



(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2023.02.20

(21) Номер заявки
201991849

(22) Дата подачи заявки
2018.02.22

(51) Int. Cl. **C12N 15/113** (2010.01)
C12N 15/82 (2006.01)
C12N 5/10 (2006.01)

(54) СИСТЕМА И СПОСОБ РЕДАКТИРОВАНИЯ ГЕНОМА

(31) **201710089494.9**

(32) **2017.02.20**

(33) **CN**

(43) **2020.03.23**

(86) **PCT/CN2018/076949**

(87) **WO 2018/149418 2018.08.23**

(71)(73) Заявитель и патентовладелец:
**СУЧЖОУ ЧИ БИОДИЗАЙН
БИОТЕХНОЛОДЖИ КОМПАНИ
ЛИМИТИД (CN)**

(72) Изобретатель:
**Гао Саиксиа, Чжанг Хуавей, Чжанг
Дингбо (CN)**

(74) Представитель:
Зуйков С.А. (RU)

(56) XIE, C. et al. "Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tunaprocessing system" PNAS, Vol. 112, No. 11, 17 March 2015 (2015-03-17), ISSN: 0027-8424, Pages 3570-3575, see page 3570, right column, the second paragraph, page 3572, left column, the first paragraph, page 3575, the left column, the third paragraph, Table S1 in Supporting Information

LEE, R. T. H. et al. "Ribozyme Mediated gRNA Generation for In Vitro and in Vivo CRISPR/Cas9 Mutagenesis" PloS One, Vol. 11, No. 11, 10 November 2016 (2016-11-10), ISSN: 1932-6203, pages 1-12 <doi:

10.1371/journal.pone.0166020>, See page 2, the third paragraph, page 3, the last paragraph to page 5, the first paragraph, page 9, the second paragraph, figure 1A

SLAYMAKER, I. M. et al. "Rationally engineered Cas9 nucleases with improved specificity" Science, Vol. 351, 01 January 2016 (2016-01-01), ISSN: 0036-8075, Issus6268, pages 84-88, see the abstract, the seventh paragraph

KLEINSTIVER, B. P. et al. "High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects" Nature, Vol. 529, 28 January 2016 (2016-01-28), ISSN: 0028-0836, pages 490-495, see the abstract, page 490, right column, the second and third paragraphs

ZHANG, D. B. et al. "Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases" Genome Biology, Vol. 18, 11 October 2017 (2017-10-11), ISSN: 1759-6653, pages 1-7 <doi: 10.1186/s 13059-017-1325-9>, see the abstract, page 6, left column, the second and third paragraphs, figure s 1 in Supplementary material

ZHANG, D. B. et al. "Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases" Genome Biology, Vol. 18, 11 October 2017 (2017-10-11), ISSN: 1759-6653, pages 1-7 <doi: 10.1186/sl3059-017-1325-9>, see the abstract, page 6, left column, the second and third paragraphs, figure s1 in Supplementary material

CN-A-107475256

(57) Предложена система редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит вариант нуклеазы Cas9, имеющей более высокую специфичность по сравнению с нуклеазой Cas9 дикого типа или ее экспрессионным конструктором, и экспрессионный конструктор, содержащий кодирующую последовательность слияния транспортной РНК (tPHK)-направляющей РНК (gPHK) или слияния рибозима и направляющей РНК (gPHK), или содержит экспрессионный конструктор, содержащий кодирующие последовательности как варианта нуклеазы Cas9, так и tPHK-gPHK или рибозима-gPHK. Также предложен способ генетической модификации клетки путем введения системы редактирования генома, которая обладает высокой эффективностью и высокой специфичностью.

Область техники, к которой относится изобретение

Настоящее изобретение относится к области генной инженерии. В частности, настоящее изобретение относится к способу редактирования генома, который обладает высокой эффективностью и высокой специфичностью. Более конкретно, настоящее изобретение относится к способу повышения эффективности сайт-направленной модификации последовательности-мишени в геноме организма при помощи высокоспецифичного варианта нуклеазы Cas9.

Предшествующий уровень техники

Наиболее распространенными средствами редактирования генома являются короткие палиндромные повторы, регулярно расположенные группами (CRISPR), и связанная с ними система (CRISPR/Cas9). В этой системе для создания двухцепочечного разрыва (DSB) белок Cas9 расщепляет специфическую последовательность ДНК под воздействием гРНК. Двухцепочечный разрыв может активировать внутриклеточные механизмы репарации нехомологичного соединения концов (NHEJ) и гомологичной рекомбинации (HR) для восстановления повреждения ДНК в клетках таким образом, что в процессе восстановления конкретная последовательность ДНК может быть отредактирована. В настоящее время наиболее часто используемым белком Cas9 является Cas9, полученный из *Streptococcus pyogenes* (SpCas9). Одним из недостатков системы редактирования генома CRISPR/Cas9 является ее низкая специфичность и ненаправленное действие, что существенно ограничивает ее применение.

В данной области техники остается потребность в способе и инструментальном средстве, которые обеспечили бы эффективное высокоспецифичное редактирование генома.

Сущность изобретения

Один объект настоящего изобретения предусматривает систему редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент, выбранный из следующих с i) по iii):

- i) вариант нуклеазы Cas9 и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние tРНК-направляющей РНК;
- ii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние tРНК-направляющей РНК; и
- iii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность, кодирующую слияние tРНК-направляющей РНК;

в котором вариант нуклеазы Cas9 имеет повышенную специфичность по сравнению с нуклеазой Cas9 дикого типа,

в котором конец 5' направляющей РНК связан с концом 3' tРНК,

в котором слияние расщепляется на конце 5' направляющей РНК после транскрибирования в клетку, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на конце 5'.

Второй объект настоящего изобретения предусматривает систему редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент, выбранный из следующих с i) по iii):

- i) вариант нуклеазы Cas9 и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК;
- ii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК; и
- iii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК;

в котором вариант нуклеазы Cas9 имеет повышенную специфичность по сравнению с нуклеазой Cas9 дикого типа,

в котором конец 5' направляющей РНК связан с концом 3' первого рибозима,

в котором первый рибозим предназначен для расщепления слияния на конце 5' направляющей РНК, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на конце 5'.

Третий объект настоящего изобретения предусматривает способ для генетического модифицирования клетки, содержащий введение в клетку системы редактирования генома по настоящему изобретению, посредством которого вариант нуклеазы Cas9 нацеливается на последовательность-мишень в геноме клетки при помощи гРНК, а результат заключается в замене, делеции и/или добавлении одного или более нуклеотидов в последовательность-мишень.

Четвертый объект настоящего изобретения предусматривает генетически модифицированный организм, который содержит генетически модифицированную клетку, полученную способом по настоящему изобретению.

Краткое описание фигур

На фиг. 1 показаны стратегии конструирования одиночной направляющей РНК (sgРНК) для последовательностей-мишеней с разными 5' концевыми нуклеотидами при использовании U3 или U6. А: путем слияния с tРНК, sgРНК может быть спроектирована без учета 5' концевого нуклеотида последовательности-мишени; В: точное расщепление слияния tРНК-sgРНК.

На фиг. 2 показана эффективность редактирования WT SpCas9 (дикий тип SpCas9), eSpCas9(1.0), eSpCas9(1.1), SpCas9-HF1 на мишенях класса (1).

На фиг. 3 показана эффективность редактирования WT SpCas9 (дикий тип SpCas9), eSpCas9(1.0), eSpCas9(1.1), SpCas9-HF1 на мишенях класса (2).

На фиг. 4 показано, что дополнительный нуклеотид на конце 5' sgРНК влияет на эффективность редактирования при использовании промотора U6.

На фиг. 5 показано, что для локуса OsMCK4 tРНК-sgРНК способно улучшить эффективность редактирования и обеспечить высокий уровень специфичности по сравнению с sgРНК.

На фиг. 6 показано, что для локуса OsCDKB2 использование tРНК-sgРНК может повысить эффективность редактирования до уровня SpCas9 дикого типа, с одновременным обеспечением высокой специфичности.

На фиг. 7 показана специфичность редактирования варианта Cas9 для нарушения комплементарности между gРНК и последовательностью-мишенью.

На фиг. 8 показано tРНК-sgРНК, улучшающее эффективность редактирования eSpCas9(1.1) и SpCas9-HF1 до уровня SpCas9 дикого типа в клетках организма человека.

На фиг. 9 показана структура последовательности вектора pUC57-U3-tРНК-sgРНК для экспрессии слияния tРНК-sgРНК.

Подробное описание изобретения

1) Определение.

В настоящем изобретении, если не указано иное, используемые в данном документе научные и технологические термины имеют значения, общепринятые у специалистов в данной области техники. Также, используемые в данном документе терминология и порядок проведения экспериментов, относящиеся к химии белков и нуклеотидов, молекулярной биологии, культивированию клеток и тканей, микробиологии, иммунологии, относятся к терминологии и стандартным способам, обычно используемым в данной области техники. Например, используемая здесь стандартная технология рекомбинации ДНК и молекулярного клонирования хорошо известна специалистам в данной области и подробно описана в следующих справочных материалах: Sambrook, J., Fritsch, E.F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989. Тем не менее, для лучшего понимания настоящего изобретения ниже приведены определения и пояснения для соответствующих терминов.

"Нуклеаза Cas9" и "Cas9" в данном документе могут использоваться взаимозаменяемо, и относятся к РНК-направленной нуклеазе, включая белок Cas9 или его фрагменты (такие как белок, содержащий активный домен Cas9 расщепления ДНК и/или домен Cas9 связывания gРНК). Cas9 является компонентом системы редактирования генома CRISPR/Cas (короткие палиндромные повторы, регулярно расположенные группами, и связанная с ними система), которая под влиянием направляющей РНК таргетируется и расщепляет последовательность-мишень ДНК с образованием двухцепочечных разрывов ДНК (DSB).

