постоянную доступность матрицы для репарации *in situ* в сайте индуцированного двухцепочечного разрыва ДНК, чтобы гарантировать высокую эффективность и прогнозируемость различных подходов геномной инженерии.

Предпосылки создания изобретения

Точное редактирование генов или геномная инженерия развилась как одна из наиболее важных областей генной инженерии, обеспечивающая целенаправленное и сайт-направленное манипулирование интересующим геномом. Непременной предпосылкой для сайт-направленной геномной инженерии являются программируемые нуклеазы, которые могут быть использованы для разрыва интересующей нуклеиновой кислоты в определенном положении для индуцирования либо двухцепочечного разрыва (DSB), либо, по меньшей мере, одного одноцепочечного разрыва. В альтернативном варианте осуществления настоящего изобретения, указанные нуклеазы могут быть химерными или мутированными вариантами, которые больше не содержат функцию нуклеазы, а действуют скорее как распознающие молекулы в комбинации с другим ферментом. Таким образом, эти нуклеазы или их варианты являются ключом к любому подходу к редактированию генов или к геномной инженерии. За последние годы было разработано много подходящих нуклеаз, в частности, эндонуклеаз с задаваемой структурой, содержащих мегануклеазы, цинк-пальцевые нуклеазы, TALE-нуклеазы, нуклеазы Аргонавты, полученные, например, из *Natronobacterium gregoryi*, и нуклеазы CRISPR, содержащие, например, нуклеазы Cas, Cpf1, CasX или CasY в качестве части системы коротких палиндромных повторов, регулярно расположенных группами (CRISPR).

CRISPR (короткие палиндромные повторы, регулярно расположенные группами) в своей природной среде первоначально эволюционировали в бактериях, где система CRISPR выполняет роль адаптивной иммунной системы для защиты от вирусной атаки. При воздействии на вирус, короткие сегменты вирусной ДНК интегрируются в локус CRISPR. РНК транскрибируется из части локуса CRISPR, которая включает вирусную последовательность. Эта РНК, которая содержит последовательность,

комплементарную вирусному геному, опосредует нацеливание эффекторного белка CRISPR на последовательность-мишень в вирусном геноме. Эффекторный белок CRISPR расщепляет и тем самым препятствует репликации вирусной мишени. За последние годы система CRISPR была успешно адаптирована к редактированию генов или к геномной инженерии также и в эукариотических клетках. В настоящее время значительное внимание в исследованиях уделяется редактированию в животных клетках и применениям в терапевтических целях для людей. Целенаправленная модификация сложных геномов животных, а также растений по-прежнему представляет собой трудную задачу.

Система CRISPR в своей природной среде описывает молекулярный комплекс, содержащий, по меньшей мере, одну малую и отдельную некодирующую РНК в комбинации с нуклеазой Cas или другой нуклеазой CRISPR, такой как, нуклеаза Cpf1 (Цетше и др., «Cpf1 представляет собой эндонуклеазу, направляемую одноцепочечной РНК, системы CRISPR-Cas Класса 2», Cell, 163, страницы 1-13, октябрь 2015 года), который может продуцировать специфичный двухцепочечный разрыв ДНК. В настоящее время системы CRISPR подразделяются на два класса, содержащие пять типов систем CRISPR, например, систему Типа II, использующую Cas9 в качестве эффектора, и систему Типа V, использующую Cpf1 в качестве эффекторной молекулы (Макарова и др., Nature Reviews Microbiology, 2015). В искусственных системах CRISPR, синтетическая некодирующая РНК и нуклеаза CRISPR, и/или, при необходимости, модифицированная нуклеаза CRISPR, модифицированная для того, чтобы она действовала в качестве никазы или была лишена какой-либо нуклеазной функции, могут использоваться в комбинации с, по меньшей мере, одной синтетической или искусственной направляющей РНК или gРНК, сочетающей функцию crРНК и/или tracrРНК (см. выше, Макарова и др., 2015). Иммунный ответ, опосредованный системой CRISPR-Cas в природных системах, требует CRISPR-РНК (crРНК), при этом, созревание этой направляющей РНК, которая контролирует специфичную активацию нуклеазы CRISPR, варьируется в значительной степени между различными системами CRISPR, которые были охарактеризованы до этого момента. Во-первых, вторгающаяся ДНК, также известная как спейсер, интегрирована между двумя соседними областями повтора на проксимальном конце локуса CRISPR. Системы CRISPR Типа II кодируют нуклеазу Cas9 в качестве ключевого фермента для этапа интерференции, причем, эти системы содержат как crРНК, так и трансактивирующую РНК (tracrРНК) в качестве направляющего мотива. Они гибридизуются и образуют области двухцепочечной (ds) РНК, которые распознаются РНКазой III и могут быть расщеплены для образования зрелых crРНК. Затем они, в свою очередь, ассоциируются с молекулой Cas для того, чтобы направить нуклеазу точно в область нуклеиновой кислоты-мишени. Рекомбинантные молекулы gРНК могут содержать как вариабельную область распознавания ДНК, так и область взаимодействия Cas, и их можно точно сконструировать, независимо от специфичной нуклеиновой кислоты-мишени и желаемой нуклеазы Cas. В качестве дополнительного механизма безопасности, в области нуклеиновой кислоты-мишени должны присутствовать PAM (мотивы, примыкающие к протоспейсеру); они представляю собой последовательности ДНК, которые следуют непосредственно из ДНК, распознанной комплексом Cas9/РНК. PAM-последовательность для Cas9, полученная из *Streptococcus pyogenes*, была описана, как «NGG» или «NAG» (Стандартный нуклеотидный код ИЮПАК) (Джинек и др., «Программируемая эндонуклеаза ДНК, направляемая двойной РНК, в адаптивном бактериальном иммунитете», Science, 2012, 337: 816-821). PAM-последовательность для Cas9, полученная из *Streptococcus aureus*, прочитывается как «NNGRRT» или «NNGRR(N)». Известны следующие варианты систем CRISPR-Cas9. Так, Cas9 *Neisseria meningitidis* расщепляет в PAM-последовательности NNNNGATT. Cas9 *Streptococcus thermophilus* расщепляет в PAM-последовательности NNAGAAW. Недавно был описан еще один мотив PAM NNNNRYAC для системы

CRISPR *Campylobacter* (WO 2016/021973 A1). В отношении нуклеаз Cpf1, было описано, что комплекс Cpf1-crРНК эффективно расщепляет ДНК-мишень с помощью короткоцепочечного Т-богатого PAM, в отличие от обычных G-богатых PAM, распознаваемых системами Cas9 (см. выше, Цетше и др.). Кроме того, используя модифицированные полипептиды CRISPR, можно получить специфичные одноцепочечные разрывы. Комбинированное использование никаз Cas с различными рекомбинантными gРНК также может индуцировать высокоспецифичные двухцепочечные разрывы ДНК с помощью двойного никования ДНК. Кроме того, используя две gРНК, можно оптимизировать специфичность связывания ДНК и, таким образом, расщепление ДНК.

В настоящее время, например, такие основанные на Cas9 системы Типа II или их вариант, или любая химерная форма, такая как, эндонуклеаза, были модифицированы для геномной инженерии. Синтетические системы CRISPR, состоящие из двух компонентов, направляющей РНК (gРНК), также называемой одноцепочечной направляющей РНК (sgРНК), и неспецифичной CRISPR-ассоциированной эндонуклеазы могут быть использованы для генерации нокаутных клеток или животных путем ко-экспрессии gРНК, специфичной для гена, который должен быть мишенью, и способной ассоциироваться с эндонуклеазой Cas9. Примечательно, что gРНК представляет собой искусственную молекулу, содержащую один домен, взаимодействующий с Cas или с любым другим эффекторным белком CRISPR, или с его вариантом, или каталитически активным фрагментом, и другой домен, взаимодействующий с интересующей нуклеиновой кислотой-мишенью, и, таким образом, представляющую собой синтетическое слияние crРНК и tracrРНК («одноцепочечная направляющая РНК» (sgРНК) или просто «gРНК»; см. выше, Джинек и др., 2012). Геномной мишенью может быть любая последовательность ДНК из, приблизительно, 20 нуклеотидов, при условии, что мишень присутствует непосредственно выше PAM. PAM-последовательность имеет огромное значение для целевого связывания, и точная последовательность зависит от вида Cas9 и, например, прочитывается как 5' NGG 3' или 5' NAG 3' (Стандартный нуклеотидный код ИЮПАК) (см. выше, Джинек и др., 2012) для *Streptococcus pyogenes,* полученного Cas9. Используя модифицированные нуклеазы Cas, целенаправленные одноцепочечные разрывы могут быть введены в интересующую последовательность-мишень. Комбинированное использование такой никазы Cas с различными рекомбинантными gРНК для высоко сайт-специфичных двухцепочечных разрывов ДНК может быть введено с помощью системы двойного никования. Использование, по меньшей мере, одной gРНК может еще больше повысить общую специфичность и уменьшить количество нецеленаправленных действий.

После экспрессии, белок Cas9 и gРНК образуют рибонуклеопротеиновый комплекс посредством взаимодействий между доменом «скаффолд» gРНК и поверхностно-экспонированными положительно заряженными бороздками на Cas9. Важно отметить, что «спейсерная» последовательность gРНК остается свободной для взаимодействия с ДНК-мишенью. Комплекс Cas9-gРНК будет связывать любую геномную последовательность с PAM, но степень, до которой спейсер gРНК соответствует ДНК-мишени, определяет, будет ли отрезана Cas9. Как только комплекс Cas9-gРНК свяжет предполагаемую ДНК-мишень, «порождающая» последовательность на 3'-конце последовательности, нацеленной на gРНК, начинает отжигать ДНК-мишень. Если порождающая последовательность и последовательность ДНК-мишени совпадают, то gРНК будет продолжать отжигать ДНК-мишень в направлении от 3'-конца до 5'-конца (относительно полярности gРНК).

В последнее время сконструированные системы CRISPR/Cpf1, в дополнение к системам CRISPR/Cas9, становятся все более важными для целенаправленной геномной инженерии (см. выше, Цетше

и др., и EP 3 009 511 A2). Система Типа V, вместе с системой Типа II, относится к системам CRISPR Класса 2 (Методы Макаровой и Кунина. Молекулярная биология, 2015, 1311:47-753). Эффекторный белок Cpf1 представляет собой большой белок (примерно,1300 аминокислот), который содержит RuvC-подобный нуклеазный домен, гомологичный соответствующему домену Cas9, а также аналог характерному богатому аргинином кластеру Cas9. Однако Cpf1 не имеет нуклеазного домена HNH, который присутствует во всех белках Cas9, и RuvC-подобный домен является прилегающим в последовательности Cpf1, в отличие от Cas9, где он содержит длинные вставки, включающие домен HNH (Чилински, 2014; Макарова, 2015). Эффекторы Cpf1 обладают определенными отличиями от эффекторов Cas9, а именно: отсутствие обязательных дополнительных трансактивирующих crРНК (tracrРНК) для обработки массива CRISPR, эффективного расщепления ДНК-мишени короткоцепочечными Т-богатыми PAM (в отличие от Cas9, где за PAM следует G-богатая последовательность), и введения ступенчатых двухцепочечных разрывов ДНК посредством Cpf1. Совсем недавно были идентифицированы дополнительные новые системы CRISPR-Cas на основе CasX и CasY, которые, из-за относительно небольшого размера эффекторного белка, представляют особый интерес для многих подходов редактирования генов или геномной инженерии (Бурштейн и др., «Новые системы CRISPR-Cas, полученные из некультивированных микробов», Nature, декабрь 2016 года). Специфичность систем CRISPR во многом определяется тем, насколько специфична последовательность, нацеленная на gРНК, для геномной мишени, по сравнению с остальной частью генома.

Царство *Plantae* содержит очень гетерогенные и разнообразные виды, принимая во внимание геномные и фенотипические отличия зеленых водорослей, бриофитов, птеридофитов и геофитов. Геномы растений и их сложность представляют собой нелегкую задачу для высокоточного редактирования генов или геномной инженерии. Например, на долю *Zea mays* (маис или кукуруза) приходится самое высокое в мире производство среди всех зерновых культур, урожай которых составил 875 миллионов тонн в 2012 году. Оно имеет крупный геном, примерно, 2,4 гигабаз (Гб) с гаплоидным числом хромосом 10 (Шнейбл и др., 2009; Чжан и др., 2009). *Triticum aestivum* (пшеница обыкновенная), например, представляет собой гексаплоид, размер генома которого оценивается, приблизительно, в 17 Гб. *Beta vulgaris* ssp. *vulgari* (сахарная свекла) имеет размер генома от, примерно, 470 мегабаз (Мб) до, примерно, 569 Мб. Специфичная архитектура и состав растительных клеток, а также особое развитие растений, требует специфичной адаптации для инструментов геномной инженерии, когда они предназначены для использования с целью модификации последовательности-мишени в растительной клетке. Поэтому инструменты геномной инженерии и ассоциированные с ними принципы, установленные для систем животных, в частности, млекопитающих, не обязательно будут работать в представляющей интерес растительной клетке, и существует необходимость в специфичных стратегиях для создания технологии с целью достижения широкого применения в растениях.

Аналогичным образом, геномы животных и, в частности, млекопитающих, являются сложными, например, содержащими 2,7 Гб для генома *Mus musculus* или 3,2 Гб для генома *Homo sapiens*. В частности, когда подходы редактирования генов на основе CRISPR или геномной инженерии предназначены для использования с целью точного редактирования генов или геномной инженерии мишеней в геноме человека, таким образом, существует настоятельная необходимость в предложении высокой специфичности, поскольку любой вид нецеленаправленного действия может быть весьма вредным.

Другим аспектом, который необходимо принципиально рассмотреть в отношении геномной инженерии, является механизм репарации, необходимый после расщепления интересующего геномного сайта-мишени, поскольку двухцепочечные разрывы (DSB) или повреждения ДНК в целом вредны для

целостности генома. DSB в геномном материале могут быть вызваны ионизирующим излучением, химическими веществами, окислением, ферментами и одноцепочечными разрывами во время репликации и представляют собой серьезную форму повреждения ДНК, которая может привести к потере генов, остановке репликации ДНК и гибели клеток. Таким образом, чрезвычайно важно, чтобы клеточный механизм мог предложить механизмы репарации двухцепочечного разрыва (DSB). Клетки обладают внутренними механизмами, благодаря чему они пытаются восстановить любое двухцепочечное или одноцепочечное повреждение ДНК. Механизмы репарации DSB были разделены на два главных базовых типа: негомологичное соединение концов (NHEJ) и гомологичная рекомбинация (HR). Механизмы репарации на основании гомологии обычно называются репарацией, направляемой гомологией (HOR).

NHEJ представляет собой ответ доминантного ядра у животных и растений, который не требует гомологичных последовательностей, но он зачастую подвержен ошибкам и, следовательно, потенциально мутагенен (Ваймен С., Канаар Р. «Репарация двухцепочечного разрыва ДНК: все хорошо, что хорошо кончается», Annual Review of Genetics, 2006; 40, 363-83). Репарация по HOR требует гомологии, но те пути HOR, которые используют интактную хромосому для репарации разорванной хромосомы, то есть, репарации двухцепочечного разрыва и синтез-зависимого отжигания цепи, очень точны. В классическом пути репарации DSB, 3'-концы захватывают интактную гомологичную матрицу, а затем служат в качестве праймера для синтеза репарации ДНК, что в конечном итоге приводит к образованию двойных структур Холлидея (dHJ). Двойные структуры Холлидея представляют собой четырехцепочечные разветвленные структуры, которые образуются, когда элонгация захваченной цепи «захватывает» и синтезирует ДНК со второго конца DSB. Отдельные HJ разрешаются посредством расщепления одним из двух путей. Синтез-зависимое отжигание цепи является консервативным и заканчивается исключительно некроссоверными событиями. Это означает, что все вновь синтезированные последовательности присутствуют в одной и той же молекуле. В отличие от пути репарации NHEJ, после захвата цепи и образования петли D в синтез-зависимом отжигании цепи, вновь синтезированная часть захваченной цепи вытесняется из матрицы и возвращается к обработанному концу незахваченной цепи с другого конца DSB. 3'-конец незахваченной цепи элонгируется и лигируется для заполнения зазора. Существует еще один путь HOR, называемый путем репарации, индуцированной разрывом, который еще не полностью охарактеризован. Основным признаком этого пути является наличие только одного инвазивного конца в DSB, который может быть использован для репарации.

Дальнейший путь HOR представляет собой одноцепочечное отжигание (SSA). SSA не является консервативным и происходит между прямыми повторами > 30 пар оснований и приводит к делециям. В последние годы микрогомологически опосредованное соединение концов (MMEJ) было признано отдельным типом репарации DSB у эукариот. Только очень короткоцепочечные (2-14 пар оснований) области гомологии необходимы для этого пути, и он, как правило, приводит к делециям, таким как SSA. Его также отличали с генетической точки зрения от путей HR и NHEJ, и в клетках млекопитающих оно выступает в качестве резерва для NHEJ (Квон, Т,. Хак, Е., и Херрин, Д.Л. (2010). «Микрогомологически опосредованная и негомологичная репарация двухцепочечного разрыва в хлоропластном геноме *Arabidopsis*». *Труды Национальной академии наук Соединенных штатов Америки*, 107(31), 13954-13959). Таким образом, HR/HOR использует гомологичное удлинение ДНК на сестринской хроматиде в качестве матрицы. Таким образом, предлагается высокая точность, но с меньшей эффективностью. NHEJ, напротив, представляет собой высокоэффективный и простой путь, благодаря которому можно воссоединить два конца независимо от выраженной гомологии, хотя эта эффективность сопровождается недостатком,

который состоит в том, что этот процесс сопровождается ошибками, и он может быть ассоциирован со вставками или делециями.

Таким образом, для подходов редактирования генов или геномной инженерии, стремящихся повлиять на пути естественной репарации, требуется физическое конструирование матрицы для репарации (RT), которая является важным параметром. Представляется возможным предложить RT либо в виде ssДНК, либо в виде частичной dsДНК. Текущие протоколы, основанные на инструментах CRISPR для редактирования генома в комбинации с матрицей для репарации (RT), основываются исключительно на отдельном предложении RT нуклеиновой кислоты, либо двухцепочечной, либо одноцепочечной, которая, в свою очередь, распознает разрыв в ДНК, подлежащий репарации исключительно путем спаривания оснований и гибридизации. Однако физическая и временная доступность RT в сайте, где индуцируется разрыв ДНК, может не контролироваться способами, доступными в настоящее время, поскольку эти способы не предлагают точного пространственного и временного предложения RT в правильной конфигурации, концентрации и, следовательно, стехиометрии в компартменте, где должна происходить репарация, в предпочтительном варианте осуществления настоящего изобретения, сразу после индукции целенаправленного разрыва ДНК для точного контролирования не только разрыва, но и события репарации.

Подобно нуклеазам CRISPR/Cas, эндонуклеазы Аргонавты («Аргонавты») участвуют в защите от чужеродных нуклеиновых кислот с помощью направляющих нуклеиновых кислот для указания последовательности-мишени, которая затем расщепляется компонентом белка Аргонавта. В частности, Аргонавт может связывать и расщеплять нуклеиновую кислоту-мишень, образуя комплекс со сконструированной или синтетической нуклеиновой кислотой, нацеленной на нуклеиновую кислоту, где расщепление нуклеиновой кислоты-мишени может вводить двухцепочечные разрывы в нуклеиновой кислоте-мишени. Также как и система Cas9, направляющие нуклеиновые кислоты Аргонавтов предлагают легкий способ для программирования специфичности последовательности эндонуклеазы. Однако молекулы короткоцепочечной ssРНК используются в качестве направляющих многими эукариотическими Аргонавтами без каких-либо ограничений на распознавание вторичных структур, таких как, структуры, которые присутствуют во взаимодействии Cas9-короткоцепочечная направляющая РНК (sgРНК, gРНК). Следовательно, копийность ssРНК в большинстве эукариотических клеток делает специфичное нацеливание эукариотических Аргонавтов, направляемых РНК, потенциальной проблемой. Напротив, некоторые прокариотические Аргонавты направляются молекулами короткоцепочечной 5'-фосфорилированной ssДНК (Шварц, Д.К. и др., Интерференция ДНК, направляемой ДНК, посредством прокариотического Аргонавта. Nature, 507, 258-261, 2014; Шварц, Д.К. и др. Аргонавт архея Pyrococcus furiosus представляет собой нуклеазу, направляемую ДНК, которая нацелена на когнатную ДНК. Nucleic Acids Research, 43, 5120-5129, 2015) и, следовательно, по своей природе обладают более низким потенциалом для ошибочного направления нуклеиновыми кислотами, полученными из клеток-хозяев, из-за дефицита молекул короткоцепочечной ssДНК, присутствующих в эукариотических клетках. Таким образом, эндонуклеазы Аргонавты, направляемые ДНК, имеют потенциал для применения в эукариотическом редактировании генома. Использование системы Аргонавтов *Natronobacterium gregoryi* (NgAgo) в растениях, однако, ранее не демонстрировалось.

В литературе отмечено, что гомологичная рекомбинация между двумя последовательностями происходит чаще, если последовательности находятся в непосредственной близости в ядре, а не разделены значительным расстоянием. Например, анализ в *Arabidopsis* скорости редактирования генов, полученных между хромосомно расположенными донорскими молекулами и мишенями, был выше в обоих случаях,

когда донор существовал на той же хромосоме, что и мишень, чем в других случаях, когда два локуса были расположены на разных хромосомах (Фаузер и др., 2012). Однако эти результаты никогда не использовались рациональным образом для оптимизации подходов к редактированию генов или в геномной инженерии на основе сайт-специфичной нуклеазы в эукариотических клетках.

EP 2 958 996 A1 стремится преодолеть проблему специфичной репарации DSB, предлагая ингибитор механизмов NHEJ в клетке для увеличения разрушения генов, опосредованного нуклеазой (например, цинк-пальцевая нуклеаза (ZFN) или эффекторная нуклеаза, подобная активаторам транскрипции (TALEN)) или системой нуклеаз (например, CRISPR/Cas). Ингибирование критической ферментативной активности этих NHEJ путей репарации ДНК с использованием низкомолекулярных ингибиторов каталитической субъединицы ДНК-зависимой протеинкиназы (ДНК-PKcs) и/или Поли-(АДФ-рибоза)-полимеразы 1/2 (PARP1/2) проводит к повышению уровня разрушения генов нуклеазами из-за того, что клетки заставляют прибегать к путям репарации, которые в большей степени подвержены ошибкам, чем классический NHEJ, таким как, альтернативный NHEJ и/или микрогомологически опосредованное соединение концов. Поэтому в процессе редактирования генома добавляется дополнительное химическое вещество, которое, однако, может быть неблагоприятным для нескольких типов клеток и анализов. Это также может повлиять на целостность генома обработанных клеток и/или регенеративный потенциал.

Ма и др. (2016, JCB, 214(5):529, «Ядерная динамика системы CRISPR-Cas9 и распознавание мишени в живых клетках») использовали 3'-модифицированную sgРНК, которая обеспечивала связывание флуоресцентного репортера на основе аптамера для изучения динамики Cas9 и динамики sgРНК в направлении теломерной мишени. Примечательно, что модификация в последовательности tracrРНК не оказала влияния на нацеливание. Только последующее усечение последовательности tracrРНК привело к дестабилизации sgРНК, независимой от модификации аптамера.

Следовательно, существует постоянная потребность в предложении подходящих инструментов CRISPR, в частности, инструментов, оптимизированных для точного редактирования растений, в частности, основных сельскохозяйственных культур, которые сочетают в себе высокоточное расщепление генома, например, путем предложения gРНК, оптимизированных для сайта-мишени в интересующей клетке, и одновременном предложении возможности опосредования высокоточной и правильной HOR и, таким образом, целенаправленной репарации DSB, что является обязательным для контроля вмешательства в редактирование генов или геномную инженерию.

Таким образом, цель состоит в том, чтобы представить новые стратегии для предложения матриц для репарации для точного редактирования генома, в частности, подходящих для эукариотических клеток, включая клетки дрожжей, животные и растительные клетки, а также подходящих для прокариотических клеток, например, для метаболической инженерии и различных других целей, или для модификации вирусных геномов, например, для ослабления вируса или для снижения вирулентности вируса. Несмотря на огромные успехи редактирования генома в биотехнологии, например, для терапевтических подходов, в генной терапии или геномной инженерии растений или микробов для развития целенаправленного признака, все еще существуют серьезные проблемы и проблемные вопросы в отношении специфичности целенаправленной модификации генома, подлежащей введению, или нецеленаправленных действий. Эта проблема, среди прочего, ассоциирована со степенью точности, которая может быть получена при индуцировании разрыва и ассоциированной репарации интересующей геномной нуклеиновой кислоты-мишени.

Поскольку любой вид подхода к редактированию генов или в геномной инженерии,

индуцирующий DSB, вводит потенциально вредный разрыв ДНК и, возможно, нежелательный механизм репарации ДНК, приводящий к нежелательным обменам нуклеиновыми кислотами, то существует постоянная потребность в разработке более эффективных способов и инструментов для достижения высокоточного и контролируемого редактирования генов или геномной инженерии, что также подразумевает использование целенаправленных матриц для репарации (RT) ДНК.

Другая проблема, зачастую ассоциированная с предложением успешной геномной инженерии без опосредования нецеленаправленными действиями – физическая доступность матрицы для репарации в сайте DSB точно в тот момент, когда осуществляется разрыв, и который, следовательно, должен быть репарирован. Как правило, событие желаемого редактирования вытесняется репарацией посредством пути негомологичного соединения концов (NHEJ) или посредством рекомбинации с эндогенной гомологичной последовательностью, как подробно описано выше. В зависимости от организма-мишени, подлежащего модификации, требуется согласованная стратегия для введения инструмента редактирования генов или геномной инженерии вместе с интересующей матрицей для репарации таким образом, чтобы все инструменты могли, при наступлении соответствующего момента, достичь компартмента в клетке, содержащей геном, то есть, в предпочтительном варианте осуществления настоящего изобретения, в ядре или любом другом несущем геном компартменте, таком как митохондрии. Один способ частичного преодоления этого ограничения заключается в амплификации матрицы для репарации и, таким образом, в увеличении копийности матрицы в ядре и, возможно, в том, чтобы сделать его более доступным для использования в репарации DSB с помощью геминивирусного вектора (см., например, Мэч, Plant Cell. 2014, идентификатор цифрового объекта:10.1105/tpc.114.122606; и Балтес и др., Plant Cell. 2014, идентификатор цифрового объекта:10.1105/tpc.113.119792). Матрица для репарации, однако, доставляется как отдельный физический объект, и поэтому не существует механизма контроля, удостоверяющего, что матрица для репарации действительно будет присутствовать в месте, где репарация ДНК необходима именно в тот момент, когда DSB вводится эндонуклеазой.

Что касается применений CRISPR, то часто предлагается использовать свободные нуклеотиды ssДНК в качестве матриц для репарации или матриц для репарации, передаваемых посредством плазмид, но при этом не раскрывается и не предлагается стратегия, которая гарантировала бы, что матрица для репарации действительно приводится в физический контакт с DSB, подлежащим репарации *in situ*, когда генерируется DSB.

Взаимодействия биотин-стрептавидин и биотин-авидин являются одними из самых стабильных в природе, с константой диссоциации Kd, составляющей 10^{-15} М. Ассоциация основана на гомотетрамерной структуре между белком авидином или белком стрептавидином (-16,5 и 13,2 кДа на субъединицу, соответственно) и повсеместно присутствующем, но с низкой копийностью, витамином биотином. Гомотетрамерные комплексы стрептавидина или авидина образуются спонтанно и способны связывать четыре молекулы биотина с низкими константами диссоциации. По меньшей мере, в двух попытках спонтанная тетрамеризация могла быть преодолена с уменьшением аффинности связывания (Лайтинен и др. 2003, «Рациональная конструкция активного мономера авидина». Journal of Biological Chemistry 278(6): 4010-4014; Манн и др., 2016, «Маркировка клеток и зависимое от близости биотинилирование с помощью сконструированного мономерного стрептавидина». TECHNOLOGY 4(3): 1-7). Кроме того, было продемонстрировано, что биотинилирование нуклеазы возможно путем включения сигнала биотинилирования в последовательность (Кэй и др. 2009, «Биотинилирование белков с высокой пропускной способностью». Methods in molecular biology (Клифтон, Нью-Джерси) 498: 185-196). BirA представляет

собой возможный биотинилирующий фермент для экспрессии бактериального белка, но биотинилирование также происходит и у высших растений (Тиссо и др., 1996, «Биотинилирование белка у высших растений: характеристика активности синтетазы голокарбоксилазы, катализирующей инкорпорацию биотина, полученной из листьев гороха (*Pisum sativum*)», Biochemical Journal 314 (Часть 2): 391-395).

Одноцепочечные вариабельные фрагменты (scFvs) представляют собой белки слияния вариабельных областей тяжелых (VH) и легких цепей (VL) иммуноглобулинов, соединенных с коротким линкерным пептидом, включающим от десяти до, примерно, 25 аминокислот, и известны как неустойчивые высокоаффинные связывающие молекулы. Двухвалентные (или бивалентные) одноцепочечные вариабельные фрагменты (di-scFvs, bi-scFvs) могут быть сконструированы путем связывания двух scFvs. Этого можно достигнуть путем продуцирования одной пептидной цепи с двумя областями VH и двумя областями VL, в результате получается тандем scFvs (Куфер и др., 2004, Тенденции в области биотехнологии, 22 (5), 238-244; Сюн и др., 2006, Конструирование и выбор белка, 19(8), 359-367).

Однако до сих пор эти полученные результаты о возможностях биотинилированных молекул и их когнатных партнеров по связыванию или о других высокоаффинных молекулярных связывающих парах, таких как, например, антитела или одноцепочечные вариабельные фрагменты и их когнатные партнеры, еще не применялись для целенаправленной геномной инженерии с использованием сайт-специфичных нуклеаз и матрицы для репарации.

На настоящий момент, своеобразные различия в доставке инструментов редактирования генов или геномной инженерии и/или матрицы для репарации, необходимых для различных клеток-мишеней, становятся очевидными. В этом отношении, растительные клетки имеют определенные отличительные признаки, включая клеточные стенки, что делает редактирование генов или геномную инженерию в растительных клетках совершенно иной задачей, по сравнению с редактированием генов или геномной инженерией, как это предусмотрено для животных клеток/клеток млекопитающих, поскольку доставка инструментов редактирования и/или репарации генома опосредуется иными способами трансформации, трансфекции и/или трансдукции, чем для других эукариотических клеток. Эти особенности, однако, должны быть приняты во внимание для достижения высокоточного редактирования генома растений. Поэтому целью настоящего изобретения было удовлетворить ярко выраженную потребность – предложить новые инструменты и способы, подходящие для высокоточного редактирования генома в эукариотических клетках, включая растительные клетки, в частности, в области CRISPR и Аргонавтов, опосредованные редактированием генома, чтобы преодолеть текущее ограничение в области редактирования генов, касающееся физической доступности матрицы для репарации в сайте и в момент репарации DSB и, таким образом, конкуренцию механизмов репарации ДНК путем негомологичного соединения концов (NHEJ) или посредством рекомбинации с (эндогенной) гомологичной последовательностью (HR/HOR). Еще одной целью настоящего изобретения является предложение упрощенного инструментария сайт-направленных нуклеаз, подходящего для любой сайт-специфичной нуклеазы и не ограниченного CRISPR, направляемыми нуклеиновыми кислотами, или нуклеазами Аргонавтов, которые могут быть использованы для сайт-направленного редактирования генома в эукариотических или прокариотических клетках, или в любом прокариотическом, эукариотическом или вирусном геноме, путем предложения молекулы или молекулярного комплекса, который объединяет свойства распознавания ДНК, расщепления и матрицы для репарации, и одновременно может быть легко доставлен в сайт-мишень, то есть, в прокариотическую клетку, эукариотический или вирусный геном, в частности, геном животной клетки, в частности, клетки млекопитающего или растительной клетки, поскольку степень точности, которая должна быть достигнута

при редактировании генома животных или растительных клеток, по-прежнему нуждается в улучшении, чтобы соответствовать непременно высоким нормативным требованиям, установленным уполномоченными органами медицинской и пищевой отраслей. Риск для нецеленаправленных интеграций искусственных молекулярных комплексов, раскрытых в настоящем документе, ниже, чем для матрицы для репарации ss- или ds-ДНК, введенной в клетку в виде свободных молекул. Кроме того, была поставлена цель предложить инструмент доставки, который специально оптимизирован для переноса специфичной конструкции редактирования генома растения с помощью способа специфичной доставки в растение. Еще одной целью было предложить подход, который может быть основан на транзиентной активности редактирования с использованием транзиентно предложенной РНК и сайт-специфичных нуклеаз, при желании, из-за чувствительности в определенных зонах применения к любой форме генетической модификации, которая интегрирует чужеродную ДНК в качестве промежуточного звена в продуктивной фазе. И последней целью настоящего изобретения является предложение способа редактирования генов или геномной инженерии, который превосходит современные способы в том, что он экономит время, необходимое для тестирования новых мишеней, поскольку для него не потребуется проведение клонирования в большом объеме и предварительное тестирование.

Краткое изложение сущности изобретения

Вышеуказанные выявленные цели по настоящему изобретению были достигнуты, благодаря решению проблемы доступности матрицы для репарации посредством доставки матрицы для репарации в сайт DSB путем непосредственного приспособления ее в виде «груза» к нуклеазному комплексу, при этом, спектр нуклеаз, подходящих для этого подхода, был значительно увеличен путем предложения искусственных молекулярных комплексов, которые основаны на какой-либо интересующей сайт-специфичной нуклеазе (SSN). Направление матрицы для репарации в двухцепочечный разрыв в момент разрыва осуществляют *in situ*, посредством предложения, по меньшей мере, одного стыковочного домена матрицы для репарации (RTDD) вместе, по меньшей мере, с одной SSN, при этом, стыковочный домен матрицы для репарации выполнен с возможностью непосредственного взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты (RT), повышая доступность матрицы для репарации (RT) для использования в репарации разрыва. Таким образом, искусственные молекулярные комплексы по настоящему изобретению не только помогают предложить матрицы для репарации, выполненные с индивидуальными параметрами, но они также могут помочь увеличить частоту и/или специфичность событий редактирования генов. Таким образом, эта идея объединяет функциональные возможности сайт-специфичной нуклеазы и матриц для репарации в единый молекулярный комплекс для одновременного расщепления генома и целенаправленной репарации в сочетании со специальный инструментами доставки и способами доставки инструмента(ов) редактирования генома и/или матрицы для репарации в представляющий интерес компартмент в клетку-мишень. Таким образом, эта система обеспечивает более высокую специфичность и, таким образом, сокращает количество нецеленаправленных действий существующих подходов редактирования, что необходимо для минимизации нецеленаправленного расщепления у крупных животных, в частности, у млекопитающих, или иногда даже более сложных геномов растений.

В частности, вышеуказанные цели были достигнуты благодаря предложению, в первом аспекте, искусственного молекулярного комплекса, содержащего (a) по меньшей мере, одну сайт-специфичную нуклеазу (SSN) или ее каталитически активный фрагмент, или последовательность нуклеиновой кислоты, кодирующую ее, и непосредственно взаимодействующий с ней (b) по меньшей мере, один стыковочный

домен матрицы для репарации (RTDD) или последовательность нуклеиновой кислоты, кодирующую его, при этом, стыковочный домен матрицы для репарации выполнен с возможностью непосредственного взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты (RT); (c) при необходимости, содержащего, по меньшей мере, один взаимодействующий домен (IA), или последовательность нуклеиновой кислоты, кодирующую его, при этом, по меньшей мере, один взаимодействующий домен непосредственно взаимодействует с, по меньшей мере, одной сайт-специфичной нуклеазой или ее каталитически активным фрагментом, и, при этом, по меньшей мере, один взаимодействующий домен выполнен с возможностью предложения, по меньшей мере, одной из функциональных возможностей, выбранных из группы, состоящей из (i) взаимодействия с, по меньшей мере, одним стыковочным доменом матрицы для репарации; и/или (ii) взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты; и/или (iii) сиквенс-специфичного взаимодействия с геномной ДНК; при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит, по меньшей мере, одну часть, комплементарную, по меньшей мере, одной геномной комплементарной последовательности, и, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты выполнена с возможностью опосредования репарации последовательности ДНК-мишени.

В одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, сайт-специфичная нуклеаза или последовательность нуклеиновой кислоты, кодирующая ее, выбрана из, по меньшей мере, одной нуклеазы CRISPR, включая нуклеазы Cas или Cpf1, TALEN, ZFN, мегануклеазу, рестрикционную эндонуклеазу, включая рестрикционную эндонуклеазу класса IIS, включая эндонуклеазу FokI или ее вариант, или две сайт-специфичные никующие эндонуклеазы, или их вариант, или каталитически активный фрагмент.

В другом варианте осуществления настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, один стыковочный домен матрицы для репарации, или последовательность нуклеиновой кислоты, кодирующая его, выбрана из, по меньшей мере, одного биотина, аптамера, ДНК, РНК или белкового красителя, содержащего флуорофоры, содержащие флуоресцеин или его вариант, малеимидов или Tetraxolium (XTT), направляющей последовательности нуклеиновой кислоты, специфично выполненной с возможностью взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, стрептавидина или его варианта, в предпочтительном варианте осуществления настоящего изобретения, из мономерного стрептавидина, авидина или его варианта, аффинной метки, в предпочтительном варианте осуществления настоящего изобретения, стрептовидиновой метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (наночастицы), антикалина, белка Agrobacterium VirD2 или его домена, пикорнавируса VPg, топоизомеразы или ее домена, белка A бактериофага ФХ174, белка A* ФХ, белка VirE2 или его домена, или дигоксигенина. Другая известная система для взаимодействия – SNAP-метка, например, слитая с dCas9, как предложено компанией New England Biolabs Inc. (www.neb.com). SNAP-метка способна связывать ряд флуорофоров, биотин и другие конъюгаты. Основная цель – разрешить визуализацию, но это было бы полезно также и для привязки матрицы для репарации.

В еще одном варианте осуществления настоящего изобретения, в соответствии с вышеупомянутым первым аспектом настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, один взаимодействующий домен или последовательность

нуклеиновой кислоты, кодирующая его, выбрана из, по меньшей мере, одного ДНК-связывающего домена, стрептавидина или его варианта, в предпочтительном варианте осуществления настоящего изобретения, из мономерного стрептавидина, авидина или его варианта, аффинной метки, сигнала биотинилирования, сайта акцептора биотина, стрептовидиновой метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (наночастицы), антикалина, биотина, аптамера, ДНК, РНК или белкового красителя, содержащего флуорофоры, содержащие флуоресцеин или его вариант, малеимидов, или Tetraxolium (XTT), направляющей последовательности нуклеиновой кислоты, специфично выполненной с возможностью взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, белка Agrobacterium VirD2 или его домена, пикорнавируса VPg, топоизомеразы или ее домена, белка А бактериофага ФХ174, белка А* ФХ, белка VirE2 или его домена, или дигоксигенина.

В еще одном варианте осуществления настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, одна сайт-специфичная нуклеаза и/или, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, один взаимодействующий домен содержит, по меньшей мере, одну последовательность ядерной локализации, последовательность пластидной локализации, в предпочтительном варианте осуществления настоящего изобретения, последовательность митохондриальной локализации или последовательность хлоропластной локализации.

В другом варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит, по меньшей мере, одну концевую часть, в предпочтительном варианте осуществления настоящего изобретения, 3'-конец, при этом, эта концевая часть не взаимодействует ни с каким-либо другим компонентом искусственного молекулярного комплекса и, таким образом, выполнена с возможностью гибридизировать, по меньшей мере, одну геномную комплиментарную последовательность для опосредования репарации последовательности ДНК-мишени, и/или, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты предложена в виде плазмиды.

В еще одном варианте осуществления настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, одна сайт-специфичная нуклеаза или ее каталитически активный фрагмент, или последовательность, кодирующая ее, выбрана из нуклеазы CRISPR, в предпочтительном варианте осуществления настоящего изобретения, из нуклеазы Cas или Cpf1, или нуклеазы FokI, или ее каталитически активного фрагмента, и, по меньшей мере, один взаимодействующий домен, или последовательность, кодирующая его, выбрана из одноцепочечного вариабельного фрагмента или мономерного стрептавидина.

Кроме того, в другом варианте осуществления настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, комплекс содержит, по меньшей мере, одну направляющую последовательность нуклеиновой кислоты, представляющую, по меньшей мере, один стыковочный домен матрицы для репарации, при этом, каждая из, по меньшей мере, одной направляющей последовательности нуклеиновой кислоты содержит (i) первую часть последовательности, которая комплементарна последовательности распознавания ДНК-мишени, и (ii) вторую часть последовательности, при этом, вторая часть последовательности выполнена с возможностью взаимодействия с, по меньшей мере, одной сайт-специфичной нуклеазой, и (iii), при этом, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты физически ассоциирована с, по меньшей мере, одной

репарационной матричной последовательностью нуклеиновой кислоты, и, таким образом, образует гибридную последовательность нуклеиновой кислоты, содержащую или состоящую из, по меньшей мере, одной РНК или ДНК и, по меньшей мере, одной дополнительной последовательности нуклеиновой кислоты ДНК, и (iv), при необходимости, содержит линкерную область между, по меньшей мере, одной направляющей последовательностью нуклеиновой кислоты и, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, в предпочтительном варианте осуществления настоящего изобретения, при этом, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с направляющей последовательностью нуклеиновой кислоты на 3'-конце направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с 5'-концом направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты расположена в пределах направляющей последовательности нуклеиновой кислоты.

В другом варианте осуществления настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты содержат нуклеотидную последовательность, выбранную из нуклеотидной последовательности природного или неприродного происхождения, включая синтетическую нуклеотидную последовательность, при необходимости, содержащую остов и/или базовые модификации, при этом, направляющая последовательность нуклеиновой кислоты содержит одноцепочечную или частично одноцепочечную нуклеотидную последовательность РНК или ДНК, и, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит одноцепочечную или двухцепочечную нуклеотидную последовательность ДНК.

В еще одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, предложен искусственный молекулярный комплекс, при этом, по меньшей мере, одна сайт-специфичная нуклеаза или последовательность, кодирующая ее, и, по меньшей мере, один взаимодействующий домен, или последовательность, кодирующая его, и/или, по меньшей мере, один стыковочный домен матрицы для репарации или последовательность, кодирующая его, соединены, по меньшей мере, одним линкерным доменом.

В одном из предложенных вариантов осуществления настоящего изобретения, по меньшей мере, одна сайт-специфичная нуклеаза или ее каталитически активный фрагмент, или последовательность, кодирующая ее, выбрана независимо из группы, состоящей из полипептида Cas из *Streptococcus spp.,* включая *Streptococcus pyogenes, Streptococcus thermophilus, Staphylococcus aureus,* или *Neisseria spp.,* включая *Neisseria meningitides, Corynebacter, Sutterella, Legionella, Treponema, Filifactor, Eubacterium, Lactobacillus, Mycoplasma, Bacteroides, Flaviivola, Flavobacterium, Sphaerochaeta, Azospirillum, Gluconacetobacter, Roseburia, Parvibaculum, Nitratifractor, Mycoplasma, Campylobacter, Candidatus* Micrarchaeum acidiphilum ARMAN-1, *Parcubacteria* (База генетических данных (GenBank): APG80656.1), *Sulfolobus* spp., включая *Sulfolobus islandicus* HVE10/4 (GenBank: ADX81770.1) или REY15A (GenBank: ADX84852.1), и *Candidatus* Parvarchaeum acidiphilum ARMAN-4, полипептида Cpf1 из археи или бактерии, включая полипептид Cpf1 *Acidaminococcus spp.*, включая *Acidaminococcus sp.* BV3L6, *Lachnospiraceae spp.,* включая *бактерию Lachnospiraceae* ND2006, бактерию *Lachnospiraceae* MC2017, бактерию *Lachnospiraceae* MA2020, *Butyrivibrio proteoclasticus,* Candidatus spp., *Methanoplasma termitum, Leptospira inadai, Moraxella bovoculi* 237, бактерию Peregrinibacteria GW2011_GWA2_33_10, бактерию Parcubacteria

GW2011_GWC2_44_17, Smithella sp. SCADC, Smithella sp. SC_K08D17, *Francisella spp.,* включая *Francisella novicida* U112, *Eubacterium eligens, Prevotella spp.,* или *Porphyromonas spp.,* или нуклеазы Аргонавта из *Natronobacterium gregoryi* (GenBank: AFZ73749.1), *Microcystis aeruginosa* (Стандартная/эталонная последовательность Национального центра биотехнологической информации (NCBI): WP_012265209.1 или Стандартная/эталонная последовательность NCBI: WP_002747795.1 или Стандартная/эталонная последовательность NCBI: WP_012265209.1), *Halogeometricum pallidum* (GenBank): ELZ29017.1), *Natrialaba asiatica* (Стандартная/эталонная последовательность NCBI: WP_006111085.1J, *Natronorubrum tibetense* (Стандартная/эталонная последовательность NCBI: WP_006090832.1), *Natrinema pellirubrum* (Стандартная/эталонная последовательность NCBI: WP_006183335.1), или *Synechococcus spp.* (Стандартная/эталонная последовательность NCBI: WP_011378069.1) или вариантов и/или функциональных фрагментов, и/или их комбинаций, включая никазы, или нуклеазы, не обладающие эндонуклеолитической активностью.

Во втором аспекте настоящего изобретения, предложен искусственный молекулярный комплекс по любому из предыдущих вариантов осуществления настоящего изобретения для использования в способе лечения заболевания, при этом, заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации. Таким образом, предложен способ лечения заболевания с использованием искусственного молекулярного комплекса по любому из предыдущих пунктов, при этом заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации.

В другом аспекте, предложено растение, растительная клетка, растительный материал или производное, или его потомство, содержащее или отредактированное, по меньшей мере, одним искусственным молекулярным комплексом в соответствии с любым из предыдущих аспектов и/или вариантов осуществления настоящего изобретения.

В еще одном аспекте, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, содержащий следующие этапы: (i) предложение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки, и/или генома, содержащего, по меньшей мере, одну геномную комплементарную последовательность и, по меньшей мере, одну последовательность ДНК-мишени в интересующей геномной области; (ii) предложение, по меньшей мере, одного искусственного молекулярного комплекса, как определено в любом из предыдущих аспектов и/или вариантов осуществления настоящего изобретения; (iii) контактирование, по меньшей мере, одного искусственного молекулярного комплекса с, по меньшей мере, одной последовательностью ДНК-мишени в подходящих условиях для достижения (a) взаимодействия, по меньшей мере, одной сайт-специфичной нуклеазы с, по меньшей мере, одной последовательностью ДНК-мишени; и (b) комплементарного спаривания оснований, по меньшей мере, одной последовательности репарационной матричной нуклеиновой кислоты с, по меньшей мере, одной геномной комплементарной последовательностью для достижения распознавания, по меньшей мере, одной комплементарной последовательности и индукции, по меньшей мере, одного разрыва ДНК, по меньшей мере, одной сайт-специфичной нуклеазой, при этом, по меньшей мере, одна последовательность репарационной матричной нуклеиновой кислоты направляет репарацию, направляемую гомологией, в сайте, по меньшей мере, одной последовательности ДНК-мишени; и (iv) получение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки, и/или генома, содержащего модификацию, по меньшей мере, одной последовательности ДНК-мишени.

В одном варианте осуществления вышеупомянутого аспекта настоящего изобретения, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, при этом, по меньшей мере, одна последовательность репарационной матричной нуклеиновой кислоты и/или, по меньшей мере, один стыковочный домен матрицы для репарации искусственного молекулярного комплекса предложен/предложены для, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки и/или генома независимо от, по меньшей мере, одной сайт-специфичной нуклеазы, по меньшей мере, одного молекулярного комплекса, и, по меньшей мере, один искусственный молекулярный комплекс собирается или частично собирается, в пределах, по меньшей мере, одного прокариотического или эукариотического, или вирусного генома и/или клетки.

В другом варианте осуществления вышеупомянутого аспекта настоящего изобретения, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, при этом, по меньшей мере, один искусственный молекулярный комплекс представляет собой собранный *ex vivo* искусственный молекулярный комплекс.

В другом варианте осуществления вышеупомянутого аспекта настоящего изобретения, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, при этом, по меньшей мере, одна эукариотическая клетка является растительной клеткой, в предпочтительном варианте осуществления настоящего изобретения, растительной клеткой из растения, выбранного из группы, состоящей из *Hordeum vulgare, Hordeum bulbusom, Sorghum bicolor, Saccharum officinarium, Zea spp.,* включая *Zea mays, Setaria italica, Oryza minuta, Oryza sativa, Oryza australiensis, Oryza alta, Triticum aestivum, Triticum durum, Secale cereale,* Triticale, *Malus domestica, Brachypodium distachyon, Hordeum marinum, Aegilops tauschii, Daucus glochidiatus, Beta spp.,* включая *Beta vulgaris, Daucus pusillus, Daucus muricatus, Daucus carota, Eucalyptus grandis, Nicotiana sylvestris, Nicotiana tomentosiformis, Nicotiana tabacum, Nicotiana benthamiana, Solarium lycopersicum, Solarium tuberosum, Coffea canephora, Vitis vinifera, Erythrante guttata, Genlisea aurea, Cucumis sativus, Marus notabilis, Arabidopsis arenosa, Arabidopsis lyrata, Arabidopsis thaliana, Crucihimalaya himalaica, Crucihimalaya wallichii, Cardamine nexuosa, Lepidium virginicum, Capsella bursa pastoris, Olmarabidopsis pumila, Arabis hirsute, Brassica napus, Brassica oleracea, Brassica rapa, Raphanus sativus, Brassica juncacea, Brassica nigra, Eruca vesicaria subsp. sativa, Citrus sinensis, Jatropha curcas, Populus trichocarpa, Medicago truncatula, Cicer yamashitae, Cicer bijugum, Cicer arietinum, Cicer reticulatum, Cicer judaicum, Cajanus cajanifolius, Cajanus scarabaeoides, Phaseolus vulgaris, Glycine max, Gossypium sp., Astragalus sinicus, Lotus japonicas, Torenia fournieri, Allium cepa, Allium fistulosum, Allium sativum, Helianthus annuus, Helianthus tuberosus* и *Allium tuberosum,* или любой сорт или подвид, принадлежащий одному из вышеупомянутых растений.