"Направляющая РНК" и "gРНК" в данном документе могут использоваться взаимозаменяемо, и, как правило, состоят из молекул crРНК и tracrРНК, образующих комплексы посредством частичного комплемента, в которых crРНК содержит последовательность, которая является достаточно комплементарной по отношению к последовательности-мишени для гибридизации и направляет комплекс CRISPR (Cas9+crРНК+tracrРНК) на специфическое связывание с последовательностью-мишенью. Однако в данной области техники известно, что может быть сконструирована одиночная направляющая РНК (sgРНК), которая содержит характеристики как crРНК, так и tracrРНК.

В данном документе термины "tРНК" и "транспортная РНК" могут использоваться взаимозаменяемо и относятся к малоразмерным молекулам РНК, которые имеют функцию переноса и транспортировки аминокислот. Молекула tРНК обычно состоит из короткой цепи, в которой имеются примерно 70-90 нуклеотидов, упакованных в форме клевера. В эукариотах гены tРНК в геноме транскрибируются в прекурсоры tРНК, которые затем перерабатываются в зрелую tРНК после удаления дополнительных последовательностей 5' и 3' с помощью рибонуклеазы P (RNase P) и рибонуклеазы Z (RNase Z).

В данном документе термин "рибозим" относится к молекуле РНК, имеющей каталитическую функцию, которая участвует в расщеплении и процессинге РНК, катализируя реакции гидролиза трансфосфатной и фосфодиэфирной связи.

Термин "геном", используемый в настоящем документе, охватывает не только хромосомную ДНК, присутствующую в ядре, но также и ДНК органелл, присутствующую в субклеточных компонентах клетки (например, митохондриях, пластадах).

В данном документе термин "организм" включает в себя любой организм, пригодный для геномного редактирования. Типовые организмы включают в себя, не ограничиваясь только перечисленным, мле-

копитающих, таких как человек, мышь, крыса, обезьяна, собака, свинья, овца, крупный рогатый скот, кошка; домашнюю птицу, такую как курица, утка, гусь; растения, включающие в себя однодольные и двудольные растения, такие как рис, кукуруза, пшеница, сорго, ячмень, соя, арахис, арабидопсис и т.п.

"Генетически модифицированный организм" или "генетически модифицированная клетка" означает организм или клетку, которая содержит в своем геноме экзогенный полинуклеотид, или модифицированный ген, или последовательность, которая контролирует экспрессию. Например, экзогенный полинуклеотид устойчиво интегрируется в геном организма или клетки и наследуется в последующих поколениях. Экзогенный полинуклеотид может быть интегрирован в геном отдельно или как часть конструкта рекомбинантной ДНК. Модифицированный ген или последовательность контроля экспрессии представляют собой последовательность в геноме организма или клетки, которая состоит из одной или нескольких дезоксирибонуклеотидных замен, делеций и присоединений.

Термин "экзогенный" в отношении последовательности означает последовательность, которая происходит от чужеродного вида или, если она происходит от того же вида, существенно изменена по сравнению с его нативной формой по составу и/или геномному локусу вследствие преднамеренного вмешательства человека.

Термины "полинуклеотид", "последовательность нуклеиновой кислоты", "нуклеотидная последовательность" или "фрагмент нуклеиновой кислоты" используются взаимозаменяемо для обозначения полимера РНК или ДНК, который является одно- или двухцепочечным и, возможно, содержит синтетические, неприродные или измененные нуклеотидные основания. Нуклеотиды (обычно встречающиеся в их 5'-монофосфатной форме) обозначаются одной буквой следующим образом: "А" для аденилата или дезоксиаденилата (для РНК или ДНК, соответственно), "С" для цитидилата или дезоксцитидилата, "G" для гуанилата или дезоксигуанилата, "U" для уридилата, "T" для дезокситимидилата, "R" для пуринов (А или G), "Y" для пиримидинов (С или T), "K" для G или T, "H" для А, или С, или T, "I" для инозина и "N" для любого нуклеотида.

Термины "полипептид", "пептид", "аминокислотная последовательность" и "белок" используются в данном документе взаимозаменяемо для обозначения полимера из аминокислотных остатков. Эти термины применяются к аминокислотным полимерам, в которых один или более аминокислотных остатков представляют собой искусственный химический аналог соответствующей аминокислоты, встречающейся в природе, а также встречающихся в природе аминокислотных полимеров. Термины "полипептид", "пептид", "аминокислотная последовательность" и "белок" также охватывают модификации, включая, но не ограничиваясь только перечисленным, гликозилирование, липидное присоединение, сульфатирование, гамма-карбоксилирование остатков глутаминовой кислоты, гидроксिलирование и АДФ-рибозилирование.

В данном документе термин "экспрессионный конструкт" относится к вектору, пригодному для экспрессии в организм представляющей интерес нуклеотидной последовательности, такой, как рекомбинантный вектор. Термин "экспрессия" относится к производству функционального продукта. Например, экспрессия нуклеотидной последовательности может относиться к транскрипции нуклеотидной последовательности (такой как транскрипция для получения мРНК или функциональной РНК) и/или трансляции РНК в прекурсор белка или зрелый белок.

"Экспрессионный конструкт" по данному изобретению может представлять собой линейный фрагмент нуклеиновой кислоты, кольцевую плазмиду, вирусный вектор или, в некоторых вариантах осуществления, РНК, которая может быть транслирована (такая как мРНК).

"Экспрессионный конструкт" по данному изобретению может содержать регуляторные последовательности и представляющие интерес нуклеотидные последовательности, которые получены из разных источников, или регуляторные последовательности и представляющие интерес нуклеотидные последовательности, полученные из одного и того же источника, но упорядоченные способом, отличающимся от того, который обычно встречается в природе.

Термины "регуляторная последовательность" или "регуляторный элемент" используются взаимозаменяемо и относятся к нуклеотидным последовательностям, расположенным перед (5' некодирующие последовательности), внутри или после (3' некодирующие последовательности) кодирующей последовательности, которые влияют на транскрипцию, процессинг или стабильность РНК, или трансляцию связанной кодирующей последовательности. Регуляторные последовательности могут включать в себя, не ограничиваясь только перечисленным, промоторы, лидерные последовательности трансляции, интроны и последовательности распознавания полиаденилирования.

Термин "промотор" относится к фрагменту нуклеиновой кислоты, который способен контролировать транскрипцию другого фрагмента нуклеиновой кислоты. В некоторых вариантах осуществления настоящего изобретения промотор представляет собой промотор, способный контролировать транскрипцию гена в клетку, независимо от того, получен он из такой клетки или нет. Промотор может являться конститутивным промотором, или тканеспецифичным промотором, или промотором, регулируемым развитием, или индуцируемым промотором.

Термин "конститутивный промотор" относится к промотору, который в большинстве обстоятельств может вызвать экспрессию гена в большинстве типов клеток. Термины "тканеспецифичный промотор" и

"тканепредпочтительный промотор" используются взаимозаменяемо и относятся к промотору, который экспрессируется преимущественно, но не обязательно исключительно, в одну ткань или орган, но который также может быть экспрессирован в одну конкретную клетку или тип клетки. Термин "промотор, регулируемый развитием" относится к промотору, активность которого определяется событиями, связанными с развитием. "Индукцируемый промотор" селективно экспрессирует последовательность ДНК, функционально связанную с ним, в ответ на эндогенный или экзогенный стимул (среда, гормоны или химические сигналы и т.д.).

В данном документе термин "функционально связанный" означает, что регуляторный элемент (например, не ограничиваясь только перечисленным, последовательность промотора, последовательность терминации транскрипции и т.д.) связан с последовательностью нуклеиновой кислоты (такой как кодирующая последовательность или открытая рамка считывания), таким образом, что транскрипция нуклеотидной последовательности контролируется и регулируется транскрипционным регуляторным элементом. В данной области техники известны технологии для функционального связывания области регуляторного элемента с молекулой нуклеиновой кислоты.

"Введение" молекулы нуклеиновой кислоты (например, плазмиды, линейного фрагмента нуклеиновой кислоты, РНК и т.д.) или белка в организм означает, что нуклеиновая кислота или белок используются для трансформации клетки организма таким образом, что нуклеиновая кислота или белок функционирует в такой клетке. При использовании в настоящем изобретении, термин "трансформация" включает в себя как стабильные, так и транзистентные трансформации. "Стабильная трансформация" относится к введению в геном экзогенной нуклеотидной последовательности, обеспечивающему в результате устойчивое наследование инородных генов. После стабильной трансформации последовательность экзогенной нуклеиновой кислоты устойчиво интегрируется в геном организма и его любых последующих поколений. "Транзистентная трансформация" относится к введению молекулы нуклеиновой кислоты или белка в клетку, выполняющих свою функцию без устойчивого наследования экзогенного гена. При транзистентной трансформации последовательность экзогенной нуклеиновой кислоты не интегрируется в геном.

2) Система редактирования генома с высокой эффективностью и высокой специфичностью.

Как сообщалось, нуклеаза Cas9 в варианте eSpCas9(1.0) (K810A/K1003A/R1060A), eSpCas9(1.1) (K848A/K1003A/R1060A), разработанная Feng Zhang и др., и нуклеаза Cas9 в варианте SpCas9-HF1 (N497A/R661A/Q695A/Q926A), разработанная J. Keith Joung и др., способны существенно снизить уровень ненаправленности геномного редактирования, т.е. обладают высокой специфичностью. Однако неожиданно авторы настоящего изобретения обнаружили, что эти три варианта нуклеазы Cas9, обладая высокой специфичностью, имеют гораздо более низкую эффективность редактирования генов по сравнению с Cas9 дикого типа.

Авторы настоящего изобретения неожиданно обнаружили, что эффективность редактирования высокоспецифичным вариантом нуклеазы Cas9 может быть повышена путем слияния конца 5' направляющей РНК с tРНК, даже до уровня дикого типа при сохранении высокой специфичности.

Не рассматривая ограничения, налагаемые какой-либо теорией, предполагается, что снижение эффективности редактирования высокоспецифичных вариантов нуклеазы Cas9 связана с тем, может ли быть точно иницирована транскрипция направляющей РНК или нет. В данной области техники промоторы, обычно используемые для получения *in vivo* направляющей РНК, включают в себя, например, промоторы snРНК U6 или U3, для которых транскрипция управляется РНК-полимеразой III. Промотор U6 требуется для иницирования транскрипции в G, и, таким образом, для последовательностей-мишеней с первым нуклеотидом A, C или T, дополнительный G будет присутствовать на конце 5' транскрибируемой sgРНК. Промотор U3 иницирует транскрипцию в A, и, таким образом, для последовательностей-мишеней с первым нуклеотидом G, C или T, дополнительный A будет присутствовать на конце 5' транскрибируемой sgРНК. Авторы изобретения обнаружили, что эффективность редактирования высокоспецифичных вариантов нуклеазы Cas9 снижается в случае, если на конце 5' sgРНК присутствует дополнительный нуклеотид. sgРНК без дополнительного нуклеотида на конце 5' легко может быть получена даже с использованием промоторов U6 или U3 путем слияния транскрипции с tРНК, благодаря механизму точного процессинга tРНК (точному удалению дополнительной последовательности 5' и 3' прекурсора tРНК для образования зрелой tРНК), без необходимости учитывать тип первого нуклеотида последовательности-мишени. Таким образом, может быть улучшена эффективность редактирования высокоспецифичных вариантов нуклеазы Cas9, а также расширен диапазон выбора последовательностей-мишеней. Кроме того, не рассматривая ограничения, налагаемые какой-либо теорией, слияние с tРНК может повысить уровень экспрессии sgРНК, что также может способствовать улучшению эффективности редактирования высокоспецифичными вариантами нуклеазы Cas9.

Таким образом, настоящее изобретение предусматривает систему редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент, выбранный из следующих с i) по iii):

- i) вариант нуклеазы Cas9 и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние tРНК-направляющей РНК;
- ii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вари-

ант нуклеазы Cas9, и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние tРНК -направляющей РНК; и

iii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность, кодирующую слияние tРНК-направляющей РНК;

в котором вариант нуклеазы Cas9 имеет повышенную специфичность по сравнению с нуклеазой Cas9 дикого типа,

в котором конец 5' направляющей РНК связан с концом 3' tРНК,

в котором слияние расщепляется на конце 5' направляющей РНК после транскрибирования в клетку, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на конце 5'.

В некоторых вариантах осуществления tРНК и клетка, подлежащая модифицированию, относятся к одному виду.

В некоторых особых вариантах осуществления tРНК кодируется следующей последовательностью аасааагсaccagtggtctagtggtagaatagfaccctgccacggtagacaccgggttcgattcccggtgtgca (SEQ ID NO:1).

Дизайн слияния tРНК-направляющей РНК находится в компетенции специалистов в данной области техники. Например, можно сослаться на работу Xie и др., PNAS, Mar 17, 2015; vol. 112, no. 11, 3570-3575.

В настоящем изобретении также рассматривается слияние направляющей РНК и рибозима. На основании того, что по данному изобретению эффективность редактирования при помощи высокоспецифичных вариантов нуклеазы Cas9 связана с точностью инициации транскрипции sgРНК, используя способность рибозима разрезать РНК в конкретном сайте, sgРНК без дополнительного нуклеотида на конце 5' можно получать путем рационального конструирования слияния РНК и рибозима, что позволяет улучшить эффективность редактирования при сохранении высокой специфичности.

Таким образом, настоящее изобретение предусматривает систему редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент, выбранный из следующих с i) по iii):

i) вариант нуклеазы Cas9 и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК;

ii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК; и

iii) экспрессионный конструкт, содержащий нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность, кодирующую слияние рибозима-направляющей РНК;

в котором вариант нуклеазы Cas9 имеет повышенную специфичность по сравнению с нуклеазой Cas9 дикого типа,

в котором конец 5' направляющей РНК связан с концом 3' первого рибозима,

в котором первый рибозим предназначен для расщепления слияния на конце 5' направляющей РНК, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на конце 5'.

В одном варианте осуществления, где конец 3' направляющей РНК связан с концом 5' второго рибозима, такой второй рибозим предназначен для расщепления слияния на конце 3' направляющей РНК, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на конце 3'.

Дизайн первого рибозима или второго рибозима находится в компетенции специалиста в данной области техники. Например, можно сослаться на работу Gao и пр., JIPB, Apr, 2014; Vol 56, Issue 4, 343-349.

В одном особом варианте осуществления первый рибозим кодируется следующей последовательностью

5'-(N)₆CTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC-3' (SEQ ID NO: 12),

в которой N независимо выбран из A, G, C и T, а (N)₆ относится к последовательности, обратной комплементарной первым 6 нуклеотидам на конце 5' направляющей РНК. В одном особом варианте осуществления второй рибозим кодируется следующей последовательностью

5'-GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAA TGGGAC-3' (SEQ ID NO: 13).

Вариант нуклеазы Cas9 по данному изобретению, который обладает более высокой специфичностью по сравнению с нуклеазой Cas9 дикого типа, может быть получен из Cas9 различных видов, например, полученных из Cas9 *Streptococcus pyogenes* (SpCas9, нуклеотидная последовательность, показанная в SEQ ID NO: 2, аминокислотная последовательность, показанная в SEQ ID NO: 3).

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 представляет собой

SEQ ID NO: 2, которая содержит аминокислотную замену на позиции 855 в SEQ ID NO: 2. В некоторых особых вариантах осуществления аминокислотная замена на позиции 855 представляет собой K855A.

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 представляет собой вариант SEQ ID NO: 2, которая содержит аминокислотные замены на позициях 810, 1003 и 1060 в SEQ ID NO: 2. В некоторых особых вариантах осуществления аминокислотными заменами являются, соответственно, K810A, K1003A и R1060A.

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 представляет собой SEQ ID NO: 2, которая содержит аминокислотные замены на позициях 848, 1003 и 1060 в SEQ ID NO: 2. В некоторых особых вариантах осуществления аминокислотными заменами являются, соответственно, K848A, K1003A и R1060A.

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 представляет собой вариант SEQ ID NO: 2, которая содержит аминокислотные замены на позициях 611, 695 и 926 в SEQ ID NO: 2. В некоторых особых вариантах осуществления аминокислотными заменами являются, соответственно, R611A, Q695A и Q926A.

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 представляет собой вариант SEQ ID NO: 2, которая содержит аминокислотные замены на позициях 497, 611, 695 и 926 в SEQ ID NO: 2. В некоторых особых вариантах осуществления аминокислотными заменами являются, соответственно, N497A, R611A, Q695A и Q926A.

В некоторых особых вариантах осуществления изобретения вариант нуклеазы Cas9 содержит последовательность аминокислоты, показанную в SEQ ID NO: 4 (eSpCas9(1.0)), SEQ ID NO: 5 (eSpCas9(1.1)) или SEQ ID NO: 6 (SpCas9-HF1).

В некоторых вариантах осуществления изобретения вариант нуклеазы Cas9 по данному изобретению дополнительно содержит последовательность ядерной локализации (NLS). Как правило, одна или более NLS в варианте нуклеазы Cas9 должны обладать достаточным количественным содержанием, чтобы управлять накоплением варианта нуклеазы Cas9 в ядре клетки в количестве, достаточном для функции редактирования генома. Как правило, количественная активность ядерной локализации определяется количеством и положением последовательностей ядерной локализации (NLS), и в варианте нуклеазы Cas9 используется одна или несколько NLS, или их комбинация.

В некоторых вариантах осуществления изобретения последовательности NLS варианта нуклеазы Cas9 по данному изобретению могут быть расположены в N-конце и/или в C-конце. В некоторых вариантах осуществления вариант нуклеазы Cas9 содержит около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более NLS. В некоторых вариантах осуществления вариант нуклеазы Cas9 содержит около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более NLS на N-конце или в его окрестности. В некоторых вариантах осуществления вариант нуклеазы Cas9 содержит около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более NLS на C-конце или в его окрестности. В некоторых вариантах осуществления вариант нуклеазы Cas9 содержит их комбинации, такие как одна или более NLS на N-конце или одна или более NLS на C-конце. Если имеется более одной NLS, каждая NLS может быть выбрана независимо от других NLS. В некоторых предпочтительных вариантах осуществления изобретения вариант нуклеазы Cas9 содержит две NLS, например, эти две NLS расположены на N-конце и C-конце, соответственно.

В целом, NLS состоит из одной или более коротких последовательностей положительно заряженного лизина или аргинина, находящихся на поверхности белка, но в данной области техники известны также и другие типы NLS. Неограничивающие примеры NLS включают в себя KKRRKV (нуклеотидная последовательность 5'-AAGAAGAGAAAGGTC-3'), PKKKRKV (нуклеотидная последовательность 5'-CCCAAGAAGAAGAGGAAGGTG-3' или CCAAAGAAGAAGAGGAAGGTT) или SGGSPKKRKV (нуклеотидная последовательность 5'-TCGGGGGGGAGCCCAAAGAAGAAGCGGAAGGTG-3').

В некоторых вариантах осуществления изобретения N-конец варианта нуклеазы Cas9 содержит NLS с аминокислотной последовательностью, показанной как PKKKRKV.

В некоторых вариантах осуществления изобретения C-конец варианта нуклеазы Cas9 содержит NLS с аминокислотной последовательностью, показанной как SGGSPKKRKV.

Кроме того, вариант нуклеазы Cas9 по настоящему изобретению также может включать в себя другие последовательности локализации, такие как последовательности цитоплазматической локализации, последовательности локализации хлоропластов, последовательности митохондриальной локализации и тому подобное, в зависимости от расположения редактируемой ДНК.

Для получения эффективной экспрессии в клетке-мишени, в некоторых вариантах осуществления изобретения нуклеотидная последовательность, кодирующая вариант нуклеазы Cas9, кодон-оптимизирована для организма, из которого происходит клетка, подлежащая редактированию генома.