В другом варианте осуществления настоящего изобретения, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, при этом, модификация, по меньшей мере, одной последовательности ДНК-мишени вызывает редактирование признака, выбранного из группы, состоящей из повышения урожайности, толерантности к абиотическому стрессу, включая стресс, вызванный засухой, осмотический стресс, тепловой стресс, стресс от холода, оксидативный стресс, стресс под действием тяжёлых металлов, стресс, вызванный засолением или переувлажнением, толерантности к биотическому стрессу, включая толерантность к насекомым, толерантность к бактериям, толерантность к вирусам, толерантность к грибкам или толерантность к нематодам, устойчивости к гербицидам, включая глифосат, глуфосинат, ингибиторы ALS и дикамбу, устойчивости к полеганию, времени цветения, устойчивости к осыпанию, цвета семян, состава эндосперма, содержания питательных веществ или метаболической

инженерии, включая редактирование генома, для обеспечения подхода к молекулярному фармингу в, по меньшей мере, одной растительной клетке.

Предложен еще один способ модификации, по меньшей мере, одной последовательности ДНК-мишени, дополнительно содержащий следующий этап: (v) идентификация и/или выбор, по меньшей мере, одного прокариотического, эукариотического или вирусного генома и/или клетки, содержащей модификацию в, по меньшей мере, одной последовательности ДНК-мишени.

В еще одном аспекте, предложен способ получения растения или растительной клетки, содержащий следующие этапы: (i) осуществление способа по любому из вышеупомянутых аспектов и/или вариантов осуществления настоящего изобретения, при этом, по меньшей мере, одна эукариотическая клетка представляет собой растительную клетку; (ii) получение, по меньшей мере, одного растения или его потомства из, по меньшей мере, одной растительной клетки из этапа (i); (iii) при необходимости: определение модификации в, по меньшей мере, одной последовательности ДНК-мишени в, по меньшей мере, одной клетке, по меньшей мере, одного растения или его потомства.

В одном варианте осуществления настоящего изобретения, предложен способ изготовления растения или растительной клетки, при этом, по меньшей мере, одно растение или растительная клетка выбраны из однодольного или двудольного растения, при этом, в предпочтительном варианте осуществления настоящего изобретение, растение выбрано из группы, состоящей из *Hordeum vulgare, Hordeum bulbusom, Sorghum bicolor, Saccharum officinarium, Zea spp.,* включая *Zea mays, Setaria italica, Oryza minuta, Oryza sativa, Oryza australiensis, Oryza alta, Triticum aestivum, Triticum durum, Secale cereale,* Triticale, *Malus domestica, Brachypodium distachyon, Hordeum marinum, Aegilops tauschii, Daucus glochidiatus, Beta spp.,* включая *Beta vulgaris, Daucus pusillus, Daucus muricatus, Daucus carota, Eucalyptus grandis, Nicotiana sylvestris, Nicotiana tomentosiformis, Nicotiana tabacum, Nicotiana benthamiana, Solarium lycopersicum, Solarium tuberosum, Coffea canephora, Vitis vinifera, Erythrante guttata, Genlisea aurea, Cucumis sativus, Marus notabilis, Arabidopsis arenosa, Arabidopsis lyrata, Arabidopsis thaliana, Crucihimalaya himalaica, Crucihimalaya wallichii, Cardamine nexuosa, Lepidium virginicum, Capsella bursa pastoris, Olmarabidopsis pumila, Arabis hirsute, Brassica napus, Brassica oleracea, Brassica rapa, Raphanus sativus, Brassica juncacea, Brassica nigra, Eruca vesicaria subsp. sativa, Citrus sinensis, Jatropha curcas, Populus trichocarpa, Medicago truncatula, Cicer yamashitae, Cicer bijugum, Cicer arietinum, Cicer reticulatum, Cicer judaicum, Cajanus cajanifolius, Cajanus scarabaeoides, Phaseolus vulgaris, Glycine max, Gossypium sp., Astragalus sinicus, Lotus japonicas, Torenia fournieri, Allium cepa, Allium fistulosum, Allium sativum, Helianthus annuus, Helianthus tuberosus* и *Allium tuberosum,* или любой сорт или подвид, принадлежащий одному из вышеупомянутых растений.

В другом аспекте, предложено использование, по меньшей мере, одного искусственного молекулярного комплекса в соответствии с любым из вышеупомянутых аспектов и/или вариантов осуществления настоящего изобретения для геномной инженерии в прокариотической, эукариотической или вирусной клетке, геноме или организме, в предпочтительном варианте осуществления настоящего изобретения, в растительной клетке или организме.

Другие аспекты и варианты осуществления настоящего изобретения могут быть получены из последующего подробного описания, чертежей, перечня последовательностей, а также из прилагаемой совокупности пунктов формулы изобретения.

Краткое описание чертежей

На Фигуре 1 от A до D (Фигура 1 A-D) показаны неограничивающие примеры возможных конфигураций и различные способы ассоциации для различных гибридных последовательностей

нуклеиновой кислоты РНК-ДНК или последовательностей нуклеиновой кислоты ДНК-ДНК, причем, часть направляющей нуклеиновой кислоты, представляет собой, по меньшей мере, один стыковочный домен матрицы для репарации (RTDD) и/или, по меньшей мере, один взаимодействующий домен (IA) по настоящему изобретению. (A) Нековалентная ассоциация путем спаривания оснований по Уотсону-Крику одноцепочечной матрицы для репарации (RT) (ssДНК) с молекулой направляющей нуклеиновой кислоты, содержащей последовательность, функционирующую как sgРНК или tracrРНК, или как gДНК. (B) Ковалентная ассоциация одноцепочечной RT (ssДНК) с молекулой направляющей нуклеиновой кислоты. Эта форма может быть получена путем последовательного синтеза молекулы направляющей нуклеиновой кислоты RTDD и частей RT в виде единой молекулы или путем лигирования отдельных частей с образованием единой молекулы. (C) Нековалентная ассоциация двухцепочечной RT (dsДНК) с молекулой направляющей нуклеиновой кислоты. (D) Ковалентная ассоциация двухцепочечной RT (dsДНК) с молекулой направляющей нуклеиновой кислоты.

На Фигуре 2 от A до C (Фигура 2 A-C) показаны неограничивающие примеры возможных местоположений, в которых RT может быть присоединена или ассоциирована с молекулой направляющей нуклеиновой кислоты в качестве, по меньшей мере, одного RTDD и/или, по меньшей мере, одного IA по настоящему изобретению. (A) Ковалентная или нековалентная ассоциация одноцепочечной или двухцепочечной RT с 3'-концом молекулы направляющей нуклеиновой кислоты. (B) Ковалентная или нековалентная ассоциация одноцепочечной или двухцепочечной RT с 5'-концом молекулы направляющей нуклеиновой кислоты. (C) Ковалентная или нековалентная ассоциация одноцепочечной или двухцепочечной RT изнутри с молекулой направляющей нуклеиновой кислоты. Участок матрицы для репарации (RT) показан в белом цвете на этом и всех последующих Фигурах.

На Фигуре 3 от A до E (Фигура 3 A-E) показан неограничивающий пример поэтапного введения редактирования в представляющую интерес геномную последовательность посредством нуклеазного комплекса сайт-специфичной нуклеазы (SSN), раскрытого в настоящем документе, используя в качестве примера один вариант осуществления ковалентной ассоциации RT с 3'-концом молекулы направляющей нуклеиновой кислоты. (A) Схематическое изображение молекулы направляющей нуклеиновой кислоты в комплексе с SSN, например, NgAgo, Cas, включая Cas9,CasX или CasY, или Cpf1. (B) Схематическое изображение комплекса, связанного с ДНК-мишенью (геномная ДНК (gДНК)), и указание сайтов разрезания (черные треугольники). (C) Схематическое изображение расщепленной ДНК-мишени. (D) Схематическое изображение расщепленной ДНК-мишени, высвобожденной посредством SSN, и взаимодействующей с матрицей для репарации (RT) путем комплементарного спаривания оснований по Уотсону-Крику. (E) Схематическое изображение репарированного сайта-мишени (gДНК), включая редактирования, скопированные из RT во время гомологичной рекомбинации. Участок матрицы для репарации (RT) показан в белом цвете на всех Фигурах.

На Фигуре 4 от A до C (Фигура 4 A-C) показан неограничивающий пример конструкции белка слияния эндонуклеазы, направляемой нуклеиновой кислотой, в виде SSN и белка или домена белка в виде взаимодействующего домена (IA) со способностью непосредственно или опосредованно связывать матрицу для репарации (RT). (A) Схематическое изображение указанного белка слияния в виде комплекса с ДНК-мишенью. (B) Схематическое изображение комплекса после введения двухцепочечного разрыва. Эндонуклеаза, направляемая нуклеиновой кислотой, отделяется от ДНК-мишени. Слитая матрица для репарации нуклеиновой кислоты образует комплекс с целевой областью способом на основе гомологии. (C) Схематическое изображение ДНК-мишени после того, как произошла гомологически направленная

репарация. Примечательно, что в представленном подходе используется более одного RTDD для того, чтобы повысить точность комплекса геномной инженерии.

На Фигуре 5 на левой панели показана очищенная нуклеаза (в данном случае, нуклеаза CRISPR), которая была слита с RTDD1 и экспрессирована в *E. coli*. Процесс запускали в денатурирующем геле в присутствии додецилсульфата натрия (SDS) с непрерывным градиентом (4-10%); и показано количество и чистота белка. Белок был окрашен в этом геле. На правой панели показана привязка. Это – 4% неденатурирующий акриламидный гель (голубой нативный электрофорез в полиакриламидном геле – Blue Native PAGE), и в настоящем документе ДНК окрашивается красителем GelRed. Слева показана меченая флуоресцеином (FAM) (RTDD2) матрица для репарации, которую инкубировали в нуклеазном буферном растворе либо без нуклеазы RTDD1, либо с нуклеазой RTDD1. Если белок присутствовал, то происходила привязка, как видно по ДНК, которую определяют на более высоком уровне молекулярной массы (стрелка).

На Фигуре 6 в строке 1 показана часть последовательности дикого типа сайта-мишени (полноразмерная последовательность представляет собой SEQ ID NO: 47), в строке 2 и строке 3 приведены примеры возникновения инсерционно-делеционной мутации (INDEL) (полноразмерные последовательности представляют собой SEQ ID NO: 48 и 49), в строке 4 показано событие репарации путём гомологичной рекомбинации (HDR) (полноразмерные последовательности представляют собой SEQ ID NO: 50), и в строке 5 показана матрица для репарации (полноразмерные последовательности представляют собой SEQ ID NO: 51).

На Фигуре 7 показано сравнение эффективности нормализованной HDR, когда матрица для репарации отсутствует (левый столбец), и когда она привязана к нуклеазе (правый столбец).

Определения

Следует отметить, что, используемые в настоящем документе формы единственного числа «a», «an» и «the» включают ссылки на множественное число, если иное явно не следует из контекста. Например, ссылка на один компонент также подразумевает композиции, состоящие из множества компонентов. Ссылки на композицию, содержащую элемент «a», включают и другие элементы, помимо названного. Иными словами, термины «a», «an» и «the» не обозначают ограничение количества, а скорее обозначают наличие, «по меньшей мере, одного» из указанных элементов. Считается, что каждый термин имеет самое широкое значение, как его понимают специалисты в данной области техники, и включает все технические эквиваленты, которые действуют аналогичным образом для достижения аналогичной цели.

Диапазоны в настоящем документе могут быть выражены следующим образом: от «примерно» или «приблизительно», или «по существу» одного конкретного значения и/или до «примерно» или «приблизительно», или «по существу» другого конкретного значения. Когда такой диапазон выражен, другие примерные варианты осуществления настоящего изобретения включают: от одного конкретного значения и/или до другого конкретного значения. Кроме того, термин «примерно» означает – в пределах допустимого диапазона погрешности для конкретного значения, определенного одним из специалистов обычной квалификации в данной области техники, которое будет частично зависеть от того, как измеряется или определяется значение, то есть, от ограничений системы измерений. Например, «примерно» может означать – в пределах допустимого стандартного отклонения, согласно практике в данной области техники. В альтернативном варианте осуществления настоящего изобретения, «примерно» может означать диапазон до ±20%, в предпочтительном варианте осуществления настоящего изобретения – до ±10%, в более предпочтительном варианте осуществления настоящего изобретения – до ±5% и в еще более предпочтительном варианте осуществления настоящего изобретения – до ±1% от данного значения. Кроме

того, в частности, в отношении биологических систем или процессов, этот термин может означать – в пределах порядка возрастания, в предпочтительном варианте осуществления настоящего изобретения – в 2-х кратном размере от значения. В тех случаях, когда конкретные значения описаны в заявке и формуле изобретения, если не указано иное, термин «примерно» является предполагаемым и в данном контексте означает – в пределах допустимого диапазона погрешности для конкретного значения.

Термин «содержащий» или «содержащий в себе», или «включающий» обозначает, что, по меньшей мере, названное соединение, элемент, частица или этап способа присутствует в композиции или изделии, или способе, но не исключает присутствия других соединений, материалов, частиц, этапов способа, даже если такие другие соединения, материал, частицы, этапы способа имеют ту же функцию, что и названное.

Используемый в настоящем документе термин «нуклеиновая кислота» означает полинуклеотид и включает одноцепочечный или двухцепочечный полимер дезоксирибонуклеотидных или рибонуклеотидных оснований. Нуклеиновые кислоты могут также включать фрагменты и модифицированные нуклеотиды. Таким образом, термины «полинуклеотид», «последовательность нуклеиновой кислоты», «нуклеотидная последовательность» или «фрагмент нуклеиновой кислоты» используются взаимозаменяемо для обозначения полимера РНК и/или ДНК, которая является одноцепочечной или двухцепочечной, при необходимости, содержащей синтетические, неприродные или измененные нуклеотидные основания. Нуклеотиды (обычно встречающиеся в своей 5'-монофосфатной форме) обозначают одним буквенным обозначением следующим образом: «A» – аденозин или дезоксиаденозин (для РНК или ДНК, соответственно), «C» – цитозин или дезоксицитозин, «G» – гуанозин или дезоксигуанозин, «U» – уридин, «T» – дезокситимидин, «R» – пурины (A или G), «Y» – пиримидины (C или T), «K» – G или T, «H» – A или C или T, «I» – инозин и «N» – любой нуклеотид. Нуклеиновая кислота может содержать нуклеотиды. Нуклеиновая кислота может быть экзогенной или эндогенной для клетки. Нуклеиновая кислота может существовать в бесклеточной среде. Нуклеиновая кислота может представлять собой геном или его фрагмент, но нуклеиновая кислота не обязательно должна кодировать ген. Нуклеиновая кислота может представлять собой ДНК, нуклеиновая кислота может представлять собой РНК. Нуклеиновая кислота может содержать, по меньшей мере, один аналог (например, измененный остов, сахар или нуклеотидное основание). Некоторые неограничивающие примеры аналогов включают: 5-бромурацил, пептидную нуклеиновую кислоту, ксенонуклеиновую кислоту, морфолиновые олигонуклеотиды, закрытые нуклеиновые кислоты, гликолевые нуклеиновые кислоты, треозо-нуклеиновые кислоты, дидезоксинуклеотиды, кордицепин, 7-деаза-ГТФ, флуорофоры (например, родамин или флуоресцеин, сцепленный с сахаром), тиолсодержащие нуклеотиды, нуклеотиды, сцепленные с биотином, аналоги на флуоресцентной основе, CpG-островки, метил-7-гуанозин, метилированные нуклеотиды, инозин, тиоуридин, 20 псевдоурдин, дигидроуридин, квеуозин и виозин. Нуклеиновая кислота по настоящему изобретению может быть соединена фосфидиэфирными сцеплениями, например, природного происхождения, или фосфоротиоатными сцеплениями, или смесью обоих.

Термины «направляющая РНК», «gРНК» или «одноцепочечная направляющая РНК», или «sgРНК» используются взаимозаменяемо в данном документе, и либо относятся к синтетическому слиянию CRISPR РНК (crРНК) и трансактивирующей crРНК (tracrРНК), либо этот термин относится к молекуле одноцепочечной РНК, состоящей только из crРНК и/или tracrРНК, или этот термин относится к gРНК, отдельно содержащий фрагмент молекулы crРНК или tracrРНК. Таким образом, фрагменты молекулы tracr и crРНК не обязательно должны присутствовать на одной ковалентно присоединенной молекуле РНК, но они

также могут присутствовать на двух отдельных молекулах РНК, которые могут ассоциироваться, или они могут быть ассоциированы посредством нековалентного или ковалентного взаимодействия для предложения gРНК по настоящему изобретению. Термины «gДНК» или «sgДНК», или «направляющая ДНК» используются в настоящем документе взаимозаменяемо, и оба относятся к молекуле нуклеиновой кислоты, взаимодействующей с нуклеазой Аргонавтом. Как gРНК, так и gДНК, раскрытые в настоящем документе, называются «направляющей нуклеиновой кислотой(ами)» или «гидовой нуклеиновой кислотой(ами)» вследствие их способности взаимодействовать с сайт-специфичной нуклеазой и содействовать нацеливанию указанной сайт-специфичной нуклеазы на геномный сайт-мишень.

Термины «редактирование генов», «редактирование генома» и «геномная инженерия» используются в настоящем документе взаимозаменяемо и относятся к стратегиям и методам целенаправленной, специфичной модификации любой генетической информации или генома живого организма. Как таковые, термины содержат редактирование генов, но также и редактирование областей, отличных от областей генома, кодирующих гены. Они также включают редактирование или инженерию ядра (если имеется), а также другую генетическую информацию клетки. Кроме того, термины «редактирование генома» и «геномная инженерия» также содержат эпигенетическое редактирование или инженерию, то есть, целевую модификацию, например, метилирование, модификацию гистонов или некодирующих РНК, возможно, вызывающих наследственные изменения в экспрессии генов.

Термины «нуклеотид» и «нуклеиновая кислота», применительно к последовательности или молекуле, используются в настоящем документе взаимозаменяемо и относятся к одноцепочечной или двухцепочечной ДНК или РНК природного или синтетического происхождения. Таким образом, термин «нуклеотидная последовательность» используется для любой последовательности ДНК или РНК независимо от ее длины, таким образом, этот термин содержит любую нуклеотидную последовательность, содержащую, по меньшей мере, один нуклеотид, а также и любой вид более крупного олигонуклеотида или полинуклеотида. Таким образом, термин(ы) относится к последовательностям природных и/или синтетических дезоксирибонуклеиновых кислот (ДНК) и/или рибонуклеиновых кислот (РНК), которые, при необходимости, могут содержать аналог синтетической нуклеиновой кислоты. Нуклеиновая кислота по настоящему изобретению, при необходимости, может быть кодон-оптимизированной. Оптимизация кодонов подразумевает, что использование кодонов ДНК или РНК адаптировано к использованию кодонов представляющей интерес клетки или организма для улучшения скорости транскрипции указанной рекомбинантной нуклеиновой кислоты в представляющей интерес клетке или организме. Специалисту в данной области техники хорошо известно, что нуклеиновая кислота-мишень может быть модифицирована в одном положении из-за вырождения кодонов, при этом, эта модификация все равно приведет к той же аминокислотной последовательности в этом положении после трансляции, которая достигается оптимизацией кодонов с учетом видоспецифичного использования кодонов целевой клетки или организма. Последовательности нуклеиновых кислот по настоящей заявке могут нести оптимизацию специфичных кодонов для следующего неограничивающего списка организмов: *Hordeum vulgare, Sorghum bicolor, Secale cereale,* Triticale, *Saccharum officinarium, Zea mays, Setaria italic, Oryza sativa, Oryza minuta, Oryza australiensis, Oryza a/ta, Triticum aestivum, Triticum durum,* Triticale, *Hordeum bulbosum, Brachypodium distachyon, Hordeum marinum, Aegilops tauschii, Malus domestica, Beta vulgaris, Helianthus annuus, Daucus glochidiatus, Daucus pusillus, Daucus muricatus, Daucus carota, Eucalyptus grandis, Erythranthe guttata, Genlisea aurea, Nicotiana sylvestris, Nicotiana tabacum, Nicotiana tomentosiformis, Nicotiana benthamiana, Solanum /ycopersicum, Solanum tuberosum, Coffea canephora, Vitis vinifera, Cucumis sativus, Marus notabilis,*

Arabidopsis thaliana, Arabidopsis lyrata, Arabidopsis arenosa, Crucihimalaya himalaica, Crucihimalaya wallichii, Cardamine flexuosa, Lepidium virginicum, Capsella bursa-pastoris, Olmarabidopsis pumila, Arabis hirsuta, Brassica napus, Brassica oleracea, Brassica rapa, Brassica juncacea, Brassica nigra, Raphanus sativus, Eruca vesicaria sativa, Citrus sinensis, Jatropha curcas, Glycine max, Gossypium ssp., Populus trichocarpa, Mus musculus, Rattus norvegicus или *Homo sapiens.*

Используемый в настоящем документе термин «ненативный» или «неприродного происхождения», или «искусственный» может относиться к последовательности нуклеиновой кислоты или полипептида, или к любой другой биомолекуле, такой как, биотин или флуоресцеин, которая не обнаружена в нативной нуклеиновой кислоте или белке. Термин «ненативный» может относится к афинным меткам. Термин «ненативный» может относится к слияниям. Термин «ненативный» может относиться к последовательности нуклеиновой кислоты или полипептида природного происхождения, которая содержит мутации, вставки и/или делеции. Ненативная последовательность может проявлять и/или кодировать активность (например, ферментную активность, активность метилтрансферазы, активность ацетилтрансферазы, активность киназы, активность убиквитина и так далее), которая также может проявляться посредством последовательности нуклеиновой кислоты и/или полипептида с которой слита ненативная последовательность. Ненативная последовательность нуклеиновой кислоты или полипептида может быть сцеплена с последовательностью нуклеиновой кислоты или полипептида (или их варианта) природного происхождения посредством генной инженерии для генерации химерной последовательности нуклеиновой кислоты и/или полипептида, кодирующей химерную нуклеиновую кислоту и/или полипептид. Ненативная последовательность может относиться к 3'-концевой гибридизируемой удлиненной последовательности.

Используемый в настоящем документе термин «нуклеотид», как правило, может относиться к комбинации: основание-сахар-фосфат. Нуклеотид может содержать синтетический нуклеотид. Нуклеотид может содержать аналог синтетического нуклеотида. Нуклеотиды могут представлять собой мономерные единицы последовательности нуклеиновой кислоты (например, дезоксирибонуклеиновой кислоты (ДНК) и рибонуклеиновой кислоты (РНК)). Термин «нуклеотид» может включать рибонуклеозидтрифосфаты, аденозинтрифосфат (ATP), уридинтрифосфат (UTP), цитозинтрифосфат (CTP), гуанозинтрифосфат (GTP), инозинтрифосфат (ITP) и дезоксирибонуклеозидтрифосфаты, такие как dATP, dCTP, diTP, dUTP, dGTP, dTTP или их производные. Такие производные могут включать и этим не ограничиваться, например, [aS]dATP, 7-деаза-dGTP и 7-деаза-dATP, а также нуклеотидные производные, которые придают устойчивость нуклеазам на молекуле нуклеиновой кислоты, содержащей их. Термин «нуклеотид», используемый в настоящем документе, может относиться к дидезоксирибонуклеозидтрифосфатам (ddNTP) и их производным. Иллюстративные примеры дидезоксирибонуклеозидтрифосфатов могут включать, но этим не ограничиваются, ddATP, ddCTP, ddGTP, ddITP и ddTTP. Нуклеотид может быть немеченым или меченым с возможностью определения посредством хорошо известных методов. Мечение также может быть выполнено с помощью квантовых точек. Определяемые метки могут включать, например, радиоактивные изотопы, флуоресцентные метки, хемилюминесцентные метки, биолюминесцентные метки и ферментные метки. Флуоресцентные метки нуклеотидов могут включать, но этим не ограничиваются, флуоресцеин, 5-карбоксифлуоресцеин (FAM), 2'7'-5-диметокси-4'5-дихлоро-6-карбоксифлуоресцеин (JOE), родамин, 6-карбоксиродамин (R6G), N,N,N',N'- тетраметил-6-карбоксиродамин (TAMRA), 6-карбокси-X-родамин (ROX), 4-(4'-диметиламинофенилазо)бензойная кислота (DABCYL), красители: Cascade Blue, Oregon Green, Texas Red, цианин и 5-(2'-аминоэтил)аминонафталин-1-сульфокислота (EDANS).

Используемый в настоящем документе термин «слияние» может относиться к белку и/или нуклеиновой кислоте, содержащей, по меньшей мере, одну нативную последовательность (например, фрагмент молекулы). Слияние может быть на N-терминальном или С-терминальном конце модифицированного белка, или на них обоих, или в пределах молекулы в виде отдельного домена. В отношении молекул нуклеиновых кислот, молекула слияния может быть присоединена на 5'- или 3'-конце, или в любом подходящем положении между ними. Слияние может быть транскрипционным и/или трансляционным слиянием. Слияние может содержать, по меньшей мере, одну одинаковую ненативную последовательность. Слияние может содержать, по меньшей мере, одну неодинаковую ненативную последовательность. Слияние может представлять собой химеру. Слияние может содержать аффинную метку нуклеиновой кислоты. Слияние может содержать штрих-код. Слияние может содержать пептидную аффинную метку. Слияние может предложить субклеточную локализацию Аргонавта (например, сигнал ядерной локализации (NLS) для нацеливания на ядро, сигнал митохондриальной локализации для нацеливания на митохондрию, сигнал хлоропластной локализации для нацеливания на хлоропласт, сигнал удержания в эндоплазматическом ретикулуме (ER) 15 и тому подобное). Слияние может предложить ненативную последовательность (например, аффинную метку), которую можно использовать для отслеживания или очистки. Слияние может представлять собой малую молекулу, такую как, биотин или краситель, такой как, красители Alexa Fluor, краситель Цианин-3, краситель Цианин-5. Слияние может предложить повышенную или пониженную стабильность. В некоторых вариантах осуществления настоящего изобретения, слияние может содержать определяемую метку, включая фрагмент молекулы, который может предложить определяемый сигнал. Подходящие определяемые метки и/или фрагменты молекулы, которые могут предложить определяемый сигнал, могут включать, но этим не ограничиваются, фермент, радиоизотоп, участника специфичной связывающейся пары; флуорофор; флуоресцентный репортер или флуоресцентный белок; квантовую точку; и тому подобное. Слияние может содержать участника пары белков для резонансного переноса энергии флуоресценции (FRET) или донорную/акцепторную пару флуорофорной/квантовой точки. Слияние может содержать фермент. Подходящие ферменты могут включать, но этим не ограничиваются, пероксидазу хрена, люциферазу, бета-25-галактозидазу и тому подобное. Слияние может содержать флуоресцентный белок. Подходящие флуоресцентные белки могут включать, но этим не ограничиваются, зеленый флуоресцентный белок (GFP) (например, GFP из Aequoria victoria, флуоресцентные белки из Anguilla japonica, или их мутант, или производное), красный флуоресцентный белок, желтый флуоресцентный белок, желто-зеленый флуоресцентный белок (например, mNeonGreen, полученный из тетрамерного флуоресцентного белка из головолохордового Branchiostoma lanceolatum), любой из множества флуоресцентных и окрашенных белков. Слияние может содержать наночастицу. Подходящие наночастицы могут включать флуоресцентные или люминесцентные наночастицы, а также магнитные наночастицы или наноалмазы, при необходимости, сцепленные с наночастицей. Любое оптическое или магнитное свойство, или характеристику наночастицы(наноцастиц) можно определить. Слияние может содержать геликазу, нуклеазу (например, FokI), эндонуклеазу, экзонуклеазу (например, 5'-экзонуклеазу и/или 3'-экзонуклеазу), лигазу, никазу, нуклеазу-геликазу (например, Cas3), ДНК-метилтрансферазу (например, Dam), или ДНК-деметилазу, гистонметилтрансферазу, гистондеметилазу, ацетилазу (включая, например, и этим не ограничиваясь, гистонацетилазу), деацетилазу (включая, например, и этим не ограничиваясь, гистондеацетилазу), фосфатазу, киназу, (ко)активатор транскрипции, (ко)фактор транскрипции, субъединицу РНК-полимеразы, репрессор транскрипции, ДНК-связывающий белок, ДНК-структурирующий белок, длинную

некодирующую РНК, репарационный белок ДНК (например, белок, участвующий в репарации или одноцепочечных, и/или двухцепочечных разрывов, например, белки, участвующие в эксцизионной репарация оснований, эксцизионной репарации нуклеотидов, репарации ошибочно спаренных оснований, NHEJ, HR, микрогомологически опосредованном соединении концов (MMEJ) и/или в другом негомологичном соединении концов (ANHEJ), таком как, например, и этим не ограничиваясь, регуляторы HR и сигналы сборки комплекса HR), маркерный белок, репортерный белок, флуоресцентный белок, лиганд-связывающий белок (например, mCherry или белок, связывающий тяжелые металлы), сигнальный пептид (например, Tat-сигнальная последовательность), нацеливающий белок или пептид, последовательность субклеточной локализации (например, последовательность ядерной локализации, последовательность хлоропластной локализации) и/или эпитоп антитела, или любую их комбинацию.

Термин «каталитически активный фрагмент», используемый в настоящем документе, применительно к аминокислотным последовательностям, обозначает сердцевинную последовательность, полученную из данной матричной аминокислотной последовательности, или последовательность нуклеиновой кислоты, кодирующую последовательность, содержащую весь или часть активного сайта матричной последовательности при условии, что полученный каталитически активный фрагмент все еще обладает активностью, характеризующей матричную последовательность, за которую отвечает активный сайт нативного фермента или его варианта. Указанные модификации подходят для получения менее громоздких аминокислотных последовательностей, все еще обладающих той же активностью, что и матричная последовательность, что делает каталитически активный фрагмент более универсальным или более стабильным инструментом, являющимся стерически менее требовательным.

Термин «вариант» любой сайт-специфичной нуклеазы, раскрытой в настоящем документе, представляет собой молекулу, содержащую, по меньшей мере, одну мутацию, делецию или вставку, по сравнению с сайт-специфичной нуклеазой дикого типа для изменения активности нуклеазы дикого типа природного происхождения. Термин «вариант» может, в качестве неограничивающего примера, представлять собой каталитически неактивный Cas9 (dCas9) или сайт-специфичную нуклеазу, которая была модифицирована для функционирования в качестве никазы.

Термин «конструкция доставки» или «вектор доставки», используемый в настоящем документе, применяется к любому биологическому или химическому средству, используемому в качестве груза для транспортировки нуклеиновой кислоты, включая гибридную нуклеиновую кислоту, содержащую РНК и ДНК, и/или интересующую аминокислотную последовательность, в клетку-мишень, в предпочтительном варианте осуществления настоящего изобретения, в эукариотическую клетку. Термин «конструкция доставки» или «вектор», используемый в настоящем документе, применяется к средству транспортировки для доставки генетической или рекомбинантной конструкции по настоящему изобретению в целевую клетку, ткань, орган или организм. Таким образом, вектор может содержать последовательности нуклеиновых кислот, при необходимости, содержащие последовательности, такие как регуляторные последовательности или последовательности локализации для доставки, непосредственно или опосредованно, в представляющую интерес клетку-мишень или в структуру-мишень растения в желаемом клеточном компартменте растения. Вектор также может быть использован для введения аминокислотной последовательности или рибонуклео-молекулярного комплекса в клетку-мишень или в структуру-мишень. Как правило, вектор, используемый в настоящем документе, может представлять собой плазмидный вектор. Кроме того, в соответствии с некоторыми предпочтительными вариантами осуществления настоящего изобретения, осуществляется непосредственное введение конструкции или последовательности, или

интересующего комплекса. Термин «непосредственное введение» подразумевает, что желаемая клетка-мишень или структура-мишень, содержащая последовательность ДНК-мишени, которая должна быть модифицирована по настоящему изобретению, непосредственно трансформируется или трансдуцируется, или трансфицируется в интересующую специфичную клетку-мишень, где материал, доставленный вектором доставки, будет оказывать свое действие. Термин «опосредованное введение» подразумевает, что введение осуществляется в структуру, например, в клетки листьев или клетки органов, или тканей, которые сами по себе не представляют фактическую клетку-мишень или структуру-мишень, представляющую интерес для трансформации, но эти структуры служат основой для системного распространения и переноса вектора, в предпочтительном варианте осуществления настоящего изобретения, содержащего генетическую конструкцию по настоящему изобретению для фактической структуры-мишени, например, для меристематической клетки или ткани, или стволовой клетки, или ткани. В случае использования термина «вектор» в контексте трансфекции аминокислотных последовательностей и/или нуклеиновых последовательностей, включая последовательности гибридных нуклеиновых кислот, в клетку-мишень, термин «вектор» подразумевает подходящие агенты для пептидной или белковой трансфекции, такие как, например, смеси ионных липидов, проникающие в клетки пептиды (CPP) или бомбардировка частицами. В контексте введения материала нуклеиновой кислоты, термин «вектор» не только подразумевает плазмидные векторы, но также и подходящие материалы-носители, которые могут служить основой для введения нуклеиновой кислоты и/или доставки аминокислотной последовательности в представляющую интерес клетку-мишень, например, посредством бомбардировки частицами. Указанный материал-носитель содержит, среди прочего, частицы золота или вольфрама. Наконец, термин «вектор» также подразумевает использование вирусных векторов для введения, по меньшей мере, одной генетической конструкции по настоящему изобретению, такой как, например, модифицированные вирусы, например, полученные из следующих штаммов вирусов: аденовирусные или адено-ассоциированные вирусные (AAV) векторы, лентивирусные векторы, вирус простого герпеса (HSV-1), вирус коровьей оспы, вирус Сендай, вирус Синдбис, альфавирусы леса Семлики, вирус Эпштейна-Барр (EBV), вирус полосатости кукурузы (MSV), вирус ложной штриховатости ячменя (BSMV), вирус мозаики костра (BMV, учетные номера: РНК 1: X58456; РНК2: X58457; РНК3: X58458), вирус штриховатости кукурузы (MSpV), вирус кукурузы rayado fino (MRFV), вирус желтой карликовости кукурузы (MYDV), вирус карликовой мозаики кукурузы (MDMV), вирусы с положительной цепью РНК семейства *Benyviridae,* например, *вирус некротической желтой прожилковатой свеклы* (учетные номера: РНК 1: NC_003514; РНК2: NC_003515; РНК3: NC_003516; РНК4: NC_003517) или семейства *Bromoviridae*, например, вирусы рода *вируса мозаики люцерны* (учетные номера: РНК1: NC_001495; РНК2: NC_002024; РНК3: NC_002025) или рода *Bromovirus*, например, BMV (см. выше), или рода *Cucumovirus*, например, *вирус огуречной мозаики* (учетные номера: РНК1: NC_002034; РНК2: NC_002035; РНК3: NC_001440) или рода *Oleavirus*, вирусы dsДНК семейства *Caulimoviridae*, в частности, семейства *Badnavirus* или *Caulimovirus*, например, различные вирусы полосатости *банана* (например, учетные номера: NC_007002, NC_015507, NC_006955 или NC_003381) или *вирус мозаики цветной капусты* (учетный номер: NC_001497), или вирусы рода *Cavemovirus, Petuvirus, Rosadnavirus, Solendovirus, Soymovirus* или *Tungrovirus*, вирусы с положительной цепью РНК семейства *Closteroviridae*, например, рода *Ampelovirus*, *Crinivirus*, например, *вирус инфекционной желтухи латука* (учетные номера: РНК 1: NC_003617; РНК2: NC_003618) или *вирус хлороза томата* (учетные номера: РНК 1: NC_007340; РНК2: NC_007341), *Closterovirus*, например, *вирус желтухи свеклы* (учетный номер: NC_001598), или *Velarivirus*, вирусы с одноцепочечной ДНК (+/-) семейства *Geminiviridae*, например, вирусы семейства *Becurtovirus*,

Begomovirus, например, *вирус золотисто-желтой мозаики бобов*, *вирус курчавости побегов табака, вирус курчавости крапчатых листьев табака, вирус хлоротической крапчатости томата*, *вирус карликовости листьев томата, вирус золотистой мозаики томата, вирус курчавости листьев томата*, *вирус крапчатости томата,* или *вирус желтой пятнистости томата,* или *Geminiviridae* рода *Curtovirus,* например, *вирус курчавости верхушки свеклы,* или *Geminiviridae* рода *Topocuvirus, Turncurtvirus* или *Mastrevirus,* например, *вирус полосатости кукурузы* (см. выше), *вирус желтой карликовости табака, вирус карликовости пшеницы,* вирусы с положительной цепью РНК семейства *Luteoviridae,* например, рода *Luteovirus,* например, *вирус желтой карликовости ячменя PAV* (учетный номер: NC_004750), или рода *Polerovirus*, например, *вирус скручивания листьев картофеля* (учетный номер: NC_001747), вирусы с одноцепочечной ДНК семейства *Nanoviridae*, содержащие род *Nanovirus* или *Babuvirus*, вирусы с двухцепочечной РНК семейства *Partiviridae*, содержащие, среди прочего, семейства *Alphapartitivirus, Betapartitivirus* или *Deltapartitivirus*, вироиды семейства *Pospiviroidae*, вирусы с положительной цепью РНК семейства *Potyviridae*, например, содержащего род *Brambyvirus, Bymovirus, Ipomovirus, Macluravirus, Poacevirus*, например, *вирус мозаики пшеницы* (учетный номер: NC_012799) или *Potyviridae* рода *Potyvirus*, например, *вирус мозаики свеклы* (учетный номер: NC_005304), *вирус карликовой мозаики кукурузы* (учетный номер: NC_003377), *вирус Y картофеля* (учетный номер: NC_001616), или *вирус мозаики маиса* (учетный номер: NC_018833), или *Potyviridae* рода *Tritimovirus*, например, *вирус полосатой мозаики костра* (учетный номер: NC_003501) или *вирус полосатой мозаики* пшеницы (учетный номер: NC_001886), вирусы с одноцепочечной РНК семейства *Pseudoviridae*, например, рода *Pseudovirus*, или *Sirevirus*, вирусы с двухцепочечной РНК семейства *Reoviridae*, например, вирус карликовости риса (учетный номер: РНК1: NC_003773; РНК2: NC_003774; РНК3: NC_003772; РНК4: NC_003761; РНК5: NC_003762; РНК6: NC_003763; РНК7: NC_003760; РНК8: NC_003764; РНК9: NC_003765; РНК10: NC_003766; РНК11: NC_003767; РНК12: NC_003768), вирусы с положительной цепью РНК семейства *Tombusviridae*, например, содержащего род *Alphanecrovirus, Aureusvirus, Betanecrovirus, Carmovirus, Dianthovirus, Gallantivirus, Macanavirus, Machlomovirus, Panicovirus, Tombusvirus, Umbravirus* отряда *Zeavirus,* например, *вирус некротический полосатости маиса* (учетный номер: NC_007729), или вирусы с положительной цепью РНК семейства *Virgaviridae*, например, вирусы рода *Furovirus, Hordeivirus,* например, *вирус ложной штриховатости ячменя* (учетный номер: РНК1: NC_003469; РНК2: NC_003481; РНК3: NC_003478), или рода *Pecluvirus, Pomovirus, Tobamovirus,* или *Tobravirus,* например, *вирус погремковости табака* (учетный номер: РНК1: NC_003805; РНК2: NC_003811), а также вирусы с отрицательной цепью РНК отряда *Mononegavirales*, в частности, семейства *Rhabdoviridae*, например, *вирус желтой ложной штриховатости ячменя* (учетный номер: KM213865) или *вирус некротической желтухи латука* (учетный номер/образец: NC_007642/ AJ867584), вирусы с положительной цепью РНК отряда *Picornavirales*, в частности, семейства *Secoviridae*, например, рода *Comovirus, Fabavirus, Nepovirus, Cheravirus, Sadwavirus, Sequivirus, Torradovirus* или *Waikavirus*, вирусы с положительной цепью РНК отряда *Tymovirales,* в частности, семейства *Alphaflexiviridae*, например, вирусы рода *Allexivirus, Lolavirus, Mandarivirus* или *Potexvirus, Tymovirales,* в частности, семейства *Betaflexiviridae*, например, вирусы рода *Capillovirus, Carlavirus, Citrivirus, Foveavirus, Tepovirus,* или *Vitivirus,* вирусы с положительной цепью РНК отряда *Tymovirales*, в частности, семейства *Tymoviridae*, например, вирусы отряда *Maculavirus, Marafivirus* или *Tymovirus*, и бактериальные векторы, такие как, например, *Agrobacterium spp.,* такие как, например, *Agrobacterium tumefaciens*. Наконец, термин «вектор» также подразумевает подходящие транспортные агенты химических веществ для введения линейных последовательностей нуклеиновых кислот (одноцепочечных или двухцепочечных) в клетку-

мишень в сочетании с физическим способом введения, включая полимерные конструкции для доставки или конструкции доставки на основе липидов.

Подходящие конструкции доставки или векторы доставки, таким образом, содержат биологические средства для доставки нуклеотидных последовательностей в клетку-мишень, включая вирусные векторы, *Agrobacterium spp.* или конструкции доставки посредством химических веществ, включая наночастицы, например мезопористые наночастицы диоксида кремния (MSNP), катионные полимеры, включая подходы на основе ПЭИ (полиэтилениминовых) полимеров или полимеры, такие как, DEAE-декстран, или нековалентное поверхностное присоединение ПЭИ для получения катионных поверхностей, липидных или полимерных пузырьков, или их комбинаций. Липидные или полимерные пузырьки могут быть выбраны, например, из липидов, липосом, систем липидной инкапсуляции, наночастиц, составов из малых частиц нуклеиновой кислоты и липидов, полимеров и полимерсомов.

Термины «генетическая конструкция» или «рекомбинантная конструкция» используются в настоящем документе для обозначения конструкции, содержащей, среди прочего, плазмиды или плазмидные векторы, космиды, искусственные дрожжи или бактериальные искусственные хромосомы (YAC и BAC), фагемиды, векторы на основе бактериальных фагов, кассету экспрессии, выделенные последовательности одноцепочечных или двухцепочечных нуклеиновых кислот, содержащие последовательности ДНК и РНК, или аминокислотные последовательности, вирусные векторы, включая модифицированные вирусы, и их комбинацию или смесь для введения или трансформации, трансфекции или трансдукции в любую прокариотическую или эукариотическую клетку-мишень, включая растение, растительную клетку, ткань, орган или материал в соответствии с настоящим изобретением. Рекомбинантная конструкция по настоящему изобретению может содержать эффекторный домен, либо в форме нуклеиновой кислоты, либо аминокислотной последовательности, при этом, эффекторный домен представляет собой молекулу, которая может оказывать действие на клетку-мишень и включает трансген, одноцепочечную или двухцепочечную молекулу РНК, включая направляющую РНК ((s)gРНК), miРНК или siРНК, или аминокислотные последовательности, включая, среди прочего, фермент или его каталитически активный фрагмент, связывающий белок, антитело, фактор транскрипции, нуклеазу, в предпочтительном варианте осуществления настоящего изобретения, сайт-специфичную нуклеазу и тому подобное. Кроме того, рекомбинантная конструкция может содержать регуляторные последовательности и/или последовательности локализации. Рекомбинантная конструкция может быть интегрирована в вектор, включая плазмидный вектор, и/или она может быть представлена изолированно от векторной структуры, например, в виде полипептидной последовательности или в виде не связанной с вектором одноцепочечной или двухцепочечной нуклеиновой кислоты. После ее введения, например, путем трансформации, генетическая конструкция может либо продолжать существовать экстрахромосомно, то есть, не быть интегрированной в геном клетки-мишени, например, в виде двухцепочечной или одноцепочечной ДНК, двухцепочечной или одноцепочечной РНК, либо в виде аминокислотной последовательности. В альтернативном варианте осуществления настоящего изобретения, генетическая конструкция или ее части могут быть стабильно интегрированы в геном клетки-мишени, включая ядерный геном или другие генетические элементы клетки-мишени, включая геном пластид, таких как, митохондрии или хлоропласты. Термин «плазмидный вектор», используемый в этой связи, относится к генетической конструкции, первоначально полученной из плазмиды. Плазмида обычно относится к кольцевому автономно реплицирующемуся экстрахромосомному элементу в виде последовательности двухцепочечной нуклеиновой кислоты. В области генной инженерии эти плазмиды обычно подвергают целевым

модификациям путем вставления, например, генов, кодирующих устойчивость к антибиотику или гербициду, гена, кодирующего последовательность нуклеиновой кислоты-мишени, последовательность локализации, регуляторную последовательность, меченую последовательность, маркерного гена, включая маркер-антибиотик или флуоресцентный маркер, и тому подобное. Структурные компоненты исходной плазмиды, как источника репликации, сохраняются. В соответствии с определенными вариантами осуществления настоящего изобретения, последовательность локализации может содержать последовательность ядерной локализации, последовательность пластидной локализации, в предпочтительном варианте осуществления настоящего изобретения, последовательность митохондриальной локализации или последовательность хлоропластной локализации. Указанные последовательности локализации доступны специалисту в области биотехнологии растений. Разнообразные плазмидные векторы для использования в различных представляющих интерес клетках-мишенях доступны на рынке, и их модификация известна специалисту в соответствующей области.

Термин «генетически модифицированный» или «генетическая манипуляция», или «под воздействием генетических манипуляций» используется в настоящем документе в широком смысле и означает любую модификацию последовательности нуклеиновой кислоты или аминокислотной последовательности, целевой клетки, ткани, органа или организма, которая осуществляется путем вмешательства человека, непосредственно или опосредованно, для воздействия на эндогенный генетический материал или транскриптом, или протеом целевой клетки, ткани, органа или организма для модификации его целенаправленно, в результате чего он отличается от своего состояния, в котором он находился без вмешательства человека, тогда как термин «редактирование генома» конкретно относится к целенаправленной манипуляции с геномом клетки-мишени. Вмешательство человека может происходить либо *in vitro*, либо *in vivo*, либо в обоих случаях. Могут быть включены дополнительные модификации, например, по меньшей мере, одна точечная мутация, например, для нацеленной белковой инженерии или для оптимизации кодонов, делеции(ям) и, по меньшей мере, одной вставки или делеции, по меньшей мере, одной молекулы нуклеиновой кислоты или аминокислоты (включая также гомологичную рекомбинацию), модификации последовательности нуклеиновой кислоты или аминокислотной последовательности, или их комбинации. Термины также должны содержать молекулу нуклеиновой кислоты или молекулу аминокислоты, или клетку-хозяина, или организм, включая растение или его растительный материал, который подобен/подобны сопоставимой последовательности, организму или материалу природного происхождения, но которые были сконструированы, по меньшей мере, посредством одного этапа целенаправленной манипуляции.

Таким образом, «целевая генетическая манипуляция» или «нацеленаправленная», или «сайт-направленное» редактирование генов или редактирование генома, в том смысле, как это используется в настоящем документе, является результатом «генетической манипуляции», которая осуществляется целевым образом, то есть, по меньшей мере, в одной конкретной позиции в клетке-мишени и в конкретных подходящих условиях для достижения желаемого действия, по меньшей мере, в одной клетке, в предпочтительном варианте осуществления настоящего изобретения, в клетке растения, подлежащей манипуляции.

Термин «трансгенный», используемый в соответствии с настоящим изобретением, относится к животному, животной клетке, ткани или органу, растению, растительной клетке, ткани, органу или материалу, который содержит ген или генетическую конструкцию, содержащую трансген, который был перенесен в растение, растительную клетку, ткан, орган или материал естественными путями или с

помощью методов генной инженерии из другого организма. Термин «трансген» содержит последовательность нуклеиновой кислоты, включая ДНК или РНК, или их комбинацию, или смесь. Следовательно, термин «трансген» не ограничен последовательностью, обычно идентифицируемой как ген, то есть, последовательностью, кодирующей белок. Это может также относиться, например, к небелковой кодирующей последовательности ДНК или РНК. Следовательно, термин «трансгенный» обычно подразумевает, что соответствующая нуклеиновая кислота, введенная в интересующую клетку, присутствует не естественным образом в соответствующей целевой прокариотической или эукариотической клетке, включая бактериальную клетку, дрожжевую клетку, грибковую клетку, животную или растительную клетку, или растение, растительную клетку, ткань, орган или материал. Термины «трансген» или «трансгенный», используемые в настоящем документе, таким образом, относятся к последовательности нуклеиновой кислоты или аминокислотной последовательности, которые взяты из генома одного организма или продуцированы синтетическим путем, и которые затем вводятся в другой организм транзиентным или стабильным способом, искусственными методами молекулярной биологии, генетики и тому подобное.

Используемый в настоящем документе термин «растение» или «растительная клетка» относится к организму растения, органу растения, дифференцированным и недифференцированным растительным тканям, растительным клеткам, семенам и их производным, и к их потомству. Растительные клетки включают без ограничения, например, клетки из семян, из зрелых и незрелых зародышей, меристематических тканей, проростков, каллюсных тканей в различных состояниях дифференцировки, листьев, цветов, корней, побегов, гаметофитов, спорофитов, пыльцы, пыльцевых трубок и микроспор, протопластов, макроводорослей и микроводорослей. Различные растительные клетки могут быть либо гаплоидными, диплоидными, тетраплоидными, гексаплоидными, либо полиплоидными.

Используемый в настоящем документе термин «субъект» может означать либо человека, либо нечеловекоподобное животное. Этот термин включает, но этим не ограничивается, млекопитающих (например, людей, иных приматов, свиней, грызунов (например, мышей и крыс или хомяков), кроликов, морских свинок, коров, лошадей, кошек, собак, овец и коз). В одном варианте осуществления настоящего изобретения, субъект представляет собой человека.

Используемые в настоящем документе термины «лечить», «лечащий» и «лечение», обычно означают получение желаемого фармакологического и/или физиологического действия. Действие может быть профилактическим, с точки зрения полного или частичного предотвращения заболевания или его симптома, и/или он может быть терапевтическим, с точки зрения частичного или полного излечения от заболевания и/или неблагоприятного эффекта, связанного с заболеванием. Используемый в настоящем документе термин «лечение» охватывает любое лечение заболевания или симптома у млекопитающего и включает: (a) предотвращение возникновения заболевания или симптома у субъекта, который может быть предрасположен к приобретению заболевания или симптома, но заболевание или симптом еще не было у него диагностировано; (b) подавление заболевания или симптома, то есть, прекращение его развития; или (c) облегчение заболевания, то есть, вызывание регрессии заболевания. Терапевтический агент можно вводить до, во время или после начала заболевания или травмы. Лечение продолжающегося заболевания, при котором лечение стабилизирует или уменьшает нежелательные клинические симптомы пациента, представляет особый интерес. Такое лечение желательно проводить до полной потери функции в пораженных тканях. Введение терапии субъекту желательно проводить во время симптоматической стадии заболевания, а в некоторых случаях – после симптоматической стадии заболевания.