Кодон-оптимизация относится к процессу модификации последовательности нуклеиновой кислоты для усиления экспрессии в представляющих интерес клетках-хозяевах путем замены по меньшей мере одного кодона (например, приблизительно или более 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 или более кодонов) нативной последовательности с кодонами, которые достаточно часто или чаще всего используются в генах этой клетки-хозяина при обеспечении нативной аминокислотной последовательности. Различные виды прояв-

ляют особое предпочтение к определенным кодомам конкретной аминокислоты. Предпочтение к кодону (отличия в использовании кодонов между организмами) часто коррелирует с эффективностью трансляции информационной РНК (мРНК), которая, в свою очередь, как полагают, зависит, помимо прочего, от свойств транскрибуемых кодонов и наличия молекул конкретной транспортной РНК (тРНК). Преобладание выбранных тРНК в клетке, как правило, является отражением кодонов, наиболее часто используемых при синтезе пептидов. Соответственно, гены могут быть адаптированы для оптимальной экспрессии гена в данном организме на основе кодон-оптимизации. Таблицы использования кодонов доступны, например, в "Базе данных использования кодонов", которая находится по адресу www.kazusa.or.jp/codon/, и эти таблицы могут быть адаптированы различными способами. См. Nakamura, Y., и др. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" *Nucl. Acids Res.* 28:292 (2000).

Организм, из которого происходит клетка, геном которой может быть отредактирован с помощью системы по настоящему изобретению, может быть, не ограничиваясь только перечисленным, млекопитающим, таким как человек, мышь, крыса, обезьяна, собака, свинья, овца, корова и кошка; домашней птицей, такой как курица, утка и гусь; растением, включая однодольные и двудольные, например, рис, кукуруза, пшеница, сорго, ячмень, соя, арахис, *Arabidopsis thaliana* и тому подобное.

В некоторых особых вариантах осуществления изобретения кодон-оптимизированной нуклеотидной последовательностью, кодирующей вариант нуклеазы Cas9, является последовательность, показанная в SEQ ID NO: 7 (eSpCas9(1.0)), SEQ ID NO: 8 (eSpCas9(1.1)) или SEQ ID NO: 9 (SpCas9-HF1).

В некоторых вариантах осуществления изобретения направляющая РНК представляет собой одиночную направляющую РНК (sgРНК). В данной области техники известны способы конструирования подходящих sgРНК в соответствии с заданной последовательностью-мишенью. См., например, Wang, Y. и др. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947-951 (2014); Shan, Q. и др. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 686-688 (2013); Liang, Z. и др. Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J Genet Genomics.* 41, 63-68 (2014).

В некоторых вариантах осуществления изобретения нуклеотидная последовательность, кодирующая вариант нуклеазы Cas9, и/или нуклеотидная последовательность, кодирующая слияние направляющей РНК, функционально связана с элементом, регулирующим экспрессию, таким как промотор.

Примеры промоторов, которые можно использовать в настоящем изобретении, включают в себя, не ограничиваясь только перечисленным, промоторы полимеразы (pol) I, pol II или pol III. Примеры промоторов pol I включают в себя промотор pol I РНК курицы. Примеры промоторов pol II включают в себя, не ограничиваясь только перечисленным, немедленно-ранний промотор цитомегаловируса (CMV), промотор с длинным концевым повтором вируса саркомы Рауса (RSV-LTR) и немедленно-ранний промотор вируса обезьян 40 (SV40). Примеры промоторов pol III включают в себя промоторы U6 и H1. Может быть использован индуцируемый промотор, такой как металлотионеин. Другие примеры промоторов включают в себя промотор бактериофага T7, промотор бактериофага T3, промотор (3-галактозидазы, промотор бактериофага Sp6 и т.д. При использовании для растений можно применять промоторы, которые включают в себя, не ограничиваясь только перечисленным, промотор 35S вируса мозаики цветной капусты, промотор Ubi-1 кукурузы, промотор пшеницы U6, промотор риса U3, промотор кукурузы U3, промотор актина риса и т.д.

3) Способ генетического модифицирования клетки.

Еще один объект изобретения предусматривает способ для генетического модифицирования клетки, содержащий: введение в клетку системы редактирования генома по настоящему изобретению, посредством которого вариант нуклеазы Cas9 нацеливается на последовательность-мишень в геноме клетки при помощи направляющей РНК, а результат заключается в замене, делеции и/или добавлении одного или более нуклеотидов в последовательность-мишень.

Разработка последовательности-мишени, которая может быть распознана и обеспечить нацеливание комплексом Cas9 и направляющей РНК, находится в пределах технических навыков специалиста средней квалификации в данной области. Как правило, последовательность-мишень представляет собой последовательность, которая является комплементарной лидерной последовательности из примерно 20 нуклеотидов, содержащихся в направляющей РНК, и конец 3' которой является смежным с мотивом, примыкающим к протоспейсеру (PAM) NGG.

Например, в некоторых вариантах осуществления изобретения последовательность-мишень имеет структуру: 5'-N_x-NGG-3', в которой N выбирается независимо от А, G, С и Т; X является целым в пределах 14≤X≤30; N_x представляет X перекрывающихся нуклеотидов, а NGG является последовательностью мотива, примыкающего к протоспейсеру (PAM). В некоторых особых вариантах осуществления изобретения, X равно 20.

По настоящему изобретению модифицируемая последовательность-мишень может быть расположена в любом месте генома, например, внутри функционального гена, такого как ген, кодирующий белок, или, например, может быть расположена в регуляторной области экспрессии гена, такой как область промотора или область энхансера, и, таким образом, может выполнить функциональную модификацию указанного гена или выполнить модификацию экспрессии гена.

Замена, делеция и/или добавление в целевой последовательности клетки могут быть обнаружены с помощью T7E1, PCR/RE или способов секвенирования, см., например, Shan Q., Wang Y., Li, J. & Gao, C. Genome editing in rice and wheat using the CRISPR/Cas system. Nat. Protoc. 9, 2395-2410 (2014).

В способе по настоящему изобретению система редактирования генома может быть введена в клетку с использованием различных способов, хорошо известных специалистам в данной области техники.

Способы введения системы редактирования генома в клетку по настоящему изобретению включают в себя, не ограничиваясь только перечисленным, трансфекцию фосфата кальция, слияние протопластов, электропорацию, трансфекцию липосом, микроинъекцию, вирусную инфекцию (такую как бакуловирус, вирус коровьей оспы, аденовирус и другие вирусы), бомбардировку частицами, ПЭГ-опосредованную трансформацию протопласта или трансформацию посредством агробактерий.

Клетка, геном которой может быть отредактирован с помощью системы по настоящему изобретению, может принадлежать, например, млекопитающим, таким как человек, мышь, крыса, обезьяна, собака, свинья, овца, корова и кошка; домашней птице, такой как курица, утка и гусь; и растениям, включая однодольные и двудольные, такие как рис, кукуруза, пшеница, сорго, ячмень, соя, арахис, *Arabidopsis thaliana* и тому подобное.

В некоторых вариантах осуществления способ по настоящему изобретению выполняется *in vitro*. Например, такая клетка является изолированной клеткой. В некоторых других вариантах осуществления способ по настоящему изобретению может быть выполнен *in vivo*. Например, такая клетка представляет собой клетку в организме, и система по настоящему изобретению может быть введена в указанную клетку *in vivo* с использованием, например, вирус-опосредованного способа. В некоторых вариантах осуществления такая клетка представляет собой зародышевую клетку. В некоторых вариантах осуществления такая клетка является соматической клеткой.

Другой объект настоящего изобретения дополнительно предусматривает генетически модифицированный организм, который содержит генетически модифицированную клетку, полученную способом по настоящему изобретению.

Такие организмы включают в себя, не ограничиваясь только перечисленным, млекопитающих, таких как человек, мышь, крыса, обезьяна, собака, свинья, овца, корова и кошка; домашнюю птицу, такую как курица, утка, и гусь; растения, включающие в себя однодольные и двудольные растения, такие как рис, кукуруза, пшеница, сорго, ячмень, соя, арахис и *Arabidopsis thaliana*.

Примеры

Материалы и способы.

Конструирование бинарных векторов экспрессии pJIT163-SpCas9, pJIT163-eSpCas9(1.0), pJIT163-eSpCas9(1.1) и pJIT163-SpCas9-HF1.

Последовательности SpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 были кодон-оптимизированы для риса, SpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 были получены путем сайт-направленного мутагенеза с использованием системы Fast MultiSite Metagenesis System (TransGen) с плазмидой pJIT163-SpCas9 (SEQ ID NO: 10) в качестве матрицы.

Конструирование вектора экспрессии sgPHK.

Последовательности-мишени sgPHK, использованные в экспериментах, показаны в табл. 1 далее.

Таблица 1. Целевой ген и последовательность-мишень sgPHK

sgPHK	Последовательность-мишень	Олиго-F	Олиго-R
OsCDKB2	AGGTCGGGGAGGGGACGTACGGG	GGCAAGGTCGGGGAGGGGACGTAC	AAACGTACGTCCCTCCCGACCT
OsMKK4	GACGTCGGCGAGGAAGGCCTCGG	GGCAGACGTCGGCGAGGAAGGCCT	AAACAGGCCTTCCTCGCCGACGTC
A1	CATGGTGGGGAAAGCTTGAGGG	GGCACATGGTGGGGAAAGCTTGGA	AAACTCCAAGCTTCCCAACCATG
A2	CCGGACGACGACGTGACGACGG	GGCACGGACGACGACGTGACGCA	AAACTCGTCGACGTGTCGTCCGG
A3	TTGAAGTCCCTTCTAGATGGAGG	GGCATTGAAGTCCCTTCTAGATGG	AAACCCATCTAGAAGGGACTTCAA
A4	ACTGCGACACCCAGATATCGTGG	GGCAACTGCGACACCCAGATATCG	AAACCGATATCTGGGTGTCGAGT
PDS	GTTGGTCTTTGCTCCTGCAGAGG	GGCAGTTGGTCTTTGCTCCTGCAG	AAACCTGCAGGAGCAAAGACCAAC

Векторы экспрессии sgPHK: pOsU3-CDKB2-sgPHK, pOsU3-MKK4-sgPHK, pOsU3-A1-sgPHK, а также pOsU3-A2-sgPHK, pOsU3-A3-sgPHK, pOsU3-A4-sgPHK и pOsU3-PDS-sgPHK сконструированы на основе pOsU3-sgPHK(Addgene ID53063), как описывалось ранее (Shan, Q. и др. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31, 686-688, 2013).