Используемый в настоящем документе термин «растительный материал» относится к любому

материалу, который может быть получен из растения на любой стадии развития. Растительный материал может быть получен либо *in planta*, либо *in vitro* из культуры растения или ткани растения, или его органа. Этот термин, таким образом, включает растительные клетки, ткани и органы, а также развитые структуры растения, а также субклеточные компоненты, такие как, нуклеиновые кислоты, полипептиды и все химические вещества растения или метаболиты, которые могут быть установлены в растительной клетке или компартменте, и/или которые могут быть продуцированы растением, или которые могут быть получены из экстракта любой растительной клетки, ткани или растения на любой стадии развития. Этот термин также содержит производное растительного материала, например, протопласт, полученный из, по меньшей мере, одной растительной клетки, содержащей растительный материал. Таким образом, этот термин также содержит меристематические клетки или меристематическую ткань растения.

Используемые в настоящем документе термины «мутация» и «модификация» используются взаимозаменяемо для обозначения делеции, вставки, добавления, замещения, редактирования, разрыва цепи и/или введения аддукта в контексте манипулирования нуклеиновой кислотой *in vivo* или *in vitro*. Делеция определяется как изменение последовательности нуклеиновой кислоты, в которой отсутствует, по меньшей мере, один нуклеотид. Вставка или добавление – это изменение последовательности нуклеиновой кислоты, которое привело к добавлению, по меньшей мере, одного нуклеотида. «Замещение» или редактирование происходит в результате замены, по меньшей мере, одного нуклеотида молекулой, которая отличается от молекулы замененного, по меньшей мере, одного нуклеотида(ов). Например, нуклеиновая кислота может быть заменена другой нуклеиновой кислотой, такой как, приведенная в качестве примера замена тимина цитозином, аденином, гуанином или уридином. Пиримидин на пиримидин (например, нуклеотидные замещения C на Tor T на C) или пурин на пурин (например, нуклеотидные замещения G на A или A на G) называются транзициями, тогда как пиримидин на пурин или пурин на пиримидин (например, G на T или G на C, или A на T, или A на C) называются трансверсиями. В альтернативной варианте осуществления настоящего изобретения, нуклеиновая кислота может быть заменена модифицированной нуклеиновой кислотой, примером которой является замена тимина на тимингликоль. Мутации могут привести к ошибочному спариванию оснований. Термин «ошибочное спаривание оснований» относится к нековалентному взаимодействию между двумя нуклеиновыми кислотами, каждая из которых находится на другой нуклеотидной последовательности или молекуле нуклеиновой кислоты, которая не соответствует правилам спаривания оснований. Например, у частично комплементарных последовательностей 5'-AGT-3 'и 5'-AAT-3' имеется ошибочное спаривание оснований G-A (транзиция).

Термин «разрыв цепи» применительно к последовательности двухцепочечной нуклеиновой кислоты, например, геномной последовательности в виде последовательности ДНК-мишени, включает одноцепочечный разрыв и/или двухцепочечный разрыв. Одноцепочечный разрыв (ник) относится к прерыванию в одной из двух цепей последовательности двухцепочечной нуклеиновой кислоты. В противоположность этому, двухцепочечный разрыв относится к прерыванию в обеих цепях последовательности двухцепочечной нуклеиновой кислоты. Разрывы цепи по настоящему изобретению могут быть введены в последовательность двухцепочечной нуклеиновой кислоты путем ферментативного разреза в интересующем положении основания нуклеиновой кислоты с использованием подходящей эндонуклеазы, включая эндонуклеазу CRISPR или ее вариант, и в качестве варианта может выступать мутированная или усеченная версия белка или эндонуклеазы дикого типа, которая по-прежнему может осуществлять ферментативную функцию белка дикого типа.

Используемый в настоящем документе термин «комплементарный» или «комплементарность»

описывает взаимоотношения между двумя ДНК, двумя РНК или, взаимоотношения гибридных последовательностей по настоящему изобретению, взимоотношения между областью нуклеиновой кислоты РНК и областью нуклеиновой кислоты ДНК. Определяемые нуклеиновыми основаниями ДНК или РНК две области нуклеиновой кислоты могут гибридизоваться друг с другом по типу ключ-замок. С этой целью принципы спаривания оснований по Уотсону-Крику имеют в качестве комплементарных оснований аденин и тимин/урацил, а также гуанин и цитозин, соответственно. Кроме того, термин «комплементарный», используемый в настоящем документе, также содержит такое понятие, как спаривание оснований не по Уотсону-Крику, такое как обратимое спаривание оснований по Уотсону-Крику, спаривание оснований по Хугстену, обратимое спаривание оснований по Хугстену и спаривание неоднозначных пар оснований, при условии, что соответствующие пары оснований могут образовывать водородные связи друг с другом, то есть, две разные цепи нуклеиновой кислоты могут гибридизоваться друг с другом на основе указанной комплементарности. Совершенная комплементарность в смысле двух последовательных удлинений, выравнивающихся 100% друг к другу по заданной длине, не требуется, поскольку специалисту в данной области техники известно, что на гибридизацию нуклеиновых кислот влияют такие факторы, как степень и длина комплементарности между нуклеиновыми кислотами, жесткость условий проведения, Tm образовавшегося гибрида и соотношение G:C в нуклеиновых кислотах и так далее. Кроме того, стерические факторы могут влиять на то, будут ли гибридизоваться две последовательности, даже если они не на 100% комплементарны друг другу. Следовательно, две комплементарные последовательности нуклеиновой кислоты по настоящему изобретению могут иметь, по меньшей мере, 70%, по меньшей мере, 71%, по меньшей мере, 72%, по меньшей мере, 73%, по меньшей мере, 74%, по меньшей мере, 75%, по меньшей мере, 76%, по меньшей мере, 77%, по меньшей мере, 78%, по меньшей мере, 79%, по меньшей мере, 80%, по меньшей мере, 81%, по меньшей мере, 82%, по меньшей мере, 83%, по меньшей мере, 84%, по меньшей мере, 85%, по меньшей мере, 86%, по меньшей мере, 87%, по меньшей мере, 88%, по меньшей мере, 89%, по меньшей мере, 90%, по меньшей мере, 91%, по меньшей мере, 92%, по меньшей мере, 93%, по меньшей мере, 94%, по меньшей мере, 95%, по меньшей мере, 96%, по меньшей мере, 97%, по меньшей мере 98% или, по меньшей мере, 99% гомологии последовательностей или комплементарности друг к другу и могут по-прежнему гибридизоваться друг с другом в условиях, примерно, средней жесткости. «Средние» условия жесткости относятся к 0,165-0,330 M NaCl в диапазоне температур от 20 до 29° C ниже Tm, при этом, Tm определяется как Tm для последовательности ДНК, и ее можно оценить с помощью обычно используемого расчета:

Tm = 81.5+16.6logio([Na+]/1.0+0.7[Na+])+0.41(%[G+C])-(500/n)-P-F, при этом,

Tm = температура плавления в ° C, $[Na^+]$ = молярная концентрация ионов натрия, %[G+C] = процент оснований G+C в последовательности ДНК, n = длина последовательности ДНК в основаниях P = поправка на температуру для % ошибочно спаренных пар оснований (~1° C на 1% ошибочных спариваний оснований) и F = поправка на концентрацию формамида (= 0,63° C на 1% [формамид]).

Используемый в настоящем документе термин «транзиентное введение» относится к транзиентному введению, по меньшей мере, одной последовательности нуклеиновой кислоты по настоящему изобретению, в предпочтительном варианте осуществления настоящего изобретения, включенной в вектор доставки или в рекомбинантную конструкцию, с помощью или без помощи вектора доставки, в структура-мишень, например, в растительную клетку, при этом, по меньшей мере, одна последовательность нуклеиновой кислоты вводится в подходящих условиях реакции, благодаря чему не происходит интеграции, по меньшей мере, одной последовательности нуклеиновой кислоты в эндогенный

материал нуклеиновой кислоты структуры-мишени, в геном в целом, поэтому, по меньшей мере, одна последовательность нуклеиновой кислоты не будет интегрирована в эндогенную ДНК клетки-мишени. Как следствие, в случае транзиентного введения, введенная генетическая конструкция не будет передаваться по наследству потомству структуры-мишени, например, прокариоту, животной или растительной клетке. По меньшей мере, одна последовательность нуклеиновой кислоты или продукты, полученные в результате ее транскрипции или трансляции, присутствуют только временно, то есть, транзиентным образом, в конститутивной или индуцибельной форме, и, таким образом, могут быть активными только в клетке-мишени для оказания своего действия в течение ограниченного времени. Следовательно, по меньшей мере одна последовательность нуклеиновой кислоты, введенная посредством транзиентного введения, не будет передаваться по наследству потомству клетки. Однако эффект, который последовательность нуклеиновой кислоты ввела транзиентным образом, потенциально может передаваться по наследству потомству клетки-мишени.

Используемый в настоящем документе термин «стабильная интеграция» или «стабильно интегрирована» относится к стабильной интеграции, по меньшей мере, одной последовательности нуклеиновой кислоты по настоящему изобретению, в предпочтительном варианте осуществления настоящего изобретения, включенной в вектор доставки или в рекомбинантную конструкцию. Интеграция может происходить либо в ядерный геном клетки-мишени, либо в любой другой геномный внеядерный материал в интересующем компартменте эукариотической клетки, например, в митохондрию или в пластиду растительной клетки. Таким образом, стабильно интегрированная, по меньшей мере, одна рекомбинантная конструкция будет передаваться по наследству потомству модифицированной таким образом клетки-мишени. В зависимости от природы генетической конструкции вся или часть генетической конструкции будет стабильно интегрирована, поскольку генетическая конструкция может содержать несколько интересующих областей, включая область-мишень, которая должна быть стабильно интегрирована, а также другие области, среди прочего, необходимые для транспортировки, доставки, обслуживания и корректной локализации генетической конструкции в растительной клетке, области которой, однако, сами не будут интегрироваться, но будут служить грузом для интересующей области, подлежащей стабильной интеграции, как это известно специалисту в данной области техники. Стабильная интеграция, по меньшей мере, одной генетической конструкции по настоящему изобретению в, по меньшей мере, одну гематопоэтическую или меристематическую клетку или ткань, следовательно, приведет к передаче по наследству модифицированной таким образом области генома структуры-мишени, то есть, области мишени ДНК-мишени, потомству модифицированной клетки посредством всех стадий развития указанной, по меньшей мере, одной гематопоэтической или меристематической клетки, что может быть благоприятным для подходов, при которых желательна целевая генетическая модификация и выход конечного типа клеток, являющихся результатом дифференциации и развития, по меньшей мере, одной гематопоэтической меристематической клетки. Достижение, например, стабильной интеграции в, по меньшей мере, одну меристематическую клетку незрелого соцветия растения может, таким образом, привести к стабильной передачи по наследству введенного генетического признака в гамету пыльцы или семяпочки в результате развития, по меньшей мере, одной меристематической клетки незрелого соцветия. Стабильная интеграция в, по меньшей мере, одну плюрипотентную гематопоэтическую клетку или любую плюрипотентную, или мультипотентную клетку аналогичным образом приведет к стабильному наследованию введенного генетического признака.

Используемый в настоящем документе термин «бомбардировка частицами», также называемый

биолистической трансфекцией или опосредованным микрочастицами переносом генов, относится к способу физической доставки для переноса покрытой микрочастицы или наночастицы, содержащей нуклеиновую кислоту или интересующую генетическую конструкцию, в клетку-мишень или ткань-мишень. Микрочастица или наночастица функционирует как снаряд, и ею обстреливают интересующую структуру-мишень под высоким давлением с использованием подходящего устройства, зачастую называемого «генная пушка». При трансформации посредством бомбардировки частицами используется микроснаряд из металла, покрытый интересующим геном, которым затем расстреливают клетки-мишени с использованием оборудования, известного как «генная пушка» (Сэндфорд и др., 1987), на высокой скорости, что позволяет достаточно быстро проникнуть в клеточную стенку ткани-мишени, но не настолько стремительно, чтобы вызвать гибель клетки. Для протопластов, у которых клеточная стенка полностью удалена, условия, естественно, отличаются. Осажденная нуклеиновая кислота или генетическая конструкция, по меньшей мере, на одной микрочастице, высвобождаются в клетку после бомбардировки и интегрируются в геном или экспрессируются транзиентным образом в соответствии с определением, приведенным выше. Ускорение микрочастиц осуществляется посредством высоковольтного электрического разряда или сжатого газа (гелия). Что касается используемых металлических частиц, то обязательно, чтобы они были нетоксичными, нереактивными, и чтобы они имели меньший диаметр, чем клетка-мишень. Чаще всего используются золото или вольфрам. Существует большой объем общедоступной информации от производителей и поставщиков генных пушек и информации, связанной с ними, которые касаются общих условий их использования.

Используемый в настоящем документе термин «производное» или «потомок», или «потомство» в контексте прокариотической или эукариотической клетки, в предпочтительном варианте осуществления настоящего изобретения, животной клетки и, в более предпочтительном варианте осуществления настоящего изобретения, растения или растительной клетки, или растительного материала по настоящему изобретению относится к потомкам такой клетки или материала, которые являются результатом естественного репродуктивного размножения, включая размножение половым и бесполым путем. Специалисту в данной области техники хорошо известно, что указанное размножение может привести к введению мутаций в геном организма в результате природных явлений, что приводит к появлению потомка или потомства, которое, в геномическом смысле, отличается от родительского организма или клетки, однако, оно все еще принадлежит к тому же роду/виду и обладает, в основном, теми же характеристиками, что и родительская рекомбинантная клетка-хозяин. Такие производные или потомки, или потомство, возникающие в результате природных явлений во время репродукции или регенерации, таким образом, включены в термин настоящего изобретения. Кроме того, термин «производное» может подразумевать, в контексте вещества или молекулы, а не в контексте клетки или организма, непосредственно или посредством модификации, опосредованно полученное от другого. Это может подразумевать последовательность нуклеиновой кислоты, полученную из клетки или растительного метаболита, полученного из клетки или материала. Таким образом, эти термины не относятся к какому-либо произвольному производному, потомку или предку, а скорее к производному, или потомку, или предку, филогенетически связанному, то есть, основанному на родительской клетке или вирусе, или на их молекуле, принимая во внимание тот факт, что эта взаимосвязь между производным, потомком или предком и «родителем» может быть явным образом определена специалистом в данной области техники.

Кроме того, используемые в настоящем документе термины «полученный», «полученный из» или «производное», в контексте биологической последовательности (нуклеиновой кислоты или аминокислоты) или молекулы, или комплекса, подразумевают, что соответствующая последовательность основана на

стандартной/эталонной последовательности, например, из списка последовательностей или учетного номера базы данных, или соответствующей структуры скаффолда, то есть, происходящей из указанной последовательности, тогда как стандартная/эталонная последовательность может содержать больше последовательностей, например, весь геном или полную последовательность, кодирующую полипротеин, вирус, тогда как последовательность, «полученная из» нативной последовательности, может содержать только один свой изолированный фрагмент или свой когерентный фрагмент. В этом контексте, можно сказать, что молекула сДНК или РНК «получена из» последовательности ДНК, служащей в качестве молекулярной матрицы. Таким образом, специалист в данной области техники может легко определить последовательность, «полученную из» стандартной/эталонной последовательности, которая, посредством выравнивания последовательности на уровне ДНК или аминокислоты будет иметь высокую идентичность с соответствующей стандартной/эталонной последовательностью, и которая будет иметь когерентные удлинения ДНК/аминокислот совместно с соответствующей стандартной/эталонной последовательностью (> 75% запрашиваемой идентичности для заданной длины выровненной молекулы при условии, что полученная последовательность является запрошенной, а стандартная/эталонная последовательность представляет собой субъект во время выравнивания последовательности). Таким образом, специалист в данной области техники может клонировать соответствующие последовательности на основе изобретения, предложенного в настоящем документе, посредством полимеразных цепных реакций и тому подобного, в интересующую подходящую векторную систему, или использовать последовательность в качестве векторного скаффолда. Термин «полученный из», таким образом, не представляет произвольную последовательность, но он представляет последовательность, соответствующую стандартной/эталонной последовательности, из которой она получена, в то время как определенные различия, например, определенные мутации, естественным образом возникающие во время репликации рекомбинантной конструкции в клетке-хозяине, не могут быть исключены и, таким образом, включены в термин «полученный из». Кроме того, некоторые удлинения последовательности из родительской последовательности могут быть последовательно соединены в последовательность, полученную из родительской последовательности. Различные удлинения будут иметь высокую (в предпочтительном варианте осуществления настоящего изобретения, более 90%) или даже 100% гомологию с родительской последовательностью. Специалисту в данной области техники хорошо известен тот факт, что последовательность искусственных молекулярных комплексов по настоящему изобретению, если она предложена или частично предложена в виде последовательности нуклеиновой кислоты, будет затем транскрибирована и, при необходимости, транслирована *in vivo* и, возможно, будет подвергнута дальнейшему расщеплению и/или обработке внутри клетки-хозяина (расщепление сигнальных пептидов, эндогенное биотинилирование и так далее), таким образом, термин «полученный из» указывает на корреляцию с последовательностью, первоначально использованной в соответствии с раскрытием настоящего изобретения.

Термины «область-мишень», «сайт-мишень», «структура-мишень», «конструкция-мишень», «нуклеиновая кислота-мишень» или «целевая клетка/ткань/организм», или «область ДНК-мишени», в контексте настоящего документа, относятся к мишени, которая может представлять собой любую область генома в любом компартменте клетки-мишени.

Используемый в настоящем документе термин «регуляторная последовательность» относится к нуклеиновой кислоте или аминокислотной последовательности, которая может направлять транскрипцию и/или трансляцию, и/или модификацию интересующей последовательности нуклеиновой кислоты.

Термины «белок», «аминокислота» или «полипептид» используются в настоящем документе взаимозаменяемо и относятся к аминокислотной последовательности, имеющей каталитическую ферментативную функцию или структурный, или функциональный эффект. Термин «аминокислота» или «аминокислотная последовательность» или «молекула аминокислоты» содержит любой природный или химически синтезированный белок, пептид, полипептид и фермент или модифицированный белок, пептид, полипептид и фермент, при этом, термин «модифицированный» содержит любую химическую или ферментативную модификацию белка, пептида, полипептида и фермента, включая усечения последовательности дикого типа до более короткой, но все еще активной части.

В соответствии с настоящим изобретением, могут быть использованы общепринятые методы молекулярной биологии, микробиологии и рекомбинантной ДНК, известные специалистам в данной области техники. Такие методы подробно описаны в литературе. См., например, Сэмбрук, Фрич и Маниатис, *«Молекулярное клонирование: инструкция по проведению лабораторных работ»,* Второе издание (1989), Cold Spring Harbor Laboratory Press, Колд Спринг Харбор, Нью-Йорк (в данном документе, «Сэмбрук и др., 1989»); *«Клонирование ДНК: практический подход»,* Тома I и II (под редакцией Д.Н. Гловера, 1985); *«Синтез олигонуклеотидов»* (под редакцией М.Дж. Гейта, 1984); *«Гибридизация нуклеиновых кислот»* (под редакцией Б.Д. Хемса и С.Дж. Хиггинса, (1985); *«Транскрипция и трансляция»* (под редакцией Б.Д. Хемса и С.Дж. Хиггинса, (1984); *«Культура животных клеток»* (под редакцией Р.Я. Фрешни, 1986); *«Иммобилизованные клетки и ферменты»* (IRL Press, (1986); Б. Пербэл, *«Практическое руководство по молекулярному клонированию»* (1984); Ф.М. Осубель и др. (под редакцией), *«Текущие протоколы в молекулярной биологии»*, John Wiley and Sons, Inc. (1994); и другие.

Всякий раз, когда настоящее изобретение относится к проценту гомологии или идентичности последовательностей нуклеиновых кислот или аминокислот, эти значения определяют значения, полученные для нуклеиновых кислот с использованием программы EMBOSS Water Pairwise Sequence Alignments (нуклеотиды) (www.ebi.ac.uk/Tools/psa/ emboss_water/nucleotide.html) или для аминокислотных последовательностей с использованием программы EMBOSS Water Pairwise Sequence Alignments (белки) (www.ebi.ac.uk/Tools/psa/emboss_water/). В инструментах, предложенных Европейским институтом молекулярной биологии (EBI) Европейской молекулярно-биологической лаборатории (EMBL) для локального выравнивания последовательностей, используется модифицированный алгоритм Смита-Уотермана (см. www.ebi.ac.uk/Tools/psa/ и Смит, Т.Ф. и Вотерман, М.С. «Идентификация общих молекулярных подпоследовательностей» *Journal of Molecular Biology*, 1981 147 (1):195-197). При проведении выравнивания используются параметры по умолчанию, определенные EMBL-EBI. Эти параметры предназначены (i) для аминокислотных последовательностей: Матрица = BLOSUM62, штраф за открытие гэпа = 10 и штраф за продолжение гэпа = 0,5 или (ii) для последовательностей нуклеиновых кислот: Матрица = ДНКполная, штраф за открытие гэпа = 10 и штраф за продолжение гэпа = 0,5.

Подробное описание вариантов осуществления изобретения

Согласно первому аспекту настоящего изобретения предложен искусственный молекулярный комплекс, содержащий (a) по меньшей мере, одну сайт-специфичную нуклеазу (SSN) или ее каталитически активный фрагмент, или последовательность нуклеиновой кислоты, кодирующую последовательность, и непосредственно взаимодействующую с ней (b) по меньшей мере, один взаимодействующий домен матрицы для репарации (RTDD) или последовательность нуклеиновой кислоты, кодирующую последовательность, при этом, взаимодействующий домен матрицы для репарации выполнен с возможностью непосредственного взаимодействия, по меньшей мере, с одной последовательностью репарационной матричной нуклеиновой

кислоты (RT); (c) при необходимости, содержащий, по меньшей мере, один взаимодействующий домен (IA) или последовательность нуклеиновой кислоты, кодирующую последовательность, при этом, по меньшей мере, один взаимодействующий домен непосредственно взаимодействует, по меньшей мере, с одной сайт-специфичной нуклеазой или ее каталитически активным фрагментом, и, при этом, по меньшей мере, один взаимодействующий домен выполнен с возможностью предложения, по меньшей мере, одной из функциональных возможностей, выбранных из группы, состоящей из (i) взаимодействия, по меньшей мере, с одним стыковочным доменом матрицы для репарации; и/или (ii) взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислоты; и/или (iii) сиквенс-специфичного взаимодействия с геномной ДНК; при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит, по меньшей мере, одну часть, комплементарную, по меньшей мере, одной геномной комплементарной последовательности, и при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты выполнена с возможностью опосредования репарации последовательности ДНК-мишени.

Таким образом, настоящее изобретение основано на сайт-специфичной нуклеазе (SSN). Эта нуклеаза характеризуется наличием нуклеазной функции и функции распознавания ДНК. Функция распознавания ДНК может быть присуща нуклеазе в виде домена, опосредующего распознавание или связывание ДНК, или ей могут помогать дополнительные направляющие молекулы, например, для CRISPR, направляемых нуклеиновой кислотой (направляемых РНК) или для нуклеаз Аргонавтов (направляемых ДНК), но настоящее изобретение не ограничивается использованием вышеупомянутых нуклеаз, направляемых нуклеиновой кислотой, и, таким образом, расширяет область применения целевой геномной инженерии для любых сайт-специфичных нуклеаз, не связанных с CRISPR или Аргонавтами. Другой частью искусственной системы по настоящему изобретению является, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты (RT), поскольку продукты и способы настоящего изобретения в основном направлены на обеспечение физической доступности RT в сайте двухцепочечного разрыва, индуцированного SSN. Кроме того, настоящее изобретение опирается на стыковочный домен матрицы для репарации (RTDD), как на часть оптимизированной молекулярной системы. Этот RTDD выполняет функцию непосредственного или опосредованного приведения SSN и, по меньшей мере, одной RT в близкий контакт для обеспечения эффективной и целевой геномной инженерии.

Таким образом, RTDD ковалентно или нековалентно ассоциирован с RT, то есть, он непосредственно взаимодействует с RT на молекулярном уровне. Одновременно, RTDD непосредственно взаимодействует, по меньшей мере, с одной SSN и, таким образом, представляет собой сцепляющую молекулу или домен между SSN и RT. Для RTDD существует несколько возможных конфигураций. В одном варианте осуществления настоящего изобретения, RTDD непосредственно ассоциирован с SSN. Например, если SSN представляет собой нуклеазу CRISPR, то RTDD может представлять собой gРНК, или, если SSN представляет собой нуклеазу Аргонавт, то RTDD может представлять собой а gДНК. В другом варианте осуществления настоящего изобретения, RTDD может быть частью самой SSN, или он может быть частью RT, причем RTDD представляет собой специфичную часть RT или SSN. В этих вариантах осуществления настоящего изобретения, SSN может содержать домен в виде части своей аминокислотной последовательности, которая может взаимодействовать с аптамером, несущим RT. Таким образом, в определенных вариантах осуществления настоящего изобретения, не существует отдельного взаимодействующего домена, поскольку сама сайт-специфичная нуклеаза содержит домен для взаимодействия с RTDD. Таким образом, RTDD может представлять собой аптамер, ассоциированный с

последовательностью нуклеиновой кислоты RT, при этом, аптамер распознается и, таким образом, может специфически взаимодействовать или с SSN, и/или с дополнительным взаимодействующим доменом.

Кроме того, ковалентные и нековалентным взаимодействия между компонентами искусственного молекулярного комплекса предусмотрены по настоящему изобретению.

В дополнительном варианте осуществления настоящего изобретения, искусственный молекулярный комплекс содержит дополнительный взаимодействующий домен (IA). В этой конфигурации, RTDD может быть ассоциирован с дополнительным взаимодействующим доменом для определенных вариантов осуществления настоящего изобретения. Взаимодействующий домен непосредственно взаимодействует, то есть, физически ассоциирован, либо ковалентно, как молекула слияния, либо нековалентно, с SSN и предлагает дополнительную функциональную возможность молекулярному комплексу. Взаимодействующий домен может быть белковым доменом, содержащим функции распознавания/связывания ДНК, то есть, он может быть доменом, который способен взаимодействовать с сайтом геномной ДНК-мишени сайт-специфичным образом, или взаимодействующий домен может быть выполнен со специальной возможностью взаимодействия с RTDD и/или RT. Например, взаимодействующий домен может содержать собственную функцию распознавания и связывания ДНК без функции самой нуклеазы для специфичного взаимодействия с геномной ДНК. В другом варианте осуществления настоящего изобретения, взаимодействующий домен может функционировать в качестве высокоспецифичного взаимодействующего партнера для RTDD, ассоциированного с RT, как более подробно описано ниже. Добавление этой дополнительной функциональной возможности распознавания ДНК или взаимодействия RTDD молекулярному комплексу путем добавления взаимодействующего домена предлагает еще один уровень специфичности для геномной инженерии в дополнение к естественной функциональной возможности исключительно SSN.

В конечном счете, в частности, RTDD, непосредственно взаимодействующий с RT, и, кроме того, взаимодействующие домены, подробно описанные в настоящем документе ниже, предлагают универсальный инструментарий для (i) приведения RT в тесный контакт с интересующей SSN и, таким образом, в непосредственную близость от двухцепочечного разрыва, индуцированного, по меньшей мере одной SSN для предложения молекулярной системы в виде искусственного молекулярного комплекса, имеющего (ii) превосходящий целевой диапазон и более высокую точность, подходящие для различных индивидуальных подходов геномной инженерии у эукариот и прокариот для достижения оптимизированных результатов для геномной инженерии, метаболической инженерии, развития признаков у растений и для терапевтического применения.

Различные аспекты и варианты осуществления настоящего изобретения, таким образом, все основаны на предложении подходящего фермента, индуцирующего двухцепочечный разрыв, или двух никаз, таких как SSN, а также надлежащим образом сконструированной репарационной матричной последовательности нуклеиновой кислоты (RT), при этом сущность настоящего изобретения заключается в том, что SSN и RT приводятся в непосредственную близость для направления события геномной инженерии целевым образом.

В одном варианте осуществления настоящего изобретения, по меньшей мере, один RTDD представляет собой CRISPR gРНК или gДНК, и он непосредственно взаимодействует или ассоциируется с матрицей для репарации для построения последовательности гибридной нуклеиновой кислоты РНК-ДНК или ДНК из RTDD и ДНК из RT.

Таким образом, «искусственный молекулярный комплекс» по настоящему изобретению

представляет собой комплекс, содержащий, по меньшей мере, один аминокислотный компонент, то есть SSN и, при необходимости, взаимодействующий домен, RTDD и матрицу для репарации на основе нуклеиновой кислоты (RT). В собранном состоянии комплекс обычно будет содержать, по меньшей мере, один компонент, содержащий аминокислоту (белок), то есть, по меньшей мере, одну SSN, и компонент, содержащий нуклеиновую кислоту, то есть RT. По меньшей мере, один RTDD и, при необходимости, по меньшей мере, один взаимодействующий домен также могут содержать аминокислоты и/или нуклеиновые кислоты в качестве строительных блоков, но, благодаря функциям указанных компонентов в молекулярном комплексе, возможен более широкий спектр молекул, включая синтетические строительные блоки или комбинации различных биомолекул, и/или синтетических молекул.

Таким образом, искусственный молекулярный комплекс по настоящему изобретению преодолевает недостаток олигонуклеотидных (RT)-ферментных (SSN) конъюгатов, заключающийся в том, что они не могут самостоятельно собраться *in vivo*, что тем самым серьезно ограничивает их пользу для редактирования генома *in vivo* путем добавления, по меньшей мере, одного дополнительного домена, опосредующего взаимодействие, то есть, RTDD и, при необходимости, IA, гарантируя тесную ассоциацию RT и SSN, и идеальную сборку молекулярного комплекса *in vivo* или в целом в физиологических условиях, *in vivo* и *in vitro* при работе с, по меньшей мере, одной интактной клеткой, несущей геномную ДНК-мишень (геномные, включая кодирующие и некодирующие, области, включая ядерную, пластидную и эписомальную ДНК-мишень и эпигенетические сайты-мишени), представляющие интерес для модификации.

В одном варианте осуществления настоящего изобретения, искусственный молекулярный комплекс может быть предложен и собран полностью *in vivo*, например, путем предложения необходимых конструкций для синтеза и последующей сборки комплекса в клетке-хозяине. В другом варианте осуществления настоящего изобретения, искусственный молекулярный комплекс может быть предложен в виде собранного *ex vivo* молекулярного комплекса, который впоследствии вводится в интересующую клетку-хозяина *in vivo*, или который приводится в контакт с интересующей геномной молекулой-мишенью *in vitro*. В еще одном варианте осуществления настоящего изобретения, части искусственного молекулярного комплекса могут быть продуцированы *ex vivo*, и части могут быть продуцированы *in vivo*, например, после введения подходящего вектора доставки, несущего плазмиду для транскрипции и/или экспрессии компонента искусственного молекулярного комплекса, и конечный искусственный молекулярный комплекс, выполняющий свою функцию, будет затем собираться *in vivo* на основе внутренней функции распознавания, опосредованной RTDD.

Таким образом, «взаимодействие» или «непосредственное взаимодействие» между любыми компонентами искусственного молекулярного комплекса по настоящему изобретению подразумевает любое ковалентное или нековалентное взаимодействие или сцепление между двумя компонентами искусственного молекулярного комплекса. Таким образом, ковалентное сцепление на уровне нуклеиновой кислоты может подразумевать фосфодиэфирное или фосфоротиоатное сцепление между нуклеотидами молекулы нуклеиновой кислоты. Кроме того, ковалентное сцепление может быть дисульфидным мостиком между аминокислотой и другой аминокислотой, и/или модифицированной молекулой нуклеиновой кислоты, однако любое сцепление природного происхождения или ковалентное сцепление искусственного происхождения может быть предусмотрена по настоящему изобретению. Нековалентные взаимодействия содержат электростатические взаимодействия, в том числе ионную, водородную связь или галогенную связь, Ван-дер-Ваальсовы силы, включая диполь-дипольные, диполь-индуцированные дипольные,

лондоновские дисперсионные силы, rr-эффекты и гидрофобные эффекты. Примечательно, что в компонентах искусственного молекулярного комплекса по настоящему изобретению может присутствовать более одного типа взаимодействия. Например, SSN, например, нуклеаза CRISPR, может взаимодействовать посредством нековалентного взаимодействия(й) с gPHK в качестве RTDD. RTDD может быть ковалентно сцеплен с матрицей для репарации RT. В другом варианте осуществления настоящего изобретения, белок слияния Аргонавта в виде SSN может быть ковалентно слит с одноцепочечным вариабельным фрагментом антитела в качестве взаимодействующего домена (IA). IA может, среди прочего, быть специфичным для флуоресцеина и, таким образом, может нековалентно взаимодействовать с флуоресцеином RTDD. Флуоресцеин и такая меченая нуклеиновая кислота матрицы для репарации RT могут быть предложены в виде синтетического ковалентного слияния. В другом варианте осуществления настоящего изобретения, ассоциация различных компонентов опосредуется нековалентными взаимодействиями, например, распознаванием посредством лейциновой «молнии» последовательности ДНК-мишени и/или аптамера (на основе нуклеиновой кислоты или аминокислоты), взаимодействующего либо с SSN, либо с IA. В одном варианте осуществления настоящего изобретения, RTDD может представлять собой аптамер, например, последовательность, предлагающую функцию аптамера в матрице для репарации. В другом варианте осуществления настоящего изобретения, удлинение направляющей нуклеиновой кислоты, позволяющее гибридизацию с матрицей для репарации, может функционировать в качестве, по меньшей мере, одной RTDD. При определении направляющей нуклеиновой кислоты как таковой в качестве RTDD, в таком варианте осуществления настоящего изобретения используется более одного RTDD. В еще одном варианте осуществления настоящего изобретения, 3'- или 5'-конец направляющей нуклеиновой кислоты, используемой для лигирования с матрицей для репарации, может быть выполнен со специальной возможностью функционирования в качестве RTDD.

Согласно настоящему изобретению, различные компоненты искусственного молекулярного комплекса могут содержать строительные белки природного происхождения и/или синтетические искусственные строительные блоки.

Таким образом, сайт-специфичная нуклеаза (SSN), в соответствии с различными вариантами осуществления настоящего изобретения, или последовательность нуклеиновой кислоты, кодирующая последовательность, может представлять собой любую нуклеазу природного происхождения или сконструированную нуклеазу, которая способна распознавать и расщеплять ДНК сайт-специфичным образом. Поскольку многие SSN будут иметь большое количество потенциальных сайтов расщепления в геноме организма или вируса, то такие SSN с определенным шаблоном расщепления или сконструированные SSN с индивидуальными шаблонами расщепления являются предпочтительными. Таким образом, SSN включают сайт-специфичные нуклеазы для методов редактирования генома, таких как сконструированные цинковые пальцы, эффекторы по типу активаторов транскрипции (TALE), (хоминг) мегануклеазы, нуклеазы, полученные из системы CRISPR, включая нуклеазы Cas или Cpf1, или нуклеазы Аргонавты, а также редкие разрезающие эндонуклеазы или две сайт-специфичные никующие эндонуклеазы, включая эндонуклеазу рестрикции класса IIS, включая Fokl или ее вариант, или две сайт-специфичные никующие эндонуклеазы, или их вариант, или их каталитически активный фрагмент, или любой вариант или каталитически активный фрагмент вышеупомянутых SSN. Следовательно, в соответствии с настоящим изобретением может присутствовать более одной SSN или последовательности нуклеиновой кислоты, кодирующей последовательность, тогда как суммарные молекулы способны индуцировать разрыв двухцепочечной ДНК-мишени или два последовательных одноцепочечных разрыва в последовательности

ДНК-мишени.

«Последовательность ДНК-мишени» по настоящему изобретению может представлять собой любую область в двухцепочечной ДНК, на геномной или плазмидной основе, где индуцируется разрыв ДНК-мишени, и который впоследствии восстанавливается с помощью матрицы для репарации (RT) по настоящему изобретению. Даже если «последовательность ДНК-мишени» происходит из эндогенной последовательности, то редактирование или конструирование указанной последовательности может быть выполнено *in vitro* путем представления соответствующей последовательности на молекуле, содержащей геномную ДНК, в предпочтительном варианте осуществления настоящего изобретения, на плазмиде. В таких вариантах осуществления настоящего изобретения, интересующий локус-мишень может содержаться в молекуле ДНК внутри клетки. Клетка может представлять собой прокариотическую клетку или эукариотическую клетку, или вирусный геном на плазмиде внутри прокариотической клетки, или эукариотическую клетку-хозяина, используемую для размножения вируса. Клетка может представлять собой клетку млекопитающего. Клетка млекопитающего может представлять собой клетку нечеловекоподобного примата, крупного рогатого скота, свиньи, грызуна или мыши. Клетка может представлять собой эукариотическую клетку не млекопитающего, например, птицы, рыбы или креветки. Клетка также представлять собой растительную клетку. Растительная клетка может представлять собой клетку культурного растения, такого как маниока, кукуруза, сорго, пшеница, соя, хлопок, сахарная свекла или рис. Растительная клетка также может представлять собой клетку водоросли, дерева или овоща. Модификация, введенная в клетку по настоящему изобретению, может быть такой, что клетка и потомство клетки изменяются для улучшения продуктивности биологических продуктов, таких как, антитело, крахмал, спирт или другой желаемой клеточной продукции. Модификация, введенная в клетку по настоящему изобретению, может быть такой, что клетка и потомство клетки включают изменение, которое изменяет полученный биологический продукт. В другом варианте осуществления настоящего изобретения, «последовательность ДНК-мишени» может представлять собой интересующий эпигеномный локус.

«Геномная комплементарная последовательность» по настоящему изобретению относится к той части последовательности RT по настоящему изобретению, которая может выравниваться посредством комплементарного спаривания оснований. Таким образом, «последовательность ДНК-мишени» и «геномная комплементарная последовательность» могут перекрываться или даже совпадать, но для определенных вариантов осуществления настоящего изобретения указанные последовательности могут быть разными, например, в случае, если, по меньшей мере, одна SSN будет иметь сайт разрезания выше или ниже части «геномной комплементарной последовательности» RT.

В любом из описанных вариантов осуществления настоящего изобретения, цепочечный разрыв может представлять собой двухцепочечный разрыв, или он может представлять собой два двухцепочечных разрыва.

В определенных вариантах осуществления настоящего изобретения, компонент SSN и, при необходимости, компонент IA искусственного молекулярного комплекса будут доставлены в клетку-хозяина или в анализирующую систему, содержащую интересующую область генома, подлежащую модификации посредством совместной доставки белка с меченым RTDD и ассоциированным олигонуклеотидом матрицы для репарации в одном варианте осуществления настоящего изобретения, или в другом варианте осуществления настоящего изобретения – в виде экспрессии белка слияния на основе плазмиды и последующего воздействия на матрицу для репарации, меченую RTDD. Дополнительный RTDD может быть совместно доставлен в случае, если предусмотрено более одного RTDD, например, один RTDD

представляет собой молекулу направляющей нуклеиновой кислоты, а другой RTDD представляет собой молекулу, например, биотина или маркера, например флуоресцеина, ассоциированного с RT. Подходы на основе плазмид или векторов по настоящему изобретению также включают подходы стабильной экспрессорной линии SSN и/или IA, и/или их белка слияния.

В одном варианте осуществления настоящего изобретения, искусственный молекулярный комплекс содержит две молекулы SSN, причем одна SSN представляет собой активную нуклеазу, а другая SSN представляет собой каталитически неактивную нуклеазо-дефицитную молекулу, при этом, неактивная SSN будет функционировать в качестве взаимодействующего партнера для RTD/RT. Указанная конфигурация искусственного молекулярного комплекса может усиливать специфичность для интересующих определенных последовательностей ДНК.

В другом варианте осуществления настоящего изобретения, белок слияния или нековалентно ассоциированный активный Cpf1 и неактивный dCas9 в качестве взаимодействующего домена могут быть предложены в виде SSN. gPHK для Cas9 в качестве RTDD может нацеливаться на матрицу для репарации или ее удлинение, образуя комплекс Cpf1-dCas9-RT. crPHK (Cpf1) нацеливается на геномный локус, определенный для двухцепочечного разреза, чтобы инициировать HDR.

Аналогичным образом, в качестве взаимодействующего домена можно использовать высокоактивный цинк-пальцевой белок, мегаTAL или неактивную мегануклеазу.

В одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, по меньшей мере, один стыковочный домен матрицы для репарации (RTDD), или последовательность нуклеиновой кислоты, кодирующая последовательность, или, по меньшей мере, один искусственный молекулярный комплекс выбран, по меньшей мере, из одного биотина, аптамера, ДНК, РНК или белкового красителя, содержащего флуорофоры, содержащие флуоресцеин или его вариант, малеимидов или Tetraxolium (XTT), направляющей последовательности нуклеиновой кислоты, специально выполненной с возможностью взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислоты, стрептавидина или его варианта, в предпочтительном варианте осуществления настоящего изобретения, из мономерного стрептавидина, авидина или его варианта, аффинной метки, в предпочтительном варианте осуществления настоящего изобретения, стрептовидиновой метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (наночастицы), антикалина, белка Agrobacterium VirD2 или его домена (см., например, SEQ ID NO: 33), пикорнавируса VPg, топоизомеразы или ее домена, белка А бактериофага ФХ174, белка А* ФХ, белка VirE2 или его домена, или дигоксигенина. Следовательно, RTDD может представлять собой молекулу природного происхождения или синтетическую молекулу, не ограниченную молекулой нуклеиновой кислоты или аминокислоты. Таким образом, RTDD представляет собой скорее специфический посредник взаимодействия искусственного молекулярного комплекса по настоящему изобретению, который может быть сконструирован универсальным образом для соединения, по меньшей мере, одной интересующей SSN и, по меньшей мере, одной матрицы для репарации, специфичной для интересующей геномной комплементарной области, и, при необходимости, несущий интересующую вставку, подлежащую введению в интересующую последовательность ДНК-мишени, расщепляемую, по меньшей мере, одной SSN. Для вариантов осуществления настоящего изобретения, использующих SSN на основе CRISPR или Аргонавтов, RTDD может представлять собой направляющую последовательность нуклеиновой кислоты. Таким образом, RTDD по настоящему изобретению может представлять собой молекулу, принадлежащую к различным классам молекул искусственного или природного происхождения. Таким образом, RTDD определяется по

его способности непосредственного взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислоты (RT) и, дополнительно, непосредственного взаимодействия, по меньшей мере, с одной SSN. Таким образом, RTDD представляет собой молекулярный линкер внутри искусственного молекулярного комплекса, предлагающий непосредственную физическую близость RT и SSN, и – благодаря его двойному взаимодействию с RT и SSN – гарантирующий ассоциацию искусственного молекулярного комплекса *in vitro* и *in vivo* посредством высокоспецифичных молекулярных взаимодействий. В определенных вариантах осуществления настоящего изобретения, могут присутствовать более одного RTDD, несущего более одной RT.

В другом варианте осуществления настоящего изобретения, искусственный молекулярный комплекс содержит взаимодействующий домен, при этом, по меньшей мере, один взаимодействующий домен или последовательность нуклеиновой кислоты, кодирующая последовательность, выбрана, по меньшей мере, из ДНК-связывающего домена, стрептавидина или его варианта, в предпочтительном варианте осуществления настоящего изобретения, из мономерного стептавидина, авидина или его варианта, аффинной метки, сигнала биотинилирования, сайта акцептора биотина, стрептовидиновой метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (наночастицы), антикалина, биотина, аптамера, ДНК, РНК или белкового красителя, содержащего флуорофоры, содержащие флуоресцеин или его вариант, малеимидов или Tetraxolium (XTT), направляющей последовательности нуклеиновой кислоты, специально выполненной с возможностью взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислотой, белка Agrobacterium VirD2 или его домена, пикорнавируса VPg, топоизомеразы или ее домена, белка A бактериофага ФХ174, белка A* ФХ, белка VirE2 или его домена, или дигоксигенина.

В частности, RTDD и взаимодействующий домен могут быть выбраны из сопоставимого и перекрывающегося класса молекул благодаря тому, что взаимодействующий домен является, при необходимости, компонентом, который может дополнительно оптимизировать специфичность или эффективность искусственного молекулярного комплекса по настоящему изобретению. Присутствие взаимодействующего домена может иметь значение для вариантов осуществления настоящего изобретения, использующих искусственные молекулярные комплексы, при этом, в качестве SSN не используют нуклеазу, направляемую нуклеиновой кислотой, или при этом, SSN несет, по меньшей мере, одну мутацию, модифицирующую внутреннюю активность SSN по распознаванию, связыванию или расщеплению ДНК. В еще одном варианте осуществления настоящего изобретения, присутствие взаимодействующего домена может быть благоприятным для использования в комбинации с любым видом SSN для дальнейшего расширения целевого диапазона, эффективности связывания и/или расщепления, скорости расщепления или точности нацеливания на интересующую последовательность ДНК-мишени, поскольку взаимодействующий домен в качестве дополнительного компонента в искусственном молекулярном комплексе может добавить расширенную функциональную возможность комплексу и, таким образом, расширить область его применения. В частности, для генной инженерии у высших эукариот, содержащих сложные геномы, присутствие дополнительного компонента, то есть, взаимодействующего домена, может, таким образом, иметь первостепенное значение для достижения улучшенной точности расщепления ДНК и – опосредованной посредством RT по настоящему изобретению – целевой репарации. В определенных вариантах осуществления настоящего изобретения, IA может представлять собой высокоспецифичного партнера по связыванию для молекулярного партнера, не участвующего в самой геномной инженерии, при этом, молекулярный партнер или когнатный партнер по связыванию представляет собой RTDD,

ассоциированный с RT. Следовательно, дополнительный уровень добавления домена IA, а также когнатного партнера RTDD, может значительно повысить специфичность связывания и доступность RT для искусственного молекулярного комплекса, что улучшит результат подхода целевой геномной инженерии.

Взаимодействующий домен (IA) по настоящему изобретению имеет несколько функциональных возможностей, выбранных из группы, состоящей из (i) взаимодействия, по меньшей мере, с одним стыковочным доменом матрицы для репарации; и/или (ii) взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислоты; и/или (iii) сиквенс-специфичного взаимодействия с геномной ДНК. Более одной из этих функциональных возможностей может быть объединено в одном конкретном IA.

Возможно, было бы предпочтительным использовать IA, который представляет собой белок или полипептид, обладающий внутренними свойствами высокой специфичности и высокой аффинностью связывания в отношении когнатного лиганда, например, синтетического лиганда, включая флуоресцеин для биомолекулы, включая биотин или дигоксигенин и их варианты, для аптамера или антигена/эпитопа. Термин «антиген», используемый в настоящем документе и обычно используемый в области иммунологии, относится к молекуле, «генерирующей антитела», то есть, к веществу, которое может вызывать адаптивный иммунный ответ. Таким образом, антиген представляет собой молекулу, связывающуюся с антигенспецифичным рецептором, либо с Т-клеточным, либо с В-клеточным рецептором, либо с их вариантом, например, с нанотелом или одноцепочечным вариабельным фрагментом антитела, биспецифичными антителами тандема di-scFv, диатела, тандема tri-scFv (трехвалентного) или триатела (трехвалентного). Антиген обычно представляет собой (поли)пептид, но он также может представлять собой полисахарид или липид, возможно, в комбинации с молекулой-носителем белка или полисахарида. Благодаря опосредованию этим внутренним свойством связывания/распознавания IA, можно выбрать интересующий IA, который будет конкретно распознавать RTDD очень специфичным образом, и IA может быть соединен или слит, ковалентно или нековалентно, с SSN. Таким образом, включение такого IA добавляет дополнительный уровень специфичности искусственному молекулярному комплексу по настоящему изобретению и гарантирует, что RT, непосредственно взаимодействующая с RTDD, будет специфически ассоциирована с комплексом SSN-IA, опосредованным высокоспецифичной ассоциацией IA - RTDD. В наиболее предпочтительном варианте осуществления настоящего изобретения, IA и когнатный RTDD имеют высокую константу аффинности или аффинность связывания и, таким образом, низкую константу диссоциации (Kd) в отношении друг друга в физиологических условиях, то есть, имеют значение Kd в диапазоне низких значений pM или, в предпочтительном варианте осуществления настоящего изобретения, в диапазоне значений nM, и в предпочтительном варианте осуществления настоящего изобретения, ниже этих значений. IA может представлять собой одновалентную, двухвалентную, трехвалентную или многовалентную молекулу, имеющую, по меньшей мере, одну специфичность (фрагмент, полученный из трехвалентного антитела), соответственно, или имеющую более одного сайта связывания (тетрамерный стрептавидин). В этих вариантах осуществления настоящего изобретения, более одного RTDD и/или RT могут присутствовать и передаваться, по меньшей мере, одной SSN искусственным молекулярным комплексом. Предпочтительными являются те IA, которые имеют низкие константы диссоциации (Kd), то есть, в свою очередь, которые имеют высокую аффинность в отношении своих когнатных лиганд. Обычно константы субпикомолярной диссоциации, полученные в результате нековалентных связывающих взаимодействий между двумя молекулами, то есть, формы типичного взаимодействия между белком и лигандом, редки. Тем не менее, имеются некоторые важные исключения.

Биотин и авидин природного происхождения связываются с константой диссоциации со значением, приблизительно, 10^{-15} М, которая представляет собой такую высокую аффинность, которая не подходит для применений, где предполагается проведение обратимого связывания. Доступные на рынке антитела или scFv могут иметь значения Kd в диапазоне от 10^{-14} М до 10^{-6} М. Для цели настоящего изобретения, пара IA-RTDD должна, таким образом, иметь низкую константу диссоциации, то есть, высокую аффинность.

Кроме того, в определенных вариантах осуществления настоящего изобретения, IA может непосредственно взаимодействовать с RT. Когда RT последовательность нуклеиновой кислоты содержит удлинение, например, аптамер на основе нуклеиновой кислоты, то эта последовательность может быть распознана когнатным партнером по связыванию, IA, который затем может взаимодействовать с RT высокоспецифичным образом. Кроме того, IA может представлять собой двухвалентную, трехвалентную или многовалентную молекулу, имеющую более одной специфичности связывания. Одна часть IA может быть выполнена с возможностью взаимодействия с RTDD, и одна часть может быть выполнена с возможностью взаимодействия с RT, в то время как IA ассоциируется с SSN, так что может быть достигнута даже более тесная ассоциация RT и SSN во время геномной инженерии.

В другом варианте осуществления настоящего изобретения, IA может представлять собой связывающую молекулу, обладающую способностью сиквенс-специфичного взаимодействия с геномной ДНК. Это добавит больше специфичности при нацеливании искусственного молекулярного комплекса на последовательность ДНК-мишени. Кроме того, это позволяет использовать модифицированные SSN, чтобы можно было предложить SSN с оптимизированной активностью расщепления, при этом, IA опосредует функцию нацеливания искусственного молекулярного комплекса на последовательность ДНК-мишени с высокой точностью, при этом, SSN и/или IA могут взаимодействовать с RTDD, взаимодействующим с RT и, таким образом, передающим RT в сайт, где будет индуцирован двухцепочечной разрыв. Таким образом, в одном варианте осуществления настоящего изобретения, IA может представлять собой ДНК-связывающий домен или ДНК-связывающий мотив, сконструированный для того, чтобы быть частью белка слияния либо на N-терминальном конце, либо на С-терминальном конце, по меньшей мере, одной нуклеазы SSN или ее варианта. Линкер на основе аминокислоты обеспечит гибкость и позволит избежать стерического несоответствия для связывания ДНК или нуклеазной активности. Потенциальные ДНК-связывающие домены могут также представлять собой цинковые пальцы (Рой и др., 2012), такие как Cys2/His2 Цинковый палец (Кубо и др., 1998), Эффекторные нуклеазы, подобные активаторам транскрипции (Хаббард и др., 2015) или инактивированные белки Аргонавты или Cas9, способные к высокоспецифичному связыванию ДНК. Любой из этих ДНК-связывающих доменов в идеале должен быть нацелен на последовательность, находящуюся вне интересующей гомологичной фланкированной последовательности, чтобы избежать стерического несоответствия взаимодействия, и, таким образом, может добавить другой уровень специфичности искусственному молекулярному комплексу по настоящему изобретению.

В другом варианте осуществления настоящего изобретения, в искусственном молекулярном комплексе можно использовать более одного домена IA, то есть, один IA, используемый в качестве связующего элемента с высокой специфичностью и аффинностью для RTDD, и один IA, используемый в качестве дополнительного ДНК-связывающего домена, причем оба IA находятся в непосредственном взаимодействии, то есть, ковалентно или нековалентно ассоциированы с, по меньшей мере, одной SSN искусственного молекулярного комплекса.

В одном варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и/или, по меньшей мере, один IA содержат сигнал биотинилирования или сайт биотинилированного акцептора, или

стреп-метку. Соответствующий сигнал/сайт может быть биотинилирован *in vitro* или *in vivo* эндогенными (BirA) или экзогенными биотинилирующими ферментами/агентами, или на этапе биотинилирования *in vitro*, и затем сигнал/сайт биотинилирования и/или стреп-метка может быть распознана и связана стрептавидином или авидином, или, в предпочтительном варианте осуществления настоящего изобретения – его модифицированным вариантом, в самом предпочтительном варианте осуществления настоящего изобретения – его мономерным вариантом, при этом, стрептавидин или авидин, или его вариант будет ассоциирован с интересующей RT. Поскольку известно, что авидин неспецифично взаимодействует с ДНК (Морпурго и др., 2004), то в предпочтительном варианте осуществления настоящего изобретения, это могли бы быть модифицированные варианты авидина, или в более предпочтительном варианте осуществления настоящего изобретения – стрептавидин или его варианты.

В частности, для SSN, не основанные на направляющих РНК/ДНК, дополнительная связывающая способность и, следовательно, способность к нацеливанию на RT мономерного стрептавида или scFv с заданной специфичностью связывания может значительно расширить диапазон подходящих SSN для геномной инженерии, если они используются в комбинации с RTDD и/или IA по настоящему изобретению.

В одном варианте осуществления настоящего изобретения, биотин может быть слит с матрицей для репарации ДНК посредством доступных на рынке наборов или как часть проведенного сторонней компанией процесса синтеза, например, RTDD. Использование последовательности модифицированного стрептавидина или авидина гарантирует, что не происходит межбелкового комплексообразования, и что одна биотинилированная матрица для репарации ДНК связана с белком. Затем матрицу для репарации сцепляют со стрептавидином или любым его вариантом в качестве взаимодействующего домена (Нимейер и др., 1999, «Функционализация ковалентных конъюгатов ДНК-стрептавидина посредством биотинилированных компонентов модулятора». Bioconjug Chem 10(5): 708-719), при этом, взаимодействующий домен непосредственно взаимодействует с SSN, например, путем предложения SSN и стрептавидина в качестве молекулы слияния. В другом варианте осуществления настоящего изобретения, SSN может содержать сигнал биотинилирования или пептид, и биотинилирование будет происходить *in vivo* в клетке-хозяине. В этом варианте осуществления настоящего изобретения, стрептавидин или авидин, или его вариант, может функционировать как сам RTDD, сцепленный с RT. Примерная последовательность, кодирующая мономерный стрептавидин (mSA), подходящая в качестве взаимодействующего домена или в качестве RTDD по настоящему изобретению, показана в SEQ ID NO: 34. Таким образом, mSA, слитый с SSN, можно понимать либо как RTDD, либо как взаимодействующий домен по настоящему изобретению. В другом варианте осуществления настоящего изобретения, SSN может нести стреп-метку, причем эта метка распознается вариантом стрептавидина, функционирующим в качестве взаимодействующего домена или в качестве RTDD, соответственно. Подходящие ферменты стрептавидина или авидина или их варианты, или векторы, кодирующие векторы, доступны для специалиста в данной области техники, например, из источников: компания IBA Lifesciences (Геттинген, Германия), Addgene (некоммерческое хранилище биологических материалов) (Кембридж, Массачусетс, США), компании Integrated DNA Technologies (Коралвилл, Айова, США) или компания GeneArt (ThermoFisher; Уолтем, Массачусетс, США). Другая примерная последовательность для конструкции мономерного стрептавидина, кодирующая mSA, подходящая в качестве IA или RTDD по настоящему изобретению, показана в SEQ ID NO: 42.

В одном варианте осуществления настоящего изобретения, взаимодействие или присоединение, или ассоциация между RTDD и SSN, и/или взаимодействующим доменом, таким образом, является результатом взаимодействия связывающей пары, выбранной из нековалентного взаимодействия

связывающей пары, выбранной из, но этим не ограничиваясь: биотина-авидина; биотина-стрептавидина; модифицированных биотином форм авидина; белка-белка; взаимодействий белка-нуклеиновой кислоты; взаимодействий лиганда-рецептора; взаимодействий лиганда-субстрата; антитела-антигена; одноцепочечного антитела-антигена; антитела или одноцепочечного антитела-гаптена, гормон-гормон-связывающего белка, рецептора-антагониста, рецептора-антагониста рецептора; IgG-белка A; фермента-кофактора фермента; фермента-ингибитора фермента; одноцепочечной ДНК-VirE2; StickyC-dsДНК; RISC (РНК-индуцированного комплекса молчания)-РНК; белка вирусной оболочки-нуклеиновой кислоты; анти-флуоресцеинового одноцепочечного вариабельного фрагмента антитела (анти-FAM scFV)-флуоресцеина; анти-дигоксигенинового (DIG) одноцепочечного вариабельного фрагмента (scFv) иммуноглобина (DIG-scFv)-дигоксигенина (DIG) и *Agrobacterium* VirD2-связывающего белка или любой их комбинации, или варианта. Примечательно, что антитела и фрагменты антител или производные типа scFv, нанотела или диатела, имеющие индивидуальные специфичности и высокие аффинности (в диапазоне pM или даже fM), являются доступными на рынке, в частности, такие антитела или фрагменты, или их варианты, связывающие классические красители, такие как, флуоресцеин или его производные.

В одном варианте осуществления настоящего изобретения, взаимодействующий домен по настоящему изобретению выбран из лейциновой «молнии», последовательности аптамера, dCas9, dCPF1, мегануклеазы, цинкового пальца или конструкции TALE. В этом варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и RT ДНК может быть приведена в непосредственное взаимодействие посредством промежуточного ДНК-связывающего домена или ДНК-связывающего мотива, сконструированного для того, чтобы быть частью белка слияния на N-терминальном и/или C-терминальном конце SSN. Линкер на основе аминокислоты обеспечит гибкость и позволит избежать стерического несоответствия для связывания ДНК или нуклеазной активности. Потенциальные ДНК-связывающие домены могут также представлять собой цинковые пальцы (Рой и др., 2012, «Прогнозирование ДНК-связывающей специфичности в цинк-пальцовых белках». Дж. Биосси 37 (3): 483- 491), такие как Cys2/His2 Цинковый палец (Кубо и др., 1998, «Семейство цинк-пальцевых белков Cys2/His2 петунии: эволюция и общий механизм распознавания последовательности-мишени». Nucleic Acids Research 26 (2): 608-615), 608-615), Эффекторные нуклеазы, подобные активаторам транскрипции (Хаббард и др., 2015, «Непрерывная направленная эволюция ДНК-связывающих белков для улучшения специфичности TALEN». Nat Methods 12(10): 939-942) или инактивированные белки Аргонавты или Cas9, способные к высокоспецифичному связыванию ДНК. Любой из этих ДНК-связывающих доменов в качестве взаимодействующих доменов может дополнительно помочь нацелить последовательность, находящуюся вне интересующей гомологичной фланкированной последовательности, чтобы избежать стерического несоответствия взаимодействия. Указанные взаимодействующие домены могут выполнять функцию усиления ДНК-связывания искусственного молекулярного комплекса по настоящему изобретению и/или могут обеспечить предоставление дополнительных стыковочных сайтов для сцепления RTDD/RT, чтобы предложить высокоспецифичный комплекс, подходящий для геномной инженерии.

Согласно определенным вариантам осуществления настоящего изобретения, по меньшей мере, одна SSN по настоящему изобретению может быть слита с ДНК-связывающим доменом, то есть, с белком или его фрагментом, или с последовательностью гена, кодирующей указанный белок или фрагмент белка, который связывает молекулы ДНК или связывает другие клеточные молекулы, включая, но этим не ограничиваясь, мальтозосвязывающий белок (MBP), S-метку, слияния ДНК-связывающего домена (DBD) LexA, слияния ДНК-связывающего домена GAL4 и слияния белка BP16 вируса простого герпеса (HSV).

В определенных вариантах осуществления настоящего изобретения, может быть более одного RTDD. Первый RTDD может представлять собой mSA или одноцепочечный вариабельный фрагмент (scFv), тогда как второй RTDD может представлять собой биотин или когнатный лиганд scFv. В одном варианте осуществления настоящего изобретения, используя систему SSN на основе CRISPR или Аргонавтов, первый RTDD представляет собой направляющую последовательность нуклеиновой кислоты, а второй RTDD представляет собой биотин или флуоресцеин, или любой другой фрагмент молекулы партнера по связыванию с высокой аффинностью, сцепленный с RT, при этом, мономерный стрептавидин или scFv, или другой когнатный партнер по связыванию белка представляет собой IA, распознающий второй RTDD и связанный с ним с высокой аффинностью. Такая конструкция искусственного молекулярного комплекса по настоящему изобретению обеспечивает максимальную гибкость для приведения RT в тесный контакт с эффекторной SSN, одновременно предлагая высокую доступность RT и отсутствие потери RT, поскольку RTDD предлагают прочные и надежные взаимодействия с RT и SSN для достижения событий точной геномной инженерии.

В одном варианте осуществления настоящего изобретения, сцепление матрицы для репарации с SSN может быть достигнуто посредством одноцепочечного вариабельного фрагмента (scFv) антитела против красителя флуоресцеина. scFv, специфически связывающий флуоресцеин и производные флуоресцеина, слит с SSN гибридно-белковым образом (Шенк и др., 2007, «Получение и применение флуоресцеин-специфичного одноцепочечного антитела». Biochimie 89(11): 1304-1311). В другом варианте осуществления настоящего изобретения, SSN может содержать молекулу флуоресцеина, взаимодействующую с ней, и когнатный флуоресцеин-специфичный scFv может быть предложен в виде слияния с RT и может связывать флуоресцеин, ассоциированный с SSN. Таким образом, scFv может функционировать как RTDD или как взаимодействующий домен по настоящему изобретению.

В других вариантах осуществления настоящего изобретения, может быть использован любой scFv с другой специфичностью связывания.

Подходящие пары scFv-лиганд выбирают из группы, состоящей из scFv и флуоресцеина (FAM) или любого производного, или варианта FAM, scFv, распознающего дигоксигенин (DIG), индивидуального scFv, распознающего эпитоп/антиген в интересующей SSN, и тому подобного. Одна примерная последовательность, кодирующая последовательность, кодирующую scFv со способностью связывания флуоресцеина, показана в SEQ ID NO: 43.

В другом варианте осуществления настоящего изобретения, последовательность аптамера сконструирована для специфичного взаимодействия, по меньшей мере, с одной SSN. В этом варианте осуществления настоящего изобретения, последовательность аптамера может быть ковалентно или нековалентно сцеплена с интересующей последовательностью матрицы для репарации, чтобы обеспечить непосредственную ассоциацию между SSN и аптамером в качестве RTDD без создания белка слияния и/или использования дополнительного взаимодействующего домена. В вариантах осуществления настоящего изобретения, в которых не используется отдельный взаимодействующий домен, RTDD, взаимодействующий с RT, содержит нуклеотидный мотив, способный специфически взаимодействовать, то есть, присоединяться или связываться с доменом, по меньшей мере, одного белка SSN или с ее специфичный доменом, выполненным с возможностью взаимодействия с RTDD: В некоторых вариантах осуществления настоящего изобретения, взаимодействие выбирают, но этим не ограничивается, из следующего: цинк-пальцевый белок-цинк-пальцевый мотив; домен распознавания рестрикционного фермента-последовательность распознавания рестрикционного фермента; ДНК-связывающий домен фактора транскрипции-мотив ДНК;

репрессор-оператор; лейциновая «молния»-промотор; спираль-петля-спираль-домен E-бокса; РНК-связывающие мотивы, содержащие домены аргинин-богатых мотивов, домены ар-белков, домены РНК-распознающих мотивов (RRM), домены K-гомологии, мотивы, связывающие двухцепочечную РНК, РНК-связывающие цинковые пальцы, и ферменты, нацеленные на РНК-когнатная специфичная последовательность РНК; белок ВИЧ-rev-Стебель I IB элемента ответа (RRE) на ВИЧ-rev; Tat-главный связывающий домен бычьего вируса иммунодефицита (BIV)-петля 1 последовательности элемента транс-действующего ответа (TAR) на BIV; Фаг лямбда, фаг21, и белки P22 N, шпильки петли бокса-B в сайтах N-использования (nut) в соответствующих им РНК.

Поскольку настоящее изобретение относится к применению Аргонавта в качестве сайт-специфичной нуклеазы, в дополнение к преимуществам молекулы направляющей ДНК, доставка эндонуклеазы NgAgo облегчается, благодаря ее малому размеру. Белок дикого типа (WT) (Номер доступа в GenBank AFZ73749) состоит из 887 аминокислот или имеет, приблизительно, 2/3 размера Cas9 *Streptococcus pyogenes*. Это упрощает клонирование и сборку вектора, может увеличивать уровни экспрессии нуклеазы в клетках и снижает проблему экспрессии белка из очень чувствительных к размеру платформ, таких как вирусы, включая вирусы ДНК или РНК. Как и для других направляемых нуклеиновой кислотой эндонуклеаз, для SSN NgAgo обычно требуется минимум два компонента для целевого мутагенеза в растительных клетках: 5'-фосфорилированная одноцепочечная направляющая ДНК и белок эндонуклеазы NgAgo. Для проведения целевых редактирований, вставок или замен последовательностей, для введения изменений в растительную клетку также может быть предложена матрица ДНК, кодирующая желаемые изменения последовательности, посредством следующих путей репарации: либо NHEJ, либо HR. Успешные события редактирования чаще всего обнаруживаются посредством фенотипических изменений (например, посредством нокаута или введения гена, что приводит к видимому фенотипу), посредством способов на основе PCR (таких как, обогащение PCR, PCR-расщепление или T7EI, или анализы эндонуклеазы Surveyor), или посредством целевого секвенирования следующего поколения (NGS; также называемого глубоким секвенированием). В одном особом варианте осуществления настоящего изобретения, модифицированная эндонуклеаза Аргонавт активна при температуре от, примерно, 20° C до, примерно, 35° C. В одном особом варианте осуществления настоящего изобретения, модифицированная эндонуклеаза Аргонавт активна при температуре от, примерно, 23° C до, примерно, 32° C. Белки Аргонавты, которые могут функционировать в качестве эндонуклеаз, могут содержать три ключевых функциональных домена: домен эндонуклеазы PIWI, *PAZ*домен и домен MID. Домен PIWI может напоминать нуклеазу. Нуклеаза может представлять собой РНКазу H или рибонуклеазу, направляемую ДНК. Домен PIWI может иметь двухвалентный катион-связывающий мотив для катализа, проявляемого другими нуклеазами, которые могут расщеплять РНК и ДНК. Двухвалентный катион-связывающий мотив может содержать четыре отрицательно заряженные, эволюционно консервативные аминокислоты. Четыре отрицательно заряженные эволюционно консервативные аминокислоты могут представлять собой аспартат-глутамат-аспартат-аспартат (DEDD). Четыре отрицательно заряженные эволюционно консервативные аминокислоты могут образовывать каталитическую тетраду, которая связывает два иона Mg2+ и расщепляет нуклеиновую кислоту-мишень на продукты, несущие 3'-гидроксильную и 5'-фосфатную группу. Домен PIWI может дополнительно содержать, по меньшей мере, одну аминокислоту, выбранную из основного остатка. Домен PIWI может дополнительно содержать, по меньшей мере, одну аминокислоту, выбранную из гистидина, аргинина, лизина и их комбинации. Гистидин, аргинин и/или лизин могут играть важную роль в катализе и/или расщеплении. Расщепление нуклеиновой кислоты-мишени Аргонавтом может происходить по одной фосфодиэфирной

связи. В некоторых случаях, по меньшей мере, один катион магния и/или марганца может способствовать расщеплению нуклеиновой кислоты-мишени, при этом, первый катион может нуклеофильно атаковать и активировать молекулу воды, а второй катион может стабилизировать состояние транзиции и замещаемую группу. Для некоторых нуклеаз Аргонавтов, длина gДНК будет предлагать аффинность между Аргонавтом и направляющей gДНК.

Подходящие белки Аргонавтов по настоящему изобретению показаны посредством SEQ ID NO: 19 и 20, или они могут содержать последовательность, имеющую в своем составе, по меньшей мере, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% , 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95 %, 96%, 97%, 98% или 99% гомологии последовательностей при условии, что гомологичная последовательность по-прежнему выполняет функцию белка Аргонавта, из которого она получена, то есть, происходит из него. Другие подходящие последовательности Аргонавтов раскрыты в предварительной заявке США № 62/345,448, которая включена в настоящий документ посредством ссылки. Другие подходящие последовательности Аргонавтов могут быть получены из последовательности в соответствии с SEQ ID NO: 21-29 или последовательности, имеющей в своем составе, по меньшей мере, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79% , 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96 %, 97%, 98% или 99% гомологии последовательностей.

Аргонавт может содержать домен, связывающий нуклеиновую кислоту. Домен, связывающий нуклеиновую кислоту, может содержать область, которая контактирует с нуклеиновой кислотой. Домен, связывающий нуклеиновую кислоту, может содержать нуклеиновую кислоту. Домен, связывающий нуклеиновую кислоту, может содержать белковый материал. Домен, связывающий нуклеиновую кислоту, может содержать нуклеиновую кислоту и белковый материал. Домен, связывающий нуклеиновую кислоту, может содержать ДНК. Домен, связывающий нуклеиновую кислоту может содержать одноцепочечную ДНК. Примеры доменов, связывающих нуклеиновую кислоту, могут включать, но этим не ограничиваются, домен спираль-поворот-спираль, цинк-пальцевый домен, домен лейциновая молния (bZIP), домен крылатая спираль, домен крылатая спираль-поворот-спираль, домен спираль-петля-спираль, домен HMG-бокс, домен Wor3, домен иммуноглобулина, домен B3 или домен TALE. Домен, связывающий нуклеиновую кислоту, может представлять собой домен белка Аргонавта. Белок Аргонавт может представлять собой эукариотический Аргонавт или прокариотический Аргонавт. Белок Аргонавт может связывать РНК или ДНК, или и РНК, и ДНК. Белок Аргонавт может расщеплять РНК или ДНК, или и РНК, и ДНК. В некоторых случаях, белок Аргонавт связывает ДНК и расщепляет ДНК. В некоторых случаях, белок Аргонавт связывает двухцепочечную ДНК и расщепляет двухцепочечную ДНК. В некоторых случаях, по меньшей мере, два домена, связывающих нуклеиновую кислоту, могут быть сцеплены друг с другом. Сцепление друг с другом множества доменов, связывающих нуклеиновые кислоты, может предложить повышенную специфичность нацеливания на полинуклеотиды. По меньшей мере, два домена, связывающих нуклеиновые кислоты, могут быть сцеплены посредством, по меньшей мере, одного линкера. Линкер может представлять собой гибкий линкер. Линкеры могут содержать 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, 30, 35, 40 или более аминокислот в длину. Линкерный домен может содержать глицин и/или серин, и в некоторых вариантах осуществления настоящего изобретения, может состоять или может состоять по существу из глицина и/или серина. Линкеры могут представлять собой линкер нуклеиновой кислоты, который может содержать нуклеотиды. Линкер нуклеиновой кислоты может сцеплять друг с другом два ДНК-связывающих домена. Линкер нуклеиновой кислоты может иметь длину не более 5, 10, 15,

20, 25, 30, 35, 40, 45 или 50, или более нуклеотидов. Линкер нуклеиновой кислоты может иметь длину, по меньшей мере, 5, 10, 15, 30, 35, 40, 45 или 50, или более нуклеотидов. Домены, связывающие нуклеиновую кислоту, могут связываться с последовательностями нуклеиновых кислот. Домены, связывающие нуклеиновую кислоту, могут связываться с нуклеиновыми кислотами посредством гибридизации. Домены, связывающие нуклеиновую кислоту, могут быть сконструированы (например, сконструированы для гибридизации с последовательностью в геноме). Домен, связывающий нуклеиновую кислоту, может быть сконструирован методами молекулярного клонирования (например, направленной эволюцией, сайт-специфичной мутацией и рациональным мутагенезом).

В определенных вариантах осуществления настоящего изобретения, SSN по настоящему изобретению будет представлять собой нуклеазу CRISPR, включая Cas или Cpf1, или нуклеазу Аргонавта, или ее вариант, или каталитически активный фрагмент. Подходящие последовательности нуклеаз CRISPR выбраны из группы, состоящей из SEQ ID NO: 19-29 или 35-41, или дополнительной последовательности, имеющей в своем составе, по меньшей мере, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% гомологии последовательностей. Кроме того, подходящие эффекторы Cas или Cpf1 могут быть получены из организма рода, включающего *Streptococcus, Campylobacter, Candidatus* Micrarchaeum acidiphilum ARMAN-1, *Parcubacteria* (GenBank: APG80656.1), *Sulfolobus* spp., включая *Sulfolobus islandicus* HVE10/4 (GenBank: ADX81770.1) или REY15A (GenBank: ADX84852.1), *Nitratifractor, Staphylococcus, Parvibaculum, Roseburia, Neisseria, Gluconacetobacter, Azospirillum, Sphaerochaeta, Lactobacillus, Eubacterium, Corynebacter, Carnobacterium, Rhodobacter, Listeria, Paludibacter, Clostridium, Lachnospiraceae, Clostridiaridium, Leptotrichia, Francisella, Legionella, Alicyclobacillus, Methanomethyophilus, Porphyromonas, Prevotella, Bacteroidetes, Helcococcus, Letospira, Desulfovibrio, Desulfonatronum, Opitutaceae, Tuberibacillus, Bacillus, Brevibacilus, Methylobacterium* или *Acidaminococcus.*, например, из *S. mutans, S. agalactiae, S. equisimilis, S. sanguinis, S. pneumonia; C. jejuni, C. coli; N. salsuginis, N. tergarcus; S. auricularis, S. carnosus; N. meningitides, N. gonorrhoeae; L. monocytogenes, L. ivanovii; C. botulinum, C. difficile, C. tetani, C. sordellii.*

В одном варианте осуществления настоящего изобретения, по меньшей мере, одна сайт-специфичная нуклеаза или ее каталитически активный фрагмент в качестве части искусственного молекулярного комплекса по настоящему изобретению, или последовательность, кодирующая последовательность, выбрана независимо из группы, состоящей из полипептида Cas из *Streptococcus spp.*, включая *Streptococcus pyogenes, Streptococcus thermophilus, Staphylococcus aureus,* или *Neisseria spp.*, включая *Neisseria meningitides, Corynebacter, Sutterella, Legionella, Treponema, Filifactor, Eubacterium, Lactobacillus, Mycoplasma, Bacteroides, Flaviivola, Flavobacterium, Sphaerochaeta, Azospirillum, Gluconacetobacter, Roseburia, Parvibaculum, Nitratifractor, Mycoplasma* и *Campylobacter, Candidatus* Micrarchaeum acidiphilum ARMAN-1, *Parcubacteria* (GenBank: APG80656.1), *Sulfolobus* spp., включая *Sulfolobus islandicus* HVE10/4 (GenBank: ADX81770.1) или REY15A (GenBank: ADX84852.1), полипептид Cpf1 из археи или бактерии, включая полипептид Cpf1 *Acidaminococcus spp.*, включая *Acidaminococcus sp.* BV3L6, *Lachnospiraceae spp.*, включая бактерию *Lachnospiraceae* ND2006, бактерию *Lachnospiraceae* MC2017, бактерию *Lachnospiraceae* MA2020, *Butyrivibrio proteoclasticus,* Candidatus spp., *Methanoplasma termitum, Leptospira inadai, Moraxella bovoculi* 237, бактерию Peregrinibacteria GW2011_GWA2_33_10, бактерию Parcubacteria GW2011_GWC2_44_17, Smithella sp. SCADC, Smithella sp. SC_K08D17, *Francisella spp.*, включая *Francisella novicida* U112, *Eubacterium eligens, Prevotella spp.*, или *Porphyromonas spp.*, или

нуклеазу Аргонавта из *Natronobacterium gregoryi* (GenBank: AFZ73749.1), *Microcystis aeruginosa* (Стандартная/эталонная последовательность Национального центра биотехнологической информации (NCBI): WP_012265209.1 или Стандартная/эталонная последовательность NCBI: WP_002747795.1 или Стандартная/эталонная последовательность NCBI: WP_012265209.1), *Halogeometricum pallidum* (GenBank): ELZ29017.1), *Natrialaba asiatica* (Стандартная/эталонная последовательность NCBI: WP_006111085.1), *Natronorubrum tibetense* (Стандартная/эталонная последовательность NCBI: WP_006090832.1), *Natrinema pellirubrum* (Стандартная/эталонная последовательность NCBI: WP_006183335.1), или *Synechococcus spp.* (Стандартная/эталонная последовательность NCBI: WP_011378069.1) или вариантов и/или функциональных фрагментов, и/или их комбинаций, включая никазы, или нуклеазы, не обладающие эндонуклеолитической активностью.

В других вариантах осуществления настоящего изобретения, в которых используется, по меньшей мере, один эффектор Cpf1 в качестве SSN, мотив, примыкающий к протоспейсеру (PAM), или PAM-подобный мотив направляет связывание комплекса эффекторного белка к интересующему локусу-мишени. В варианте осуществления настоящего изобретения, в котором используется, по меньшей мере, один эффектор Cpf1 в качестве SSN, PAM представляет собой 5' TTN, где N представляет собой A/C/G или T. В другом предпочтительном варианте осуществления настоящего изобретения, PAM представляет собой 5' TTTV, где V представляет собой A/C или G. В некоторых вариантах осуществления настоящего изобретения, PAM представляет собой 5'-TTN, где N представляет собой A/C/G или T, и PAM расположен выше 5'-конца протоспейсера. В некоторых вариантах осуществления настоящего изобретения, PAM представляет собой 5 'CTA, и PAM расположен выше 5'-конца протоспейсера или локуса-мишени. В некоторых вариантах осуществления настоящего изобретения, предлагается расширенный целевой диапазон для нуклеаз редактирования генома, направляемых РНК, где Т-богатые PAM семейства Cpf1 обеспечивают нацеливание и редактирование АТ-богатых геномов.

В некоторых вариантах осуществления настоящего изобретения, фермент CRISPR сконструирован и может содержать, по меньшей мере, одну мутацию, которая уменьшают или устраняют нуклеазную активность. Аналогичным образом, в настоящем изобретении рассматриваются способы использования, по меньшей мере, двух никаз, в частности подход с двойной никазой или с двумя никазами для генерирования разрыва двухцепочечной ДНК-мишени.

В вариантах осуществления настоящего изобретения, где используются комплексы эффекторного белка Cpf1 в искусственном молекулярном комплексе по настоящему изобретению, может использоваться эффектор Cpf1, имеющий, по меньшей мере, один компонент нуклеиновой кислоты неприродного происхождения, сконструированный или модифицированный, или оптимизированный, или кодируемый белок. В предпочтительном варианте осуществления настоящего изобретения, компонент нуклеиновой кислоты комплекса может содержать направляющую последовательность, сцепленную с последовательностью прямого повтора, при этом, последовательность прямого повтора содержит, по меньшей мере, одну петлю стебля или оптимизированную вторичную структуру. В предпочтительном варианте осуществления настоящего изобретения, прямой повтор имеет минимальную длину из 16 нуклеотидов и одну петлю стебля. В других вариантах осуществления настоящего изобретения, прямой повтор имеет длину более 16 нуклеотидов, в предпочтительном варианте осуществления настоящего изобретения, более 17 нуклеотидов, и имеет более одной петли стебля или оптимизированной вторичной структуры. В предпочтительном варианте осуществления настоящего изобретения, прямой повтор может быть модифицирован таким образом, чтобы он сдержал, по меньшей мере, один аптамер РНК,

связывающийся с белком. В предпочтительном варианте осуществления настоящего изобретения, может быть включен, по меньшей мере, один аптамер, такой как часть оптимизированной вторичной структуры. Такие аптамеры могут обладать способностью связывать белок оболочки бактериофага. Белок оболочки бактериофага может быть выбран из группы, содержащей Qß, F2, GA, fr, JP501, MS2, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, Fl, ID2, NL95, TW19, AP205, φCb5, φCb8r, φCb12r, φCb23r, 7s и PRR1. В предпочтительном варианте осуществления настоящего изобретения, белок оболочки бактериофага представляет собой MS2.

В определенных вариантах осуществления настоящего изобретения, также предложено, чтобы, по меньшей мере, одна мутация или, по меньшей мере, две мутации находились в каталитически активном фрагменте, по меньшей мере, одного эффекторного белка SSN, содержащего домен RuvC. В некоторых вариантах осуществления настоящего изобретения, домен RuvC может содержать домен RuvCI, RuvCII или RuvCIII, или каталитически активный домен, который является гомологичным домену RuvCI, RuvCII или RuvCIII и так далее, или любому соответствующему домену, как описано в любом из способов, описанных в настоящем документе. Эффекторный белок SSN может содержать, по меньшей мере, один гетерологичный функциональный домен. По меньшей мере, один гетерологичный функциональный домен искусственного молекулярного комплекса может содержать, по меньшей мере, один домен сигнала ядерной локализации (NLS). По меньшей мере, один гетерологичный функциональный домен может содержать, по меньшей мере, два домена NLS. По меньшей мере, один домен NLS может быть расположен на или рядом, или вблизи терминального конца эффекторного белка (например, Cpf1), и, в случае, по меньшей мере, двух NLS, каждый из двух может быть расположен на или рядом, или вблизи терминального конца эффекторного белка (например, Cpf1). По меньшей мере, один гетерологичный функциональный домен может содержать, по меньшей мере, один домен активации транскрипции. В предпочтительном варианте осуществления настоящего изобретения, домен активации транскрипции может содержать VP64. По меньшей мере, один гетерологичный функциональный домен может содержать, по меньшей мере, один домен репрессии транскрипции. В предпочтительном варианте осуществления настоящего изобретения, домен репрессии транскрипции содержит домен KRAB или домен SID (например, SID4X). По меньшей мере, один гетерологичный функциональный домен может содержать, по меньшей мере, один нуклеазный домен в качестве SSN. В одном варианте осуществления настоящего изобретения, SSN может содержать Fok1 или каталитически активный фрагмент или его вариант.

В одном предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, одна сайт-специфичная нуклеаза или вариант, или ее каталитически активный фрагмент искусственного молекулярного комплекса по настоящему изобретению, или последовательность, кодирующая последовательность, выбрана из нуклеазы CRISPR, в предпочтительном варианте осуществления настоящего изобретения, из нуклеазы Cas или Cpf1, или нуклеазы FokI, или ее каталитического фрагмента, и, по меньшей мере, один взаимодействующий домен, или последовательность, кодирующая последовательность, выбрана из одноцепочечного вариабельного фрагмента или мономерного стрептавидина.

В одном варианте осуществления настоящего изобретения, искусственный молекулярный комплекс по настоящему изобретению содержит, по меньшей мере, одну один SSN, полученную из CRISPR или Аргонавта, или вариант, или ее каталитически активный фрагмент, и, по меньшей мере, одну направляющую последовательность нуклеиновой кислоты, представляющую собой, по меньшей мере, один стыковочный домен матрицы для репарации, при этом, каждая из, по меньшей мере, одной направляющей

последовательности нуклеиновой кислоты содержит (i) часть первой последовательности, которая комплементарна последовательности распознавания ДНК-мишени, и (ii) часть второй последовательности, при этом, вторая часть последовательности выполнена с возможностью взаимодействия, по меньшей мере, с одной сайт-специфичной нуклеазой, и (iii), при этом, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты физически ассоциирована с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, и, таким образом, образует гибридную последовательность нуклеиновой кислоты, содержащую или состоящую из, по меньшей мере, одной РНК или ДНК и, по меньшей мере, одной дополнительной последовательности нуклеиновой кислоты ДНК, и (iv), при необходимости, содержит линкерную область между, по меньшей мере, одной направляющей последовательностью нуклеиновой кислоты и, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, в предпочтительном варианте осуществления настоящего изобретения, при этом, репарационная матричная последовательность нуклеиновой кислоты матрицы ассоциирована с направляющей последовательностью нуклеиновой кислоты на 3'-конце направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с 5'-концом направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты расположена в пределах направляющей последовательности нуклеиновой кислоты.

По меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты в соответствии с различными аспектами и вариантами осуществления настоящего изобретения, содержит нуклеотидную последовательность, выбранную из нуклеотидной последовательности природного или неприродного происхождения, включая синтетическую нуклеотидную последовательность, при необходимости, содержащую остов и/или базовые модификации, при этом, направляющая последовательность нуклеиновой кислоты содержит одноцепочечную или частично одноцепочечную нуклеотидную последовательность РНК или ДНК, и, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит одноцепочечную или двухцепочечную нуклеотидную последовательность ДНК.

В определенных вариантах осуществления настоящего изобретения, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты (RT) искусственного молекулярного комплекса по настоящему изобретению содержит, по меньшей мере, одну концевую часть, в предпочтительном варианте осуществления настоящего изобретения, 3'-концевую, при этом, эта концевая часть не взаимодействует ни с каким другим компонентом искусственного молекулярного комплекса и, таким образом, выполнена с возможностью гибридизации с, по меньшей мере, одной геномной комплементарной последовательностью, чтобы опосредовать репарацию последовательности ДНК-мишени, и/или, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты предложена в виде плазмиды. Таким образом, чтобы иметь возможность доступа к, по меньшей мере, одной геномной комплементарной последовательности, RT должна быть предложена в конфигурации, обеспечивающей оптимальное спаривание оснований с, по меньшей мере, одной геномной комплементарной последовательностью. Эта конфигурация будет варьироваться в зависимости от природы RT и в зависимости от способа предложения RT. В некоторых вариантах осуществления настоящего изобретения, используется, по меньшей мере, одна RT, которая может быть ковалентно или нековалентно присоединена к RTDD, например, к gРНК или gДНК.

В определенных вариантах осуществления настоящего изобретения, где используется молекула,

содержащая, по меньшей мере, одно удлинение РНК в виде RTDD, или используется РНК, кодирующая интересующий белок, РНК может быть представлена вместе с защитной или протекторной молекулой, или цепью, и эта протекторная молекула будет отжигать, по меньшей мере, частично, РНК, представляющую собой фактическую эффекторную молекулу искусственного молекулярного комплекса для защиты эффекторной молекулы РНК от деградации в пределах клетки.

Подходящие конфигурации для искусственного молекулярного комплекса по настоящему изобретению показаны на **Фигурах 1-4**. Искусственные молекулярные комплексы, использующие «гибридную последовательность нуклеиновой кислоты» в виде RTDD и RT по настоящему изобретению, показаны на **Фигуре 1 A-D** и **Фигуре 2 A-C**, но этим не ограничиваются. В зависимости от SSN, матрица для репарации (RT) может находиться в виде ssДНК или dsДНК и, в случае использования белка CRISPR или Аргонавта – используется в виде SSN, может быть присоединена, по меньшей мере, к одной направляющей нуклеиновой кислоте (sgРНК или gРНК, или gДНК) на 3'-конце, 5'-конце ковалентным или нековалентным образом, или она может лежать в пределах gРНК, например, образуя шпилечную вторичную структуру определенного размера и формы. Эта конструкция позволяет обеим, и gРНК, и части RT, выполнять свои функции, не нарушая взаимодействия, по меньшей мере, одной интересующей gРНК с интересующей нуклеазой CRISPR или Аргонавтом, и одновременно располагать RT в непосредственной близости от сайта разрыва ДНК, индуцированного, по меньшей мере, одной парой CRISPR/gРНК Аргонавта/gДНК.

В определенных вариантах осуществления настоящего изобретения, где используются нуклеазы CRISPR, искусственный молекулярный комплекс будет содержать гибридную последовательность нуклеиновой кислоты, содержащую или состоящую из, по меньшей мере, одной РНК и, по меньшей мере, одной последовательности нуклеиновой кислоты ДНК, или просто гибридной последовательности нуклеиновой кислоты РНК/ДНК по настоящему изобретению, таким образом, он представляет собой молекулу, содержащую химерную РНК и ДНК, которая содержит две функциональные возможности. Во-первых, он содержит фрагмент молекулы направляющей нуклеиновой кислоты (gРНК), содержащей рибонуклеиновую кислоту. Эта gРНК содержит две части нуклеотидной последовательности, причем, одна нуклеотидная последовательность необходима для взаимодействия с интересующем полипептидом CRISPR, также как и другая нуклеотидная последовательность, содержащая домен-мишень, при этом, домен-мишень способен гибридизоваться путем спаривания оснований с интересующей комплементарной последовательностью ДНК-мишени, примыкающей к последовательности PAM в противоположной цепи, эта комплементарная последовательность ДНК-мишени, таким образом, представляет собой первую последовательность ДНК-мишени по настоящему изобретению. Во-вторых, гибридная последовательность нуклеиновой кислоты РНК/ДНК содержит фрагмент молекулы репарационной матричной последовательности нуклеиновой кислоты, который может содержать желаемое редактирование, подлежащее введению в интересующую последовательность ДНК-мишени. Кроме того, репарационная матричная последовательность нуклеиновой кислоты может содержать дополнительную гомологичную последовательность непосредственно выше и ниже последовательности ДНК-мишени, то есть, левое и правое плечи гомологии. Длина и позиция связывания каждого плеча гомологии зависит от размера вводимого изменения и может быть отрегулирована для оптимальной эффективности. Например, вполне вероятно, что матрица для репарации с комплементарностью, специфичной для расщепленной цепи ДНК, сначала высвобождаемой Cas9 (как описано в работе Ричардсона и др., Nature Biotechnology. 2016, идентификатор цифрового объекта:10.1038/nbt.3481), может продуцировать самую эффективную

репарацию. Матрица для репарации может представлять собой нуклеотидную последовательность одноцепочечной или двухцепочечной ДНК в зависимости от конкретного применения.

Матрица для репарации может содержать полиморфизмы относительно геномной ДНК для нарушения связывания нуклеазой, в противном случае матрица для репарации становится подходящей мишенью для расщепления полипептида CRISPR. Например, PAM может быть мутирован таким образом, что он больше не будет присутствовать, но на кодирующую область гена это не повлияет, что соответствует молчащей мутации, не изменяющей закодированную аминокислотную последовательность. В другом варианте осуществления настоящего изобретения, где полипептид CRISPR с дефицитом нуклеазы используется в искусственном молекулярном комплексе в качестве SSN, возможно присутствие последовательности PAM в пределах репарационной матричной последовательности. В одном варианте осуществления настоящего изобретения, последовательность RTDD/RT содержит, по меньшей мере, одну последовательность направляющей нуклеиновой кислоты и, по меньшей мере, одну репарационную матричную последовательность нуклеиновой кислоты, но гибрид RTDD/RT также может содержать дополнительные присоединенные к нему фрагменты молекулы, подходящие для редактирования генома, как более подробно описано ниже. В другом варианте осуществления настоящего изобретения, гибридная последовательность RTDD/RT состоит из, по меньшей мере, одной направляющей последовательности нуклеиновой кислоты и, по меньшей мере, одной репарационной матричной последовательности нуклеиновой кислоты.

Было установлено, что оптимальный размер RT может существовать в зависимости от используемой SSN, которая предлагает баланс эффективности нуклеазы благодаря размеру плеча гомологии для эффективности HR-опосредованной репарации DSB.

В одном варианте осуществления настоящего изобретения, направляющая последовательность нуклеиновой кислоты или gРНК предложена в виде одной последовательности нуклеиновой кислоты РНК, объединяющей элемент tracrРНК и элемент crРНК. В другом варианте осуществления настоящего изобретения, например, при работе с системой CRISPR Типа V с использованием полипептида Cpf1 или варианта, или его каталитически активного фрагмента, эта gРНК содержит элемент crРНК. В еще одном варианте осуществления настоящего изобретения, gРНК может быть предложена в виде более одной последовательности нуклеиновой кислоты РНК, имитирующей естественную ситуацию во многих системах CRISPR, при которой crРНК и tracrРНК, если они обе необходимы, предложены на двух отдельных молекулах РНК. Таким образом, в определенных вариантах осуществления настоящего изобретения, такое расположение позволяет иметь два элемента (tracrРНК и crРНК) в отдельных цепях РНК, как в природе. В одном варианте осуществления настоящего изобретения, предложена отдельная молекула нуклеиновой кислоты РНК, предлагающая crРНК, и предложена отдельная молекула нуклеиновой кислоты РНК, то есть, представлено более одного RTDD. Либо фрагмент молекулы crРНК, либо фрагмент молекулы tracrРНК могут быть ассоциированы с репарационной матричной (RT) последовательностью нуклеиновой кислоты. Например, предложение гибрида tracrРНК:RT или crРНК:RT может быть предпочтительным, когда химический синтез *ex vivo* tracrРНК:RT или crРНК:RT выбран из-за более короткой длины соответствующей молекулы, по сравнению с гибридом gРНК:RT, при этом, gРНК состоит из одной молекулы РНК, объединяющей функцию crРНК и tracrРНК.

Таким образом, последовательность RTDD/RT по настоящему изобретению подходит для точного редактирования генома в любых интересующих типах клеток, включая прокариотические клетки и эукариотические клетки, включая грибковые клетки, животные и растительные клетки, а также любого

интересующего генома в условиях *in vitro* и представляет собой подходящий физически соединенный инструмент, обеспечивающий одновременную пространственно-временную доступность матрицы для репарации и SSN во время редактирования генома.

В соответствии со всеми аспектами и вариантами осуществления настоящего изобретения, по меньшей мере, один RTDD и, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты ассоциированы друг с другом. Термин «ассоциированный с» или «в ассоциации», согласно настоящему изобретению, должен толковаться в широком смысле и, следовательно, согласно настоящему изобретению он подразумевает, что RTDD, например, молекула рРНК или биотина, FAM или дигоксигенин, предложен в физической ассоциации с матрицей для репарации ДНК, причем, ассоциация имеет либо ковалентную, либо нековалентную природу, что по существу увеличивает доступность матрицы для репарации для гомологичной рекомбинации. Вместо беспорядочной амплификации матрицы для репарации или предложения избыточного количества матрицы для репарации, которая, однако, не сцеплена с RTDD, репарационная матричная нуклеотидная последовательность, таким образом, представлена в DSB вместе с SSN искусственного молекулярного комплекса в интересующей последовательности ДНК-мишени, что, в свою очередь, значительно улучшает прогнозируемость и специфичность подхода к редактированию генома.

В другом варианте осуществления настоящего изобретения, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты присоединена, по меньшей мере, к одной последовательности RTDD посредством ковалентных и/или нековалентных связей, или присоединений. В соответствии с этим вариантом осуществления настоящего изобретения, гибридный комплекс RTDD и RT может быть предложен в виде синтезированной *in vitro* молекулы, которая затем может быть ассоциирована с, по меньшей мере, одной представляющей интерес SSN, либо *in vitro*, либо *in vivo* в интересующей клетке-мишени, либо в рамках интересующего анализа *in vitro*. В предпочтительном варианте осуществления настоящего изобретения, клетка представляет собой эукариотическую клетку, включая грибковую, животную или растительную клетку. Клетка также может представлять собой прокариотическую клетку. Кроме того, клетка может представлять собой прокариотическую или эукариотическую клетку-хозяина, несущую либо на плазмиде, либо интегрированную в геном, гетерологичную последовательность-мишень другого организма или вируса. В этом варианте осуществления настоящего изобретения, клетка функционирует в качестве хозяина для осуществления геномной инженерии на гетерологичной последовательности, предложенной в указанной клетке-хозяине.

В одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, по меньшей мере, одна репарационная матричная (RT) последовательность нуклеиновой кислоты ковалентно присоединена, по меньшей мере, к одному RTDD. Ковалентное присоединение или ковалентная связь представляет собой химическую связь, которая включает совместное использование электронных пар между атомами молекул или последовательностей, ковалентно присоединенных друг к другу.

В другом варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты нековалентно присоединена к, по меньшей мере, одной последовательности RTDD. Нековалентное взаимодействие отличается от ковалентной связи тем, что оно не включает обмен электронами, а скорее включает более дисперсные вариации электромагнитных взаимодействий между молекулами/последовательностями или в пределах молекулы/последовательности. Таким образом,

нековалентные взаимодействия или присоединения включают электростатические взаимодействия, Ван-дер-Ваальсовы силы, ТТ-эффекты и гидрофобные эффекты. Особое значение в контексте молекул нуклеиновой кислоты имеют водородные связи в виде электростатического взаимодействия. Водородная связь (Н-связь) представляет собой особый тип диполь-дипольного взаимодействия, которое включает взаимодействие между частично положительным атомом водорода и высоко электроотрицательным, частично отрицательным атомом кислорода, азота, серы или фтора, нековалентно связанным с указанным атомом водорода.

Используемый в настоящем документе термин «гибридизация» относится к спариванию комплементарных нуклеиновых кислот, то есть, ДНК и/или РНК, с использованием любого процесса, с помощью которого цепь нуклеиновой кислоты соединяется с комплементарной цепью посредством спаривания оснований с образованием гибридизированного комплекса. На гибридизацию и силу гибридизации (то есть, на силу ассоциации между нуклеиновыми кислотами) влияют такие факторы, как степень и длина комплементарности между нуклеиновыми кислотами, жесткость условий, вовлеченных в процесс, Tm образованного гибрида и соотношение G:C в пределах нуклеиновых кислотах. Термин «гибридизированный комплекс» относится к комплексу, образованному между двумя последовательностями нуклеиновых кислот в результате образования водородных связей между комплементарными основаниями G и C, и между комплементарными основаниями A и T/U. Гибридизованный комплекс или соответствующая гибридная конструкция может образовываться между двумя молекулами нуклеиновой кислоты ДНК, между двумя молекулами нуклеиновой кислоты РНК или между молекулой нуклеиновой кислоты ДНК и молекулой нуклеиновой кислоты РНК. Для всех констелляций, молекулы нуклеиновой кислоты могут представлять собой молекулы нуклеиновой кислоты природного происхождения, сгенерированные *in vitro* или *in vivo*, и/или молекулы искусственной или синтетической нуклеиновой кислоты. Гибридизация, как подробно описано выше, например, пары оснований Уотсона-Крика, которые могут образовываться между последовательностями ДНК, РНК и ДНК/РНК, обусловлена шаблоном специфичных водородных связей, которая, таким образом, представляет собой нековалентную форму присоединения по настоящему изобретению.

Что касается нековалентных ассоциаций по настоящему изобретению, то, по меньшей мере, один RTDD искусственного молекулярного комплекса по настоящему изобретению и, по меньшей мере, одна репарационная матричная последовательность по настоящему изобретению могут ассоциироваться друг с другом путем спаривания оснований РНК-ДНК.

Другой формой нековалентного взаимодействия является ассоциация, по меньшей мере, одной репарационной матричной последовательности с, по меньшей мере, одним компонентом, либо RTDD, либо RTDD, содержащимся в SSN посредством электрических зарядов.

Что касается ковалентной ассоциации или присоединения, то, по меньшей мере, один RTDD и, по меньшей мере, одна репарационная матричная последовательность соединены в виде непрерывной молекулы, продуцированной либо *in vivo*, либо *in vitro*. Ковалентное и нековалентное присоединение также может быть комбинированным, например, путем предложения ковалентно присоединенной RTDD/репарационной матричной последовательности, которая также может содержать дополнительную репарационную матричную последовательность нуклеиновой кислоты, нековалентно присоединенную к ковалентно присоединенной RTDD/репарационной матричной последовательности. Этот подход особенно подходит в случае, если ковалентно присоединенная RTDD/репарационная матричная последовательность, по меньшей мере, частично продуцирована *in vivo*, и нужно добавить другую матрицу для репарации,

продуцированную либо *in vivo*, либо *in vitro*, к уже существующему RTDD/репарационному матричному комплексу.

Как также явствует из работы Нишимасу и др., см. выше, gРНК выполнена с возможностью взаимодействия с полипептидом CRISPR или с вариантом, или его каталитически активным фрагментом, как было раскрыто в настоящем изобретении, если gРНК содержит, по меньшей мере, одну часть, обычно содержащую гетеродуплексную конфигурацию, которая распознается полипептидом CRISPR либо зависимым от последовательности образом, то есть, посредством взаимодействия с основаниями РНК, содержащими A, U, G и C, либо независимым от последовательности образом, то есть, посредством взаимодействия остова фосфата нуклеотидной последовательности gРНК с полипептидом CRISPR.