Конструирование векторов экспрессии tPHK-sgPHK.

Векторы экспрессии tPHK-sgPHK конструируются на основе вектора pUC57-U3-tPHK-sgPHK. (SEQ ID NO: 11, фиг. 6). Линейный вектор получают после расщепления pUC57-U3-tPHK-sgPHK с BsaI, соответствующие олиго-F и олиго-R гибридизированы и соединены в линейный вектор, а последующие этапы аналогичны конструированию векторов экспрессии sgPHK.

Таблица 2. Целевые гены и олигонуклеотидные последовательности для конструирования векторов экспрессии tPHK-sgPHK

sgPHK	Последовательность-мишень	Олиго-F	Олиго-R
OsCDKB2	AGGTCGGGGAGGGGACGTACGGG	TGCAAGGTCGGGGAGGGGACGTAC	AAACGTACGTCCCCTCCCCGACCT
OsMKK4	GACGTCGGCGAGGAAGGCCTCGG	TGCAGACGTCGGCGAGGAAGGCCT	AAACAGGCCTTCCTCGCCGACGTC
A1	CATGGTGGGGAAGCTTGGAGGG	TGCACATGGTGGGGAAGCTTGGGA	AAACTCCAAGCTTCCCCACCATG
A2	CCGGACGACGACGTCGACGACGG	TGCACCGGACGACGACGTCGACGA	AAACTCGTCGACGTCGTCGTCCGG
A3	TTGAAGTCCCTTCTAGATGGAGG	TGCATTGAAGTCCCTTCTAGATGG	AAACCCATCTAGAAGGGACTTCAA
A4	ACTGCGACACCCAGATATCGTGG	TGCACTGCGACACCCAGATATCG	AAACCGATATCTGGGTGTCGCAGT
PDS	GTTGGTCTTTGCTCCTGCAGAGG	TGCAGTTGGTCTTTGCTCCTGCAG	AAACCTGCAGGAGCAAAGACCAAC

Анализ протопласта.

В исследовании использован сорт риса "nipponbare". Трансформация протопластов осуществляется как описано ниже. Трансформация осуществляется для 10 мкг каждой плазмиды путем трансфекции, опосредованной полиэтиленгликолем. Протопласты были собраны после 48 ч, а ДНК была извлечена в ходе анализа PCR-RE.

Приготовление и трансформация рисового протопласта.

1) Для изоляции протопластов использовалась оболочка листьев рассады, прорезанная острым лезвием на части шириной 0,5 мм.

2) Сразу после разреза образец был перенесен в 0,6 М раствор маннитола и помещен в темноту на 10 мин.

3) Раствор маннитола был удален путем фильтрации, а продукты были перенесены в раствор для ферментализации и подвергнуты вакуумированию в течение 30 мин.

4) Ферментализация проводилась в течение 5-6 часов, в темноте и при осторожном встряхивании (обесцвечивание в шейкере, скорость 10).

5) По завершении ферментализации был добавлен равный объем W5 и выполнено горизонтальное встряхивание в течение 10 с, чтобы высвободить протопласты.

6) Протопласты были отфильтрованы в пробирку для центрифуги с круглым дном объемом 50 мл с нейлоновой мембраной 40 мкм и промыты раствором W5.

7) Для осаждения протопластов была выполнена горизонтальная обработка на центрифуге при 250 g в течение 3 мин с удалением супернатанта.

8) Протопласты были повторно приведены во взвешенное состояние путем добавления 10 мл W5, а затем обработаны на центрифуге при 250g в течение 3 мин с удалением супернатанта.

9) Соответствующее количество раствора MMG было добавлено для повторного приведения протопластов во взвешенное состояние с концентрацией 2×10^6 /мл.

Примечание: все предыдущие этапы осуществлялись при комнатной температуре.

10) 10-20 мкг плазмид, 200 мкл протопластов (около 4×10^5 клеток) и 220 мкг свежего раствора ПЭГ были добавлены в пробирку для центрифуги объемом 2 мл, перемешаны и помещены при комнатной температуре в темноту на 10-20 мин, чтобы вызвать трансформацию.

11) После завершения трансформации были медленно добавлены 880 мкл раствора W5, для перемешивания пробирки осторожно переворачивали вверх дном, а затем обрабатывали на центрифуге при 250 g в течение 3 мин с удалением супернатанта.

12) Полученные продукты были повторно приведены во взвешенное состояние путем добавления 2 мл раствора WI, перенесены на планшет с шестью ячейками и культивированы при комнатной температуре (или 25°C) в темноте. Для выделения геномной ДНК протопласта продукты необходимо культивировать в течение 48 часов.

Идентификация мутаций с помощью глубокого секвенирования.

Анализ методом глубокого секвенирования выполняется с применением методики Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., Liu, J., Zhang, H., Liu, C., Ran, Y., и др. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nature Communications 8, 14261.

Пример 1. Сравнение возможностей редактирования WT SpCas9 и ее вариантов для сайтов-мишеней.

WT SpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 были соответственно сконструированы в транзистентный вектор экспрессии pJIT163, а экспрессии WT SpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 управляются промотором гена убиквитина кукурузы. sgPHK были сконструированы в векторе pOsU3-sgPHK, а экспрессия sgPHK управляется промотором OsU3. Протопласты риса были трансформированы, и ДНК протопласта экстрагировали для анализа PCR-RE с целью оценки эффективности мутации. Для сравнения различий в способности к редактированию SpCas9 дикого типа и eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 выбраны пять целевых сайтов (A1, A2, A3, A4 и PDS, см. фиг. 2 и 3).

Транскрипцию с нуклеотидом А должен инициировать промотор OsU3, следовательно, конструи-

рование векторов экспрессии sgPHK для сайтов-мишеней можно разделить в соответствии с двумя условиями следующим образом.

(1) Если первый нуклеотид на конце 5' требуемой последовательности-мишени sgPHK (20 п.о.) представляет собой любой из G/T/C, поскольку промотор U3 инициирует транскрипцию с А, на конец 5' транскрибируемой sgPHK будет добавлен дополнительный А, и, кроме того, транскрибируемая sgPHK может не полностью совпадать с последовательностью-мишенью. Вектор экспрессии sgPHK может быть сконструирован как U3+AN₂₀ на фиг. 1, где N₂₀ представляет собой последовательность-мишень, А - это дополнительный нуклеотид на конце 5'.

(2) Если первым нуклеотидом на конце 5' требуемой последовательности-мишени sgPHK (20 п.о.) является А, он может быть использован промотором U3 для инициирования транскрипции, и, следовательно, на конце 5' транскрибируемой sgPHK не будет иметься никаких дополнительных нуклеотидов. Вектор экспрессии sgPHK может быть сконструирован как U3+AN₁₉ на фиг. 1, где AN₁₉ представляет собой последовательность-мишень.

Выбранные сайты-мишени A1, A2, A3 и PDS относятся к сайтам-мишеням класса (1), а сайт-мишень A4 относится к сайтам-мишеням класса (2).

Результаты эксперимента показывают (фиг. 2), что эффективность редактирования eSpCas9 (1.0), eSpCas9 (1.1) и SpCas9-HF1 для сайтов-мишеней класса (1) чрезвычайно низка. Отличия в эффективности редактирования eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 и эффективности редактирования WT SpCas9 не являются значительными для сайтов-мишеней класса (2). Это показывает, что дополнительный нуклеотид на конце 5' sgPHK, полученный в результате транскрипции, может снизить эффективность редактирования eSpCas9 (1.0), eSpCas9 (1.1) и SpCas9-HF1.

Аналогично промотору OsU3, промотор U6 кукурузы (TaU6) должен инициировать транскрипцию с нуклеотидом G, и, следовательно, конструирование векторов экспрессии sgPHK для сайтов-мишеней можно разделить в соответствии с двумя условиями следующим образом.

(1) Если первый нуклеотид на конце 5' требуемой последовательности-мишени sgPHK (20 п.о.) представляет собой любой из A/T/C, поскольку промотор U6 инициирует транскрипцию с G, на конец 5' транскрибируемой sgPHK будет добавлен дополнительный нуклеотид G и, кроме того, транскрибируемая sgPHK может не полностью совпадать с последовательностью-мишенью.

(2) Если первым нуклеотидом на конце 5' требуемой последовательности-мишени sgPHK (20 п.о.) является G, он может быть использован промотором U6 для инициирования транскрипции, при этом на конце 5' транскрибируемой sgPHK не имеется никаких дополнительных нуклеотидов.

Сайт-мишень OsPDS относится к сайтам-мишеням класса (2). Управление транскрипцией sgPHK GN₁₉ и GN₂₀ для сайта-мишени OsPDS осуществлялось при помощи промотора TaU6, где GN₂₀ может имитировать сайты-мишени класса (1), а именно сайты с дополнительным нуклеотидом G на конце 5' sgPHK.

Таблица 3. Целевой ген и олигонуклеотидные последовательности для конструирования векторов экспрессии TaU6-sgPHK

sgPHK	Последовательность-мишень	Олиго-F	Олиго-R
OsPDS-GN ₁₉	GTTGGTCTTTGCTCCTGCAGAGG	GGCGTTGGTCTTTGCTCCTGCAG	AAACCTGCAGGAGCAAGACCAA
OsPDS-GN ₂₀	GTTGGTCTTTGCTCCTGCAGAGG	GGCGGTTGGTCTTTGCTCCTGCAG	AAACCTGCAGGAGCAAGACCAAC

Результаты показывают (фиг. 2), что один дополнительный нуклеотид G на конце 5' sgPHK значительно снижает эффективность редактирования eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1.

Пример 2. Увеличение эффективности редактирования вариантов Cas9 путем слияния tPHK-sgPHK.

В соответствии с результатом примера 1, важным фактором, влияющим на эффективность редактирования eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1, является то, точно ли инициируется sgPHK или нет.