В соответствии с определенными вариантами осуществления первого аспекта, а также других аспектов настоящего изобретения, последовательность ДНК-мишени расположена в геноме клетки, в предпочтительном варианте осуществления настоящего изобретения, прокариотической или эукариотической клетки, в более в предпочтительном варианте осуществления настоящего изобретения, грибковой, животной или растительной клетки, при этом, геном содержит ядерный геном, а также другие части генома, включая геном пластид.

«Последовательность ДНК-мишени» определяет область генома, где должно произойти целенаправленное редактирование генома. Ввиду того, что RTDD и репарационная матричная последовательность нуклеиновой кислоты по существу имеют разные функциональные возможности, может существовать более одной области ДНК-мишени, которая может быть различной для различных компонентов искусственного молекулярного комплекса по настоящему изобретению. Таким образом, одна последовательность ДНК-мишени может определять область интересующей области ДНК-мишени, причем, часть последовательности RTDD, при этом RTDD представляет собой gРНК, является комплементарной, при этом другая последовательность ДНК-мишени определяет область интересующей области ДНК-мишени с которой будет связываться SSN и/или взаимодействующий домен. По меньшей мере, одна часть репарационной матричной последовательности нуклеиновой кислоты является комплементарной, и она определяется как геномная комплементарная последовательность, причем, указанная последовательность также представляет собой дополнительную последовательность ДНК-мишени. Области ДНК-мишени могут быть одинаковыми или, в предпочтительном варианте осуществления настоящего изобретения, разными, но, в то же время, возможно, перекрывающимися областями в интересующей последовательности ДНК-мишени.

Пространственное отношение между сайтом-мишенью RTDD и/или SSN, и/или взаимодействующим доменом и сайтом гомологии для репарационной матричной последовательности (RT) нуклеиновой кислоты может быть переменным. Два сайта могут быть идентичными, могут полностью или частично перекрываться, или они могут быть разделены любым количеством нуклеотидов в интересующем геноме. RT может иметь гомологию с обеими цепями геномной ДНК, при необходимости, представленными в виде двухцепочечной конструкции, например плазмиды, или в виде любой цепи отдельно, независимо от того, какая цепь является мишенью. Эффективная матрица для репарации может быть выполнена с возможностью обладания комплементарности, специфичной для расщепленной цепи ДНК, сначала высвобождаемой SSN, например, Cas9 (как описано в работе Ричардсона и др., Nature Biotechnology. 2016, идентификатор цифрового объекта:10.1038/nbt.3481).

Прогнозируется, что взаимодействие между RTDD и RT и, таким образом, непосредственная близость SSN и RT в соответствии с искусственным молекулярным комплексом настоящего изобретения

преодолеет в целом низкую эффективность гомологически направленной репарационной (HDR)/гомологичной рекомбинации (HR), поскольку это гарантирует физическую доступность репарационной матричной нуклеотидной последовательности, присутствующей стехиометрическим образом по отношению, по меньшей мере, к одной SSN *in situ* в месте, где целенаправленный разрыв геномной цепи вводится, по меньшей мере, одним полипептидом SSN в последовательность ДНК-мишени.

Термин «репарационная матричная (RT) последовательность нуклеиновой кислоты» как часть искусственного молекулярного комплекса по настоящему изобретению, таким образом, подразумевает нуклеотидную последовательность, которая может представлять собой последовательность одноцепочечной или двухцепочечной ДНК, которая способна предложить матрицу для модификации и/или репарации разрыва ДНК.

В одном варианте осуществления настоящего изобретения, искусственный молекулярный комплекс представляет собой предварительно собранный комплекс *in vitro*, при этом SSN, RTDD и RT, и, при необходимости, компонент или часть взаимодействующего домена предложены либо ковалентно присоединенными друг к другу, либо нековалентно ассоциированными. В одном варианте осуществления настоящего изобретения, последовательность RTDD/RT предварительно собирается, и SSN и, при необходимости, взаимодействующий домен отдельно доставляется в клетку-мишень либо в виде транскрибируемой ДНК, либо транслируемой РНК-конструкции, либо непосредственно в виде аминокислотной последовательности и последовательности RTDD/RT, и SSN, и, при необходимости, взаимодействующий домен образуют комплекс в пределах клетки-мишени. В другом варианте осуществления настоящего изобретения, последовательность RTDD/PT, а также SSN и, при необходимости, взаимодействующий домен собираются *in vitro,* и нуклеопротеиновый комплекс, при необходимости, содержащий дополнительные молекулы, например, биотин или FAM, или дигоксигенин, затем вводится в интересующую клетку-мишень или в систему *in vitro*, содержащую, по меньшей мере, одну интересующую нуклеотидную последовательность ДНК-мишени, подлежащую модификации.

Введение функционального предварительно собранного искусственного молекулярного комплекса в клетку-мишень приводит к целенаправленному двухцепочечному разрыву и одновременной репарации, и сайт-специфичной модификации, благодаря тому, что активность, по меньшей мере, одной сайт-специфичной нуклеазы (SSN) немедленно сопровождается последующей гомологичной рекомбинацией в сайте последовательности ДНК-мишени по настоящему изобретению посредством репарационной матричной последовательности нуклеиновой кислоты ДНК, сцепленной с RTDD, при этом, RTDD непосредственно взаимодействует с SSN. Поэтому недостатки, в связи со слабой доступностью RT или событиями неспецифичного NHEJ (см. выше – Предпосылки создания изобретения), препятствующие событию высокоспецифичного и контролируемого редактирования генома, могут быть одновременно уменьшены, поскольку искусственный молекулярный комплекс может достичь сайта-мишени скоординированным образом с адекватным стехиометрическим составом матрицы для репарации и нуклеазы. Еще одним преимуществом является то, что потенциал для нецелевой интеграции матрицы для репарации снижается из-за ее физической ассоциации с белком, а также с RTDD комплекса, при этом, SSN и/или RTDD не могут быть интегрированы в геном *per se*.

Термин «целенаправленная репарация, направляемая гомологией» по настоящему изобретению содержит любой тип изменений, которые могут быть введены посредством репарационной матричной последовательности в соответствии с настоящей заявкой, которая может независимо содержать вставки последовательности, редактирования, по меньшей мере, одного положения последовательности, делеции

или перестройки, причем, предпочтительной стратегией для подходов к редактированию генома у высших эукариот в настоящее время являются вставки, делеции или редактирования, поскольку эти стратегии позволяют осуществлять целенаправленный нокин или нокаут интересующей последовательности в пределах последовательности ДНК-мишени, или сайт-специфичную модификацию, по меньшей мере, одной последовательности.

Пример целенаправленной репарации, направляемой гомологией, опосредованной искусственным молекулярным комплексом с использованием нуклеазы CRISPR в качестве SSN, образованным *ex vivo* или *in vivo* при взаимодействии с гибридной последовательностью нуклеиновой кислоты по настоящему изобретению, можно установить на **Фигуре 3 A-E**, иллюстрирующей хронологическую последовательность распознавания, связывания, расщепления ДНК и последующую репарацию для примерного комплекса (RTDD)/репарационной матричной (RT) SSN/направляющей нуклеиновой кислоты и для данной эндогенной последовательности ДНК-мишени.

В одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, одна последовательность RTDD содержит нуклеотидную последовательность, выбранную из нуклеотидной последовательности природного или неприродного происхождения, включая синтетическую нуклеотидную последовательность, при необходимости, содержащую остов и/или базовые модификации, при этом, направляющая последовательность нуклеиновой кислоты содержит одноцепочечную или частично одноцепочечную нуклеотидную последовательность РНК, и, при этом, последовательность репарационной матричной нуклеиновой кислоты содержит одноцепочечную или двухцепочечную нуклеотидную последовательность ДНК.

Проблема у любого подхода к редактированию генома CRISPR заключается в том, что gРНК и функциональный полипептид CRISPR в виде SSN должны транспортироваться в ядро или любой другой компартмент, содержащий геномную ДНК, то есть, последовательность ДНК-мишени, функциональным (не деградированным) образом. Поскольку РНК менее стабильна, чем полипептид или двухцепочечная ДНК, и имеет более высокий функциональный цикл, особенно в связи с тем, что она может быть легко деградирована посредством нуклеаз, в некоторых вариантах осуществления настоящего изобретения, в соответствии с первым аспектом настоящего изобретения gРНК в виде RTDD и/или репарационная матричная последовательность нуклеиновой кислоты ДНК содержит, по меньшей мере, один нуклеотид неприродного происхождения. Предпочтительные модификации остова по настоящему изобретению, повышающие стабильность gРНК и/или репарационную матричную последовательность нуклеиновой кислоты ДНК, выбирают из группы, состоящей из модификации фосфоротиоата, модификации метилфосфоната, модификации закрытой нуклеиновой кислоты, модификации 2'O-(2-метоксиэтил), модификации дифосфоротиоата и модификации пептидной нуклеиновой кислоты. Примечательно, что все указанные модификации остова по-прежнему позволяют образовывать комплементарное спаривание оснований между двумя цепями нуклеиновой кислоты, но являются более устойчивыми к расщеплению эндогенными нуклеазами. В зависимости от нуклеазы, используемой по настоящему изобретению, может потребоваться не модифицировать те нуклеотидные положения gРНК, которые участвуют в независимом от последовательности взаимодействии с полипептидом CRISPR. Указанная информация может быть получена из имеющейся структурной информации, доступной для комплексов нуклеазы CRISPR/gРНК.

В определенных вариантах осуществления настоящего изобретения, в соответствии с первым аспектом настоящего изобретения, предусматривается, что RTDD и/или репарационная матричная (RT)

последовательность нуклеиновой кислоты ДНК, и/или взаимодействующий домен содержат/содержит модификацию нуклеотида и/или основания, в предпочтительном варианте осуществления настоящего изобретения, в выбранных, а не во всех положениях нуклеотидной последовательности. Эти модификации выбраны из группы, состоящей из добавления акридина, амина, биотина, cascade blue, холестерина, Cy3, Cy5, Cy5.5, Daboyl, дигоксигенина, динитрофенила, Edans, 6-FAM, флуоресцеина, 3'-глицерила, HEX, IRD-700, IRD-800, JOE, фосфата псоралена, родамина, ROX, тиола (SH), спейсеров, TAMRA, TET, AMCA-S", SE, BODIPY®, Marina Blue®, Pacific Blue®, Oregon Green®, Rhodamine Green®, Rhodamine Red®, Rhodol Green® and Texas Red®. В предпочтительном варианте осуществления настоящего изобретения, указанные добавки включены на 3'-конце или 5'-конце последовательности нуклеиновой кислоты, используемой в качестве RT и/или RTDD, и/или взаимодействующего домена, как часть искусственного молекулярного комплекса по настоящему изобретению. Эта модификация имеет положительные эффекты, заключающиеся в том, что клеточная локализация RTDD и/или взаимодействующего домена, и/или репарационной матричной последовательности нуклеиновой кислоты ДНК в клетке может быть визуализирована для изучения распределения, концентрации и/или доступности соответствующей последовательности. Кроме того, представляется возможным изучить взаимодействие искусственного молекулярного комплекса по настоящему изобретению с эндогенными молекулами. Способы изучения таких взаимодействий или визуализации нуклеотидной последовательности, модифицированной или меченой, как подробно описано выше, доступны специалисту в соответствующей области техники.

В любом варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, по меньшей мере, одна сайт-специфичная нуклеаза и/или, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты, и/или, по меньшей мере, один взаимодействующий домен, и/или, по меньшей мере, один RTDD содержит, по меньшей мере, одну последовательность ядерной локализации (NLS), последовательность пластидной локализации (PLS), в предпочтительном варианте осуществления настоящего изобретения, последовательность митохондриальной локализации или последовательность хлоропластной локализации. Следовательно, по меньшей мере, один из компонентов искусственного молекулярного комплекса содержит последовательность для нацеливания комплекса на ядерный геном. В определенных вариантах осуществления настоящего изобретения, RTDD также может нести, по меньшей мере, одну последовательность локализации. В предпочтительном варианте осуществления настоящего изобретения, SSN и/или взаимодействующий домен искусственного молекулярного комплекса будет содержать, по меньшей мере, одну NLS или, по меньшей мере, одну PLS, или он будет содержать как, по меньшей мере, одну NLS так и, по меньшей мере, одну последовательность PLS. Эта, по меньшей мере, одна последовательность NLS или PLS будет транспортировать весь искусственный молекулярный комплекс в ядро. NLS-меченые или PLS-меченые белки могут быть получены в виде NLS-меченых или PLS-меченых молекул слияния.

Для вариантов осуществления настоящего изобретения, в которых искусственный молекулярный комплекс по настоящему изобретению используется для целей *in vitro*, например, для модификации генома или части генома, на плазмиде или любом другом векторе *in vitro*, последовательность локализации может не потребоваться. Последовательности локализации помогают нацеливать искусственный молекулярный комплекс, по меньшей мере, на одну интересующую последовательность ДНК-мишени в соответствующем компартменте в интересующей клетке-мишени. В соответствии с определенными вариантами осуществления настоящего изобретения, последовательность локализации может содержать

последовательность ядерной локализации, последовательность пластидной локализации, в предпочтительном варианте осуществления настоящего изобретения, последовательность митохондриальной локализации или последовательность хлоропластной локализации. Следовательно, по меньшей мере, одна SSN и/или, по меньшей мере, один RTDD, и/или, по меньшей мере, один взаимодействующий домен будут содержать, по меньшей мере, одну соответствующую последовательность локализации, в предпочтительном варианте осуществления настоящего изобретения, последовательность ядерной локализации (NLS), для направления комплекса в ядерный геном клетки. В некоторых вариантах осуществления настоящего изобретения, SSN и/или RT, и/или, по меньшей мере, один взаимодействующий домен, и/или RTDD могут содержать, примерно, или более, примерно, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, или более NLS на или рядом с N-терминальным концом (для пептидов и белков), примерно, или более, примерно, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, или более NLS на или рядом с C-терминальным концом (для пептидов и белков), или их комбинацию (например, по меньшей мере, одну NLS на N-терминальном конце и, по меньшей мере, одну NLS на C-терминальном конце для пептидов и белков). Неаминокислотные компоненты искусственного молекулярного комплекса будут нести последовательность локализации, например, на 5'-конце и/или 3'-конце, как это имеет место для последовательностей нуклеиновых кислот. Кроме того, последовательность локализации, в предпочтительном варианте осуществления настоящего изобретения, синтетическая последовательность локализации, также может быть добавлена в любом положении в пределах молекулы при условии, что она не будет нарушать взаимодействий в пределах молекулярного комплекса и/или связывания, расщепления и репарации искусственных молекулярных комплексов настоящего изобретения. Когда присутствует более одной NLS, то каждая может быть выбрана независимо от других, таким образом, одна NLS может присутствовать, по меньшей мере, в одной копии и/или в комбинации с, по меньшей мере, одной другой NLS, присутствующей, по меньшей мере, в одной копии.

В предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и/или взаимодействующий домен будет содержать последовательность локализации и может содержать не более 6 NLS. В некоторых вариантах осуществления настоящего изобретения, рассматривается расположение NLS рядом с N-терминальным концом или C-терминальным концом аминокислотного компонента искусственного молекулярного комплекса, когда ближайшая аминокислота NLS находится в пределах, примерно, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 или более аминокислот вдоль полипептидной цепи от N-терминального конца или C-терминального конца. Неограничивающие примеры NLS включают последовательность NLS, полученную из: NLS большого T-антигена вируса SV40, которая имеет аминокислотную последовательность PKKKRKV (SEQ ID NO: 1); NLS из нуклеоплазмина (например, двойная NLS нуклеоплазмина с последовательностью KRPAATKKAGQAKKKK (SEQ ID NO: 2)); c-myc NLS, имеющей аминокислотную последовательность PAAKRVKLD (SEQ ID NO: 3) или RQRRNELKRSP (SEQ ID NO: 4); hRNPAI M9 NLS, имеющей последовательность NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 5); последовательности RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 6) домена 188 из импортин-альфа; последовательностей VSRKRPRP (SEQ ID NO: 7) и PPKKARED (SEQ ID NO: 8) белка миомы T; последовательности PXPKKKPL (SEQ ID NO: 9) человеческого p53, при этом, «L» в позиции 8 SEQ ID NO: 9 не является обязательной; последовательности SALIKKKKKMAP (SEQ ID NO: 10) c-abl IV мыши; последовательностей DRLRR (SEQ ID NO: 11) и PKQKKRK (SEQ ID NO: 12) вируса гриппа NS1; последовательности RKLKKKIKKL антигена вируса гепатита дельта (SEQ ID NO: 13); последовательности REKKKFLKRR (SEQ ID NO: 14) белка Mx1 мыши; последовательности

KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 15) человеческой поли(АДФ-рибоза)полимеразы; и последовательности RKCLQAGMNLEARKTKK (SEQ ID NO: 16) рецепторов стероидных гормонов (человеческих) – глюкокортикоидов. В некоторых вариантах осуществления настоящего изобретения, сигнал локализации может представлять собой сигнал пластидной локализации, например сигнал пластидной или митохондриальной локализации. Подходящие сигналы пластидной локализации выбирают из группы, состоящей из транзитных пептидов хлоропласта или нацеленных на митохондрии пептидов. Кроме того, пептиды, полученные из Tat-белка ВИЧ, или из последовательностей, кодирующих последовательности, могут быть подходящими для нацеливания интересующей конструкции или молекулы в интересующую клетку и/или субклеточный компартмент. Подходящие Tat-пептиды получают из YGRKKRRQRRR (SEQ ID NO: 17) или содержат мотив GRKKR (SEQ ID NO: 18). В другом примерном варианте осуществления настоящего изобретения, последовательность, полученная из Cox4p митохондрии дрожжей (SEQ ID NO: 30) или последовательность, полученная из лидерной последовательности митохондриальной малатдегидрогеназы человека (MLS) (SEQ ID NO: 31) или полученная из синтазы липоевой кислоты Arabidopsis (Стандартная/эталонная последовательность Национального центра биотехнологической информации (NCBI) ID: NP 179682.1, обозначенная в настоящем документе как SEQ ID NO: 32), может использоваться для локализации искусственного молекулярного комплекса по настоящему изобретению в митохондриальную матрицу для модификации митохондриальной ДНК.

В конкретных вариантах осуществления настоящего изобретения, может представлять интерес нацеливание искусственного молекулярного комплекса на хлоропласт. Во многих случаях, это нацеливание может быть достигнуто посредством присутствия удлинения N-терминального конца, называемого транзитным пептидом хлоропласта (ТПХ) или транзитным пептидом пластиды. Хромосомные трансгены из бактериальных источников должны иметь последовательность, кодирующую последовательность ТПХ, слитую с последовательностью, кодирующей экспрессированный полипептид, если экспрессированный полипептид должен быть компартментализирован в растительной пластиде (например, в хлоропласте). Соответственно, локализация экзогенного полипептида в хлоропласт зачастую осуществляется посредством оперативного сцепления полинуклеотидной последовательности, кодирующей последовательность ТПХ, с 5'-областью полинуклеотида, кодирующего экзогенный полипептид, то есть, по меньшей мере, с одной SSN по настоящему изобретению. ТПХ удаляется на этапе обработки во время транслокации в пластиду. Однако на эффективность обработки может влиять аминокислотная последовательность ТПХ и близлежащие последовательности на NH2-терминальном конце пептида. Другими вариантами нацеливания на хлоропласт, которые были описаны, являются сигнальная последовательность cab-m7 кукурузы (Патент США 7,022,896, WO 97/41228), сигнальная последовательность глутатионредуктазы гороха (WO 97/41228) и ТПХ, описанный в US 2009/029861 A1.

Различные локализационные последовательности по настоящему изобретению могут кодироваться на плазмиде или экспрессионной кассете, кодирующей, по меньшей мере, одну локализационную последовательность для оперативного сцепления локализационной последовательности с соответствующей молекулой, или локализационные последовательности могут быть присоединены к белку, нуклеиновой кислоте или другой биомолекуле с образованием искусственных молекулярных комплексов по настоящему изобретению синтетическим путем.

В еще одном дополнительном варианте осуществления настоящего изобретения, по меньшей мере, один из сигналов ядерного экспорта может использоваться в дополнение или вместо, по меньшей мере, одной локализационной последовательности.

В вариантах осуществления настоящего изобретения, где искусственный молекулярный комплекс доставляется в клетку с помощью, по меньшей мере, одного вектора доставки в виде последовательности нуклеиновой кислоты, сигнал локализации может быть ковалентно присоединен к, по меньшей мере, одной SSN и/или взаимодействующему домену, кодирующему последовательность ковалентным образом в виде последовательности нуклеиновой кислоты, кодирующей сигнал локализации.

В одном варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и/или полипептидный взаимодействующий домен может быть ковалентно или нековалентно ассоциирован с флуоресцентным репортерным геном или белком. Этот репортер может быть доставлен в виде ДНК, в виде mPHK, в виде независимого белка или в виде белка слияния, сцепленного, по меньшей мере, с одной SSN и/или полипептидным взаимодействующим доменом.

Молекула RTDD/RT по настоящему изобретению может быть получена несколькими способами. Это можно сделать посредством химического синтеза, добавляя основания РНК, где это необходимо в процессе синтеза, и основания ДНК, где это необходимо в процессе синтеза. В альтернативном варианте осуществления настоящего изобретения, RTDD и/или RT могут быть синтезированы независимо друг от друга, и затем молекулы могут быть ассоциированы друг с другом, как описано выше. Другим вариантом является использование T4 РНК лигазы или другого фермента, способного лигировать нуклеиновые кислоты в РНК, в предпочтительном варианте осуществления настоящего изобретения, в одноцепочечную РНК. В данном случае, компоненты РНК и ДНК генерируют независимо любым способом, смешивают, подвергают воздействию фермента в соответствии с протоколом от изготовителя, и они будут ковалентно сцеплены посредством лигирования, то есть, получат ковалентное присоединение. Другие стратегии ковалентного связывания RTDD с RT включают сцепление каждого из них с другими группами сцепляющих химических веществ или комплексами, например, с пептидом. Этот тип подхода особенно уместен, когда гибридная последовательность RTDD/RT должна быть определена позже в пределах клетки, или когда необходимо наделить гибридную последовательность нуклеиновой кислоты дополнительной функцией. Химическая модификация последовательности нуклеиновой кислоты RTDD и/или RT может иметь большое значение для стабилизации последовательности RTDD/RT и предотвращения деградации клеточными ферментами для достижения одновременной высокой доступности последовательности RTDD/RT и, по меньшей мере, одной сайт-специфичной нуклеазы в интересующем сайте ДНК-мишени.

В вариантах осуществления настоящего изобретения, где RTDD представляет собой gPHK, а SSN представляет собой нуклеазу CRISPR, в предпочтительном варианте осуществления настоящего изобретения, нуклеазу Cas или Cpf1, может присутствовать более одного RTDD. Было обнаружено, что одновременное использование нескольких gPHK будет увеличивать активацию или репрессию генов на основе CRISPR и может значительно снизить появление аллелей, устойчивых к генным драйвам. Следовательно, gPHK в виде RTDD могут быть представлены в виде одного необработанного транскрипта, и затем gPHK будут вырезаны из предшественника в ядре с помощью транскрипции РНК-полимеразой II, одновременно устраняя экспорт gPHK в цитоплазму (Порт и Буллок, Nat. Methods, 2016, том 13, № 10, 852-854). В этих вариантах осуществления настоящего изобретения, gPHK могут быть представлены в виде плазмид tPHK-gPHK, таким образом, механизм обработки эндогенных tPHK будет высвобождать множество функциональных gPHK.

В определенных вариантах осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, по меньшей мере, одна сайт-специфичная нуклеаза или последовательность, кодирующая последовательность, и, по меньшей мере, один взаимодействующий

домен, или последовательность, кодирующая последовательность, и/или, по меньшей мере, один стыковочный домен матрицы для репарации, или последовательность, кодирующая последовательность, соединены, по меньшей мере, одним линкерным доменом. Эта линкерная последовательность может служить в качестве молекулярного спейсера для достижения оптимальных геометрических параметров последовательности RTDD и репарационной матричной последовательности нуклеиновой кислоты, а также SSN и, при необходимости, компонента взаимодействующего домена искусственного молекулярного комплекса, в результате чего все отдельные компоненты могут полностью выполнять свою функцию. Длина и состав областей линкера или привязки могут быть важным аспектом конструкции, например, для некоторых пар RTDD и RT. В одном варианте осуществления настоящего изобретения, в частности 5'-конец левого плеча гомологии RT может содержать линкерную область. Привязка или линкер могут принимать различные формы. На основании настоящего изобретения, специалист в данной области техники может, начиная с левого или правого плеча гомологии RT, позволить этой части RT выступить в качестве привязки или гибкого линкера, чтобы позволить перемещение RT к хромосомной мишени, и в качестве гомологи для опосредования реакции HR, зная обычные параметры конструкции для матриц для репарации, которые в настоящее время широко используются для редактирования генома.

В вариантах осуществления настоящего изобретения, где искусственный молекулярный комплекс содержит, по меньшей мере, одну SSN, а также, по меньшей мере, один взаимодействующий домен (IA), SSN и IA могут быть соединены подходящим линкером.

Конструктивные параметры, которые следует учитывать, включают геометрические параметры гомологии матрицы для репарации относительно сайта разрезания полипептида CRISPR в виде SSN, цепи в интересующем сайте ДНК-мишени, которому гомологична матрица для репарации, размера матрицы для репарации, который может влиять на то, будет ли введен линкер, и какую длину он будет иметь. Линкерная последовательность может быть использована как для ковалентных, так и для нековалентных ассоциаций gPHK и матрицы для репарации. На основании настоящего изобретения и на основании информации, предоставленной в работах Нишимасу (см. выше), Цай и др. (Nature Biotechnology, 32, 569-576, (2014)) или Шечнера и др. (Nature Methods, 12 (7), 664-670 (2015), идентификатор цифрового объекта:10.1038/nmeth.3433), специалист в данной области техники может, таким образом, определить подходящую линкерную область для гибридной последовательности нуклеиновой кислоты, чтобы определить специфичную последовательность между gPHK как RTDD и RT или между различными gPHK и/или RT, в случае использования нескольких гибридных нуклеиновых кислот, таким образом, фрагменты молекулы как gPHK, так и RT могут полностью выполнять свою функцию без каких-либо стерических ограничений. По меньшей мере, одна линкерная область может содержать до 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 или 100 дополнительных нуклеотидов, чтобы надлежащим образом отделить, по меньшей мере, одну gPHK от RT, или чтобы оптимизировать расположение gPHK и/или RT. В определенных вариантах осуществления настоящего изобретения, линкерная последовательность может содержать до 150, 200, 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500 или, по меньшей мере, 4700 или 5000 нуклеотидов для достижения лучшего расположения gPHK и/или RT.

Для нековалентной ассоциации любых содержащих RTDD последовательностей нуклеиновой кислоты с RT, один подход заключается в предложении частично комплементарных последовательностей в RTDD и RT, чтобы две молекулы естественным образом ассоциировались путем спаривания оснований нуклеиновой кислоты.

Возможны и другие способы нековалентной ассоциации, такие как использование электрических

зарядов молекул, чтобы вызвать достаточную ассоциацию RT с каким-либо компонентом искусственного молекулярного редактирующего комплекса. В другом варианте осуществления настоящего изобретения, по меньшей мере, один компонент искусственного молекулярного комплекса может содержать метку и партнера по связыванию, то есть, RTDD и SSN, или RTDD и взаимодействующий домен, и/или его часть RT, соответственно, содержать соответствующий партнер по связыванию метки, таким образом, достигается нековалентное взаимодействие, при необходимости, в дополнение к спариванию оснований между RTDD или взаимодействующим доменом и RT, и ассоциации между RTDD и полипептидом SSN для увеличения взаимодействия и, следовательно, стабильности искусственного молекулярного комплекса.

В клетках человека, Cas9, нагруженный gPHK, обладающей дополнительной последовательностью из 28 пар оснований на 3'-конце плюс ассоциированный c187 аминокислотами (21,4 кД) белок Csy4, сохранял, по меньшей мере, 90% активности в индукции DSB, по сравнению со стандартными регуляциями gPHK (Цай и др., Nature Biotechnology, 32, 2014). Это указывает на довольно существенную толерантность Cas9 как SSN для груза, привязанного к 3'-концу sgPHK и надлежащего потенциала функции структуры для растянутой молекулы sgPHK. Толерантность Cas9 частично обеспечивается гибкостью свободного 3'-конца последовательности нуклеиновой кислоты, которая в стандартной gPHK заканчивается шпилькой, которая удерживается вне архитектуры белка Cas9 и на поверхности, приблизительно перпендикулярной поверхности, удерживающей активный сайт (Нишимасу и др., 2014; Андерс и др., 2014). Кроме того, Шечнер и др. («Мультиплексируемое, локус-специфичное нацеливание длинных РНК с CRISPR-Дисплеем», Nature Methods, 12 (7), 664-670 (2015), идентификатор цифрового объекта:10.1038/nmeth.3433) показывают, что длинные некодирующие молекулы ssPHK могут быть транскрипционно присоединены к 5'-концу или 3'-концу sgPHK или во внутренней петле sgPHK без потери сиквенс-специфичной нацеленной активности белка dCas9 в геноме клетки человека. ssPHK размером до 4,8 т.п.н. были размещены рибонуклеопротеиновым комплексом с сохранением сиквенс-специфичной нацеленной активности.

В одном варианте осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с RTDD, например, с направляющей последовательностью нуклеиновой кислоты, на 3'-конце направляющей последовательности нуклеиновой кислоты, и/или репарационная матричная последовательность нуклеиновой кислоты ассоциирована с 5'-концом RTDD, например, с направляющей последовательностью нуклеиновой кислоты, и/или репарационная матричная последовательность нуклеиновой кислоты расположена в пределах RTDD и, таким образом, образует отдельную функциональную часть RTDD.

Неожиданно авторы настоящего изобретения обнаружили, что gPHK в виде RTDD, несущая 3'-локализованную последовательность матрицы для репарации (RT) ДНК, либо одноцепочечной, либо двухцепочечной RT, могла свободно взаимодействовать с гомологичной последовательностью, поскольку она доставляется к мишени посредством полипептида CRISPR, например, Cas9 из системы CRISPR Типа II или Cpf1 из системы CRISPR Типа V, или другого эффектора полипептида CRISPR. Аналогичные наблюдения были сделаны, когда использовалась gPHK в виде RTDD, несущая 5'-локализованную последовательность матрицы для репарации ДНК, либо одноцепочечной, либо двухцепочечной RT, либо обеих, 3'- и 5'-локализованной RT. Таким образом, 3'- или 5'-локализованные означает, что RT либо ковалентно присоединена к 3'-концу или 5'-концу gPHK, либо это может означать, что RT гибридизирована, то есть, нековалентно ассоциирована, с областью, соответствующей последовательности, присоединенной к 3'- и/или 5'-области gPHK. Кроме того, RT могла быть ковалентно включена в петли стебля gPHK, или она

могла быть нековалентно ассоциирована указанными петлями стебля gРНК для достижения функциональной гибридной конструкции нуклеиновой кислоты, при этом, RTDD и RT непосредственно взаимодействуют. Таким образом, было обнаружено, что ДНК, ассоциированная с gРНК в различных положениях gРНК, как описано выше, обладает хорошей толерантностью, и поэтому эта новая форма гибридного комплекса подходит для объединения двух ключевых аспектов принципа редактирования генов: (1) точность нацеливания, опосредованного RTDD/gРНК, и (2) эффективная и сайт-направленная репарация, опосредованная RT. Кроме того, существует синергетический эффект, заключающийся в том, что gРНК и RT приводятся в непосредственную близость для повышения стабильности и доступности гибридной конструкции вместе с полипептидом CRISPR в качестве интересующей SSN в интересующем сайте ДНК-мишени.

Практически нет ограничений в отношении длины этой растянутой нуклеотидной последовательности матрицы для репарации, доставляемой как часть искусственного молекулярного комплекса по настоящему изобретению, в случае, если RT присоединена к 3'- или 5'-концу нуклеиновой кислоты на основе RTDD, например, gРНК или gДНК. Длина RT, независимая от вида RTDD, скорее всего определяется целевой модификацией, подлежащей введению. Типичные последовательности RT могут иметь длину от, примерно, 20 до 8000 пар оснований или даже более, например, наблюдается от 20 до 5000 пар оснований, от 30 до 8000 пар оснований,от 30 до 5000 пар оснований, от 40 до 8000 пар оснований, от 40 до 5000 пар оснований, от 50 до 8000 пар оснований, от 50 до 5000 пар оснований, 60 до 8000 пар оснований, от 60 до 5000 пар оснований, от 70 до 8000 пар оснований, от 70 до 5000 пар оснований, от 80 до 8000 пар оснований, от 80 до 5000 пар оснований, от 90 до 8000 пар оснований, от 90 до 5000 пар оснований, от 100 до 8000 пар оснований, от 100 до 5000 пар оснований одноцепочечной и/или двухцепочечной ДНК без значительной потери частоты разрезания, по меньшей мере, одной SSN. Как известно специалисту в данной области техники, длина RT-матрицы сильно зависит от вида модификации/вставки, подлежащей осуществлению/введению. В случае, если предполагается введение большей последовательности нуклеиновой кислоты, кодирующей интересующий белок, то длина RT-последовательности будет следующей: длина конструкции нуклеиновой кислоты, кодирующей интересующий белок, плюс два достаточно длинных плеча гомологии, расположенных слева и справа от последовательности. Таким образом, в принципе, нет верхнего предела в 1500 пар оснований, но RT может иметь до 5000 или даже больше пар оснований (bp). Например, более крупные вставки, которые в настоящее время вводятся с использованием плазмидной ДНК в качестве матрицы для репарации и продуцирования матрицы для репарации в сайте-мишени, используют левый и правый плечи гомологии размером 800 пар оснований и более, таким образом, общая длина матрицы для репарации может составлять несколько 1000 пар оснований. Длина вставок нуклеиновой кислоты должна быть рассчитана таким образом, чтобы не ингибировать интересующую сайт-специфичную нуклеазу, которая может быть определена в предварительных экспериментах.

Различные компоненты молекулярного комплекса по настоящему изобретению, то есть, по меньшей мере, одна SSN, по меньшей мере, один RTDD и, по меньшей мере, одна RT и, при необходимости, по меньшей мере, один взаимодействующий домен, ассоциированы функциональным образом.

Термин «ассоциированный функциональным образом» подразумевает, что компоненты искусственного комплекса приводятся в контакт, и таким образом, SSN и RTDD могут взаимодействовать друг с другом, в предпочтительном варианте осуществления настоящего изобретения, в виде нековалентной ассоциации, как подробно описано выше. По меньшей мере, одна последовательность RTDD,

взаимодействующая, по меньшей мере, с одной последовательностью RT, собирается независимо, либо до, после, либо одновременно с контактом, по меньшей мере, с последовательностью RTDD с, по меньшей мере, одной соответствующей SSN, или вариантом, или ее каталитически активным фрагментом, представляющим интерес. В одном варианте осуществления настоящего изобретения, весь комплекс, при необходимости, содержащий, по меньшей мере, один взаимодействующий домен, ассоциируется *in vitro* перед его введением в клетку-мишень, содержащую, по меньшей мере, одну интересующую область ДНК-мишени, подлежащей редактированию. В другом варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и, при необходимости, взаимодействующий домен вводится, по меньшей мере, в одну ячейку-мишень до или после, по меньшей мере, одной взаимодействующей последовательности RTDD/RT. Полипептид SSN может быть введен в клетку-мишень посредством трансфекции полипептидной последовательности или путем трансфекции, или трансформации, по меньшей мере, одной клетки-мишени с РНК, кодирующей, по меньшей мере, один полипептид SSN, или путем введения конструкции для доставки, кодирующей, по меньшей мере, один полипептид SSN, который может быть транскрибирован и транслирован в ячейке-мишени. Аналогичным образом, в определенных вариантах осуществления настоящего изобретения, последовательность(и) RTDD и репарационная матричная последовательность(и) нуклеиновой кислоты могут быть предложены одновременно в виде предложенной *in vitro* и собранной конструкции. В альтернативном варианте осуществления настоящего изобретения, последовательность RTDD и/или репарационная матричная последовательность нуклеиновой кислоты может быть трансфицирована или трансформирована в клетку-мишень с помощью подходящего вектора доставки. В предпочтительном варианте осуществления настоящего изобретения, весь искусственный молекулярный комплекс собирается *in vitro* и затем вводится в интересующую клетку-мишень для обеспечения лучшего пространственного и стехиометрического контроля конструкции редактирования генома. В другом предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и, при необходимости, полипептид взаимодействующего домена вводят в клетку-мишень до последовательностей RTDD/RT, а затем, впоследствии, по меньшей мере, одну из последовательностей RTDD/RT вводят в интересующую клетку-мишень. Последовательный порядок мог бы быть предпочтительным для определенных подходов с использованием, например, gРНК в качестве RTDD из-за по сути низкой стабильности РНК, по сравнению с полипептидом, таким образом, введенная gРНК будет немедленно связана и стабилизирована посредством SSN, то есть, для определенных вариантов осуществления настоящего изобретения, полипептид CRISPR уже присутствует в клетке. Не желая быть связанными какой-либо теорией, сборка *ex vivo* направляющей последовательности нуклеиновой кислоты и репарационной матричной последовательности нуклеиновой кислоты также может повысить стабильность конструкции, по сравнению только с одной направляющей РНК.

В настоящее время существует множество способов трансформации растений для введения генетического материала в форме генетической конструкции в интересующую растительную клетку, содержащую биологические и физические средства, известные специалисту в области биотехнологии растений. Общим биологическим средством является трансформация посредством *Agrobacterium spp.*, которое использовалось в течение многих десятилетий для множества различных растительных материалов. Опосредованная вирусным вектором трансформация растений представляет собой дополнительную стратегию для введения генетического материала в интересующую клетку. Физические средства, находящие применение в биологии растений, представляют собой бомбардировку частицами, также называемую биолистической трансфекцией или опосредованным микрочастицами переносом генов, который относится к

способу физической доставки для переноса покрытой микрочастицы или наночастицы, содержащей нуклеиновую кислоту или интересующую генетическую конструкцию в клетку-мишень или ткань-мишень. Средства физического введения подходят для введения нуклеиновых кислот, то есть, РНК и/или ДНК, и белков. Аналогичным образом, существуют способы специфичной трансформации или трансфекции для специфичного введения интересующей конструкции нуклеиновой кислоты или аминокислотной конструкции в растительную клетку, включая электропорацию, микроинъекцию, наночастицы и проникающие в клетку пептиды (CPP). Кроме того, существуют способы трансфекции на химической основе для введения генетических конструкций и/или нуклеиновых кислот, и/или белков, содержащие, среди прочего, трансфекцию посредством фосфата кальция, трансфекцию с использованием липосом, например, катионных липосом, или трансфекцию посредством катионных полимеров, включая DEAD-декстран или полиэтиленимин, или их комбинации. Указанные способы доставки и средства доставки или грузы, таким образом, по своей природе отличаются от инструментов доставки, которые используются для других эукариотических клеток, включая клетки животных и млекопитающих, и каждый способ доставки должен быть специально отрегулирован и оптимизирован, чтобы можно было ввести интересующую конструкцию для опосредования редактирования генома в конкретный компартмент интересующей клетки-мишени полностью функциональным и активным образом. Вышеуказанные методы доставки, по отдельности или в комбинации, могут быть использованы для вставки, по меньшей мере, одного искусственного молекулярного комплекса или, по меньшей мере, одного его субкомпонента, то есть, по меньшей мере, одной SSN, по меньшей мере, одного RTDD, по меньшей мере, одной RT и, при необходимости, по меньшей мере, одного IA, или последовательностей, кодирующих вышеупомянутые субкомпоненты по настоящему изобретению в клетку-мишень *in vivo* или *in vitro*.

В определенных вариантах осуществления настоящего изобретения, режимы доставки искусственного молекулярного комплекса по настоящему изобретению могут быть выбраны из PEG-опосредованной доставки комплекса SSN-(IA)-RTDD-RT, PEG-опосредованной доставки плазмиды, кодирующей SSN-(IA)-RTDD, RTDD, например, представляющего собой gРНК или gДНК, и параллельной доставки RT, бомбардировки комплекса SSN-(IA)-RTDD-RT, бомбардировки белка, кодирующего плазмиду (SSN и, при необходимости, IA)-RTDD, например, gРНК/gДНК, и параллельной доставки RT, доставки опосредованной пептидом, проникающим в клетку (CPP), комплекса SSN-(IA)-RTDD-RT, липофекции комплекса SSN-(IA)-RTDD-RT, липофекции белка, кодирующего плазмиду (SSN и, при необходимости, IA)-RTDD, например, gРНК/gДНК, и параллельной доставки RT или стабильной экспрессии белка (SSN и, при необходимости, IA) и транзиентной доставки RTDD или, для некоторых RTDD, плазмиды, кодирующей RTDD, и параллельной доставки rtДНК.

В определенных вариантах осуществления настоящего изобретения, crРНК часть gРНК содержит петлю стебля или оптимизированную структуру петли стебля или оптимизированную вторичную структуру. В другом варианте осуществления настоящего изобретения, зрелая crРНК содержит петлю стебля или оптимизированную структуру петли стебля в последовательности прямого повтора, при этом, петля стебля или оптимизированная структура петли стебля важна для активности расщепления. В определенных вариантах осуществления настоящего изобретения, зрелая crРНК, в предпочтительном варианте осуществления настоящего изобретения, содержит одну петлю стебля. В определенных вариантах осуществления настоящего изобретения, последовательность прямого повтора, в предпочтительном варианте осуществления настоящего изобретения, содержит одну петлю стебля. В некоторых вариантах осуществления настоящего изобретения, активность расщепления эффекторного белкового комплекса

модифицируется путем введения мутаций, которые влияют на дуплексную структуру РНК петли стебля. В предпочтительных вариантах осуществления настоящего изобретения, могут быть введены мутации, которые поддерживают дуплекс РНК петли стебля, благодаря чему сохраняется активность расщепления эффекторного белкового комплекса. В других предпочтительных вариантах осуществления настоящего изобретения, могут быть введены мутации, которые нарушают дуплексную структуру РНК петли стебля, в результате чего активность расщепления эффекторного белкового комплекса полностью исчезает.

Размер, по меньшей мере, одной репарационной матричной последовательности нуклеиновой кислоты по настоящему изобретению как части искусственного молекулярного комплекса по настоящему изобретению может варьироваться. Он может находиться в диапазоне от, примерно, 20 пар оснований до, примерно, 5000 пар оснований или даже до 8000 пар оснований, в зависимости от последовательности ДНК-мишени, подлежащей модификации.

Матрицы HOR, используемые для создания специфичных мутаций или вставок в интересующую область ДНК-мишени, требуют определенного количества гомологии, окружающей последовательность-мишень, которая будет модифицирована. Лучше всего, если сайты вставки модификации находятся на расстоянии не более 100 пар оснований от DSB, на который воздействует SSN или партнер по слиянию, то есть, взаимодействующий домен, в случае SSN с дефицитом нуклеазы, например, полипептида CRISPR, в идеале на расстоянии менее 10 пар оснований, если это возможно, и общая длина плеча гомологии является важным фактором, который следует учитывать при его конструировании. Более длинные расстояния будут работать, но эффективность, вероятно, будет ниже, и может потребоваться введение маркера отбора, чтобы гарантировать наличие желаемой модификации, подлежащей введению в интересующую последовательность ДНК-мишени.

Согласно различным аспектам настоящего изобретения, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты может представлять собой молекулу нуклеиновой кислоты одноцепочечной или двухцепочечной ДНК. По меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты может быть предложена в виде, по меньшей мере, одной линейной, ss- или ds-ДНК молекулы. Однако может оказаться целесообразным использовать, по меньшей мере, одну одноцепочечную или двухцепочечную репарационную матричную последовательность нуклеиновой кислоты, которую продуцируют *ex vivo*, когда молекулярный комплекс должен быть собран *ex vivo*, что особенно подходит для повышения доступности функционального комплекса SSN-RTDD-RT, поскольку все компоненты могут быть введены одновременно в корректной стехиометрии, чтобы повысить специфичность подхода к редактированию генома.

Синтез более крупных последовательностей нуклеиновой кислоты, одноцепочечной или двухцепочечной, может быть осуществлен с использованием общепринятых способов известного уровня техники. Отмечено, что для некоторых вариантов осуществления настоящего изобретения, могут также подходить частично одноцепочечные и/или частично двухцепочечные репарационные матричные последовательности нуклеиновой кислоты. Возможна любая комбинация одноцепочечной и/или двухцепочечной последовательности нуклеиновой кислоты и любой вид введения, либо одновременно, либо до, либо после введения полипептидных компонентов искусственного молекулярного комплекса. В одном варианте осуществления настоящего изобретения, предусматривается введение молекулярного комплекса, согласно второму аспекту, в клетку-мишень, при этом, клетка-мишень содержит дополнительный плазмидный вектор, кодирующий матрицу для репарации или дополнительную репарационную матричную последовательность, поскольку использование более одной репарационной матричной последовательности

нуклеиновой кислоты благоприятно для определенных подходов к редактированию генома, при этом, искусственный молекулярный комплекс может затем собираться *in vivo* после предложения различных компонентов. В целом, высокая физическая доступность репарационной матричной последовательности нуклеиновой кислоты в этом сайте в пределах клетки-мишени, где расположена область ДНК-мишени, имеет исключительную важность для обеспечения события высокоточного редактирования генома. В некоторых вариантах осуществления настоящего изобретения, матрицы для репарации одноцепочечной (ss) ДНК особенно пригодны для достижения правильного баланса, поддерживающего молекулярную массу настолько низкой, насколько это возможно, предлагая при этом достаточную длину для гомологичных взаимодействий для достижения оптимальной гомологично направленной репарации.

В одном варианте осуществления настоящего изобретения, в соответствии с любым аспектом настоящего изобретения, по меньшей мере, одна SSN представляет собой полипептид CRISPR и выбрана независимо из группы, состоящей из полипептида Cas из *Streptococcus spp.*, включая *Streptococcus pyogenes, Streptococcus thermophilus, Staphylococcus aureus,* или *Neisseria spp.*, включая *Neisseria meningitides, Corynebacter, Sutterella, Legionella, Treponema, Filifactor, Eubacterium, Lactobacillus, Mycoplasma, Bacteroides, Flaviivola, Flavobacterium, Sphaerochaeta, Azospirillum, Gluconacetobacter, Roseburia, Parvibaculum, Nitratifractor, Mycoplasma, Campylobacter, Candidatus* Micrarchaeum acidiphilum ARMAN-1, *Parcubacteria* (GenBank: APG80656.1), *Sulfolobus* spp., включая *Sulfolobus islandicus* HVE10/4 (GenBank: ADX81770.1) или REY15A (GenBank: ADX84852.1), или, при этом, полипептид CRISPR выбран из полипептида Cpf1 из археи или бактерии, включая полипептид Cpf1 *Acidaminococcus spp.*, включая *Acidaminococcus sp.* BV3L6, *Lachnospiraceae spp.*, включая *Lachnospiraceae bacterium* ND2006, *Francisella spp.*, включая *Francisella novicida* U112, *Eubacterium eligens*, *Prevotella spp.*, или *Porphyromonas spp.*, или их варианты и/или функциональные фрагменты, и/или комбинации, включая полипептидные никазы CRISPR, или полипептид CRISPR, лишенный эндонуклеолитической активности.

В одном варианте осуществления настоящего изобретения, последовательности RTDD/RT по настоящему изобретению могут использоваться с никазой SSN, например, с никазой Cas9, мутантным для минимизации нецеленаправленных мутаций, при этом, используются парные направляющие РНК, каждая из которых специфична для мутанта никазы, полученного из Cas9.

В некоторых вариантах осуществления настоящего изобретения, по меньшей мере, одна SSN и, при необходимости, по меньшей мере, один взаимодействующий домен предложен в виде экспрессированного *in vitro*, транслированного или синтезированного полипептида. В некоторых вариантах осуществления настоящего изобретения, используется вектор доставки, кодирующий, по меньшей мере, один полипептид CRISPR, при этом, вектор доставки может дополнительно содержать регуляторные последовательности или сигналы локализации. Полипептид SSN, который мутирован по отношению к соответствующему ферменту дикого типа таким образом, что мутированный фермент SSN, не обладает способностью расщеплять одну или обе цепи полинуклеотида-мишени, содержащего последовательность-мишень, которую также содержат различные варианты осуществления настоящего изобретения. Например, можно использовать замещение аспартат-на-аланин (D10A) в каталитическом RuvC I домене Cas9 из S. *pyogenes*, которое преобразовывает Cas9 из эндонуклеазы, которая расщепляет обе цепи интересующей области ДНК-мишени, в никазу, расщепляющую одноцепочечную ДНК. Другие примеры мутаций, которые превращают полипептид Cas9 в никазу, включают, без ограничения, H840A, N854A и N863A. В качестве дополнительного примера, по меньшей мере, два каталитических домена Cas9 (RuvC I, RuvC II и RuvC III или домен HNH) могут быть мутированы для продуцирования мутированного Cas9, по существу,

лишенного всей активности расщепления ДНК. В некоторых вариантах осуществления настоящего изобретения, мутация D10A комбинируется, по меньшей мере, с одной мутацией H840A, N854A или N863A для продуцирования фермента Cas9, по существу, не обладающего всей активностью расщепления ДНК. В некоторых вариантах осуществления настоящего изобретения, считается, что фермент SSN, по существу, не обладает всей активностью расщепления ДНК, когда активность расщепления ДНК мутированного фермента составляет, примерно, не более 25%, 10%, 5%, 1%, 0,1%, 0,01% или менее активности расщепления ДНК немутрированной формы фермента; примером может служить случай, когда активность расщепления ДНК мутированной формы является нулевой или незначительной, по сравнению с немутрированной формой дикого типа. Если фермент не представляет собой Cas9 из S. *pyogenes*, то мутации могут быть сделаны в любом или во всех остатках, соответствующих положениям 10, 762, 840, 854, 863 и/или 986 *Sp*Cas9 (которые могут быть установлены, например, стандартными инструментами сравнения последовательностей). В частности, любые или все следующие мутации являются предпочтительными в Cas9 из S. *pyogenes*: D10A, E762A, H840A, N854A, N863A и/или D986A; а также консервативное замещение любой из заменяемых аминокислот также предусмотрено по настоящему изобретению. Такие же или консервативные замещения этих мутаций в соответствующих положениях в других Cas9 также возможны для определенных вариантов осуществления настоящего изобретения, в частности, для D10 и H840 в Cas9 из S. *pyogenes*. Однако в других Cas9 также возможны остатки, соответствующие D10 и H840 Cas9 из S. *pyogenes*. «Ортологи» или «ортологичные» данных белков CRISPR также могут быть использованы в практическом применении изобретения. Ортологи представляют собой гены разных видов, которые произошли из общего предкового гена путем видообразования. Обычно ортологи сохраняют ту же функцию в ходе эволюции. В наиболее предпочтительном варианте осуществления настоящего изобретения, фермент Cas9 в виде SSN происходит или получен из Cas9 S. *pyogenes* или Cas9 S. *aureus*, или Cas9 дикого типа из S. *thermophilus*, белковая последовательность которого указана в базе данных SwissProt под учетным номером G3ECR1. Аналогичным образом, Cas9 S. *pyogenes* или Cas9 S. *aureus* включены в SwissProt под учетным номером Q99ZW2.

В одном варианте осуществления настоящего изобретения, направляющая РНК в виде последовательности RTDD по настоящему изобретению служит в качестве РНК, имеющей оптимальную активность, то есть, свойства распознавания, в отношении выбранного фермента SSN или полипептида определенной длины, фермент SSN, таким образом, может быть усечен до каталитически активного фрагмента SSN дикого типа, делая его меньше по длине, чем соответствующий фермент дикого типа, путем усечения молекул нуклеиновой кислоты, кодирующих фермент SSN, который может быть транскрибирован или транслирован *in vitro* или *in vivo*, или путем предложения синтезированного полипептида SSN. Также возможно получение химерных ферментов SSN, при этом, различные части фермента заменяются или обмениваются между различными ортологами для получения химерных ферментов, имеющих индивидуальную специфичность.

«Вариант» или «функциональный фрагмент» по настоящему изобретению, таким образом, содержит любую SSN и/или взаимодействующий домен, и/или белок RTDD, или его усеченную версию, полученную из SSN дикого типа, и/или взаимодействующий домен, и/или белок RTDD, то есть, имеет степень гомологии последовательностей с ферментом дикого типа, но он мутирован (модифицирован) каким-либо образом, как описано в настоящем документе. Например, ферментная активность нуклеазы, полученной из Cas9, приводит к образованию двухцепочечных разрывов в последовательностях сайта-мишени, которые гибридизуются с 20 нуклеотидами направляющей последовательности и имеют примеры

последовательностей с мотивом, примыкающим к протоспейсеру (PAM), включая NGG/NRG или PAM, которые могут быть определены, как описано в настоящем документе, после 20 нуклеотидов последовательности-мишени. Эту ферментную функцию можно варьировать, генерируя варианты SSN, обладающие никазной активностью, или вариант с неактивной нуклеазой. Кроме того, вариант с SSN и/или взаимодействующим доменом, и/или полипептидом RTDD по настоящему изобретению может быть кодон-оптимизированным для адаптации SSN и/или взаимодействующего домена, и/или полипептида RTDD к использованию кодонов клетки-мишени, в предпочтительном варианте осуществления настоящего изобретения, эукариотической клетки, в предпочтительном варианте осуществления настоящего изобретения, животной или растительной клетки.

В предпочтительных вариантах осуществления настоящего изобретения, компоненты искусственного молекулярного комплекса, в частности, компоненты SSN или IA, или их каталитически активные фрагменты, по-прежнему проявляющие каталитическую функцию полипептида дикого типа, и/или дополнительные компоненты могут быть кодон-оптимизированными, и/или полипептид SSN, и/или взаимодействующий домен, и/или RTDD, и/или RT могут быть связаны с меченой последовательностью, чтобы идентифицировать местоположение последовательности-мишени и/или искусственного молекулярного комплекса. Метка может быть выбрана из группы, состоящей из полигистидина (His)-метки, глутатиона-S-трансферазы (GST)-метки, тиоредоксина-метки, FLAG-метки, метки, имеющей флуоресцентные свойства, например, выбранной из метки (E)GFP ((усиленного) зеленого флуоресцентного белка), DsRed-метки, mCherry-метки, (t)dтомат-метки, mNeonGreen-метки и тому подобного или из стрептавидина, или стреп-метки, метки мальтоза-связывающего белка (MBP), транзитного пептида, позволяющего нацеливаться на субклеточный компартмент, включая митохондрии или ядро, SNAP-метки и/или метки секреции, позволяющей секретировать аминокислотную последовательность, присоединенную к ней, причем, искусственная аминокислота, обычно не встречается в природе, или из комбинации вышеупомянутых меток. Белковый компонент искусственного молекулярного комплекса, например, SSN и/или домен взаимодействия может содержать любую дополнительную белковую последовательность и, при необходимости, линкерную последовательность между любыми двумя доменами. Примеры белковых доменов, которые могут быть слиты с любым компонентом, по меньшей мере, одного искусственного молекулярного комплекса, включают, без ограничения, эпитопные метки, последовательности репортерных генов и белковые домены, обладающие, по меньшей мере, одной из следующих активностей: метилазная активность, деметилазная активность, активность активации транскрипции, активность репрессии транскрипции, активность фактора высвобождения транскрипции, активность модификации гистонов, активность расщепления РНК и активность связывания нуклеиновых кислот. Неограничивающие примеры эпитопных меток включают метки гистидина (His), метки V5, метки FLAG, метки гемагглютинина (HA) гриппа, метки Myc, VSV-G-метки и метки тиоредоксина (Trx). Примеры репортерных генов включают, но этим не ограничиваются, глутатион-S-трансферазу (GST), пероксидазу хрена (HRP), хлорамфениколацетилтрансферазу (CAT), бета-галактозидазу, бета-глюкуронидазу, люциферазу, зеленый флуоресцентный белок (GFP), HcRed, DsRed, голубой флуоресцентный белок (CFP), желтый флуоресцентный белок (YFP) и аутофлуоресцентные белки, включая синий флуоресцентный белок (BFP). Фермент CRISPR может быть слит с последовательностью гена, кодирующей белок или фрагмент белка, который связывают молекулы ДНК или другие клеточные молекулы, включая, но этим не ограничиваясь, мальтозосвязывающий белок (MBP), S-метку, слияния ДНК-связывающего домена (DBD) Lex A, слияния ДНК-связывающего домена GAL4 и слияния белка BP16 вируса простого герпеса (HSV).

В одном варианте осуществления настоящего изобретения, по меньшей мере, один компонент искусственного молекулярного комплекса может быть модифицированным, функционирующим в качестве ДНК-никазы и/или полипептида SSN, или ее каталитически активного фрагмента, может присутствовать в форме молекулы слияния с другим функциональным фрагментом молекулы, в предпочтительном варианте осуществления настоящего изобретения, с функциональным полипептидным фрагментом молекулы, имеющим ферментную функцию, в предпочтительном варианте осуществления настоящего изобретения, с функциональным фрагментом молекулы, имеющим функцию моделирования хроматина и/или стимулирующую гомологичную рекомбинацию, и/или модифицирующую транскрипцию. При анализе, по меньшей мере, одной модифицированной клетки в пределах ткани многоклеточного организма предпочтительными являются такие метки и маркерные белки, в частности, флуоресцентные белковые метки, которые имеют яркую флуоресценцию, благодаря чему их можно определить даже в более глубоких слоях тканей, состоящих из нескольких видов клеток. Подходящие флуоресцентные белки доступны на рынке, и специалист в данной области техники может легко их выбрать для конкретной цели.

В соответствии с различными вариантами осуществления настоящего изобретения, или SSN и/или взаимодействующий домен, и/или полипептид(ы) RTDD, и/или RTDD, и/или RT последовательность(и) содержит, по меньшей мере, одну последовательность ядерной локализации, и/или последовательность пластидной локализация, например, последовательность митохондриальной локализации или последовательность хлоропластной локализация, для эффективного нацеливания полипептида SSN на клеточный компартмент, содержащий интересующую последовательность геномной ДНК, подлежащую модификации. Требования, предъявляемые к последовательностям таких последовательностей локализации, известны специалисту в области молекулярной биологии. Чтобы не затруднять функцию полипептида SSN или RT нуклеотидной последовательности, последовательность локализации слита, то есть, ковалентно связана, с частью N-терминального конца или С-терминального конца или, соответственно, с 5'-концом или 3'-концом соответствующей молекулы.

В одном варианте осуществления настоящего изобретения, по меньшей мере, один полипептид SSN и, при необходимости, по меньшей мере, один взаимодействующий домен, если он представляет собой полипептидную последовательность, предложен в виде полипептидной последовательности, продуцируемой *ex vivo*, либо с использованием рекомбинантных технологий для продуцирования белка, либо путем синтеза соответствующей аминокислотной последовательности. В другом варианте осуществления настоящего изобретения, полипептид SSN и, при необходимости, по меньшей мере, один взаимодействующий домен представлены в виде последовательности РНК, которая может быть транслирована в соответствующую аминокислотную последовательность при введении интересующей клетки-мишени. В еще одном варианте осуществления настоящего изобретения, полипептид SSN и, при необходимости, по меньшей мере, один полипептид взаимодействующего домена вставляют в качестве конструкции ДНК, либо выполненной с возможностью стабильной экспрессии, либо транзиентной экспрессии в интересующей клетке, в результате чего, по меньшей мере, один полипептид SSN и, при необходимости, по меньшей мере, один полипептид взаимодействующего домена затем транскрибируют и транслируют в интересующую клетку-мишень конститутивным или индуцибельным образом. Подходящие конструкции ДНК и ассоциированные способы введения, по меньшей мере, одного полипептида SSN и, при необходимости, по меньшей мере, одного полипептида взаимодействующего домена по настоящему изобретению в клетку-мишень известны специалисту в данной области техники, однако конкретные пути введения, по меньшей мере, одного полипептида SSN и, при необходимости, по меньшей мере, одного полипептида взаимодействующего

домена в соответствии с определенными вариантами осуществления настоящего изобретения, специально адаптированными к применению в растительных клетках, более подробно описаны ниже.

Искусственный молекулярный комплекс или его части, то есть, по меньшей мере, один полипептид SSN, по меньшей мере, один RTDD и, по меньшей мере, одна RT и, при необходимости, по меньшей мере, один взаимодействующий домен, должны быть введены в интересующую клетку-мишень с использованием подходящей конструкции для доставки. Естественно, тип конструкции для доставки может варьироваться в зависимости от того, полностью ли молекулярный комплекс собирается *in vitro* и затем вводится в клетку-мишень, или же различные компоненты молекулярного комплекса отдельно вводятся в клетку, и комплекс затем собирается посредством нековалентных взаимодействий внутри интересующей клетки-мишени. Введение обычно происходит с использованием подходящей конструкции доставки.

Термин «конструкция доставки» или «вектор (доставки)», используемый в настоящем документе в соответствии с различными вариантами осуществления различных аспектов настоящего изобретения, относится к любым биологическим или химическим, или нехимическим, или основанным на частицах средствам и/или способам, используемым в качестве груза для транспортировки нуклеотидной и/или аминокислотной последовательности, представляющей интерес, в целевую эукариотическую клетку. Подходящие конструкции доставки содержат биологические средства для доставки нуклеотидных последовательностей в клетку-мишень, включая вирусные векторы, *Agrobacterium spp.*, проникающие в клетки пептиды (CPP) или химические конструкции доставки, включая наночастицы, липидные или полимерные пузырьки, фосфат кальция или их комбинации. Липидные или полимерные пузырьки могут быть выбраны, например, из липидов, липосом, систем липидной инкапсуляции, наночастиц, например, наночастиц мезопористого диоксида кремния, составов из малых частиц нуклеиновой кислоты и липидов, полимеров, например, катионных полимеров, таких как DEAE-декстран, или полиэтиленимина и полимерсомов. В одном варианте осуществления настоящего изобретения, полимер выбирают из группы, состоящей из линейных полимеров, разветвленных полимеров, дендримеров (высокоразветвленных органических соединений) и полисахаридов. В другом варианте осуществления настоящего изобретения, система липидной инкапсуляции содержит, по меньшей мере, один фосфолипид, холестерин, полиэтиленгликоль (PEG)-липид и липофильное соединение, которое доставляет частицу в ткань-мишень. В следующем варианте осуществления настоящего изобретения, конструкция доставки может представлять собой наночастицу мезопористого диоксида кремния.

Используемые в настоящем документе физические способы введения и пригодные для предложения, по меньшей мере, одного молекулярного комплекса или, по меньшей мере, одной гибридной последовательности нуклеиновой кислоты РНК/ДНК по настоящему изобретению относятся к электропорации, микроинъекции, бомбардировке частицами, сонопорации, магнитофекции или импалефекции с использованием удлиненных наноструктур и массивов таких наноструктур, таких как, углеродные нановолокна или кремниевые нанопроволоки, которые были функционализированы плазмидной ДНК, и к химическим способам, и могут основываться на использование микрочастиц или наночастиц, или химических веществ, включая полиэтиленгликоль (PEG).

Например, для варианта осуществления настоящего изобретения, где компоненты искусственного молекулярного комплекса ассоциированы *ex vivo*, вектор доставки может быть на основе липидов или полимерным вектором. Векторы на основе липидов или полимерные векторы могут быть выбраны, например, из липидов, липосом, систем липидной инкапсуляции, микрочастиц, вискеров, наночастиц, малых нуклеиновых кислото-липидных частиц, полимеров и полимерсомов. В некоторых вариантах осуществления

настоящего изобретения, полимер может быть выбран из группы, состоящей из линейных полимеров, разветвленных полимеров, дендримеров и полисахаридов. В другом варианте осуществления настоящего изобретения, система липидной инкапсуляции содержит, по меньшей мере, один фосфолипид, холестерин, полиэтиленгликоль (PEG)-липид и липофильное соединение, которое доставляет частицу в ткань-мишень.

Для клеток млекопитающих, модификация иммунных клеток *ex vivo* для различных терапевтических целей приобрела большой интерес в течение последнего десятилетия для борьбы с некоторыми опухолевыми заболеваниями путем адоптивно перенесенных специфично модифицированных лимфоцитов, в предпочтительном варианте осуществления настоящего изобретения, Т-клеток. В частности, CD8+ Т-клеточные лимфоциты являются интересными мишенями в этом отношении. Было описано, что иммунные ответы, полученные из единичных наивных Т-клеток, единичных первичных и единичных вторичных центральных Т-клеток памяти, достигали сходного размера и фенотипического разнообразия, подвергались сопоставимой стохастической вариации и, в конечном итоге, могли восстанавливать иммунокомпетентность против другой смертельной инфекции посредством бактериального патогена, измеренного путем использования карты судьбы *in vivo* CD8+ Т-клеток и их потомков в трех поколениях последовательного одноклеточного адоптивного переноса и реэкспансии, вызванной инфекцией (Греф и др., Immunity, 41, 116-126, 2014). После развития *de nova* Т-клеток тимуса из гемопоэтических клеток, полностью зрелые антигенспецифичные Т-клетки могут сохраняться у индивидуума в течение длительного периода времени, при этом, антиген может представлять собой чужеродный антиген, например, антиген, экспрессированный на вирусе или раковой клетке. Таким образом, целевая модификация таких эффекторных Т-клеток или их предшественников представляет собой важную стратегию для предложения подходящих Т-клеток для иммунотерапии. Наивные Т-клетки дифференцируются посредством стадии, называемой стволовыми Т-клетками памяти, которые дают начало центральным Т-клеткам памяти и эффекторным Т-клеткам памяти и, наконец, эффекторным Т-клеткам, при этом, эффекторные Т-клетки представляют собой терминально дифференцированные клетки, которые, в конечном итоге, могут распознать и уничтожить клетку-мишень. Эффекторные Т-клетки памяти и эффекторные Т-клетки представляют собой подмножества Т-клеток, которые обладают способностью перемещаться к периферическим тканям. В настоящее время предлагается другое подмножество, Т-клетки памяти, находящиеся в тканях, которые больше не циркулируют (см. например , Фарбер и др., Nature Reviews Immunology, 14, 24-35, 2014).

Кроме того, иммунотерапия раковых заболеваний предложила некоторые из первых примечательных клинических случаев, показывающих, что адоптивный перенос Т-клеток, экспрессирующих рекомбинантные опухолереактивные рецепторы, может вылечить иные устойчивые к лечению злокачественные новообразования (Брентьенс и др., 2013; Групп и др., 2013; Портер и др., 2011), и что использование сконструированных Т-клеток в терапии адоптивного переноса показало значительные перспективы в лечении раковых заболеваний, особенно гематологических раковых заболеваний. Все больше и больше генетически модифицированных Т-клеток определенного подмножества и фенотипической композиции используются для умножения успеха в иммунотерапии рака (см. Риддель и др., Cancer J., 20(2), 141-144, 2014). Применение Т-клеток, модифицированных химерным рецептором антигена, в качестве терапии гематологических злокачественных новообразований, а также солидных опухолей становится все более распространенным. С этой целью Т-клетки модифицируют для экспрессии химерных рецепторов антигена, нацеленных на опухоль (CAR) (см., например, Ануратапан и др., Molecular Therapy, 22, 623-633, 2014). Также становится все более важным и так называемое второе поколение CAR, например, CAR,

нацеленные на CD19, которые включают сигнальные домены CD2B или 4-1 BB для перенацеливания и перепрограммирования Т-клеток для увеличения их противоопухолевой эффективности (см., например, Шаукье и др., Nature Reviews Drug Discovery, 14, 499-509, 2015).

Поэтому гибридные последовательности нуклеиновой кислоты РНК/ДНК по настоящему изобретению представляют собой важный инструмент для модификации, по меньшей мере, одной клетки млекопитающих *in vivo* или *ex vivo*, в предпочтительном варианте осуществления настоящего изобретения, для лечения заболевания. Например, лимфоцитарная клетка, в более предпочтительном варианте осуществления настоящего изобретения, Т-клетка или природная клетка-киллер (NK) любой стадии развития будет изменять ген, экспрессируемый Т-клеткой или NK-клеткой, чтобы влиять на пролиферацию Т-клетки или NK-клетки, выживание и/или функционирование с высокой точностью, чтобы избежать нежелательных эффектов, которые могут быть вредными для терапевтического применения модифицированной клетки или клеточной популяции.

В определенных вариантах осуществления настоящего изобретения, искусственный молекулярный комплекс по настоящему изобретению, таким образом, подходит для использования в способе лечения заболевания, при этом заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации. Таким образом, предложен способ лечения заболевания с использованием искусственного молекулярного комплекса по любому из предыдущих пунктов, при этом заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации. Терапевтический способ лечения может содержать редактирование генов или генома, или генную терапию.

Подходящие клетки, в частности, для терапевтических подходов или для модификации вирусного генома, включают эукариотические (например, животные) и прокариотические клетки и/или клеточные линии. Неограничивающие примеры таких клеток или клеточных линий, полученных из таких клеток, включают COS, CHO (например, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (например, HEK293-F, HEK293-H, HEK293-T) и клетки perC6, а также клетки насекомых, такие как, *Spodopterafugiperda* (Sf) или грибковые клетки, такие как, *Saccharomyces, Pichia* и *Schizosaccharomyces*. В некоторых вариантах осуществления настоящего изобретения, клеточная линия представляет собой клеточную линию CHO, MDCK или HEK293. Подходящие клетки также включают стволовые клетки, такие как, например, нечеловекоподобные эмбриональные стволовые клетки, индуцированные плюрипотентные стволовые клетки, гематопоэтические стволовые клетки, нейрональные стволовые клетки и мезенхимальные стволовые клетки.

В одном аспекте, настоящее изобретение предлагает способ лечения пациента, нуждающегося в этом, содержащий индуцирование редактирования генов путем трансформации/трансфекции пациенту компонентов искусственного молекулярного комплекса, рассматриваемых в настоящем документе, или любого из векторов, рассматриваемых в настоящем документе, и введения источника энергии индуктора пациенту. Настоящее изобретение охватывает применение такого полинуклеотида или вектора в производстве лекарственного средства, например, такого лекарственного средства для лечения пациента или для такого способа лечения пациента. Настоящее изобретение охватывает полинуклеотид, рассматриваемый в настоящем документе, или любой из векторов, рассматриваемых в настоящем документе, для использования в способе лечения пациента, нуждающегося в этом, содержащем индуцирование

редактирования генов, при этом, способ дополнительно содержит введение источника энергии индуктора пациенту. В одном аспекте, в способе также предлагается матрица для репарации, например, доставляемая вектором, содержащим указанную матрицу для репарации.

В одном варианте осуществления настоящего изобретения, также рассматриваются минимальные неприматные лентивирусные векторы на основе вируса инфекционной анемии у лошадей (EIAV), в частности, для генной терапии с использованием искусственных молекулярных комплексов настоящего изобретения (см., например, Балагаан, J Gene Med 2006; 8: 275-285). В другом варианте осуществления настоящего изобретения, также рассматривается RetinoStat®, вектор лентивирусной генной терапии на основе вируса инфекционной анемии у лошадей, который экспрессирует ангиостатические белки эндостатин и ангиостатин, доставляемые посредством субретинальной инъекции для лечения влажной формы возрастной макулярной дегенерации (см., например, Бинли и др., HUMAN GENE THERAPY 23: 980-991 (сентябрь 2012)), и этот вектор может быть модифицирован для системы настоящего изобретения SSN-RTDD-RT. В настоящее время раскрыты лентивирусные векторы для применения в лечении болезни Паркинсона, см., например, Патентную заявку США № 2012/0295960 A1 и Патентную заявку США № 7,303,910 B2. Также были раскрыты лентивирусные векторы для лечения глазных заболеваний, см., например, Патентные заявки США №№ 2006/0281180, 2009/0007284, 2011/0117189, 2009/0017543, 2007/0054961 и 2010/0317109. Также были раскрыты лентивирусные векторы для доставки в мозг, см., например, Патентные заявки США №№ 2011/0293571, 2011/0293571, 2004/0013648, 2007/0025970, 2009/0111106 и Патент США № 7,259,015.

В другом варианте осуществления настоящего изобретения, искусственный молекулярный комплекс или его компоненты могут вводиться в липосомы, такие как стабильная частица нуклеиновая кислота-липид (SNALP) (см., например, Морриссей и др., Nature Biotechnology, Том 23, № 8, август 2005 года). Рассматриваются ежедневные внутривенные впрыски в дозах, примерно, 1, 3 или 5 мг/кг/день специфичной системы CRISPR-Cas, нацеленной на SNALP. Ежедневное лечение может продолжаться в течение, примерно, трех дней и затем еженедельно в течение, примерно, пяти недель. В другом варианте осуществления настоящего изобретения, также рассматривается возможность введения специфичной инкапсулированной SNALP посредством внутривенной инъекции в дозах, примерно, 1 или 2,5 мг/кг (см., например, Зиммерманн и др., Nature Letters, Том 441, 4 мая 2006 года). Состав SNALP может включать липиды 3-N-[(wметоксиполи(этиленгликоль) 2000) карбамоил]-1,2-димиристилокси-пропиламин (PEG-C-DMA), 1,2-дилинолейлокси-N,N-диметил-3-аминопропан (DLinDMA), 1,2 - дистеароил-sn-глицеро-3-фосфохолин (DSPC) и холестерин, в молярном процентном соотношении 2:40:10:48 (см., например, Зиммерманн и др., Nature Letters, Том 441, 2006 год). В другом варианте осуществления настоящего изобретения, стабильные частицы нуклеиновая кислота-липид (SNALP) доказали свою эффективность в доставке молекул в высоко васкуляризированные опухоли печени, полученные из линии HepG2, но не в низко васкуляризированные опухоли печени, полученные из линии HCT-116 (см., например, Ли, Gene Therapy (2012) 19, 775-780). Липосомы SNALP могут быть приготовлены путем соединения D-Lin-DMA и PEG-C-DMA с дистеароилфосфатидилхолином (DSPC), Холестерином и siРНК, используя соотношение 25:1 липиды/siРНК и молярное соотношение 48/40/10/2 Холестерин/D-Lin-DMA/DSPC/PEG-C-DMA. Полученные липосомы SNALP имеют размер, примерно, 80-100 нм.

В еще одном варианте осуществления настоящего изобретения, SNALP может содержать синтетический холестерин (Sigma-Aldrich, Сент-Луис, Миссури, США), дипальмитоилфосфатидилхолин (Avanti Polar Lipids, Алабастер, Алабама, США), 3-N-[(w-метоксиполиэтиленгликоль)2000)карбамоил]-1,2-

димирестилоксипропиламин и катионный 1,2-дилинолейлокси-3-N,Nдиметиламинопропан (см., например, Гейзберт и др., Lancet 2010; 375: 1896-905). Может быть рассмотрена дозировка, примерно, 2 мг/кг общего состава SSN/RTD/RT на дозу, вводимая, например, в виде болюсной внутривенной инфузии.

Аналогичным образом, искусственные молекулярные комплексы по настоящему изобретению могут представлять собой полезный инструмент для модификации генетического материала в клетках домашнего скота или в клетках других животных. Например, коррекция генетических заболеваний или редактирование для получения благоприятных характеристик, таких как мясо, молоко, например, молоко с пониженным содержанием лактозы, или производство яиц в животноводстве или птицеводстве.

Таким образом, в одном варианте осуществления настоящего изобретения, предложен способ получения популяции иммунных клеток животного, содержащий введение конструкции по настоящему изобретению, по меньшей мере, в одну представляющую интерес иммунную клетку, *in vivo* или *ex vivo*, для лечения заболевания, в предпочтительном варианте осуществления настоящего изобретения, аутоиммунного заболевания, например, диабета I Типа или ревматоидного артрита, или пролиферативного заболевания, такого как рак, например, глиома, меланома, нейробластома, рак толстой кишки, легких, молочной железы и предстательной железы, раковые заболевания с множественной лекарственной устойчивостью, а также раковые заболевания, связанные с мутированным геном p53.

Предпочтительными тканями большинства видов растений, образующих мишени для редактирования генома, являются незрелые зародыши, эмбриогенный каллус, меристемы интактных растений, пыльца, пыльцевые трубки или яйцеклетки, суспензионные клетки или другие типы клеток с регенеративным потенциалом. Для некоторых растений предпочтительными тканями могут быть протопласты или листья. Любая клетка, которую можно обработать и затем регенерировать в целое растение, может считаться предпочтительной тканью или клеткой. Протоколы для подготовки ткани, регенерации и доставки ДНК различаются в зависимости от вида, типа ткани, способа доставки и других факторов. Общий способ доставки представляет собой бомбардировку клеток частицами золота или вольфрама, покрытыми ДНК или белками. Другими способами доставки являются трансформация, опосредованная полиэтиленгликолем (PEG), электропорация, вирусная инфекция, инъекция непосредственно в клетки и трансформация, опосредованная *Agrobacterium.* В некоторые растения доставка может быть произведена в оплодотворенные яйцеклетки путем разрезания рыльца пестика вскоре после оплодотворения и нанесения жидкости с реагентами редактирования в разрезанную пыльцевую трубку. Для животных клеток, в предпочтительном варианте осуществления настоящего изобретения, клеток млекопитающих, электропорация, то есть, технология трансфекции, основанная на мгновенном создании малых пор в клеточных мембранах путем применения электрического импульса, может представлять собой подходящий подход для введения, по меньшей мере, одного молекулярного комплекса по настоящему изобретению. Несколько специфичных протоколов клеточного типа об успешной непосредственной трансфекции со множеством различных типов клеток, включая первичные клетки млекопитающих, стволовые клетки и труднотрансфицируемые клеточные линии, доступны специалисту в данной области техники, которые подходят в качестве инструментов доставки для, по меньшей мере, одного молекулярного комплекса по настоящему изобретению. Важно отметить, что комбинация, по меньшей мере, двух способов или агентов, подходящих для доставки, может предложить превосходные результаты в зависимости от типа клетки, геном которой подлежит редактированию и, таким образом, включена в объем настоящего изобретения.

В одном варианте осуществления настоящего изобретения, суперзаряженные белки могут быть

использованы для доставки искусственного молекулярного комплекса или его компонентов по настоящему изобретению. Суперзаряженные белки представляют собой класс сконструированных белков или белков природного происхождения с необычайно высоким положительным или отрицательным суммарным теоретическим зарядом и могут быть использованы для доставки искусственного молекулярного комплекса(ов) или его компонента(ов), или молекулы(молекул) нуклеиновой кислоты, кодирующей его. Как суперотрицательно заряженные, так и суперположительно заряженные белки проявляют удивительную способность противостоять термически или химически индуцированной агрегации. Суперположительно заряженные белки также способны проникать в клетки млекопитающих. Ассоциирование груза с этими белками, такими как плазмидная ДНК, РНК или другие белки, может обеспечить функциональную доставку этих макромолекул в клетки млекопитающих как *in vitro*, так и *in vivo*. Лаборатория Дэвида Лю сообщила о создании и характеристики суперзаряженных белков в 2007 году (Лоренс и др., 2007, Journal of the American Chemical Society 129, 10110-10112).

Невирусная доставка РНК и плазмидной ДНК представляет особый интерес для переноса искусственного молекулярного комплекса в клетки млекопитающих, и представляет ценность как для научных исследований, так и для терапевтических применений (Акинк и др., 2010, Nature Biotechnology 26, 561-569). Очищенный белок +36 GFP (или другой суперположительно заряженный белок, например, +48 GFP) смешивают с РНК в соответствующей бессывороточной среде и позволяют образоваться комплексу перед добавлением в клетки. Включение сыворотки на этой стадии ингибирует образование суперзаряженных комплексов белок-РНК и снижает эффективность лечения. Было установлено, что следующий протокол эффективен для различных клеточных линий (МакНотон и др., 2009, Proceedings of the National Academy of Sciences, США 106, 6111-6116) (Однако для оптимизации процедуры по конкретным клеточным линиям следует проводить предварительные эксперименты, варьирующие дозу белка и РНК): (1) За один день до обработки, произвести посев клеток в количестве 1 x 10^5 (например, HEK293, количество зависит от типа клеток) на лунку в 48-луночном планшете. (2) В день обработки, доводят очищенный белок +36 GFP в бессывороточной среде до окончательной концентрации 200 нМ. Добавляют РНК до окончательной концентрации 50 нМ. Перемешивают на вортексе и инкубируют при комнатной температуре в течение 10 мин. (3) Во время инкубации, аспирируют среду из клеток и промывают один раз посредством PBS. (4) После инкубации +36 GFP и РНК, добавляют комплексы белок-РНК к клеткам. (5) Инкубируют клетки с комплексами при температуре 37° С в течение 4 часов. (6) После инкубации, аспирируют среду и промывают три раза посредством PBS с добавлением 20 ед/мл гепарина. Инкубируют клетки в сывороточной среде в течение еще 48 или более часов, в зависимости от анализа на активность. (7) Анализируют клетки с помощью иммуноблота, qPCR, фенотипического анализа или другого подходящего способа.

Другим предпочтительным способом доставки искусственного молекулярного комплекса является сборка *in vitro* гибридной нуклеиновой кислоты RTDD-RT, а затем загрузка этого гибрида в полученный *in vitro* и, при необходимости, очищенный полипептид SSN перед нанесением его на интересующие клетки-мишени. Однако другими полезными способами доставки могла бы быть доставка полипептида SSN и, при необходимости, взаимодействующего домена, например мономерного стрептавидина, scFv с заданной специфичностью, или ДНК-связывающего домена, или дополнительного нуклеазного домена в виде мРНК, или в виде генетической конструкции ДНК, при необходимости, содержащей дополнительные регуляторные элементы, в, по меньшей мере, одну клетку-мишень для транскрипции и/или экспрессии *in vivo*, вместе с применением гибридной нуклеиновой кислоты одновременно, до или особенно после доставки полипептида

SSN. В случае нековалентной ассоциации RTDD с компонентом RT, эти молекулы также могут быть доставлены отдельно; в случае, если RTDD представляет собой gРНК, то gРНК может быть доставлена как РНК или как кассета экспрессии ДНК, которая может быть транскрибирована *in vivo*. В тех случаях, когда, по меньшей мере, один полипептид SSN или, по меньшей мере, одна gРНК доставляется в виде кассеты экспрессии, возможно будет предпочтительным экспрессировать их из вирусного РНК- или ДНК-репликона или вирусного вектора, в частности, когда клетка-мишень представляет собой растительную клетку.

В предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, один искусственный молекулярный комплекс ассоциирован *ex vivo*, различные компоненты комплекса, то есть, по меньшей мере, SSN, при необходимости, содержащая, по меньшей мере, один взаимодействующий домен, по меньшей мере, один RTDD и, по меньшей мере, одну RT репарационную матричную нуклеиновую кислоту, синтезируют химическим или рекомбинантным образом *ex vivo/in vitro*, а затем различные компоненты очищают, в предпочтительном варианте осуществления настоящего изобретения, перед сборкой. Дополнительный этап очистки может быть выполнен после сборки, по меньшей мере, одного искусственного молекулярного комплекса по настоящему изобретению. Способы очистки нуклеиновых кислот, включая ДНК и РНК, или полипептиды, или рибонуклео - и рибонуклеопротеиновые комплексы, легко доступны специалисту в данной области техники. Предложение высокочистого и стехиометрического молекулярного комплекса, который, при необходимости, может быть проанализирован *in vitro,* обеспечивает предложение точного инструмента редактирования генома с высокой эффективностью.

Для вариантов осуществления настоящего изобретения, основанных на молекулах не на основе нуклеиновой кислоты или аминокислоты в виде RTDD или взаимодействующих доменов, например, биотина (витамина H) или его производного, флуоресцеина или дигоксигенина, или любого другого когнатного партнера по связыванию для SSN-RTDD или для взаимодействия домена, взаимодействующего с RTDD, или для взаимодействия домена, взаимодействующего с SSN, было бы предпочтительным, чтобы RT синтезировали *ex vivo*, и затем сцепляли бы RT химическим образом с соответствующей молекулой.

В другом варианте осуществления, в соответствии с различными аспектами настоящего изобретения, в дополнение к RTDD/RT, в целях дальнейшего повышения эффективности события целенаправленного редактирования генома, может быть использована обычная репарационная матричная последовательность нуклеиновой кислоты, либо в виде плазмиды, либо в виде олигонуклеотида нуклеиновой кислоты. Как правило, решающий фактор в вопросе, применять ли плазмиду или другую матрицу для репарации двухцепочечной ДНК, или использовать ли одноцепочечный олигонуклеотид в качестве матрицы для репарации, зависит от размера предполагаемой модификации, подлежащей введению. Специалист в данной области техники может легко определить дальнейшую обычную матрицу для репарации, которая может быть использована в дополнение к гибридной конструкции нуклеиновой кислоты по настоящему изобретению. Эти обычные матрицы для репарации могут быть введены, по меньшей мере, в одну интересующую клетку-мишень вектором доставки, например геминивирусным вектором, в случае, если клетка-мишень представляет собой растительную клетку, или путем прямой трансфекции или путем введения, о котором также подробно описано в настоящем документе в отношении введения последовательности RTD/RT по настоящему изобретению.

В одном аспекте, настоящее изобретение предлагает наборы, содержащие любой, по меньшей мере, один из элементов, раскрытых в настоящем документе. В некоторых вариантах осуществления настоящего изобретения, набор содержит векторную систему, как описано в настоящем документе, и инструкции по использованию набора. Элементы могут быть предложены индивидуально или в

комбинации, и они могут быть предложены в любом подходящем контейнере, таком как флакон, бутылка или пробирка. Наборы могут включать gРНК и несвязанную протекторную цепь для стабилизации gРНК. Наборы могут включать gРНК в качестве RTDD, непосредственно взаимодействующего с представляющей интерес RT и, при необходимости, с дополнительной протекторной цепью, связанной, по меньшей мере, частично с направляющей последовательностью. Таким образом, наборы могут включать gРНК в виде частично двухцепочечной нуклеотидной последовательности. В некоторых вариантах осуществления настоящего изобретения, набор включает инструкции, по меньшей мере, на одном языке, например, более, чем на одном языке. Инструкции могут быть отдельно составлены для применений и способов, описанных в настоящем документе.

В следующем аспекте, в соответствии с настоящим изобретением, таким образом, предлагается набор, содержащий, по меньшей мере, один компонент и, в предпочтительном варианте осуществления настоящего изобретения, все компоненты, по меньшей мере, одного искусственного молекулярного комплекса настоящего изобретения, при этом, по меньшей мере, один молекулярный комплекс может быть предложен в виде предварительно собранного комплекса, или, в предпочтительном варианте осуществления настоящего изобретения, при этом, по меньшей мере, один молекулярный комплекс может быть предложен в виде его отдельных компонентов, содержащих, по меньшей мере, один полипептид SSN, или экспрессируемую последовательность, кодирующую последовательность, по меньшей мере, одну последовательность RTDD и, по меньшей мере, одну репарационную матричную последовательность нуклеиновой кислоты. Раздельное предложение различных компонентов молекулярного комплекса, в предпочтительном варианте осуществления настоящего изобретения, в виде высушенного или лиофилизированного порошка для последовательностей нуклеиновых кислот, гарантирует более высокую стабильность последовательностей нуклеиновых кислот, в частности, конструкции RTDD/RT, в частности, если последовательности РНК являются гораздо менее стабильными, чем полипептиды, предложенные в наборе. По меньшей мере, один белок SSN и, при необходимости, по меньшей мере, один взаимодействующий домен, взаимодействующий с ним или связанный с ним, могут быть доставлены в подходящем для хранения буферном растворе, например, содержащим 300 мМ NaCI, 10 мМ Трис-HCI, 0,1 мМ EDTA, 1 мМ OTT, 50% глицерина, pH 7,4 при 25° C для полипептида Cas9. Набор может дополнительно содержать подходящий рабочий буферный раствор, включающий подходящие ионы, например, Mg2+ для фермента Cas9, необходимого для активности соответствующего полипептида CRISPR.

В некоторых вариантах осуществления настоящего изобретения, набор содержит, по меньшей мере, один реагент для использования в процессе, использующем, по меньшей мере, один из элементов, описанных в настоящем документе. Реагенты могут быть предоставлены в любом подходящем контейнере. Например, набор может предложить, по меньшей мере, один рабочий буферный раствор или буферный раствор для хранения. Реагенты могут быть предложены в виде, подходящем для конкретного анализа, или в виде, требующем добавления, по меньшей мере, одного другого компонента перед использованием (например, в виде концентрата или в лиофилизированном виде). Буферный раствор может представлять собой любой буферный раствор, включая, но этим не ограничиваясь, буферный раствор карбоната натрия, буферный раствор бикарбоната натрия, боратный буферный раствор, Трис-буферный раствор, буферный раствор 3-(N-морфолин)пропансульфониевой кислоты, буферный раствор (4-(2-гидроксиэтил)-1-пиперазинэтансульфоновой кислоты) и их комбинации. В некоторых вариантах осуществления настоящего изобретения, буферный раствор является щелочным. В некоторых вариантах осуществления настоящего изобретения, буферный раствор имеет pH от, примерно, 7 до, примерно, 10. В некоторых вариантах

осуществления настоящего изобретения, набор содержит, по меньшей мере, один олигонуклеотид, соответствующий направляющей последовательности для вставки в вектор, с тем чтобы функционально связать направляющую последовательность и регуляторный элемент. В некоторых вариантах осуществления настоящего изобретения, набор содержит матричный полинуклеотид гомологичной рекомбинации. В некоторых вариантах осуществления настоящего изобретения, набор содержит, по меньшей мере, один вектор и/или, по меньшей мере, один полинуклеотид, описанный в настоящем документе. Преимущество этого комплекта состоит в том, что он может предложить все элементы систем настоящего изобретения.

В альтернативном варианте осуществления настоящего изобретения, набор может содержать компоненты SSN в виде лиофилизированной mPHK или в виде лиофилизированного белка, соответственно. В другом варианте осуществления настоящего изобретения, в соответствии с этим аспектом, набор может содержать дополнительный компонент, предлагающий подходящий носитель или систему доставки в дополнение к компоненту, содержащему компонент(ы) SSN в виде молекулярного комплекса. В следующем варианте осуществления настоящего изобретения, в соответствии с этим аспектом, по меньшей мере, один полипептид SSN и, по меньшей мере, одна последовательность RTDD/RT представлены в виде, по меньшей мере, двух компонентов, причем более одной SSN и/или RTDD/RT являются взаимно совместимыми. По меньшей мере, один полипептид SSN может быть представлен в виде вектора для трансформации или трансфекции в интересующую клетку, тогда как, по меньшей мере, одна последовательность RTDD/RT может быть представлена в виде отдельного компонента. Таким образом, набор по настоящему изобретению может подходить для одновременного или последующего использования различных компонентов в случае наличия более одного компонента. При необходимости, набор, в соответствии с этим аспектом, может содержать инструкции по использованию, в частности, инструкции по использованию, отдельно составленные для клетки-мишени, подлежащей редактированию. В другом предпочтительном варианте осуществления настоящего изобретения, в соответствии с этим аспектом настоящего изобретения, набор специально разработан для предложения набора для развития признака у конкретного интересующего растения, включая специфичные инструменты для достижения желаемой модификации признака. Согласно этому варианту осуществления настоящего изобретения, набор содержит специфичную матрицу для репарации, которая выполнена с возможностью переноса интересующего признака или лечения интересующего заболевания, или модификации интересующей последовательности ДНК-мишени в интересующий локус ДНК-мишени в клетке, в предпочтительном варианте осуществления настоящего изобретения, в клетке млекопитающего или в растительной клетке. Кроме того, набор содержит подходящий фермент SSN или две никазы SSN, ассоциированные как комплекс, по меньшей мере, с одним RTDD, при этом, RTDD содержит, по меньшей мере, одну часть первой последовательности, непосредственно взаимодействующую, по меньшей мере, с одной частью второй последовательности SSN, выполненной с возможностью непосредственного взаимодействия, по меньшей мере, с одной (RT) репарационной матричной последовательностью нуклеиновой кислоты, и, при этом, по меньшей мере, один RTDD выполнен с возможностью быть ассоциированным или, иметь возможность ассоциироваться с матрицей для репарации, несущей интересующий специфичный признак.

Набор, в соответствии с одним вариантом осуществления настоящего изобретения, представляет собой как растительную клетку, так и специфичный признак, и использование указанного набора позволяет быстро нацеливать и модифицировать интересующий локус геномной ДНК для достижения развития признака. В одном варианте осуществления настоящего изобретения, RTDD представляет собой gPHK, при

этом, компоненты gРНК уже предназначены для взаимодействия с мотивами PAM и интересующим ферментом CRISPR, а предложенная матрица для репарации представляет последовательность, подлежащую вставке или модификации удобным образом.

В одном аспекте, в соответствии с настоящим изобретением, предложено растение, растительная клетка, растительный материал или производное, или его потомство, содержащее или отредактированное, по меньшей мере, одним искусственным молекулярным комплексом по настоящему изобретению. В другом аспекте, в соответствии с настоящим изобретением, предложено растение, растительная клетка, растительный материал или производное, или его потомство, которое было модифицировано, по меньшей мере, одним искусственным молекулярным комплексом.

В еще одном аспекте по настоящему изобретению, предложен способ модификации, по меньшей мере, одной последовательности ДНК-мишени, содержащий следующие этапы: (i) предложение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки, и/или генома, содержащего, по меньшей мере, одну геномную комплементарную последовательность и, по меньшей мере, одну последовательность ДНК-мишени в представляющей интерес геномной области; (ii) предложение, по меньшей мере, одного искусственного молекулярного комплекса, как определено в настоящем документе; (iii) контактирование, по меньшей мере, одного искусственного молекулярного комплекса, по меньшей мере, с одной последовательностью ДНК-мишени в подходящих условиях для достижения (a) взаимодействия, по меньшей мере, одной сайт-специфичной нуклеазы с, по меньшей мере, одной последовательностью ДНК-мишени; и (b) комплементарного спаривания оснований, по меньшей мере, одной репарационной матричной последовательности нуклеиновой кислоты с, по меньшей мере, одной геномной комплементарной последовательностью для достижения распознавания, по меньшей мере, одной комплементарной последовательности и индукции, по меньшей мере, одного разрыва ДНК, по меньшей мере, одной сайт-специфичной нуклеазой, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты направляет гомологично направленную репарацию в сайте, по меньшей мере, одной последовательности ДНК-мишени; и (iv) получение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки, и/или генома, содержащего модификацию, по меньшей мере, одной последовательности ДНК-мишени.

Благодаря тому, что искусственный молекулярный комплекс может быть использован в интересующей клетке любого типа, можно сконструировать пару SSN/RTDD/RT для модификации любой геномной области, включая эписомальную или эпигенетическую область интересующего организма, содержащую интересующие последовательности прокариотической, эукариотической или вирусной ДНК-мишени или эпигенетические последовательности. Для вариантов осуществления настоящего изобретения, модифицирующих геном вируса, целесообразно переносить вирусный геном или его соответствующую часть в интересующий вектор и размножать, и модифицировать вирусный геном в подходящей клетке-хозяине (например, в прокариотической или эукариотической клетке), несущей вектор, содержащий вирусный геном или его соответствующую часть.

«Прокариотическая» клетка, используемая в настоящем документе, относится к одноклеточному организму, который не имеет связанного с мембраной ядра (кариона), митохондрий или любой другой связанной с мембраной органеллы и содержит археи и бактерии.

Вирусный геном может быть получен из любого вируса, содержащего геном, кодируемый РНК или ДНК.

В одном варианте осуществления настоящего изобретения, по меньшей мере, одна репарационная

матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, один стыковочный домен матрицы для репарации искусственного молекулярного комплекса предложен/предложены для, по меньшей мере, одной прокариотической или эукариотической клетки независимо от, по меньшей мере, одной сайт-специфичной нуклеазы, по меньшей мере, одного молекулярного комплекса, и, по меньшей мере, один искусственный молекулярный комплекс собирается или частично собирается, в пределах, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки и/или генома.

В одном варианте осуществления настоящего изобретения, по меньшей мере, одна последовательность RTDD/RT искусственного молекулярного комплекса предложена для, по меньшей мере, одной прокариотической или эукариотической клетки независимо от, по меньшей мере, одного полипептида SSN, по меньшей мере, одного молекулярного комплекса, и, по меньшей мере, один искусственный молекулярный комплекс собирается или частично собирается, в пределах, по меньшей мере, одной прокариотической или эукариотической клетки.

По меньшей мере, один искусственный молекулярный комплекс, как подробно описано выше, может быть предложен в виде собранного *in vitro* комплекса, который затем вводится, по меньшей мере, в одну интересующую клетку-мишень. В альтернативном варианте осуществления настоящего изобретения, некоторые или все из, по меньшей мере, одного полипептида SSN и/или, по меньшей мере, одной последовательности RTDD последовательности, и/или, по меньшей мере, одной репарационной матричной последовательности нуклеиновой кислоты могут быть вставлены как генетическая конструкция РНК или ДНК и могут быть продуцированы *in vivo*, так что окончательная сборка, по меньшей мере, одного молекулярного комплекса происходит *in vivo*. В предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, один молекулярный комплекс ассоциирован *ex vivo* и, по меньшей мере, один молекулярный комплекс, содержащий, по меньшей мере, один полипептид SSN, по меньшей мере, одну направляющую последовательность нуклеиновой кислоты и, по меньшей мере, одну репарационную матричную последовательность нуклеиновой кислоты, затем одновременно доставляется, по меньшей мере, в одну клетку посредством подходящего вектора доставки, позволяющего функциональное введение, по меньшей мере, одного молекулярного комплекса в, по меньшей мере, одну клетку-мишень, содержащую, по меньшей мере, одну представляющую интерес последовательность ДНК-мишени.

В другом предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, одна SSN и, при необходимости, по меньшей мере, один взаимодействующий домен предложены в виде белка слияния на плазмиде для продуцирования в пределах клетки, содержащей последовательность ДНК-мишени, подлежащей модификации. Дополнительные компоненты искусственного молекулярного комплекса могут быть продуцированы *ex vivo*. Например, индуцируемая векторная система может быть использована для продуцирования, по меньшей мере, одной SSN и, при необходимости, по меньшей мере, одного взаимодействующего домена. Как только достигается достаточный уровень экспрессии, комплекс RTDD/RT может быть введен в клетку-мишень, и искусственный молекулярный комплекс по настоящему изобретению может быть собран *in situ*.

В другом варианте осуществления настоящего изобретения, полный, по меньшей мере, искусственный молекулярный комплекс представляет собой собранный *ex vivo* искусственный молекулярный комплекс.

«Подходящие условия» или «подходящие условия реакции», упомянутые в настоящем документе в контексте способов по настоящему изобретению, относятся к условиям, которые позволяют как рост, так и развитие клетки или организма, включая прокариотические или эукариотические клетки,

трансформируемые или производимые, и к условиям, необходимым для достижения либо стабильной интеграции, либо транзиентного введения интересующей генетической конструкции в, по меньшей мере, одну интересующую клетку или организм. Условия, способствующие росту прокариот или бактерий, и/или трансформации, известны специалисту в данной области техники (см. также: Грин и Сэмбрук, Молекулярное клонирование, Инструкция по проведению лабораторных работ, 2012, Cold Spring Harbor Laboratory Press). Условия, способствующие росту животных клеток и/или введению генетического материала в животные клетки, в частности, млекопитающих, доступны специалисту в данной области техники в отношении различных клеточных линий (см. выше, Грин и Сэмбрук). Условия, способствующие росту и развитию растений или растительных клеток, включая, среди прочего, температуру, свет, воду, кислород, минеральные питательные вещества и состав почвы, которые могут варьироваться для различных видов растений, могут быть легко определены специалистом в данной области техники на основании изобретения, предложенного в настоящем документе. Дальнейшие подходящие условия для достижения стабильной интеграции или транзиентного введения, по меньшей мере, одного интересующего молекулярного комплекса зависят от способа трансформации, выбранного для введения, по меньшей мере, одного интересующего молекулярного комплекса, стадии развития растительного материала или растительной клетки, подлежащей трансформации, и, по меньшей мере, одного интересующего молекулярного комплекса, подлежащего введению. Указанные подходящие условия могут быть определены специалистом в данной области техники в свете настоящего изобретения, определяющего подходящие условия для способов в комбинации с примерными молекулярными комплексами и подходящими векторами доставки, и методами доставки, раскрытыми и заявленными в настоящем документе.