Согласно предыдущему отчету, слияние tPHK с концом 5' sgPHK может активировать экспрессию sgPHK и приводить к точному расщеплению на конце 5' sgPHK, и, таким образом, избегать наличия дополнительного нуклеотида на конце 5' sgPHK. (см. Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci USA. 2015 Mar 17; 112(11):3570-5. doi: 10.1073/pnas.1420294112. Epub 2015 Mar 2.) sgPHK для каждого сайта-мишени в примере 1 была слита с tPHK и экспрессирована под контролем промотора OsU3. Эксперименты были проведены по способу Примера 1 с tPHK-sgPHK вместо sgPHK. Как показано на фиг. 2, для сайтов-мишеней A1, A2, A3 и PDS эффективность редактирования eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 значительно повышается с использованием tPHK-sgPHK вместо sgPHK.

Пример 3. Влияние слияния tPHK-sgPHK на специфичность редактирования вариантов Cas9 3.1 Сайт-мишень OsMCK4 риса.

Для конструирования sgPHK и tPHK-sgPHK был выбран сайт-мишень GACGTCGGCGAG-GAAGGCCTCGG в гене риса MCK4. Этот сайт-мишень имеет два возможных нецелевых сайта, как показано на фиг. 5. Вектор для экспрессирования sgPHK или tPHK-sgPHK и векторы для экспрессирования WTSpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 были соответствующим образом совместно транс-

формированы в протопласты риса. Через два дня после трансформации был извлечен протопласт ДНК, а геномные фрагменты сайта-мишени и нецелевых сайтов были амплифицированы с использованием специфичных праймеров. Частота мутаций трех этих сайтов была проанализирована с использованием технологии секвенирования второго поколения.

Результаты эксперимента показаны на фиг. 5.

Когда были использованы sgPНК, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 по сравнению с WT SpCas9 демонстрируют чрезвычайно низкий нецелевой эффект, но имеют значительно более низкую эффективность редактирования.

Когда были использованы tPНК-sgPНК, эффективность редактирования каждой группы была увеличена, однако, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 способны обеспечить относительно высокую специфичность. В частности, для SpCas9-HF1 могут быть обнаружены только две крайне низкоуровневых мутации для обоих нецелевых сайтов. Следовательно, комбинация tPНК-sgPНК и SpCas9-HF1 особенно подходит для редактирования генома с высокой эффективностью и высокой специфичностью.

3.2 Сайт-мишень OsCDKB2 риса.

Для конструирования sgPНК был выбран сайт-мишень AGGTCGCGGGAGGGGACGTACGGG в гене риса OsCDKB2. Этот сайт-мишень имеет три возможных нецелевых сайта, как показано на фиг. 6. Вектор для экспрессирования sgPНК или tPНК-sgPНК и векторы для экспрессирования WTSpCas9, eSpCas9(1.0), eSpCas9(1.1) и SpCas9-HF1 были соответствующим образом совместно трансформированы в протопласты риса. Через два дня после трансформации был извлечен протопласт ДНК, а геномные фрагменты сайта-мишени и нецелевых сайтов были амплифицированы с использованием специфичных праймеров. Частота мутаций четырех этих сайтов была проанализирована с использованием технологии глубокого секвенирования.

Результаты эксперимента показаны на фиг. 6. Эффективность редактирования eSpCas9 (1.0), eSpCas9 (1.1) и SpCas9-HF1 сайтов-мишеней заметно повышается при использовании tPНК-sgPНК вместо sgPНК. В частности, эффективность редактирования SpCas9-HF1 может быть восстановлена до уровня дикого типа с обеспечением высокой специфичности. Поскольку эта последовательность-мишень начинается с нуклеотида А, с помощью которого промотор U3 может точно инициировать транскрипцию, повышенная эффективность редактирования может быть результатом повышенного уровня экспрессии sgPНК из-за слияния с tPНК.

Пример 4. Специфичность редактирования вариантов Cas9 по несоответствию между gPНК и последовательностью-мишенью.

При конструировании sgPНК для сайта-мишени GACGTCCGCGAGGAAGGCCCTCGG в гене риса МКК4, были искусственно введены несовпадения двух смежных оснований (пурин для пурина и пиримидин для пиримидина). Было выявлено редактирование при условии, что sgPНК может не полностью соответствовать сайту-мишени. Если редактирование может быть обнаружено, оно считается нецелевым. Эксперименты проводились способом, аналогичным описанному в примере 3.1.

Результаты эксперимента показаны на фиг. 7. При использовании tPНК-sgPНК варианты SpCas9 показали более высокую чувствительность к несоответствиям между gPНК и последовательностью-мишенью (в особенности, когда несоответствие ближе к одному из концов).

Пример 5. Эффективность и специфичность редактирования вариантов Cas9 на почке эмбриона человека, 293 клетки.

sgPНК были сконструированы для последовательности-мишени CGTGAGTGCAGTGTGTCCGTTGTGG в гене человека VEGFA. U6:sgPНК-GN₁₉ и U6:tPНК-sgPНК-N₂₀ показывают, что sgPНК, транскрибированные при помощи промотора U6, имеют длину 20 нуклеотидов и полностью совпадают с последовательностью-мишенью; U6:sgPНК-GN₂₀ показывает, что sgPНК, транскрибированная при помощи промотора U6, имеет длину 21 нуклеотид и содержит дополнительный G на конце 5'.

Результаты исследования T7E1 показывают (фиг. 8), что WT Cas9 демонстрирует сходную эффективность редактирования при использовании разных стратегий транскрибирования sgPНК. Однако эффективность редактирования eSpCas9(1.1) и SpCas9-HF1 была значительно снижена в том случае, когда sgPНК содержит дополнительный нуклеотид на конце 5'. А при использовании слияний tPНК-sgPНК эффективность редактирования eSpCas9(1.1) и SpCas9-HF1 повышалась до уровня WT Cas9 или даже больше.

Что касается специфичности редактирования, WT Cas9 привела к нецелевому редактированию на обоих сайтах нецелевой 1 и нецелевой 2, а при использовании слияний tPНК-sgPНК eSpCas9(1.1) и SpCas9-HF1 не привели к нецелевому редактированию.

Таблица 4. Целевой ген и олигонуклеотидные последовательности
для конструирования векторов экспрессии sgРНК

sgРНК	Последовательность-мишень	Олиго-F	Олиго-R
VEGFA-GN ₁₉	GGTGAGTGAGTGTGT GCGTGTGG	CACCGGTGAGTGAGTGTGTG CGTG	AAACCACGCACACACTCACTC ACC
VEGFA-GN ₂₀	GGTGAGTGAGTGTGT GCGTGTGG	CACCGGTGAGTGAGTGTGTG GCGTG	AAACCACGCACACACTCACTC ACCC
VEGFA-tRNA-N ₂₀	GGTGAGTGAGTGTGT GCGTGTGG	CACCGaacaagcaccagtggtagtgtagg tagaatagtaccctgccacgtacagaccgg gttcgattcccggtggtgcaGGTGAGT GAGTGTGTGCGTG	AAACCACGCACACACTCACTC ACCTgcaccagccgggaatgaacccgggtc tgtaccgtggcagggtactattctaccactagacc actggtgcttgttC

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Система редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент из следующих с i) по iii):

i) вариант нуклеазы Cas9 и экспрессионная конструкция, содержащая нуклеотидную последовательность РНК, кодирующую слияние tРНК и направляющей РНК;

ii) экспрессионная конструкция, содержащая нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и экспрессионная конструкция, содержащая нуклеотидную последовательность РНК, кодирующую слияние tРНК и направляющей РНК; и

iii) экспрессионная конструкция, содержащая нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность РНК, кодирующую слияние tРНК и направляющей РНК;

при этом вариант нуклеазы Cas9 имеет повышенную специфичность редактирования по сравнению с нуклеазой Cas9 дикого типа,

при этом 5'-конец направляющей РНК соединен с 3'-концом tРНК,

при этом РНК расщепляется на 5'-конце направляющей РНК после ее транскрибирования в клетке, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на 5'-конце.

2. Система редактирования генома для сайт-направленной модификации последовательности-мишени в геноме клетки, которая содержит по меньшей мере один элемент из следующих с i) по iii):

i) вариант нуклеазы Cas9 и экспрессионная конструкция, содержащая нуклеотидную последовательность РНК, кодирующую слияние рибозима и направляющей РНК;

ii) экспрессионная конструкция, содержащая нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и экспрессионная конструкция, содержащая нуклеотидную последовательность РНК, кодирующую слияние рибозима и направляющей РНК; и

iii) экспрессионная конструкция, содержащая нуклеотидную последовательность, кодирующую вариант нуклеазы Cas9, и нуклеотидную последовательность, кодирующую слияние рибозима и направляющей РНК;

при этом вариант нуклеазы Cas9 имеет повышенную специфичность редактирования по сравнению с нуклеазой Cas9 дикого типа,

при этом 5'-конец направляющей РНК соединен с 3'-концом первого рибозима,

при этом первый рибозим предназначен для расщепления РНК на 5'-конце направляющей РНК, вследствие чего формируется направляющая РНК, которая не несет дополнительный нуклеотид на 5'-конце.

3. Система по п.1, в которой tРНК и клетка, подлежащая модифицированию, получены от одного вида.

4. Система по п.1, в которой tРНК кодируется последовательностью, показанной в SEQ ID NO: 1.

5. Система по п.1 или 2, в которой вариант нуклеазы Cas9 содержит последовательность аминокислоты, показанную в SEQ ID NO: 4, SEQ ID NO: 5 или SEQ ID NO: 6.

6. Система по п.1 или 2, в которой нуклеотидная последовательность, кодирующая вариант нуклеазы Cas9, является кодон-оптимизированной для организма, из которого происходит клетка, подлежащая модифицированию.

7. Система по п.1 или 2, в которой направляющая РНК является одиночной направляющей РНК (sgРНК).

8. Система по п.1 или 2, в которой нуклеотидная последовательность, кодирующая слитую направляющую РНК, функционально связана с промотором pol III, при этом промотор pol III выбран из группы, состоящей из промотора U3, промотора U6 или промотора H1.