В одном варианте осуществления настоящего изобретения в соответствии с вышеуказанным способом настоящего изобретения, по меньшей мере, одна эукариотическая клетка представляет собой растительную клетку, в предпочтительном варианте осуществления настоящего изобретения, растительную клетку из растения, выбранного из группы, состоящей из *Hordeum vulgare, Hordeum bulbusom, Sorghum bicolor, Saccharum officinarium, Zea spp.,* включая *Zea mays, Setaria italica, Oryza minuta, Oryza sativa, Oryza australiensis, Oryza alta, Triticum aestivum, Triticum durum, Secale cereale,* Triticale, *Malus domestica, Brachypodium distachyon, Hordeum marinum, Aegilops tauschii, Daucus glochidiatus, Beta spp., включая Beta vulgaris, Daucus pusillus, Daucus muricatus, Daucus carota, Eucalyptus grandis, Nicotiana sylvestris, Nicotiana tomentosiformis, Nicotiana tabacum, Nicotiana benthamiana, Solanum lycopersicum, Solanum tuberosum, Coffea canephora, Vitis vinifera, Erythrante guttata, Genlisea aurea, Cucumis sativus, Marus notabilis, Arabidopsis arenosa, Arabidopsis lyrata, Arabidopsis thaliana, Crucihimalaya himalaica, Crucihimalaya wallichii, Cardamine nexuosa, Lepidium virginicum, Capsella bursa pastoris, Olmarabidopsis pumila, Arabis hirsute, Brassica napus, Brassica oleracea, Brassica rapa, Raphanus sativus, Brassica juncacea, Brassica nigra, Eruca vesicaria subsp. sativa, Citrus sinensis, Jatropha curcas, Populus trichocarpa, Medicago truncatula, Cicer yamashitae, Cicer bijugum, Cicer arietinum, Cicer reticulatum, Cicer judaicum, Cajanus cajanifolius, Cajanus scarabaeoides, Phaseolus vulgaris, Glycine max, Gossypium sp., Astragalus sinicus, Lotus japonicas, Torenia fournieri, Allium cepa, Allium fistulosum, Allium sativum, Helianthus annuus, Helianthus tuberosus* и *Allium tuberosum* или из любого сорта или подвида, принадлежащий одному из вышеупомянутых растений.

В отношении растительных клеток, рассматриваемых в качестве мишеней, например, специалисту в данной области техники доступны различные способы трансформации и/или трансфекции. Для протопластов маиса, например, подходящий способ раскрыт в работе Шина, Дж. 2002 («Анализ транзиентной экспрессии с использованием протопластов мезофилла кукурузы»). Для протопластов

Arabidopsis, соответствующий протокол доступен из: идентификатор цифрового объекта.org/10.1038/ngrot.2007.199 или его можно получить по адресу: http://www.nature.com/nprot/journal/v2/n7/full/nprot.2007.199.html. Для протопластов табака и других двудольных растений подходящий протокол доступен по адресу: www.plantphysiol.org/cgi/doi/10.1104/pp.112.205179. Специалист в данной области техники, обладающий знаниями, предоставляемыми настоящим изобретением, и осведомленный об упомянутых протоколах, может, таким образом, определить подходящий способ введения молекулярного комплекса по настоящему изобретению в протопласт растения, полученный из однодольного или двудольного растения.

Протопласты очень подходят для тестирования технологий редактирования генов и реагентов, но для регенерации растений, отредактированных генами, они не всегда являются предпочтительным типом клеток, так как имеется очень мало видов растений, которые эффективно регенерируют из протопласта. В этом случае, предпочтительными тканями для большинства видов растений являются незрелые зародыши, эмбриогенный каллус, оплодотворенные зародыши, меристемы интактных растений, пыльца, пыльцевая трубка или яйцеклетки, эмбриогенные суспензионные клетки или другие типы клеток с регенеративным потенциалом. Обычный способ физической доставки представляет собой бомбардировку клеток частицами золота или вольфрама, покрытыми ДНК или белком, в то время как в обычном биологическом способе используют *Agrobacterium* или (модифицированный) вирусный вектор, как описано в настоящем документе.

«Меристематическая(ие) клетка(и)» по настоящему изобретению относится к такому типу ткани растения, который также упоминается как меристема или камбий, или образовательная ткань. Подобно стволовым клеткам животных организмов, меристематические клетки растений, представляющие собой недифференцированные клетки, обладают внутренней способностью развиваться и дифференцироваться в специализированные типы клеток, в зависимости от генетической предрасположенности и других факторов окружающей среды и развития. В растительных организмах, меристемы присутствуют не только во время развития зародыша, но они также могут быть обнаружены в течение всего жизненного цикла растения, таким образом, целенаправленная генетическая модификация меристематических клеток или тканей по настоящему изобретению не ограничивается зародышами растений или проростками, но она также может проводиться на более крупных проростках и более зрелых растениях, например, при нацеливании на меристемы, которые создают основу для репродуктивных органов растения, например, метелки или колос у маиса.

Согласно одному варианту осуществления настоящего изобретения, в соответствии с различными аспектами настоящего изобретения, меристематическая клетка может представлять собой зрелую или незрелую растительную клетку зародыша растения или проростка растения, содержащую, по меньшей мере, одну меристематическую клетку или меристематическую ткань.

Для некоторых подходов к редактированию генома может быть желаема стабильная интеграция молекулярного комплекса, кодирующего экспрессионную кассету(ы), где трансгенный организм, несущий желаемую интересующую конструкцию или ее часть, может передавать по наследству стабильно вставленную конструкцию потомству интересующей растительной клетки, изначально трансформированной или трансфицированной. Упомянутая стабильная интеграция может происходить в любой геномной области организма, в предпочтительном варианте осуществления настоящего изобретения, в эукариотическом организме, включая ядерный геном, а также внеядерный геном, включая геном пластид.

Транзиентное введение может быть желаемым, если желаем определенный эффект при введении представляющего интерес молекулярного комплекса или его части, но конструкция *per se* не должна

изначально передаваться по наследству потомству клетки. Из-за регуляторных причин, такой подход может быть особенно подходящим для определенных применений, в частности, с растительными клетками, тканями, органами или материалом в качестве структуры, содержащей последовательность ДНК-мишени, подлежащую модификации.

Используемый в настоящем документе термин «целенаправленная интеграция» или «функциональная интеграция» относится к интеграции интересующей генетической конструкции, по меньшей мере, в одну клетку, которая обеспечивает транскрипцию и/или трансляцию, и/или каталитическую активность, и/или связывающую активность, включая связывание молекулы нуклеиновой кислоты с другой молекулой нуклеиновой кислоты, включая ДНК или РНК, или связывание белка с целевой структурой, по меньшей мере, в одной клетке. Там, где это уместно, функциональная интеграция происходит в определенном клеточном компартменте, по меньшей мере, одной клетки, включая ядро, цитозоль, митохондрий, хлоропласт, вакуоль, мембрану, клеточную стенку и тому подобное. Следовательно, термин «функциональная интеграция», в отличие от термина «стабильная интеграция», подробно описанного выше, подразумевает, что представляющий интерес молекулярный комплекс вводится в, по меньшей мере, одну клетку посредством любых способов трансформации, трансфекции или трансдукции с помощью биологических способов, включая трансформацию *Agrobacterium*, или физических способов, включая бомбардировку частицами, а также последующим этапом, при этом, молекулярный комплекс оказывает свое действие в пределах или на, по меньшей мере, одной клетке, в которую он был введен. В зависимости от природы генетической конструкции, подлежащей введению, указанное действие, естественно, может варьироваться и включать, по отдельности или в комбинации, среди прочего, транскрипцию ДНК, закодированную генетической конструкцией, в рибонуклеиновую кислоту, трансляцию РНК в аминокислотную последовательность, активность молекулы РНК в пределах клетки, содержащую активность направляющей РНК или miРНК, или siРНК для использования в интерференции РНК, и/или связывающую активность, включая связывание молекулы нуклеиновой кислоты с другой молекулой нуклеиновой кислоты, включая ДНК или РНК, или связывание белка со структурой-мишенью, в, по меньшей мере, одной клетке, или включать интеграцию последовательности, доставляемой посредством вектора или генетической конструкции, либо транзиентным, либо стабильным образом. Указанное действие может также включать каталитическую активность аминокислотной последовательности, представляющей собой фермент или его каталитически активную часть в пределах, по меньшей мере, одной клетки и тому подобное. Указанное действие, достигаемое после функциональной интеграции молекулярного комплекса по настоящему изобретению, может зависеть от присутствия регуляторных последовательностей или локализационных последовательностей, которые составляют интересующую генетическую конструкцию, как это известно специалисту в данной области техники.

Как подробно описано выше, способы по настоящему изобретению, нацеленные на плюрипотентные или мультипотентные клетки, предлагают преимущество, заключающееся в том, что как трансформация, так и дальнейшее развитие трансформированной, по меньшей мере, одной клетки, в частности, меристематической клетки, могут происходить *in planta*, что устраняет необходимость в трудоемких этапах культивирования *in vitro* для регенерации растения или его растительного материала. Однако в определенных вариантах осуществления настоящего изобретения, может быть подходящим эксплантировать или иссекать растительную клетку, ткань, орган или материал для дальнейшего культивирования, скрининга или тестирования, в зависимости от конкретных потребностей. Специалистам в данной области техники доступно несколько способов культивирования *in vitro* растительной клетки, ткани,

органа или материала.

Таким образом, может быть желаема стабильная интеграция, когда трансгенное растение, несущее желаемую интересующую конструкцию или ее часть, стабильно встраивается, и вставленная конструкция или ее часть передается по наследству потомству интересующей изначально трансформированной клетки растения. Упомянутая стабильная интеграция может происходить в любой геномной области растения, включая ядерный геном, а также внеядерный геном, включая геном пластид растительной клетки. Кроме того, искусственные молекулярные комплексы по настоящему изобретению могут быть использованы для создания эпигенетической модификации. В другом аспекте, настоящее изобретение предлагает способ функциональной оценки и скрининга генов. Таким образом, искусственный молекулярный комплекс по настоящему изобретению может использоваться для точной доставки функциональных доменов, для активации или репрессии генов, или для изменения эпигенетического состояния путем точного изменения сайта метилирования на конкретном интересующем локусе. Способ по настоящему изобретению может быть использован для создания растения, животного или клетки, которые можно использовать для моделирования и/или изучения интересующих генетических, или эпигенетических условий, например, посредством модели интересующих мутаций или модели заболевания. Используемый в настоящем документе термин «заболевание» относится к заболеванию, расстройству или показанию у субъекта. Например, способ по настоящему изобретению может быть использован для создания животного или клетки, которая содержит модификацию в, по меньшей мере, одной последовательности нуклеиновой кислоты, ассоциированной с заболеванием, или растения, животного или клетки, в которой экспрессия, по меньшей мере, одной последовательности нуклеиновой кислоты, ассоциированной с заболеванием, изменена. Такая последовательность нуклеиновой кислоты может кодировать белковую последовательность, ассоциированную с заболеванием, или может представлять собой контрольную последовательность, ассоциированную с заболеванием. Соответственно, следует понимать, что в вариантах осуществления настоящего изобретения растение, субъект, пациент, организм или клетка могут представлять собой нечеловеческое существо, пациента, организм или клетку. Таким образом, настоящее изобретение относится к растению, животному или клетке, продуцируемой настоящими способами, или к их потомству. Потомство может быть клоном продуцированного растения или животного, или может быть результатом полового размножения путем скрещивания с другими особями того же вида для интрогрессии дальнейших желаемых признаков их потомству. Клетка может быть предложена *in vivo* или *ex vivo* в случаях многоклеточных организмов, в частности, животных или растений. В случае, когда клетка культивируется, может быть установлена клеточная линия, если соблюдаются соответствующие условия культивирования, и, в предпочтительном варианте осуществления настоящего изобретения, если клетка соответствующим образом адаптирована к этой цели (например, стволовая клетка). Бактериальные клеточные линии, продуцируемые по настоящему изобретению, также предусмотрены. Следовательно, клеточные линии также предусмотрены.

Транзиентное введение может быть желаемым в том случае, если, желаем определенный эффект, например эффект молчания, целевая манипуляция, содержащая нокин или нокаут, путем введении представляющей интерес генетической конструкции или ее части, но конструкция *per se* не должна передаваться по наследству потомству клетки, изначально трансформированной.

В еще одном варианте осуществления вышеуказанного аспекта в соответствии с настоящим изобретением, введение, по меньшей мере, одного представляющего интерес молекулярного комплекса или его частей, включая гРНК и/или RT, проводят с использованием средства, выбранного из группы, состоящей

из устройства, подходящего для бомбардировки частицами, включая генную пушку, включая ручную генную пушку (например, Helios® Gene Gun System, BIO-RAD) или стационарную генную пушку, трансформации, включая трансформацию с использованием *Agrobacterium spp.* или с использованием вирусного вектора, микроинъекции, электропорации, технологии вискеров, включая технологию вискеров карбида кремния, и химического вещества, например, с использованием фосфата кальция, дендримеров, липосом или катионных полимеров, и нехимического вещества, например, с использованием электропорации, сонопорации, оптической трансфекции с использованием лазера, слияния протопластов, импалфекции, гидродинамической доставки генов ДНК путем впрыска доставляемой конструкции в орган, в предпочтительном варианте осуществления настоящего изобретения, в печень животного, в предпочтительном варианте осуществления настоящего изобретения, грызуна, трансфекции или их комбинации.

В определенных вариантах осуществления настоящего изобретения, по меньшей мере, одна эукариотическая клетка представляет собой меристематическую растительную клетку, и растительную клетку после введения искусственного молекулярного комплекса по настоящему изобретению далее культивируют в подходящих условиях до достижения стадии развития зрелости соцветия, чтобы получить растение или растительный материал, содержащий интересующую модификацию, опосредованную, по меньшей мере, одним молекулярным комплексом по настоящему изобретению. Специалисту в данной области техники доступны, например, несколько протоколов для продуцирования прорастающей и жизнеспособной пыльцы из культивируемых *in vitro* метелок маиса, например, в работе Пэредди Д.Р. и др. (1992) Созревание пыльцы маиса in vitro. Plant Cell Rep 11 (10):535-539. идентификатор цифрового объекта:10.1007/BF00236273, Стэплтона А.Е. и др. (1992) Незрелые вторичные колоски кукурузы развиваются и продуцируют пыльцу в культуре. Plant Cell Rep 11 (5-6):248-252 или Пэредди Д.Р. и др. (1989) Получение нормальной, прорастающей и жизнеспособной пыльцы из культивированных *in vitro* метелок маиса. Theor Appl Genet 77 (4):521-526. Эти протоколы, среди прочего, основаны на эксцизии метелки, поверхностной стерилизации и культивировании в среде с кинетином для стимулирования роста и созревания метелки. После формирования вторичных колосков можно проводить непрерывный сбор пыльников. После экструзии пыльники высушивают до тех пор, пока не выйдет пыльца. В альтернативном варианте осуществления настоящего изобретения, пыльники могут быть иссечены, и пыльца ссыпается в жидкую среду, которая впоследствии используется для опыления колосовых цветорасположений.

Используемый в настоящем документе термин «зрелость соцветия» относится к состоянию, когда незрелое соцветие растения, содержащее, по меньшей мере, одну меристематическую клетку, достигло стадии развития, когда достигнуто зрелое соцветие, то есть, тычинковое соцветие (мужское) или пестиковое соцветие (женское), и, таким образом, присутствует гамета пыльцы (мужская) или семяпочки (женская), или обе гаметы. Указанная стадия репродуктивной фазы растения особенно важна, поскольку полученный растительный материал может быть непосредственно использован для опыления другого растения или для оплодотворения пыльцой другого растения.

В следующем варианте осуществления настоящего изобретения, модификация, по меньшей мере, одной последовательности ДНК-мишени представляет собой подход к редактированию генома, выбранный из группы, состоящей из повышения урожайности, толерантности к абиотическому стрессу, включая стресс, вызванный засухой, осмотический стресс, тепловой стресс, стресс от холода, оксидативный стресс, стресс под действием тяжёлых металлов, стресс, вызванный засолением или переувлажнением, толерантности к биотическому стрессу, включая толерантность к насекомым, толерантность к бактериям, толерантность к

вирусам, толерантность к грибам или толерантность к нематодам, устойчивости к гербицидам, включая глифосат, глуфосинат, ингибиторы ацетолактатсинтазы (ALS) и дикамбу, устойчивости к полеганию, времени цветения, устойчивости к осыпанию, цвета семян, состава эндосперма, содержания питательных веществ, модификации фенотипического маркера или метаболической инженерии, включая редактирование генома, для обеспечения подхода к молекулярному фармингу, по меньшей мере, в одной растительной клетке. Фенотипические маркеры могут быть предпочтительными целями для подходов совместного редактирования, например, для мониторинга эффективности редактирования.

В другом варианте осуществления настоящего изобретения, развитие признака осуществляют для прокариотической клетки или вирусного генома, например, для предложения прокариотической клетки модифицированным подходящим образом метаболическим путем для продуцирования интересующего продукта или для предложения аттенюированного вирусного генома.

В другом варианте осуществления настоящего изобретения, в соответствии с вышеуказанным способом настоящего изобретения, модификация, по меньшей мере, одной последовательности ДНК-мишени представляет собой подход к редактированию генома для модификации *ex vivo* иммунной клетки, по меньшей мере, в одной эукариотической клетке, в предпочтительном варианте осуществления настоящего изобретения, в клетке млекопитающего, в предпочтительном варианте осуществления настоящего изобретения, в лейкоците млекопитающего, для получения модифицированной клетки, подходящей для лечения вирусного заболевания или для иммунотерапии, особенно иммунотерапии рака.

В одном предпочтительном варианте осуществления настоящего изобретения, вышеуказанный способ по настоящему изобретению представляет собой способ модификации эукариотической клетки, в предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, одной растительной клетки, целенаправленным образом для предложения генетически модифицированного, в предпочтительном варианте осуществления настоящего изобретения, нетрансгенного растения, при этом, способ может, среди прочего, представлять собой способ для развития признака. Например, может быть введено высоко сайт-специфичное замещение 1,2, 3 или более нуклеотидов в кодирующей последовательности гена растения, чтобы продуцировать замещения, по меньшей мере, одной аминокислоты, которая придает толерантность, по меньшей мере, одному гербициду, такому как глифосат, глуфосинат, дикамба или гербицид, ингибирующему ацетолактатсинтазу (ALS). Кроме того, в другом варианте осуществления настоящего изобретения, замещения, по меньшей мере, одной аминокислоты в кодирующей последовательности гена растения, имеющего сайт связывания нуклеотидов с лейцин-богатыми повторами (NBS-LRR), изменят спектр распознавания патогена белка для оптимизации устойчивости растений к болезням. В еще одном дополнительном варианте осуществления настоящего изобретения, небольшая энхансерная последовательность или сайт связывания факторов транскрипции могут быть модифицированы в эндогенном промоторе гена растения или могут быть введены в промотор гена растения, чтобы изменить профиль экспрессии или силу гена растения, регулируемую промоутером. Профиль экспрессии может быть изменен посредством различных модификаций, введений или делеций в других областях, таких как интроны, 3'-нетранслируемые области, *цис-* или транс-энхансерные последовательности. В еще одном дополнительном варианте осуществления настоящего изобретения, геном клетки растения, в предпочтительном варианте осуществления настоящего изобретения, меристематической клетки растения, можно модифицировать таким образом, чтобы растение, полученное из модифицированной меристематической клетки, могло продуцировать химическое вещество или соединение, представляющее интерес для агрономической или фармацевтической индустрии, например, инсулин или аналоги инсулина,

антитела, белок с интересующей ферментной функцией, или любое другое фармацевтически значимое соединение, подходящее в качестве лекарственного средства, в качестве пищевой добавки или в качестве продукта здравоохранения.

В дополнительном аспекте, редактирование признака в соответствии со способами по настоящему изобретению предлагает способ редактирования признака для достижения лечения заболевания и/или состояния, и/или предотвращения заражения насекомыми/инвазии растения, содержащего модифицирующий хромосомный или внехромосомный генетический материал указанного растения посредством использования любого из вышеуказанных способов. Неограничивающие примеры заболеваний и/или состояний, поддающихся лечению с помощью изобретенных способов, включают антракноз – стеблевую гниль, аспергиллез – сухую гниль колосовых цветорасположений, сухую гниль колосовых цветорасположений кукурузы обыкновенной, сухую гниль колосовых цветорасположений кукурузы, бурую ржавчину кукурузы, диплодиоз – сухую гниль колосовых цветорасположений, диплодиоз – полосатость листьев, диплодиоз – стеблевую гниль, ложную мучнистую росу, глазковую пятнистость, фузариоз – сухую гниль колосовых цветорасположений, фузариоз – стеблевую гниль, гиббереллу – сухую гниль колосовых цветорасположений, гиббереллу – стеблевую гниль, бактериальное увядание Госса и пятнистость листьев, серую пятнистость листьев, головню метелок, гельминтоспориозную пятнистость листьев кукурузы, физодермоз кукурузы, питиозную корневую гниль, южную пятнистость листьев кукурузы, южную ржавчину и бактериальное увядание Стюарта и пятнистость, и их комбинации.

Неограничивающие примеры насекомых, вызывающих, прямо или косвенно, заболевания и/или состояния, поддающиеся лечению с помощью изобретенных способов, включают совку луговую, хрущика азиатского садового, совку-ипсилон, коричневого мраморного щитника, коричневого щитника, огневку обыкновенную, долгоносика кукурузного, совку кукурузную, тлю кукурузную листовую, злакового корневого червя, злакового корневого червя, питающегося кистью нитей рыльца, огневку кукурузную европейскую, совку травяную, жука Colaspis виноградного, огневку хмелевой лозы, хрущика японского, осмотр-поиск совки травяной, зерновика, личинку мухи ростковой, южного кукурузного листоеда, огневку кукурузную юго-восточную, клещика паутинного, жука-носорога сахарного тростника, западную фасолевую совку, личинку майского хруща и проволочников, и их комбинации. Изобретенные способы также подходят для предотвращения заражений и/или инвазий растения любым таким насекомым(и).

Неограничивающими примерами признаков, которые могут быть введены этим способом, являются устойчивость или толерантость к насекомым-вредителям, таким как, личинки, повреждающие корни, стеблевые точильщики, совки, жуки, тля, кобылочки, долгоносики, клещи и щитники. Это можно было бы сделать путем модификации генов растений, например, для увеличения внутренней устойчивости растения к насекомым-вредителям или для снижения его привлекательности для указанных вредителей. Другими признаками могут быть устойчивость или толерантность к нематодам, бактериальным, грибковым или вирусным патогенам, или к их переносчикам. Другими признаками могли бы быть более эффективное использование питательных веществ, например, более интенсивное использование азота, повышение или внедрение эффективности в фиксации азота, боле интенсивная эффективность фотосинтеза, например, преобразование С3 растения в С4. А также, другими признаками могли бы быть повышение толерантности к абиотическим стрессорам, таким как, температура, подача воды, соленость, pH, толерантность к экстремальным воздействиям солнечного света, эффективность использования азота, эффективность использования фосфора, эффективность использования воды и урожайность сельскохозяйственных культур или биомассы. Дополнительными признаками могли бы быть характеристики, относящиеся к вкусу,

внешнему виду, профилям питательных веществ или витаминов съедобных или обеспечивающих питание участков растения, или могут быть связаны с длительностью хранения или качеством этих участков. Наконец, признаки могут быть связаны с агрономическими качествами, такими как, устойчивость к полеганию, осыпание, время цветения, созревание, прорастание, сбор урожая, структура растения, сила роста, размер, урожайность и другие характеристики. Для достижения вышеуказанной модификации признака, способ по настоящему изобретению содержит модификацию хромосомного или внехромосомного генетического материала растения или растительной клетки с использованием любого из вышеупомянутых способов.

В одном варианте осуществления настоящего изобретения в соответствии с вышеуказанным способом по настоящему изобретению, клетка-мишень представляет собой прокариотическую клетку, и модификация содержит, по меньшей мере, одну модификацию интересующей геномной области-мишени, по меньшей мере, одной прокариотической клетки, при этом, модификация подходит для модуляции или увеличения устойчивости бактерии к биотическому или абиотическому стрессу, включая устойчивость к антибиотикам, или, при этом, модификация подходит для улучшения устойчивости бактериофагов, по меньшей мере, одной прокариотической клетки. В другом варианте осуществления настоящего изобретения, модификация содержит вставку интересующего гена в сайт-мишень ДНК, по меньшей мере, одной интересующей прокариотической клетки, например, вставку последовательности, кодирующей флуоресцентный маркерный белок или другой селектируемый маркер в, по меньшей мере, один интересующий сайт-мишень ДНК. В другом варианте осуществления настоящего изобретения, модификация содержит нокаут, то есть, удаление, по меньшей мере, одного интересующего сайта-мишени ДНК в, по меньшей мере, одной прокариотической клетке. Поскольку прокариотические клетки не будут далее дифференцироваться, но могут непосредственно передавать по наследству, по меньшей мере, одну интересующую вводимую модификацию своему потомству, и поскольку прокариотические клетки обычно имеют очень короткое время генерации, по сравнению с эукариотическими клетками, то модификация, вводимая, по меньшей мере, одним комплексом RTDD/RT в виде, по меньшей мере, одного искусственного молекулярного комплекса по настоящему изобретению, может быть выполнена быстро, и полученная в результате популяция модифицированных клеток может быть получена и проанализирована за очень короткий период времени.

В определенных вариантах осуществления настоящего изобретения, вышеупомянутый способ по настоящему изобретению может дополнительно содержать следующий этап: (v) идентификацию и/или выбор, по меньшей мере, одной прокариотической или эукариотической клетки, содержащей модификацию в, по меньшей мере, одной последовательности ДНК-мишени, или идентификацию модификации в отношении вирусного генома, размноженного в прокариотической или эукариотической клетке.

Способы анализа или идентификации модификации по настоящему изобретению, осуществленной в геноме, по меньшей мере, одной прокариотической или эукариотической клетки, или вирусного генома, известны специалисту в данной области техники и содержат, но этим не ограничиваются, полимеразную цепную реакцию (PCR), включая, среди прочего, количественную PCR в реальном времени, мультиплексную PCR, RT-PCR, вложенную PCR, аналитическую PCR и тому подобное, микроскопию, включая микроскопию яркого и темного поля, дисперсионное окрашивание, фазовый контраст, флуоресцентный, конфокальный, дифференциальный интерференционный контраст, деконволюцию, электронную микроскопию, УФ-микроскопию, ИК-микроскопию, сканирующую зондовую микроскопию, анализ метаболита клетки, анализ измененного спектра устойчивости модифицированной клетки, анализ

РНК, протеомный анализ, функциональные анализы для определения функциональной интеграции, например, интересующего маркерного гена или трансгена, или нокаута, блоттинг по Саузерну, секвенирование, включая глубокое секвенирование и его комбинации. Затем клетки, содержащие желаемую модификацию, можно выбрать для дальнейшего культивирования или любого другого последующего этапа изготовления.

В еще одном аспекте в соответствии с настоящим изобретением, предложен способ получения растения или растительной клетки, содержащий следующие этапы: (i) осуществление способа модификации, по меньшей мере, одной последовательности ДНК-мишени в эукариотической клетке, как подробно описано выше, при этом, по меньшей мере, одна эукариотическая клетка представляет собой растительную клетку; (ii) получение, по меньшей мере, одного растения или его потомства из, по меньшей мере, одной растительной клетки из этапа (i); и (iii) при необходимости: определение модификации в, по меньшей мере, одной последовательности ДНК-мишени в, по меньшей мере, одной клетке, по меньшей мере, одного растения или его потомства.

Подходящие растительные клетки, ткани, органы и материалы для выполнения этого аспекта подробно описаны выше. Термин «изготовление» по настоящему изобретению следует толковать широко, и он содержит любую форму генетической манипуляции, выполняемой с генетическим материалом растения или растительной клетки. Предложение, по меньшей мере, одного искусственного молекулярного комплекса, содержащего, по меньшей мере, одну последовательность RTDD/PT, содержащую, по меньшей мере, один стыковочный домен матрицы для репарации и, по меньшей мере, одну репарационную матричную нуклеиновую кислоту и, по меньшей мере, один полипептид SSN, при необходимости, содержащий взаимодействующий домен, может иметь место таким образом, чтобы позволить транзиентное действие или стабильную интеграцию, или их сочетание, сочетание различных компонентов, как описано выше. В предпочтительном варианте осуществления настоящего изобретения, по меньшей мере, один искусственный молекулярный комплекс или его различные компоненты предложены транзиентным образом так, чтобы не происходило интеграции какого-либо из этих эффекторных компонентов как таковой, включая последовательность, кодирующую направляющую нуклеиновую кислоту РНК, последовательность, кодирующую репарационную матричную нуклеиновую кислоту ДНК, и последовательность, кодирующую полипептид CRISPR, в геном интересующей клетки-мишени.

В одном варианте осуществления настоящего изобретения по вышеуказанному способу изготовления, в соответствии с настоящим изобретением, по меньшей мере, одно растение или растительная клетка выбраны из однодольного или двудольного растения, в предпочтительном варианте осуществления настоящего изобретения, при этом, растение выбирают из группы, состоящей из *Zea spp.*, включая *Zea mays, Nicotiana benthamiana,* или *Beta spp,* включая *Beta vulgaris,* или *Secale ssp.,* включая *Secale cereal,* или *Triticum ssp., включая Triticum aestivum.*

Как подробно описано в настоящем изобретении, способы по настоящему изобретению являются подходящими и могут быть адаптированы к клеткам-мишеням, принадлежащим ко всем царствам живых организмов, поскольку основной смысл использования конструкции RTDD/RT, которая ассоциирована функциональным образом в комбинации с подходящей сайт-специфичной нуклеазой, взаимодействующей с RTDD, заключается в биологических видах и независимости от клеток, при условии наличия механизма гомологичной рекомбинации для репарации ДНК в клетке, который, однако, продиктован ковалентным или нековалентным взаимодействием, по меньшей мере, на gРНК и, по меньшей мере, на одной RT. То, что необходимо определить отдельно для каждой клетки-мишени и каждой мишени – это: (i) сайт-специфичная

нуклеаза или ее каталитически активный фрагмент, и может ли быть подходящим использование взаимодействующего домена, например, в качестве белка слияния; (ii) подходящая пара взаимодействующих доменов RTDD-SSN или RTDD, позволяющая осуществить непосредственное взаимодействие указанных компонентов путем распознавания когнатных партнеров по связыванию; и (iii) подходящая RT и ее связь с RTDD, при этом, конструкция RT имеет отношение к введению выполненной по индивидуальным параметрам репарации в интересующей последовательности ДНК-мишени, расщепляемой, по меньшей мере, одной SSN искусственного молекулярного комплекса и при необходимости (для нуклеаз CRISPR); (iv) полипептид gРНК и CRISPR, которые должны быть совместимы, как подробно описано выше; (v) соответствие представляющей интерес gРНК сайту PAM в пределах интересующей области ДНК-мишени; и (vi) последовательность ДНК-мишени и целенаправленная модификация, подлежащая введению. Для любого общедоступного секвенированного генома, конструкция подходящих последовательностей нуклеиновых кислот, таким образом, может быть выполнена *in silico* на основе раскрытия настоящего изобретения.

В еще одном аспекте в соответствии с настоящим изобретением, предложено использование, по меньшей мере, одной последовательности RTDD/RT по настоящему изобретению или использование искусственного молекулярного комплекса по настоящему изобретению для редактирования генома в прокариотической или эукариотической клетке. В одном варианте осуществления этого аспекта настоящего изобретения, использование относится к эукариотической клетке, в предпочтительном варианте осуществления настоящего изобретения, к грибковой, животной или растительной клетке, или организму, или к вирусному организму, размножаемому в прокариотической или эукариотической клетке.

Согласно различным аспектам и вариантам осуществления настоящего изобретения, эукариотическая клетка или способ, или использование для модификации эукариотических клеток, включая стволовые клетки, явно не включает какой-либо процесс клонирования человека, процесс модификации генетической идентичности зародышевой линии человека или использование человеческих эмбрионов, или способ, требующий уничтожения человеческих эмбрионов для получения клеток из них. В частности, клетки зародышевой линии человека или человеческие эмбрионы, таким образом, специально исключаются в качестве клеток-мишеней или организмов, подлежащих модификации искусственными молекулярными комплексами или способами по настоящему изобретению.

Настоящее изобретение дополнительно описано со ссылкой на следующие неограничивающие примеры.

Примеры

Настоящее изобретение дополнительно проиллюстрировано следующими неограничивающими примерами.

Пример 1: Гибридная последовательность нуклеиновой кислоты в виде пары RTD/RT, подходящей для комбинирования с Cas или Cpf 1, или полипептидом Аргонавтом

В одном эксперименте, имеющая хвост sgРНК или sgДНК гибридизуются как посредством комплементарного спаривания оснований, так и посредством лигирования РНК-ДНК или ДНК-ДНК с одноцепочечной матрицей для репарации. Для ковалентной ассоциации, синтезированные олигонуклеотиды ДНК ковалентно лигируют с 3'-концом олигонуклеотидов РНК/ДНК с использованием протокола изготовителя ssРНК-лигазы. Для нековалентной ассоциации, олигонуклеотиды РНК/ДНК и ДНК с частично комплементарной последовательностью смешивают и позволяют образоваться комплексу посредством спаривания оснований по Уотсону-Крику. Успешная гибридизация может быть установлена посредством

анализов задержки электрофоретического сдвига в геле. Обработка аликвот гибридной нуклеиновой кислоты ферментами РНКазы и ДНКазы перед анализом задержки электрофоретического сдвига в геле показывает, что часть гибридной нуклеиновой кислоты состоит из РНК, причем, некоторые ДНК для этих экспериментов используют sgРНК. Затем гибрид нуклеиновой кислоты связывается в комплекс с рекомбинантным белком Cas9 или другой нуклеазой, полученной из CRISPR или Аргонавта. Успешное связывание в комплекс может быть проверено посредством обработки протеиназой K, РНКазой, ДНКазой и контрольной обработки, а также наблюдением соответствующих шаблонов задержки электрофоретического сдвига в геле. Рекомбинантные полипептиды Cas были продуцированы и впоследствии очищены либо силами внешней коммерческой организации, либо за счет своих внутренних ресурсов. Различные архитектуры гибридных последовательностей нуклеиновых кислот между направляющей последовательностью нуклеиновой кислоты в виде RTDD и тестируемой (RT) репарационной матричной последовательностью нуклеиновой кислоты показаны на **Фигурах 1** и **2**.

Пример 2: Разрезание *in vitro* ДНК-мишени посредством комплекса белка Cas9 с гибридной нуклеиновой кислотой РНК-ДНК

В одном эксперименте, была протестирована функциональная возможность белка Cas в виде сайт-специфичной эндонуклеазы с использованием описанной технологии гибридной нуклеиновой кислоты. Линеаризованную плазмиду, содержащую, по меньшей мере, один сайт-мишень для sgРНК, смешивали с комплексом Cas9-sgРНК-RT, как описано в настоящем изобретении. После инкубации в условиях, подходящих для нуклеазной активности, включая правильный pH, температуру и кофакторы, и тому подобное, которые известны специалисту в данной области техники для различных нуклеаз CRISPR и их вариантов, плазмиду ДНК-мишени запускали в агарозном геле и наблюдали за размерами полос, указывающих на разрезание ожидаемого сайта-мишени. Расщепление *in vitro* ДНК-мишени указывало на то, что RT, ассоциированная с sgРНК в виде «груза», не вмешивалась в нормальную функцию комплекса Cas9 в виде сайт-специфичной эндонуклеазы.

Пример 3: Редактирование *in vivo* посредством белка Cas9, связанного в комплекс с гибридной нуклеиновой кислотой РНК-ДНК

Чтобы продемонстрировать, что ген-мишень может быть отредактирован *in vivo* доставленным комплексом, содержащим белок Cas9 и гибридную нуклеиновую кислоту РНК-ДНК, нефункциональный ген tdTomato, содержащийся в трансформированной плазмиде, репарировали путем обмена одного нуклеотида для восстановления флуоресцентного сигнала из гена tdTomato. Чтобы определить оптимальное использование для редактирования путем предложения матрицы для репарации ssДНК с комплементарностью относительно цепи-мишени или цепи, не являющейся мишенью, сравнивали комплексы, несущие матрицы для репарации любой цепи.

Комплекс из полипептида и гибридной нуклеиновой кислоты РНК/ДНК-Cas, полученный в Примере 1, использовали для репарации эписомальной плазмидной мишени, кодирующей ген tdTomato одноточечной мутацией от A до T, что создает ранний стоп-сигнал в положении кодона 51. Эта плазмида была введена в систему протопластов кукурузы вместе с комплексом редактирования, содержащим белок Cas9 и гибридную нуклеиновую кислоту РНК-ДНК посредством доставки, опосредованной PEG или электропорацией. Одноцепочечная матрица для репарации затем сцепляется с sgРНК посредством комплементарного спаривания оснований. Матрица для репарации комплементарна 80 парам оснований в нижней области и 40 парам оснований в области, находящейся выше сайта разреза. В этом случае, успешное редактирование приводит к тому, что некоторые клетки демонстрируют флуоресцентный фенотип tdTomato

из-за репарации гена tdTomato в, по меньшей мере, одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных матриц для репарации может быть легко оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 4: Редактирование *in vivo* белком Cas9, связанным в комплекс с RT, присоединенной к компоненту РНК ковалентным присоединением, или ассоциированной комплементарным спариванием оснований

Для демонстрации редактирования молекулами гибридной нуклеиновой кислоты, изготовленными различными способами, были использованы оптимальные условия, определенные в Примере 3, для оценки репарации одной и той же эписомальной плазмидной мишени посредством ковалентного сцепления гибридных нуклеиновых кислот или нековалентного спаривания оснований матрицы для репарации с sgРНК.

В случае использования маркера, в частности, флуоресцентного маркера, успешное редактирование приведет к тому, что некоторые клетки будут демонстрировать флуоресцентный фенотип из-за репарации гена, кодирующего флуоресценцию, такого как ген tdTomato, по меньшей мере, в одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных матриц для репарации может быть оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 5: Редактирование *in vivo* белком Cas9, связанным в комплекс с гибридной нуклеиновой кислотой, образованной путем связывания RT с 5'-концом или 3'-концом sgРНК

В одном примере, способ, описанный в Примере 3, может быть использован для идентификации предпочтения для матрицы для репарации, гибридизованной или сцепленной с 5'-концом или 3'-концом sgРНК. В этом случае, можно использовать предпочтительную ковалентность сцепления, определенную в Примере 4. На основе результатов, представленных в работе Цай и др. («Димерные РНК-направлемые нуклеазы FokI системы CRISPR для высокоспецифичного редактирования генома», Nature Biotechnology, 32, 569-576 (2014), идентификатор цифрового объекта:10.1038/nbt.2908) и далее в работе Шечнера и др. («Мультиплексное, локус-специфичное нацеливание длинных РНК с CRISPR-Дисплеем», Nature Methods, 12(7), 664-670 (2015), идентификатор цифрового объекта:10.1038/nmeth.3433), 3' слияние, как ожидается, будет предпочтительным.

Успешное редактирование приводит к тому, что некоторые клетки демонстрируют флуоресцентный фенотип, например, фенотип tdTomato, из-за репарации гена tdTomato в, по меньшей мере, одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных матриц для репарации может быть оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 6: Определение оптимальной длины линкера между sgРНК и матрицей для репарации для редактирования *in vivo* белком Cas9, связанным в комплекс с гибридной нуклеиновой кислотой

В одном примере, увеличение длины линкера с шагом 50 пар оснований вплоть до длины 500 пар оснований между sgРНК и матрицей для репарации было использовано для идентификации оптимальных условий гомологичной рекомбинации для репарации мишени, описанной в Примере 3. Использование набора длин линкеров поможет определить нужную гибкость, необходимую внутри гибрида для преодоления геометрических параметров цепей белка-мишени. Это особенно необходимо при работе с различными нуклеазами комплекса CRISPR и, следовательно, специфичными gРНК и индивидуальными

матрицами для репарации (RT) для координации взаимодействия молекулярного комплекса и гарантии того, что комплекс CRISPR также и в присутствии RT может оказывать свое действие. Условия Примера 3 использовали вместе с оптимизированными параметрами, определенными в Примерах 3-5. Линкер представлял собой ДНК с комплементарностью относительно последовательности, находящейся рядом с геном-мишенью.

Успешное редактирование приведет к тому, что некоторые клетки будут демонстрировать флуоресцентный фенотип tdTomato в случае, если используется маркер tdTomato, из-за репарации гена tdTomato в, по меньшей мере, одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных длин линкеров может быть оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки. Аналогичным образом, может быть использован любой другой выбираемый интересующий маркер, включая любой флуоресцентный маркер, подходящий для интересующего типа клеток, антибиотические маркеры, меченые последовательности, регуляторные последовательности и тому подобное.

Пример 7: Определение оптимальной конфигурации матрицы для репарации для редактирования *in vivo* белком Cas9, связанным в комплекс с гибридной нуклеиновой кислотой

Для демонстрации редактирования посредством одноцепочечных и двухцепочечных матриц для репарации был использован анализ *in vivo*, описанный в Примере 3, для относительного сравнения двух конфигураций. Ожидается, что одноцепочечные матрицы для репарации будут иметь лучшее основание при более низкой молекулярной массе и зафиксированных более высоких скоростях редактирования посредством коротких олигонуклеотидов ssДНК, чем посредством коротких олигонуклеотидов dsДНК. Однако использование двухцепочечной матрицы для репарации может потребоваться в тех случаях, когда требуется отредактировать или вставить большие последовательности. В этом примере могут быть использованы оптимальные условия Примеров 4 и 6.

Событие успешного редактирования приводит к тому, что некоторые клетки демонстрируют флуоресцентный фенотип, например, фенотип tdTomato, из-за репарации гена tdTomato в, по меньшей мере, одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных матриц для репарации может быть оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 8: Редактирование *in vivo* хромосомной мишени белком Cas9, связанным в комплекс с гибридной нуклеиновой кислотой РНК-ДНК

В одном примере, способ, оптимизированный в соответствии с Примерами 3-7, может быть использован для совершения редактирования в хромосомном гене-мишени. В этом случае, использовалось трансгенное растение кукурузы со стабильной вставкой кассеты tdTomato с ранним стоп-кодоном, чтобы продемонстрировать полезность изобретения для хромосомной мишени. Успешное редактирование привело к тому, что некоторые клетки демонстрировали флуоресцентный фенотип tdTomato из-за репарации гена tdTomato, интегрированного в геномную ДНК. Эффективность редактирования была оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 9: Вставка *in vivo* генной кассеты в хромосомную мишень белком Cas9, связанным в комплекс с гибридной нуклеиновой кислотой РНК-ДНК

Чтобы продемонстрировать полезность изобретения для вставки всей длины гена в хромосомную мишень, репортерный флуоресцентный ген tdTomato и терминатор были интегрированы в ген кукурузы hmg13, что привело к флуоресцентному сигналу tdTomato из-за экспрессии, управляемой эндогенным

промотором для hmg13. Результаты могли продемонстрировать, что длинные вставки могут быть сделаны с использованием изобретенного способа и помогут оптимизировать условия для указанной вставки.

Успешные результаты редактирования в некоторых клетках, демонстрирующих флуоресцентный фенотип tdTomato из-за вставки гена *tdTomato* в мишень hmg13 и последующая экспрессия белка tdTomato могут быть подтверждены. Таким образом, соответствующая эффективность редактирования для каждого типа тестируемых клеток может быть оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 10: Использование пептида, проникающего в клетку, для доставки в растительные клетки белка Cas9, связанного в комплекс с гибридной нуклеиновой кислотой РНК-ДНК

Оптимальная система, идентифицированная в Примерах 8 или 9, была использована в этом примере для тестирования эффективности трансформации на основе PEG, по сравнению с трансформацией посредством пептида, проникающего в клетку (CPP). Предыдущие публикации и заявки заявляют, что использование для доставки пептидов, проникающих в клетку, позволит ввести в клетки, имеющие клеточную стенку, белок Cas9, связанный в комплекс с гибридной нуклеиновой кислотой РНК-ДНК. Таким образом, CPP были использованы в белке слияния Cas или сцеплены с Cas, хотя дисульфидная связь образовалась между N-концевым цистеином на белке Cas и N-концевым цистеином на CPP. Свободные CPP можно также использовать для содействия импорту комплекса нуклеиновой кислоты Cas посредством транзиентного связывания на цепи нуклеиновой кислоты. Первичные CPP могут включать пептид TAT ВИЧ (см., например, SEQ ID NO: 17 и 18) или последовательность, полученную из них, и/или последовательность (Arg)g. Эффективность может быть протестирована с помощью использования оптимизированного способа из Примеров 3-9 посредством успешной экспрессии tdTomato в системе протопластов.

Пример 11: Другие нуклеазы CRISPR

Как подробно описано выше, гибридные последовательности нуклеиновых кислот по настоящему изобретению подходят для различных нуклеаз CRISPR различных систем CRISPR. Для любой эффекторной нуклеазы, например, Cas9 или Cpf1, оптимальные условия и длины gРНК и RT должны быть оценены, как подробно описано в Примерах 1-10 выше, чтобы достичь оптимальных результатов для интересующего события редактирования генома для каждого интересующего типа клеток. Кроме того, первые эксперименты с никазами Cas9 были проведены так же, как описано выше, с использованием более одной gРНК и любой, по меньшей мере, одной отдельной RT, ассоциированной с, по меньшей мере, одной из gРНК. Первые результаты демонстрируют, что это, по-видимому, перспективный подход для точного редактирования генома в эукариотических клетках.

Пример 12: Конструкции из животных клеток

Изобретенный способ может быть использован в эукариотических клетках при условии, что они способны к гомологичной рекомбинации. В одном первом примере, мышиные Т-клетки или предшественники Т-клеток были модифицированы *in vitro* для их модуляции, чтобы они подходили для иммунотерапии рака. Можно было бы продемонстрировать, что гибридные конструкции нуклеиновой кислоты по настоящему изобретению, когда они специально оптимизированы (оптимизация кодонов) и сконструированы (PAM, сайты-мишени) для системы животных, могут быть использованы для высокоточного редактирования генома в интересующем эукариотическом типе животной клетки. Таким образом, модификация экспрессированного гена, регулирующего пролиферацию или функцию Т-клетки, с использованием способа, описанного в настоящем изобретении, может быть использована для терапии, в частности, у млекопитающего, и, более конкретно, для лечения заболевания или расстройства у субъекта

путем модификации интересующего типа клетки посредством конструкций по настоящему изобретению.

Пример 13: Трансформация/трансфекция обработанной незрелой ткани метелки

Как подробно описано выше, различные физические/механические, а также биологические средства для трансформации растительных клеток, тканей, органов или целых растений, или их частей были описаны для введения генетического материала в растение или структуру-мишень растения. Эти способы также подходят для введения, по меньшей мере, одной гибридной последовательности нуклеиновой кислоты РНК/ДНК и/или, по меньшей мере, одной gРНК, и/или, по меньшей мере, одной матрицы для репарации, и/или, по меньшей мере, одного полипептида CRISPR по настоящему изобретению. После обработки и получения таким образом меристематической клетки, например, ткани метелки из мужского растения маиса, могут быть применены следующие способы для трансформации этой ткани:

Что касается биологических средств, то растительные ткани или их клетки могут быть трансформированы посредством трансформации, опосредованной *Agrobacterium*, включая *Agrobacterium tumefaciens* или *Agrobacterium rhizogenes.* Этот вид трансформации хорошо известен специалисту в данной области техники (см., например, Джоунс, Х.Д. и др., «Обзор методологий и протокола по *Agrobacterium*-опосредованной трансформации пшеницы», Plant Methods, 2005; или Фрейм, Б.Р. и др., «*Agrobacterium tumefaciens*-опосредованная трансформация зародышей кукурузы с использованием стандартной бинарной векторной системы», Plant, 2002). С этой целью культуру *Agrobacterium*, содержащую интересующую конструкцию, культивируют, например, в течение ночи при температуре 28° С в жидкой бульонной среде Луриа, содержащей подходящий антибиотик, 10 мМ MES и 200 мМ ACE. На следующий день, ночную культуру центрифугируют при 4400 об/мин в течение 15 мин, и супернатант отбрасывают. Затем, осадок снова центрифугируют в течение 15 мин при 4400 об/мин в течение 2 мин, а оставшийся супернатант отбрасывают. Осадок ресуспендируют (5 мл H2O, 10 мМ MES, 10 мМ MgCh + 20пМ ACE). Оптическая плотность при 600 нм регулируется до 1,5. Возможно, разбавленная суспензия может быть использована в дальнейшем.

Другой возможностью для трансформации меристематических клеток или тканей растения биологическими средствами является использование вирусных векторов. Вирусные векторы имеют то преимущество, что они могут быть введены либо как ДНК, либо как РНК, и в интересующую структуру-мишень растения. Кроме того, вирусные векторы или вирусы растений обладают способностью распространяться в различные клетки и ткани.

Для целей настоящего изобретения, вирусные частицы, *in vitro* транскрипты вирусов или *Agrobacteria*, несущие вирус, кодирующий Т-ДНК, могут быть введены в интересующую структуру-мишень растения посредством фильтрации (вакуумной и невакуумной). Альтернативные эксперименты могут быть проведены с использованием сока растения. С этой целью, либо табак, либо шпинат могут быть инфицированы представляющим интерес вирусом для последующего выделения указанного интересующего вируса из сока растения для заражения другой структуры-мишени растения, особенно меристематических клеток или тканей из различных растений, соком растения, содержащим вирус.

Несмотря на биологические средства трансформирования интересующих структур метелки, могут быть использованы дополнительные физические/механические средства для трансформации в дополнение к бомбардировке частицами.

Одним из подходящих способов является микроинъекция. Микроинъекция может быть использована для любого вида тестируемой меристематической структуры, преимущественно с использованием микроскопа с микроманипулятором. Из-за размера определенной меристематической

структуры, такой как, мерисистемы метелки или колосового цветорасположения, микроинъекцию можно проводить при микроскопическом контроле или, в случае, когда структуры-мишени достаточно велики, без помощи микроскопа. Инъекция может быть проведена с использованием множества способов для введения множества различных молекул-мишеней в интересующую структуру-мишень растения, включая двухцепочечную плазменную ДНК, линейную двухцепочечную ДНК, РНК и белки, а также вирусные частицы в жидком растворе. Эти различные молекулы могут быть применены с помощью микро- или наноигл, которые помогают проводить инъекции молекул-мишеней в меристематическую клетку или интересующую структуру. Молекулы-мишени сначала наносятся на иглу, которая затем вставляется в меристематическую клетку или интересующую структуру.

Другим подходящим средством является бомбардировка частицами, например, с использованием системы доставки частиц, причем этот способ дополнительно раскрыт выше.

Дальнейшим развитием этой технологии является использование комбинации вискеров карбида кремния (SiC) (например, Silar® Silicon Carbide Whisker) и микроинъекции. С этой целью, двухцепочечная (при необходимости, плазмидная) ДНК, линейная двухцепочечная ДНК, РНК, белок или молекулярный рибонуклеиновый комплекс по настоящему изобретению, или вирусные частицы осаждаются на вискере карбид кремния для инъекции посредством микроинъекционной иглы в меристематическую структуру или интересующую клетку. Этот метод имеет то преимущество, что не только можно трансфицировать одну клетку, но также имеется возможность проникнуть в разные клетки параллельно из-за распространения вискеров. Кроме того, клетки меньше разрушаются, так как игле не нужно проникать в клетку, а вискеры довольно малы по размеру.

Пример 14: Средство для детектирования модификации

Любая транзиентная или стабильная модификация, введенная, по меньшей мере, в одну последовательность ДНК-мишени по настоящему изобретению, может быть детектирована с использованием средства флуоресцентной детекции, в частности: в случае использования флуоресцентного репортера. Поскольку ткани метелки, такие как, пыльники и сухая пыльца, имеют сильную автофлуоресценцию, для этих клеток и тканей следует использовать другие средства. Таким образом, детекция может быть выполнена и подтверждена другими молекулярными способами, такими как, PCR, включая обогащение PCR, PCR-расщепление, комбинация обогащения PCR с PCR-расщеплением, количественная PCR или секвенирование, или RT-PCR, включая глубокое секвенирование или секвенирование следующего поколения, или Южный или Северный блоттинг по Саузерну. Уровни белка могут быть проанализированы с помощью Вестерн-блоттинга и тому подобного. В случае, если фенотипически детектируемый признак был введен, по меньшей мере, в одну интересующую клетку, то также возможно выполнить анализ, чтобы детектировать, присутствует ли указанный признак, например, устойчивость, флуоресценция, морфологический мутантный фенотип или любой другой признак, или отсутствует в, по меньшей мере, одной модифицированной клетке или ее потомстве, или производном. Вышеуказанные способы детекции известны специалисту в данной области техники.

Как правило, проведение анализа события стабильной интеграции в различных растениях-мишенях и их клетках может осуществляться следующим образом: Во-первых, ДНК и/или РНК экстрагируют из различных материалов, включая, например, метелку, пыльник или ткань/клетки пыльцы, трансформированных различными конструкциями, кодирующими флуоресцентный белок, например, красный флуоресцентный белок. В целом, образцы могут быть проанализированы посредством количественной PCR (qPCR). Из приведенных выше образцов, несколько образцов будут показывать четкий,

то есть, очень интенсивный (красный) флуоресцентный сигнал, который указывает на положительное событие, и который затем может быть выбран. Из этих образцов будет генерироваться сДНК, включая контролируемые реакции без обратной транскриптазы, чтобы исключить, что более поздние результаты не ассоциированы с нерасщепленной ДНК. Из образцов с положительным сигналом ДНК, используемых для измерения транскрипции, несколько образцов могут показать четкую транскрипцию, а другие – потенциальную транскрипцию (на границе того, что может быть четко измерено).

Пример 15: Белок слияния Cas9 и scFv

В одном эксперименте, белок слияния нуклеазы Cas9 в виде SSN и одноцепочечное антитело, направленное против флуоресцеина, в виде взаимодействующего домена можно экспрессировать *in vitro* или *in vivo* и подвергнуть воздействию FAM-меченого олигонуклеотида для действия в качестве матрицы для репарации. RT была синтезирована и ковалентно сцеплена с FAM в качестве стыковочного домена матрицы для репарации. Эффективность редактирования измерялась флуоресцентным сигналом, указывающим на репарационные или основанные на последовательности измерения частоты репарации, как подробно описано выше. Таким образом, пара взаимодействующих с SSN доменов Cas9 и scFv со специфичной аффинностью к выбранному лиганду, например, FAM, может быть получена и очищена отдельно и затем может быть поперечно сцеплена или соединена, или SSN и взаимодействующий домен (IA) могут быть продуцированы в виде молекулы слияния. В зависимости от анализа, молекула SSN-IA может быть трансфицирована в клетку или добавлена в анализ в виде белка, или конструкция может быть введена в клетку-мишень на векторе (индуцируемом или активном, конститутивным образом), подлежащим транскрибированию и трансляции *in vivo*. Кроме того, последовательность, кодирующая SSN-IA, может быть введена в клетку-мишень, содержащую интересующую последовательность ДНК-мишени, подлежащую модификикации как конструкция РНК, подлежащая трансляции *in vivo*. Примерные молекулы слияния SSN-IA по настоящему изобретению, сочетающие функциональную возможность SSN, полученной из CRISPR, с чрезвычайно высокой аффинностью связывания специализированного белка с его когнатным партнером, показаны с помощью SEQ ID NO: 44 (конструкция слияния Cas/mSA) и SEQ ID NO: 45 (конструкция слияния Cas/scFv(FAM)).

На **Фигурах 4 A-C** схематично показан подход к геномной инженерии с использованием слияний SSN и мономерного стрептавидина или scFv в качестве IA. Примечательно, что использование мономерного стрептавидина или scFvs, или любого другого IA, или RTDD не ограничивается использованием нуклеазы CRISPR или Аргонавта.

Пример 16: Связывание нуклеиновой кислоты белком слияния Cas9, сцепленного с scFv

Чтобы продемонстрировать способность белка слияния из Примера 15 связывать одноцепочечную или двухцепочечную матрицу для репарации, описанный анализ связывания будет повторен с олигонуклеотидом, меченым флуоресцеином (FAM). Олигонуклеотиды, меченые FAM, могут быть получены промышленным путем. Успешное взаимодействие может быть протестировано путем совместной миграции белка, ДНК и флуоресцентного красителя, и соответствующего увеличения молекулярной массы. Функциональная возможность части нуклеазы белка слияния будет протестирована с использованием анализа на расщепление *in vitro* специфичной направляющей РНК и линеаризованной плазмиды, содержащей соответствующую мишень. После инкубации в условиях, подходящих для нуклеазной активности, включая правильный pH, температуру и кофакторы, и тому подобное, которые известны специалисту в данной области техники в отношении различных нуклеаз CRISPR и их вариантов, плазмиду ДНК-мишени запускали в агарозном геле и наблюдали за размерами полос, указывающих на разрезание в

предполагаемом сайте-мишени. Расщепление *in vitro* ДНК-мишени указывало на то, что RT, ассоциированная с нуклеазой, не вмешивалась в нормальную функцию комплекса Cas9 как сайт-специфичная эндонуклеаза.

Пример 17: Белок слияния Cas9 и mSA2

В одном эксперименте, белок слияния нуклеазы Cas9 и метка модифицированного стрептавидина (на основе SEQ ID NO: 34) были экспрессированны и подвергнуты воздействию меченого биотином олигонуклеотида, действовавшего в качестве матрицы для репарации, причем биотин, действовал в качестве RTDD, а олигонуклеотид представлял собой RT. Эффективность редактирования измерялась флуоресцентным сигналом, указывающим на репарационные или основанные на последовательности измерения частоты репарации.

Пример 18: Связывание нуклеиновой кислоты белком слияния Cas9, сцепленного с mSA2

Чтобы продемонстрировать способность белка слияния из Примера 17 связывать одноцепочечную или двухцепочечную матрицу для репарации, описанный анализ связывания был повторен с меченым биотином олигонуклеотидом. Меченые биотином олигонуклеотиды могут быть получены промышленным путем или сгенерированы с использованием терминальной дезоксинуклеотидилтрансферазы. Успешное взаимодействие может быть протестировано путем совместной миграции белка и ДНК, и соответствующего увеличения молекулярной массы. Функциональная возможность части нуклеазы белка слияния будет протестирована с использованием анализа на расщепление *in vitro* специфичной направляющей РНК и линеаризованной плазмиды, содержащей соответствующую мишень. После инкубации в условиях, подходящих для нуклеазной активности, включая правильный pH, температуру и кофакторы, и тому подобное, которые известны специалисту в данной области техники в отношении различных нуклеаз CRISPR и их вариантов, плазмиду ДНК-мишени запускали в агарозном геле и наблюдали за размерами полос, указывающих на разрезание в предполагаемом сайте-мишени. Расщепление *in vitro* ДНК-мишени указывало на то, что RT, ассоциированная с нуклеазой, не вмешивалась в нормальную функцию комплекса Cas9 как сайт-специфичная эндонуклеаза.

Пример 19: Редактирование *in vivo* эписомальной мишени посредством белка слияния Cas9, связанного в комплекс с репарационной матричной нуклеиновой кислотой, меченой FAM или биотином, для восстановления функциональной возможности гена

Чтобы продемонстрировать, что ген-мишень может быть отредактирован *in vivo* доставленным комплексом, содержащим белок Cas9 и гибридную нуклеиновую кислоту, меченую FAM или биотином, нефункциональный ген tdTomato, содержащийся в трансформированной плазмиде, репарировали путем обмена одного нуклеотида для восстановления флуоресцентного сигнала из гена tdTomato. Чтобы определить оптимальное использование для редактирования путем предложения матрицы для репарации ssДНК с комплементарностью относительно цепи-мишени или цепи, не являющейся мишенью, сравнивали комплексы, несущие матрицы для репарации любой цепи.

Белок слияния, связанный в комплекс с нуклеиновой кислотой из Примера 16 или 18, соответственно, использовали для репарации эписомальной плазмидной мишени, кодирующей ген tdTomato одноточечной мутацией от A до T, что создает ранний стоп-сигнал в положении кодона 51. Эта плазмида была введена в систему протопластов кукурузы вместе с комплексом редактирования, содержащим белок слияния Cas9-ScFV или Cas9-mSA2 и нуклеиновую кислоту, меченую FAM или биотином, посредством доставки, опосредованной PEG. В этом случае, успешное редактирование приводит к тому, что некоторые клетки демонстрируют флуоресцентный фенотип tdTomato из-за репарации гена tdTomato в, по меньшей

мере, одной плазмиде, содержащейся в них. Таким образом, относительная эффективность редактирования с использованием различных матриц для репарации может быть легко оценена путем измерения копийности флуоресцентных клеток, полученных в результате каждой обработки.

Пример 20: Редактирование *in vivo* хромосомной мишени посредством белка слияния Cas9, связанного в комплекс с репарационной матричной нуклеиновой кислотой, меченой FAM или биотином, для интеграции последовательности ДНК в специфичный локус

Чтобы продемонстрировать, что ген-мишень может быть отредактирован *in vivo* доставленным комплексом, содержащим белок Cas9 и нуклеиновую кислоту, меченую FAM или биотином, специфичная, известная последовательность ДНК будет интегрирована в специфичный сайт в пределах геномной ДНК.

Белок слияния Cas9 и одноцепочечный вариабельный фрагмент с аффинностью к флуоресцеину (Пример 16) или белок слияния Cas9 и модифицированный стрептавидин (Пример 18) экспрессировали и подвергали воздействию меченой репарационной матричной ДНК и использовали для интеграции известной последовательности ДНК в геномный локус. Успешное редактирование будет проанализировано путем репарации сигнала флуоресцентной индикации или молекулярных анализов в сайте-мишени.

Пример 21: Связывание нуклеиновой кислоты белком слияния Аргонавтом, сцепленным с scFv

Чтобы продемонстрировать способность связывания репарационной матричной нуклеиновой кислоты с нуклеазой не из комплекса CRISPR, был проведен анализ связывания, показывающий увеличение массы в исследовании совместной миграции с использованием олигонуклеотида репарационной нуклеиновой кислоты, меченой FAM, и белка слияния нуклеазы Аргонавта (см. SEQ ID NO: 46), и одноцепочечного вариабельного фрагмента с аффинностью к FAM (см. SEQ ID NO: 43 и 45). Аналогичным образом, Аргонавт SSN может быть сцеплен с мономерным стрептавидином (см. SEQ ID NO: 42 и 44) в качестве связующего комплекса для RT. Функциональная возможность части нуклеазы белка слияния была протестирована с использованием анализа на расщепление *in vitro* специфичной направляющей нуклеиновой кислоты и линеаризованной плазмиды, содержащей соответствующую мишень. После инкубации в условиях, подходящих для нуклеазной активности, включая правильный pH, температуру и кофакторы, и тому подобное, которые известны специалисту в данной области техники в отношении различных нуклеаз, не из комплекса CRISPR, и их вариантов, плазмиду ДНК-мишени запускали в агарозном геле и наблюдали за размерами полос, указывающих на разрезание в предполагаемом сайте-мишени. Расщепление *in vitro* ДНК-мишени указывало на то, что RT, ассоциированная с нуклеазой, не вмешивалась в нормальную функцию комплекса Аргонавта как сайт-специфичная эндонуклеаза.

Пример 22: Редактирование *in vivo* хромосомной мишени посредством белка слияния Аргонавта, связанного в комплекс с репарационной матричной нуклеиновой кислотой, меченой FAM, для интеграции последовательности ДНК в специфичный локус

Чтобы продемонстрировать, что ген-мишень может быть отредактирован *in vivo* доставленным комплексом, содержащим белок Аргонавт нуклеазы, не содержащейся в комплексе CRISPR, и нуклеиновую кислоту, меченую FAM или биотином, специфичная, известная последовательность ДНК будет интегрирована в специфичный сайт в пределах геномной ДНК.

Белок слияния нуклеазы Аргонавта и одноцепочечный вариабельный фрагмент с аффинностью к флуоресцеину (см. Пример 21) экспрессировали и подвергали воздействию меченой репарационной матричной ДНК и использовали для интеграции известной последовательности ДНК в геномный локус. Успешное редактирование будет проанализировано путем репарации сигнала флуоресцентной индикации или молекулярных анализов в сайте-мишени.

Пример 23: Белок слияния нуклеазы CRISPR (типа Cas9 или Cpf1) и RTDD1

Чтобы продемонстрировать, что стратегия привязки работает, очищенную нуклеазу CRISPR типа Cas9 или Cpf1, слили с RTDD1, в этом случае она привязана к одноцепочечному вариабельному фрагменту (SEQ ID NO: 54) и экспрессировали в бактериях *E. coli.* Процесс запускали в денатурирующем геле в присутствии додецилсульфата натрия (SDS) с непрерывным градиентом (4-10%); и показали количество и чистоту белка. Белок был окрашен в этом геле. На правой панели Фигуры 5 показана привязка. Это – 4% неденатурирующий акриламидный гель (голубой нативный электрофорез в полиакриламидном геле – Blue Native PAGE), и в настоящем документе ДНК окрашивается красителем GelRed. Меченую флуоресцеином FAM (RTDD2-) матрицу для репарации инкубировали в нуклеазном буфере либо без RTDD1 нуклеазы, либо с RTDD1 нуклеазой, как показано на левой панели **Фигуры 5**. Если белок присутствовал, то происходила привязка, как видно по ДНК, которую детектируют на более высоком уровне молекулярной массы (стрелка на **Фигуре 5**).

Пример 24: Детектирование событий HDR

Чтобы продемонстрировать, что секвенирование следующего поколения, более конкретно, глубокое секвенирование ампликонов, способно детектировать событие HDR в сайте-мишени, закодированную нуклеазу (в данном случае, это была нуклеаза CRISPR), слитую с вариантом стрептавида, трансформировали на плазмиде вместе с матрицей для репарации. Матрица для репарации имела 5' биотиновую метку и была доставлена в виде одноцепочечного олигонуклеотида. Через двадцать четыре часа после трансформации собрали протопласты и извлекли ДНК. Сайт-мишень амплифицировали с использованием набора праймеров, которые были сконструированы таким образом, чтобы не перекрываться с плечами гомологии матрицы для репарации. В строке 4 **Фигуры 6** показано правильное событие HDR. Событие заменяет последовательность AAGGTGCTCGGCCCCGAGCTC (SEQ ID NO: 52; кодирующую аминокислотную последовательность KVLGPEL) на AAGTGGTCCAGCGCCGCGACCTAGCTC (SEQ ID NO: 53; кодирующую аминокислотную последовательность KWSSAAT-L). SEQ ID NO: 51 представляет собой полную матрицу для репарации, демонстрирующую, что плечи гомологии не простираются за пределы ампликона, что означает, что артефакты PCR с оставшейся матрицей для репарации маловероятны.

Пример 25: Привязка матрицы для репарации повышает эффективность HDR

Для этого эксперимента, компоненты из Примера 24 были трансформированы в протопласты листьев кукурузы. В случае привязки, нуклеазу (в данном случае, это была нуклеаза CRISPR) слили с нативной последовательностью стрептавидина. В любом случае, нуклеазу доставляли в виде плазмиды. Матрицу для репарации ДНК доставляли в виде олигонуклеотида с 5' биотиновой меткой. Через двадцать четыре часа после трансформации собрали протопласты и извлекли ДНК. Сайт-мишень амплифицировали с использованием набора праймеров, которые были сконструированы таким образом, чтобы не перекрываться с плечами гомологии матрицы для репарации. Глубокое секвенирование ампликонов (см. Пример 24) и последующий вычислительный анализ обеспечивают количественное определение события INDEL и HDR в сайте-мишени. Частота HDR была нормализована до частоты 15 INDEL как мера возникновения двухцепочечного разрыва. Средняя частота HDR увеличилась с 0,92% (± 0,06%) без привязки до 1,26% (± 0,06%), когда матрица для репарации привязана к нуклеазе (**Фигура 7).**

Формула изобретения

1. Искусственный молекулярный комплекс, содержащий

(a) по меньшей мере, одну сайт-специфичную нуклеазу (SSN) или ее каталитически активный фрагмент, или последовательность нуклеиновой кислоты, кодирующую ее, и непосредственно взаимодействующий с ней,

(b) по меньшей мере, один стыковочный домен матрицы для репарации (RTDD), или последовательность нуклеиновой кислоты, кодирующую его, отличающийся тем, что стыковочный домен матрицы для репарации выполнен с возможностью непосредственного взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты (RT);

(c) при необходимости, содержащий, по меньшей мере, один взаимодействующий домен (IA) или последовательность нуклеиновой кислоты, кодирующую его, при этом, по меньшей мере, один взаимодействующий домен непосредственно взаимодействует с, по меньшей мере, одной сайт-специфичной нуклеазой или ее каталитически активным фрагментом, и, при этом, по меньшей мере, один взаимодействующий домен выполнен с возможностью предложения, по меньшей мере, одной из функциональных возможностей, выбранных из группы, состоящей из

(i) взаимодействия с, по меньшей мере, одним стыковочным доменом матрицы для репарации; и/или

(ii) взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты; и/или

(iii) сиквенс-специфичного взаимодействия с геномной ДНК;

при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит, по меньшей мере, одну часть, комплементарную, по меньшей мере, одной геномной комплементарной последовательности, и, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты выполнена с возможностью опосредования репарации последовательности ДНК-мишени.

2. Искусственный молекулярный комплекс по пункту 1, отличающийся тем, что сайт-специфичная нуклеаза или последовательность нуклеиновой кислоты, кодирующая ее, выбрана из, по меньшей мере, одной нуклеазы CRISPR, включей нуклеазы Cas или Cpf1, TALEN, ZFN, мегануклеазы, рестрикционной эндонуклеазы, включая Fokl или ее вариант, или двух сайт-специфичных никующих эндонуклеаз, или варианта, или их каталитически активного фрагмента.

3. Искусственный молекулярный комплекс по любому из пунктов 1 или 2, отличающийся тем, что, по меньшей мере, один стыковочный домен матрицы для репарации, или последовательность нуклеиновой

кислоты, кодирующая его, выбран из, по меньшей мере, одного из биотина, аптамера, ДНК, РНК или белкового красителя, содержащего флуорофоры, содержащие флуоресцеин, или его варианта, малеимидов, или Tetraxolium (XTT), направляющей последовательности нуклеиновой кислоты, выполненной с возможностью взаимодействия с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты, стрептавидина или его варианта, предпочтительно, мономерного стептавидина, авидина или его варианта, аффинной метки, предпачтительно, стрептавидин-метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (нанотела), антикалина, белка Agrobacterium VirD2 или его домена, пикорнавируса VPg, топоизомеразы или ее домена, белка А бактериофага ФХ174, белка А* ФХ, белка VirE2 или его домена, или дигоксигенина.

4. Искусственный молекулярный комплекс по любому из предыдущих пунктов отличающийся тем, что, по меньшей мере, один взаимодействующий домен или последовательность нуклеиновой кислоты, кодирующая ее, выбрана, по меньшей мере, из ДНК-связывающего домена, стрептавидина или его варианта, предпочтительно из мономерного стептавидина, авидина или его варианта, аффинной метки, сигнала биотинилирования, сайта акцептора биотина, стрептовидиновой метки, антитела, одноцепочечного вариабельного фрагмента (scFv), однодоменного антитела (наночастицы), антикалина, биотина, аптамера, ДНК, РНК или красителя для белка, содержащего флуорофоры, содержащие флуоресцеин или его вариант, малеимидов или Tertaxolium (XTT), направляющей последовательности нуклеиновой кислоты, выполненной с возможностью взаимодействия, по меньшей мере, с одной репарационной матричной последовательностью нуклеиновой кислоты, белка Agrobacterium VirD2 или его домена, пикорнавируса VPg, топоизомеразы или ее домена, белка А бактериофага ФХ174, белка А* ФХ, белка VirE2 или его домена, или дигоксигенина.

5. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна сайт-специфичная нуклеаза и/или, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, один взаимодействующий домен содержит, по меньшей мере, одну последовательность ядерной локализации, последовательность пластидной локализации, предпочтительно последовательность митохондриальной локализации или последовательность хлоропластной локализации.

6. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит, по меньшей мере, одну концевую часть, предпочтительно 3'-конец, при этом, эта концевая часть не взаимодействует ни с каким-либо другим компонентом искусственного молекулярного комплекса и, таким образом, выполнена с возможностью гибридизации, по меньшей мере, с одной геномной комплиментарной последовательностью для опосредования репарации последовательности ДНК-мишени, и/или, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты предложена в виде плазмиды.

7. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна сайт-специфичная нуклеаза или ее каталитически активный фрагмент, или последовательность, кодирующая ее, выбрана из нуклеазы CRISPR, предпочтительно из нуклеазы Cas или

Cpf1, или нуклеазы FokI, или ее каталитического фрагмента, и, по меньшей мере, один взаимодействующий домен, или последовательность, кодирующая его, выбрана из одноцепочечного вариабельного фрагмента или мономерного стрептавидина.

8. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что он содержит, по меньшей мере, одну направляющую последовательность нуклеиновой кислоты, представляющую собой, по меньшей мере, один стыковочный домен матрицы для репарации, при этом, каждая из, по меньшей мере, одной направляющей последовательности нуклеиновой кислоты содержит

(i) первую часть последовательности, которая комплементарна последовательности, распознающей ДНК-мишень, и

(ii) вторую часть последовательности, при этом, вторая часть последовательности выполнена с возможностью взаимодействия, по меньшей мере, с одной сайт-специфичной нуклеазой, и,

(iii) при этом, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты физически ассоциирована с, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты и, таким образом, образует гибридную последовательность нуклеиновой кислоты, содержащую или состоящую из, по меньшей мере, одной РНК или ДНК и, по меньшей мере, одной дополнительной последовательности нуклеиновой кислоты ДНК, и,

(iv) при необходимости, содержащую линкерную область между, по меньшей мере, одной направляющей последовательностью нуклеиновой кислоты и, по меньшей мере, одной репарационной матричной последовательностью нуклеиновой кислоты,

предпочтительно, при этом, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с направляющей последовательностью нуклеиновой кислоты на 3'-конце направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты ассоциирована с 5'-концом направляющей последовательности нуклеиновой кислоты, и/или, при этом, репарационная матричная последовательность нуклеиновой кислоты расположена в пределах направляющей последовательности нуклеиновой кислоты.

9. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, одна направляющая последовательность нуклеиновой кислоты содержит нуклеотидную последовательность, выбранную из нуклеотидной последовательности природного или неприродного происхождения, включая синтетическую нуклеотидную последовательность, при необходимости, содержащую остов и/или базовые модификации, при этом, направляющая последовательность нуклеиновой кислоты содержит одноцепочечную или частично одноцепочечную нуклеотидную последовательность РНК или ДНК, и, при этом, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты содержит одноцепочечную или двухцепочечную нуклеотидную последовательность ДНК.

10. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна сайт-специфичная нуклеаза или последовательность, кодирующая ее, и, по меньшей мере, один взаимодействующий домен, или последовательность, кодирующая его, и/или, по меньшей мере, один стыковочный домен матрицы для репарации, или последовательность, кодирующая его, соединены, по меньшей мере, одним линкерным доменом.

11. Искусственный молекулярный комплекс по любому из предыдущих пунктов, отличающийся тем, что, по меньшей мере, одна сайт-специфичная нуклеаза или ее каталитически активный фрагмент, или последовательность, кодирующая ее, выбрана независимо из группы, состоящей из полипептида Cas из *Streptococcus spp.,* включая *Streptococcus pyogenes, Streptococcus thermophilus, Staphylococcus aureus,* или *Neisseria spp.,* включая *Neisseria meningitides, Corynebacter, Sutterella, Legionella, Treponema, Filifactor, Eubacterium, Lactobacillus, Mycoplasma, Bacteroides, Flaviivola, Flavobacterium, Sphaerochaeta, Azospirillum, Gluconacetobacter, Roseburia, Parvibaculum, Nitratifractor, Mycoplasma* и *Campylobacter, Candidatus* Micrarchaeum acidiphilum ARMAN-1, *Parcubacteria* (GenBank: APG80656.1), *Sulfolobus* spp., включая *Sulfolobus islandicus* HVE10/4 (GenBank: ADX81770.1) или REY15A (GenBank: ADX84852.1), полипептида Cpf1 из археи или бактерии, включая полипептид Cpf1 *Acidaminococcus spp.*, включая *Acidaminococcus sp.* BV3L6, *Lachnospiraceae spp.,* включая бактерию *Lachnospiraceae* ND2006, бактерию *Lachnospiraceae* MC2017, бактерию *Lachnospiraceae* MA2020, *Butyrivibrio proteoclasticus,* Candidatus spp., *Methanoplasma termitum, Leptospira inadai, Moraxella bovoculi* 237, бактерию Peregrinibacteria GW2011_GWA2_33_10, бактерию Parcubacteria GW2011_GWC2_44_17, Smithella sp. SCADC, Smithella sp. SC_K08D17, *Francisella spp.,* включая *Francisella novicida* U112, *Eubacterium eligens, Prevotella spp.,* или *Porphyromonas spp.,* или нуклеазы Аргонавта из *Natronobacterium gregoryi* (GenBank: AFZ73749.1), *Microcystis aeruginosa* (Стандартная/эталонная последовательность Национального центра биотехнологической информации (NCBI): WP_012265209.1 или Стандартная/эталонная последовательность NCBI: WP_002747795.1 или Стандартная/эталонная последовательность NCBI: WP_012265209.1), *Halogeometricum pallidum* (GenBank): ELZ29017.1), *Natrialaba asiatica* (Стандартная/эталонная последовательность NCBI: WP_006111085.1), *Natronorubrum tibetense* (Стандартная/эталонная последовательность NCBI: WP_006090832.1), *Natrinema pellirubrum* (Стандартная/эталонная последовательность NCBI: WP_006183335.1), или *Synechococcus spp.* (Стандартная/эталонная последовательность NCBI: WP_011378069.1) или их вариантов и/или функциональных фрагментов, и/или комбинаций, включая никазы, или нуклеазы, не обладающие эндонуклеолитической активностью.

12. Искусственный молекулярный комплекс по любому из предыдущих пунктов для использования в способе лечения заболевания, отличающийся тем, что заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации.

13. Способ лечения заболевания с использованием искусственного молекулярного комплекса по любому из предыдущих пунктов, отличающийся тем, что заболевание характеризуется, по меньшей мере, одной геномной мутацией, и искусственный молекулярный комплекс выполнен с возможностью нацеливания и репарации, по меньшей мере, одной геномной мутации.

14. Растение, растительная клетка, растительный материал или производное, или его потомство, содержащий или отредактированный, по меньшей мере, одним искусственным молекулярным комплексом по любому из пунктов 1-11.

15. Способ модификации, по меньшей мере, одной последовательности ДНК-мишени, содержащий следующие этапы:

(i) предложение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки, и/или генома, содержащего, по меньшей мере, одну геномную комплементарную последовательность и, по меньшей мере, одну последовательность ДНК-мишени в интересующей геномной области;

(ii) предложение, по меньшей мере, одного искусственного молекулярного комплекса, как определено в любом из пунктов 1-11;

(iii) контактирование, по меньшей мере, одного искусственного молекулярного комплекса с, по меньшей мере, одной последовательностью ДНК-мишени в подходящих условиях для достижения

(a) взаимодействия, по меньшей мере, одной сайт-специфичной нуклеазы с, по меньшей мере, одной последовательностью ДНК-мишени; и

(b) комплементарного спаривания оснований, по меньшей мере, одной репарационной матричной последовательности нуклеиновой кислоты с, по меньшей мере, одной геномной комплементарной последовательностью для достижения распознавания, по меньшей мере, одной комплементарной последовательности и индукции, по меньшей мере, одного разрыва ДНК, по меньшей мере, одной сайт-специфичной нуклеазой, отличающийся тем, что, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты направляет гомологично направленную репарацию на сайт, по меньшей мере, одной последовательности ДНК-мишени; и

(iv) получение, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки и/или генома, содержащего модификацию в, по меньшей мере, одной последовательности ДНК-мишени.

16. Способ по пункту 15, отличающийся тем, что, по меньшей мере, одна репарационная матричная последовательность нуклеиновой кислоты и/или, по меньшей мере, один стыковочный домен матрицы для репарации искусственного молекулярного комплекса предложен/предложены для, по меньшей мере, одной прокариотической или эукариотической клетки независимо от, по меньшей мере, одной сайт-специфичной нуклеазы, по меньшей мере, одного молекулярного комплекса, и, по меньшей мере, один искусственный молекулярный комплекс собирается или частично собирается, в пределах, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки и/или генома.

17. Способ по пункту 15, отличающийся тем, что, по меньшей мере, один искусственный молекулярный комплекс представляет собой собранный *ex vivo* искусственный молекулярный комплекс.

18. Способ по любому из пунктов 15-17, отличающийся тем, что, по меньшей мере, одна эукариотическая клетка представляет собой растительную клетку, в предпочтительном варианте осуществления настоящего изобретения, растительную клетку из растения, выбранного из группы, состоящей из *Hordeum vulgare, Hordeum bulbusom, Sorghum bicolor, Saccharum officinarium, Zea spp., включая Zea mays, Setaria italica, Oryza minuta, Oryza sativa, Oryza australiensis, Oryza alta, Triticum aestivum, Triticum durum, Secale cereale,* Triticale, *Malus domestica, Brachypodium distachyon, Hordeum marinum, Aegilops tauschii, Daucus glochidiatus, Beta spp.,* включая *Beta vulgaris, Daucus pusillus, Daucus muricatus, Daucus carota, Eucalyptus grandis, Nicotiana sylvestris, Nicotiana tomentosiformis, Nicotiana tabacum, Nicotiana benthamiana, Solanum lycopersicum, Solanum tuberosum, Coffea canephora, Vitis vinifera, Erythrante guttata, Genlisea aurea, Cucumis sativus, Marus notabilis, Arabidopsis arenosa, Arabidopsis lyrata, Arabidopsis thaliana, Crucihimalaya himalaica, Crucihimalaya wallichii, Cardamine nexuosa, Lepidium virginicum, Capsella bursa pastoris, Olmarabidopsis pumila, Arabis hirsute, Brassica napus, Brassica oleracea, Brassica rapa, Raphanus sativus, Brassica juncacea, Brassica nigra, Eruca vesicaria subsp. sativa, Citrus sinensis, Jatropha curcas, Populus trichocarpa, Medicago truncatula, Cicer yamashitae, Cicer bijugum, Cicer arietinum, Cicer reticulatum, Cicer judaicum, Cajanus cajanifolius, Cajanus scarabaeoides, Phaseolus vulgaris, Glycine max, Gossypium sp., Astragalus sinicus, Lotus japonicas, Torenia fournieri, Allium cepa, Allium fistulosum, Allium sativum, Helianthus annuus, Helianthus tuberosus* и *Allium tuberosum,* или любого сорта или подвида, принадлежащего одному из вышеупомянутых растений.

19. Способ по пункту 18, отличающийся тем, что модификация, по меньшей мере, одной последовательности ДНК-мишени вызывает редактирование признака, выбранного из группы, состоящей из повышения урожайности, толерантности к абиотическому стрессу, включая стресс, вызванный засухой, осмотический стресс, тепловой стресс, стресс от холода, оксидативный стресс, стресс под действием тяжёлых металлов, стресс, вызванный засолением или переувлажнением, толерантности к биотическому стрессу, включая толерантность к насекомым, толерантность к бактериям, толерантность к вирусам, толерантность к грибам или толерантность к нематодам, устойчивости к гербицидам, включая глифосат, глуфосинат, ингибиторы ALS и дикамбу, устойчивости к полеганию, времени цветения, устойчивости к осыпанию, цвета семян, состава эндосперма, содержания питательных веществ или метаболической инженерии, включая редактирование генома, для обеспечения подхода молекулярного фарминга, по меньшей мере, в одной растительной клетке.

20. Способ по любому из пунктов 15-19, отличающийся тем, что, дополнительно содержащий следующий этап:

(v) идентификация и/или выбор, по меньшей мере, одной прокариотической, эукариотической или вирусной клетки и/или генома, содержащего модификацию, по меньшей мере, одной последовательности ДНК-мишени.

21. Способ изготовления растения или растительной клетки, содержащий следующие этапы:

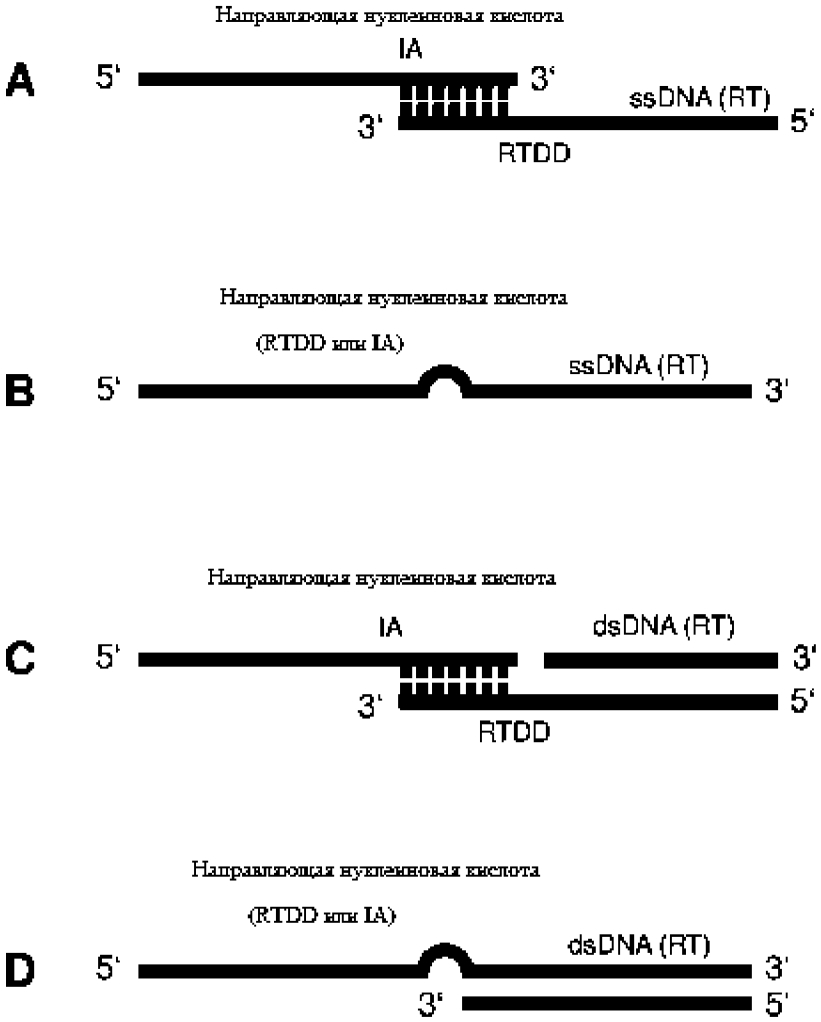
(i) выполнение способа по любому из пунктов 15-20, отличающийся тем, что, по меньшей мере, одна эукариотическая клетка представляет собой растительную клетку;

(ii) получение, по меньшей мере, одного растения или его потомства из, по меньшей мере, одной растительной клетки из этапа (i);

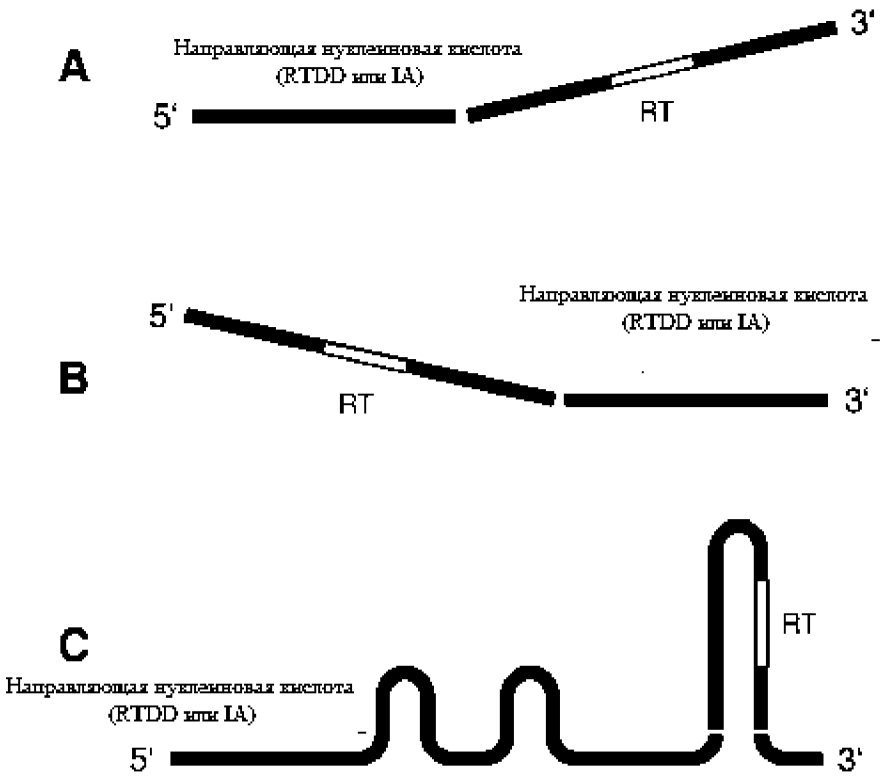
(iii) при необходимости: определение модификации, по меньшей мере, одной последовательности ДНК-мишени в, по меньшей мере, одной клетке, по меньшей мере, одного растения или его потомства.

22. Способ по пункту 21, отличающийся тем, что, по меньшей мере, одно растение или растительная клетка выбрана из однодольного или двудольного растения, предпочтительно растение выбрано из группы, состоящей из *Zea spp.*, включая *Zea mays, Nicotiana benthamiana,* или *Beta spp,* включая *Beta vulgaris,* или *Secale ssp.,* включая *Secale cereal,* или *Triticum ssp.,* включая *Triticum aestivum.*

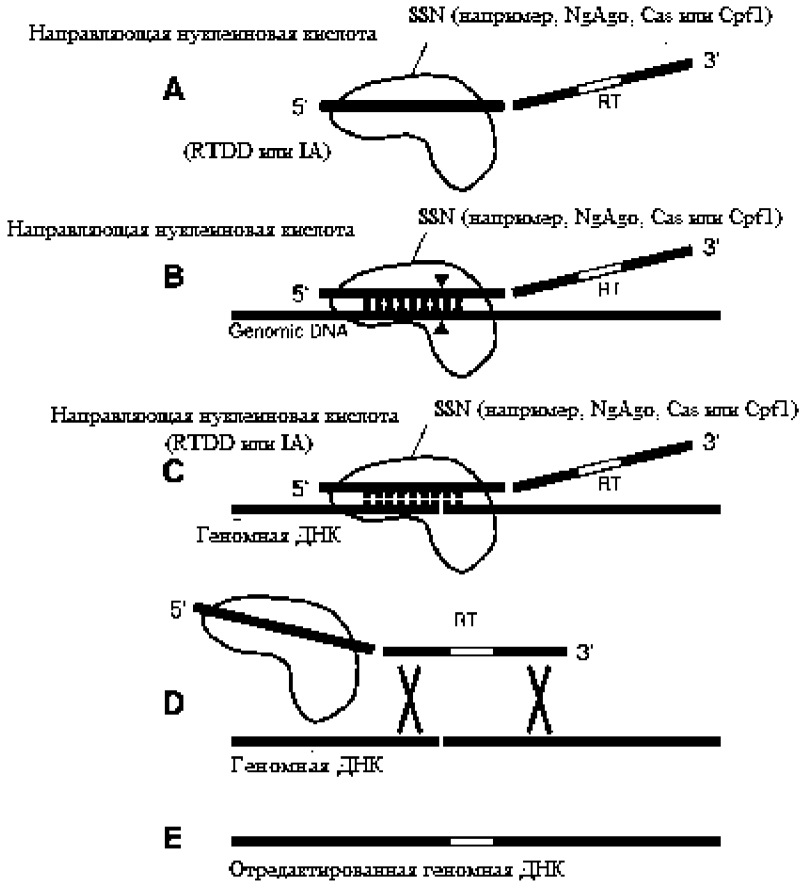
23. Использование, по меньшей мере, одного искусственного молекулярного комплекса по любому из пунктов 1-11 для геномной инженерии в прокариотической, эукариотической или вирусной клетке и/или геноме, и/или организме, предпочтительно в растительной клетке или организме.


Направляющая нуклеиновая кислота
A
IA
5'
3'
ssDNA (RT)
3'
5'
RTDD
Направляющая нуклеиновая кислота
(RTDD или IA)
B
ssDNA (RT)
5'
3'
Направляющая нуклеиновая кислота
C
IA
dsDNA (RT)
5'
3'
3'
5'
RTDD
Направляющая нуклеиновая кислота
(RTDD или IA)
D
dsDNA (RT)
5'
3'
3'
5'

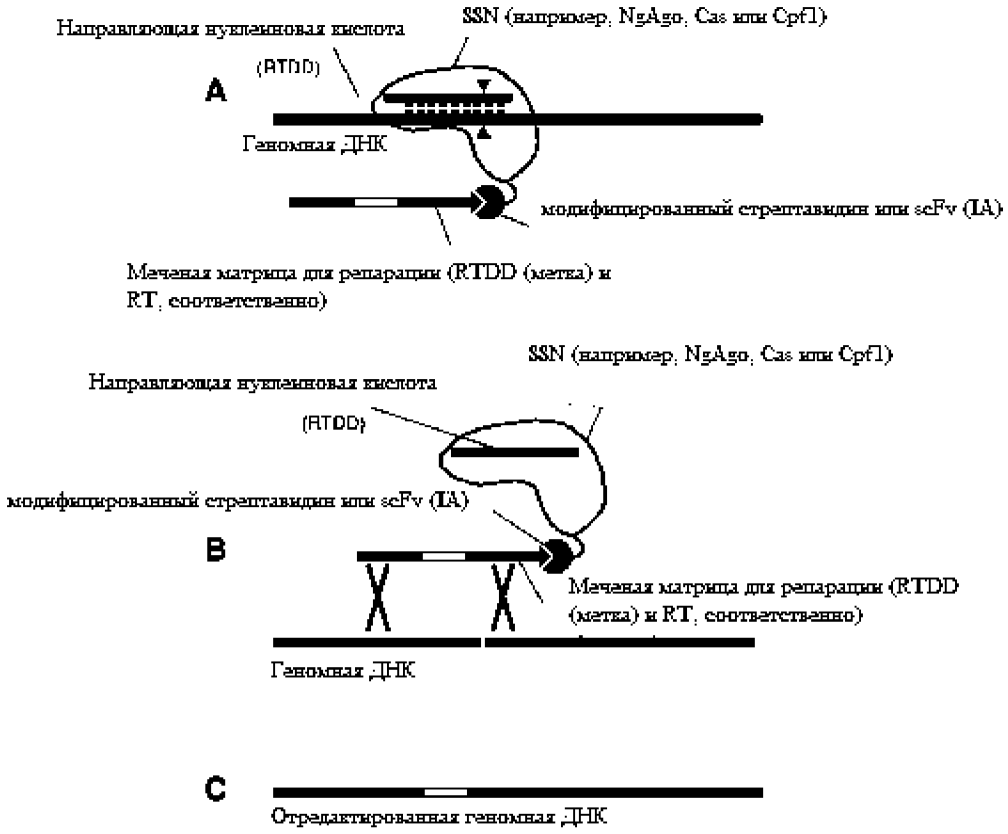
Фиг. 1


A
Направляющая нуклеиновая кислота
(RTDD или IA)
5'
3'
RT
B
5'
Направляющая нуклеиновая кислота
(RTDD или IA)
RT
3'
C
Направляющая нуклеиновая кислота
(RTDD или IA)
5'
RT
3'

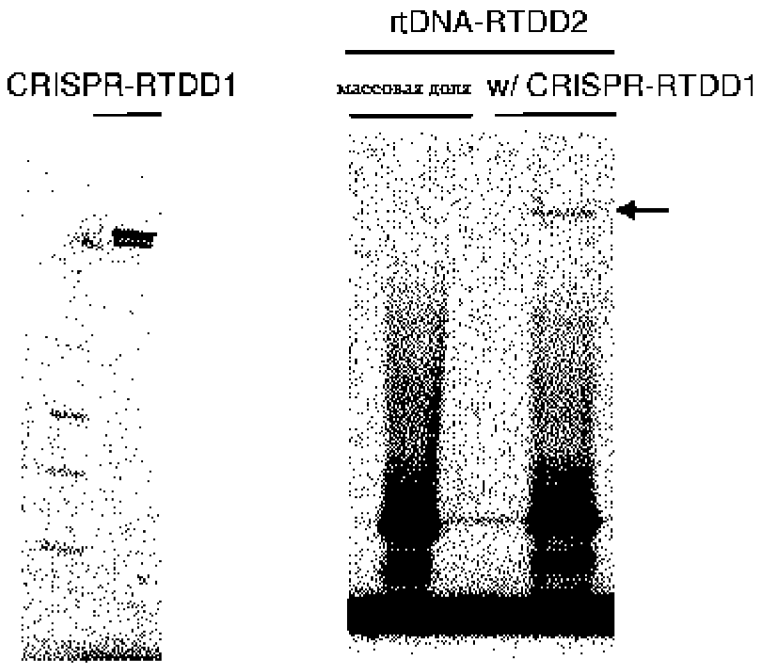
Фиг. 2


Направляющая нуклеиновая кислота
SSN (например, NgAgo, Cas или Cpf1)
A
5'
3'
RT
(RTDD или IA)
Направляющая нуклеиновая кислота
SSN (например, NgAgo, Cas или Cpf1)
B
5'
3'
RT
Genomic DNA
Направляющая нуклеиновая кислота
(RTDD или IA)
SSN (например, NgAgo, Cas или Cpf1)
C
5'
3'
RT
Геномная ДНК
5'
RT
3'
D
Геномная ДНК
E
Отредактированная геномная ДНК

Фиг. 3

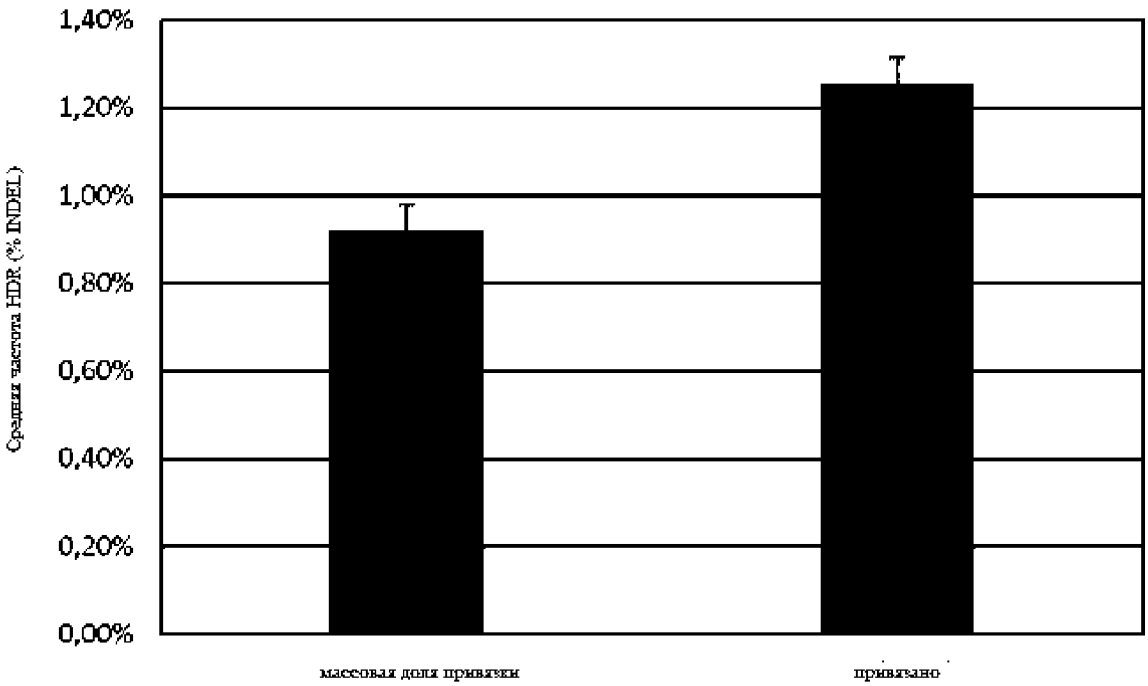

Направляющая нуклеиновая кислота
SSN (например, NgAgo, Cas или Cpf1)
(RTDD)
A
Геномная ДНК
модифицированный стрептавидин или scFv (IA)
Меченая матрица для репарации (RTDD (метка) и RT, соответственно)
SSN (например, NgAgo, Cas или Cpf1)
Направляющая нуклеиновая кислота
(RTDD)
модифицированный стрептавидин или scFv (IA)
B
Меченая матрица для репарации (RTDD (метка) и RT, соответственно)
Геномная ДНК
C
Отредактированная геномная ДНК

Фиг. 4



Фиг. 5

Фиг. 6


1,40%
1,20%
1,00%
0,80%
0,60%
0,40%
0,20%
0,00%
Средняя частота HDR (% INDEL)
массовая доля привязки
привязано

Фиг. 7