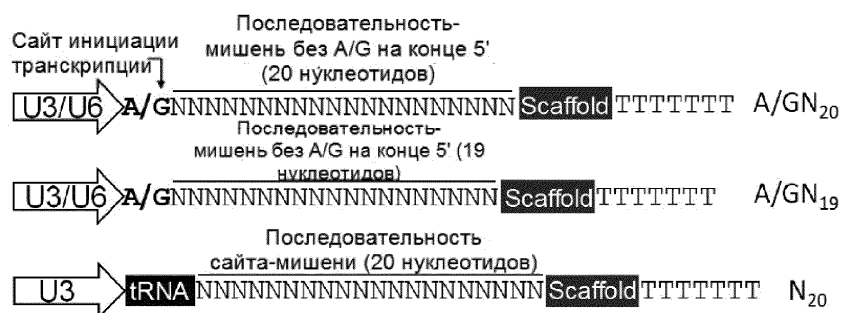
9. Способ генетического модифицирования клетки, содержащий введение в клетку системы по любому из пп.1-8, посредством которого вариант нуклеазы Cas9 нацеливается на последовательность-

мишень в геноме клетки при помощи направляющей РНК, а результат заключается в замене, делеции и/или добавлении одного или более нуклеотидов в последовательность-мишень.

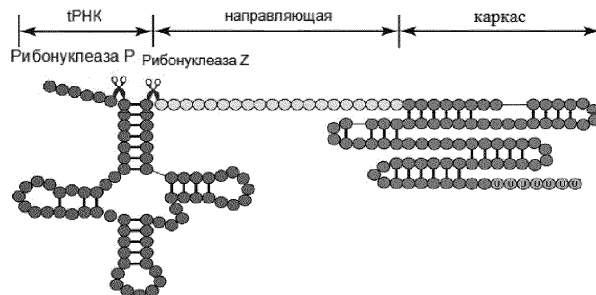
10. Способ по п.9, в котором такая клетка получена от млекопитающих, таких как человек, мышь, крыса, обезьяна, собака, свинья, овца, корова, кошка; домашней птицы, такой как курица, утка, гусь; растений, включающих в себя однодольные и двудольные растения, такие как рис, кукуруза, пшеница, сорго, ячмень, соя, арахис и *Arabidopsis thaliana*.

11. Способ по любому из пп.9-10, в котором система введена в клетку способом, выбранным из следующего: трансфекция фосфата кальция, слияние протопластов, электропорация, трансфекция липосом, микроинъекция, вирусная инфекция (такая как бакуловирус, вирус коровьей оспы, аденовирус и другие вирусы), бомбардировка частицами, ПЭГ-опосредованная трансформация протопласта или трансформация посредством агробактерий.

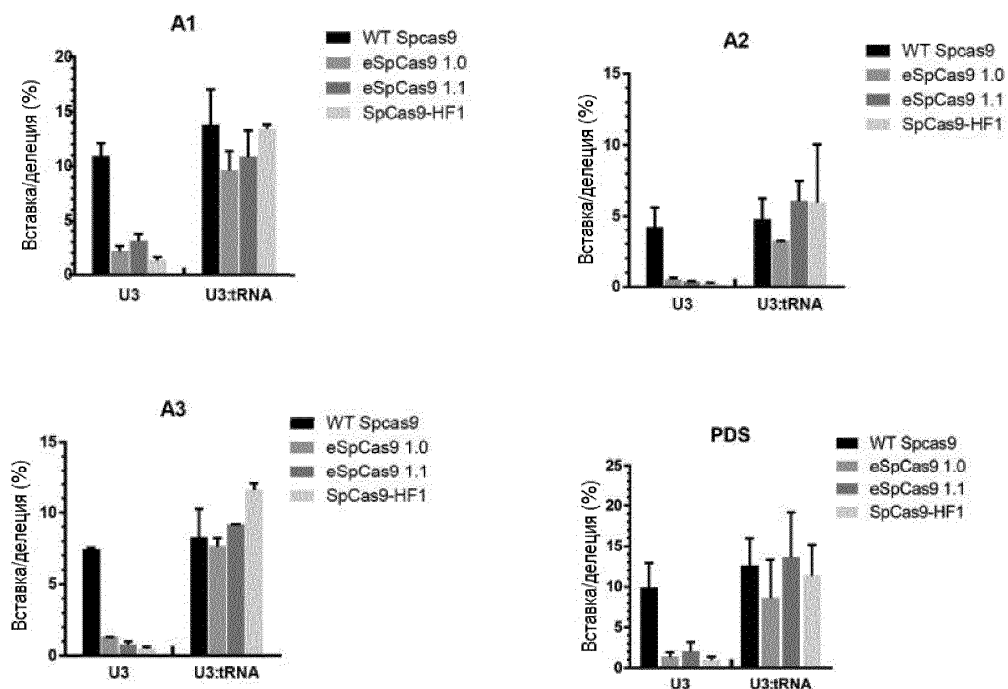
A



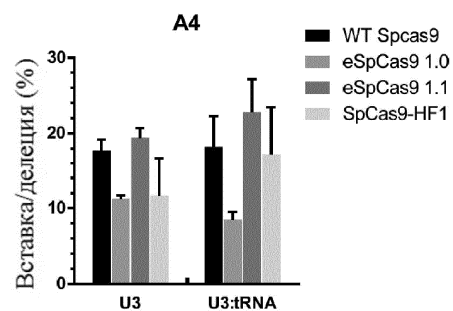
B



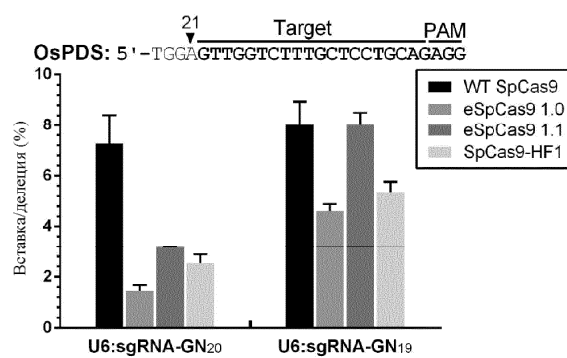
Фиг. 1



Фиг. 2

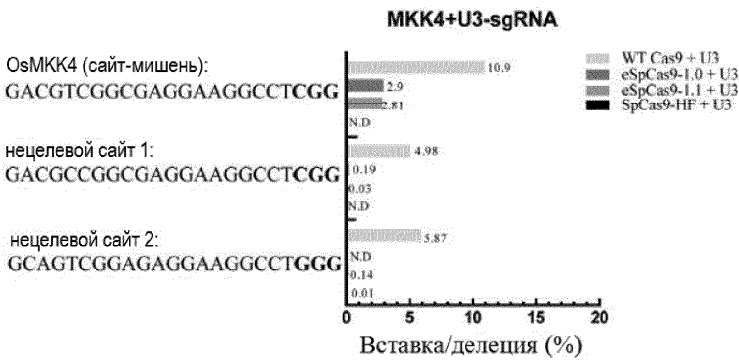


Фиг. 3

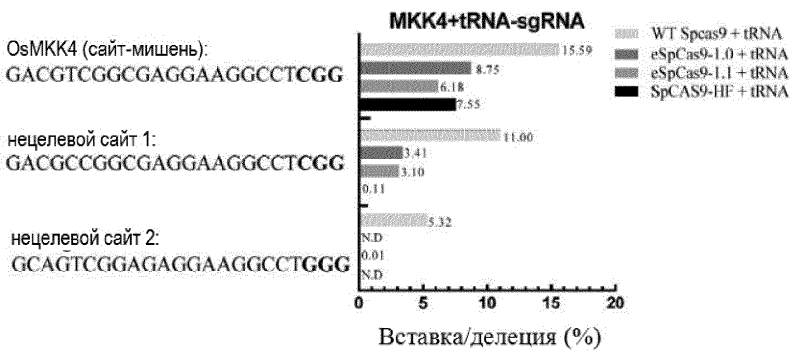


Фиг. 4

A

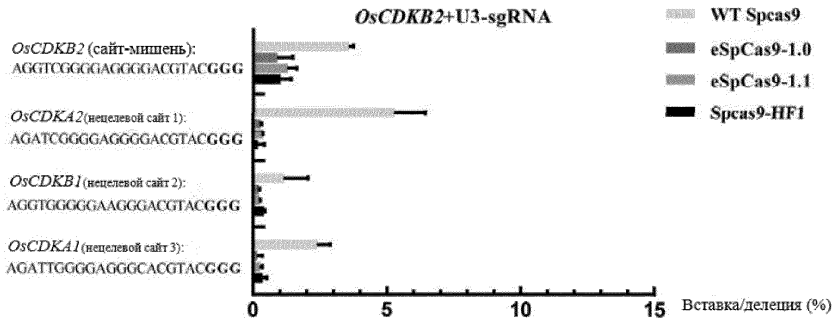


B

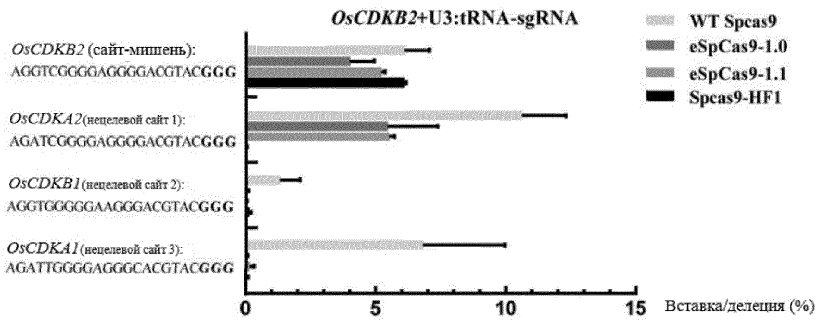


Фиг. 5

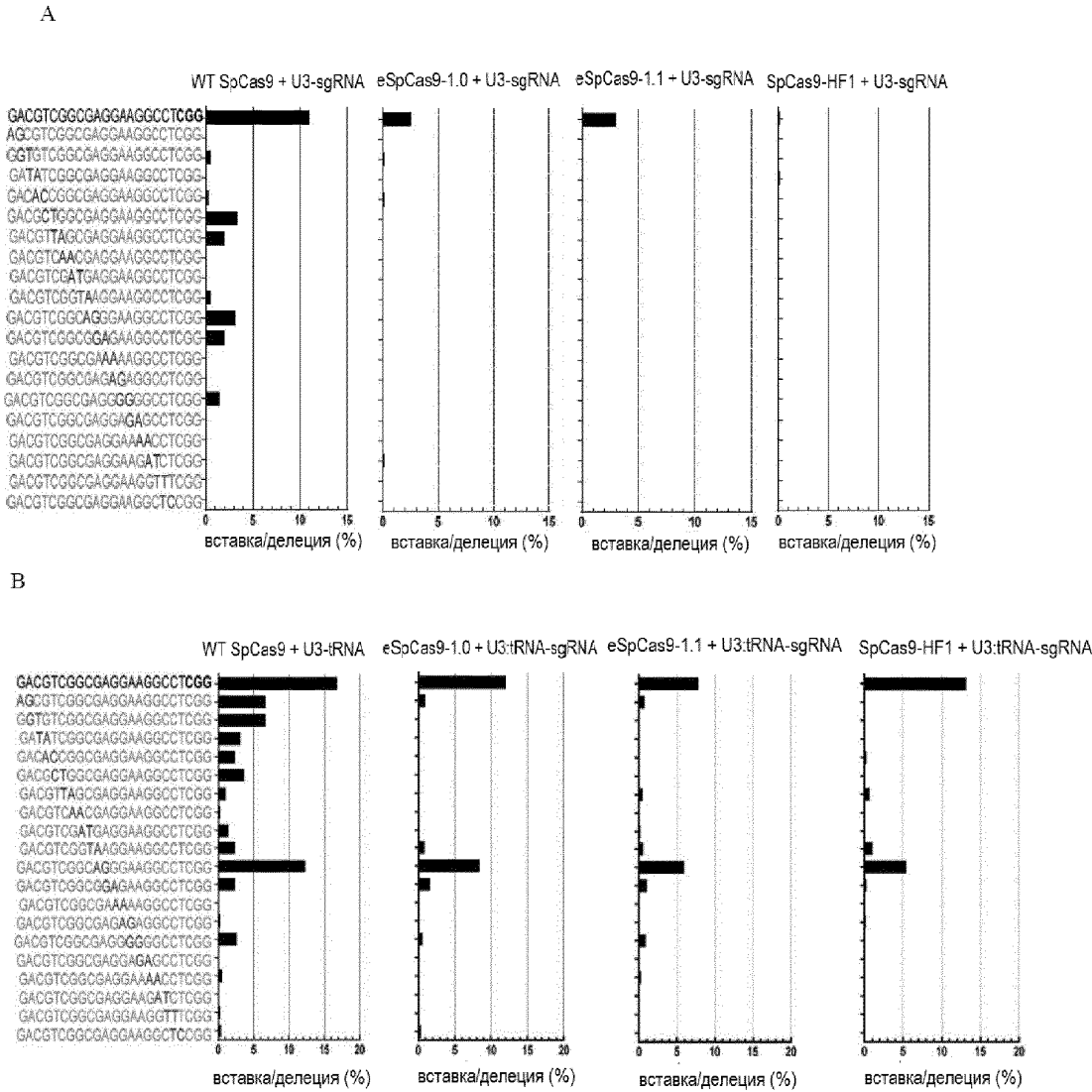
A



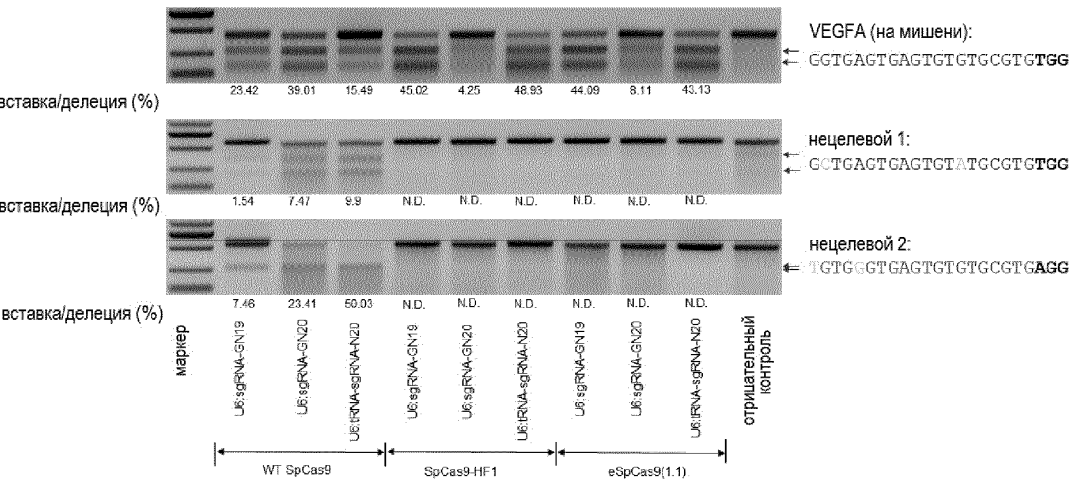
B



Фиг. 6



Фиг. 7



Фиг. 8

pUC57-U3-tPHK-sgPHK последовательность вектора
(Промотор OsU3 выделен жирным шрифтом; последовательность tPHK показана в рамке; каркас sgPHK показан жирным курсивом; сайт рестрикции BsaI подчеркнут)

tcgcgctttcgggatgacgggtgaaacctctgacacatgcagctcccgagacgggtcacagcttgctgtgaagcggatgccgggagcagacaa
gcccgtcaggcgcgctcagcgggtgttgccgggtgtcggggctggcttaactatcgccatcagagcagattgtactgagagtcaccagatgcg
gtgtgaaataccgcacagatgcgtaaggagaaaaatccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaaggcgatcggtgc
gggctcttcgtattacgccagctggcgaaagggggatgtgtcgcaaggcgattaagtgggtaacgccagggtttccagtcacgacgttgtaa
aacgacggccagtgccctgcaggtcgacgattaaggaatctttaacatacgaacagatcacttaagttcttctgaagcaacttaagttatc
aggcatgatggatcttgagggaatcagatgtgcagtcagggacatagcacaagacaggcgcttcttactggtgctaccagcaaatgctg
gaagccgggaacactgggtacgtcggaaccacgtgatgtgaagaagtaagataaactgtaggagaaaagcatttcgtagtgggcatg
aagccttcaggacatgtattgcagatgggcccgcattacgcaattggacgacaacaagactagtattagtagccctcggtatcca
catagatcaaaactgatttaaaagagttgtgcagatgatccgtggcaacaaagcaccagtggtctagtggtagaatagtacctgccacggtac
agaccgggttcgattccggctgtgcaagagaccgatcccatggctcgaggggtcctggttttagagctagaaatgaagttaaaataagg
ctagtcggtatcaacttgaaaaagtgccacggatcggtgctttttttcacataatctctagaggatccccggcgtaatcatggtcatagctgtttcc
tgtgtgaaatgttatccgctcacaattccacacaacatacagccgggaagcataaagtgtaaagcctgggggtgctaatgagtgaactaacatt
aattgcgttgcgctcactgcccgcttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcccaacgcgaggggagagcggtttgcgt
attggcgctcttccgcttctcgtcactgactcgtcgtcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaggcggtataacggtt
atccacagaatcaggggataacgcaggaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccggttctgctggcgtt
ttccataggtcctcgccccctgacgagcatcacaataacgacgtcgaagtcagaggtggcgaaccgcacaggactataagatacaggcggtt
tccccctggaagctccctcgtgcgtctctcgttccgacctgcccgttaccggatactgtcgcctttctccctcgggaagcgtggcgctttctcat
agctcacgctgtatgtatcagttcgggtgaggtcgttcgtccaaagctgggctgtgtgcacgaacccccgttcagcccagccgtgcgccttacc
cggtaactatcgtcttgatccaacccggtaagacacgacttatccactggcagcagccactggtaacaggattagcagagcgaggtatgtagg
cgggtgtacagagttcttgaagtggtggcctaactacgggtacactagaagaacagattttggtatctgcgctctgtgaagccagttacctcggaaa
aagagttgtagctcttgatccggcaacaaccaccgctggtagcgggtgtttttgttgcaagcagcagattacgcgcagaaaaaaggatctc
aagaagatcctttgatctttctacggggtcgtacgctcagtggaacgaaaactcaggttaagggtttgtcatgagattatcaaaaaggatctcac
ctagatccttttaataaaaaatgaagttttaaatcaatcaagtatatatgagtaaaacttggtctgacagttaccaatgcttaacagtgaggaacctatct
cagcgatctgtctatttcgttcacatagttgctgactccccgtcgtgtagataactacgatacgggagggttaccatctgccccagtgctgcaat
gataccgcgactccacgctcaccggctccagatttatcagcaataaacagccagccgggaaggccgagcgagaaagtgtcctgcaactttat
ccgctccatccagctctattaattgttgcgggaagctagagtaagtagttcgccagttaatagtttgcgaacgttgttgcattgctacaggcatcgt
ggtgtcacgctcgtcgtttggtatggcttcattcagctccgggtcccaacgatcaaggcgagttacatgatccccatgttggcaaaaaagcggttag
ctccttcggctccgcatgttgcagaagtaagttggccgagtggttatcactcatggttatggcagcactgcataattcttactgtcatgccatccgt
aagatcgtttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgcttgccttgcctcaatacgggataa
taccgcgccacatagcagaactttaaagtgtcatcattggaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttc
gatgtaaccactcgtgcacccaactgatcttcagcatctttacttaccagcgtttctgggtgagcaaaaacaggaaggcaaaatcccgcaaaaa
agggaataaggcgacacggaaatgtgaatactcatacttctcttttaataattattgaagcatttatcagggttattgtctcatgagcgatacatatt
tgaatgtatttagaaaaataaacaataagggttcgcgcacatttccccgaaaagtgccacctgacgtctaagaacattattatcatgacattaacc
tataaaaataggcgtatcacgagcccttctc

Фиг. 9

