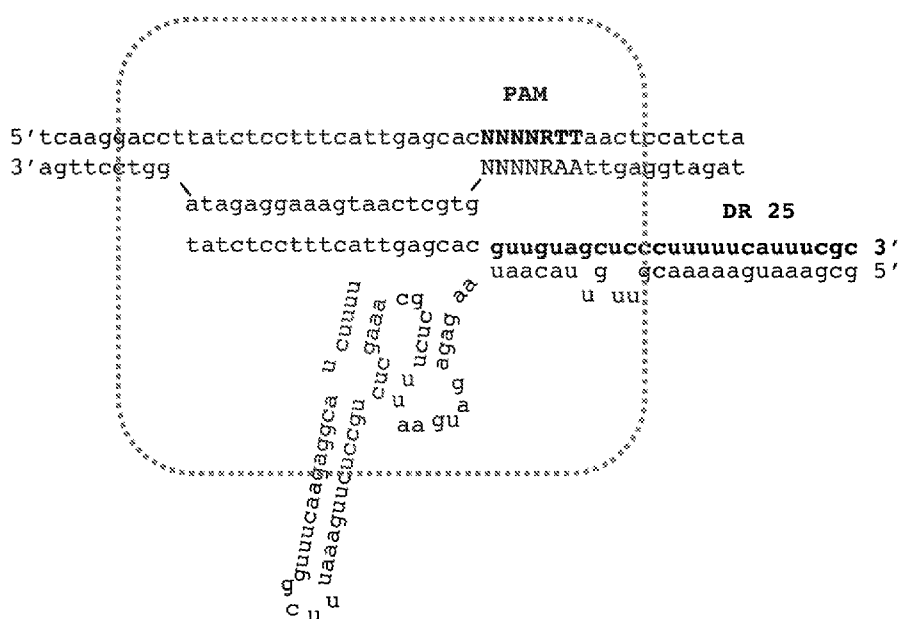


(10) Номер международной публикации
WO 2021/096391 A1

- (72) **Изобретатели: СЕВЕРИНОВ, Константин Викторович (SEVERINOV, Konstantin Viktorovich);** ул. Дмитрия Ульянова, 4/2, кв. 368 Москва, 188301, Moscow (RU). **ШМАКОВ, Сергей Анатольевич (SHMAKOV, Sergey Anatolievich);** Московская обл., ул. Пионерская, 14, кв. 15. Воскресенск, 140200, Voskresensk (RU). **АРТАМОНОВА, Дарья Николаевна (ARTAMONOVA, Daria Nikolaevna);** ул. Маршала Тимошенко, 9-59. Москва, 121359, Moscow (RU). **ГОРЯНИН, Игнатий Игоревич (GORYANIN, Ignaty Igorevich);** ул. Кировоградская 16, здание 1, кв. 351. Москва, 117587, Moscow (RU). **МУШАРОВА, Ольга Сергеевна (MUSHAROVA, Olga Sergeevna);** ул. Рогова, 7, стр. 2, кв. 289 Москва, 123098, Moscow

(54) Название изобретения: ПРИМЕНЕНИЕ CAS9 БЕЖА ИЗ БАКТЕРИИ PASTEURELLA PNEUMOTROPICA



Фиг. 11

(57) **Резюме:** Настоящее изобретение описывает новую бактериальную нуклеазу системы CRISPR-Cas9 из бактерии *P. pneumotropica*, а также ее применение для образования строго специфичных двунитевых разрывов в молекуле ДНК. Данная

[продолжение на следующей странице]

WO 2021/096391 A1

(RU). **АНДРЕЕВА, Юлия Валерьевна (ANDREEVA, Julia Valerevna)**; Московский, ул. Москвитина 3-2-108 Москва, 142784, Moscow (RU). **ЗЮБКО, Татьяна Игоревна (ZYUBKO, Tatiana Igorevna)**; ул. Яхтенная 3, стр. 1 Санкт-Петербург, 197374, Sankt-Peterburg (RU). **ФЁДОРОВА, Яна Витальевна (FEDOROVA, Iana Vitalevna)**; Ленинградская обл., ул. Куприна 40, 14 Гатчина, 188301, Gatchina (RU). **ХОДОРКОВСКИЙ, Михаил Алексеевич (KHODORKOVSKIY, Mikhail Alekseevich)**; пр. Косыгина, д. 23, к. 1, кв. 268 Санкт-Петербург, 195426, Sankt-Peterburg (RU). **ПОБЕГАЛОВ, Георгий Евгеньевич (POBEGALOV, George Evgenevich)**; ул. Партизана Германа, д.12, кв.317 Санкт-Петербург, 198205, Sankt-Peterburg (RU). **АРСЕНИЕВ, Анатолий Николаевич (ARSENIEV, Anatoliy Nikolaevich)**; пр.Новочеркасский 25, к.1, кв.36 Санкт-Петербург, 195112, Sankt-Peterburg (RU). **СЕЛКОВА, Полина Анатольевна (SELKOVA, Polina Anatolevna)**; рес. Удмуртская, ул. Верхняя, дом 15, кв. 72 Воткинск, 427439, Votkinsk (RU). **ВАСИЛЬЕВА, Александра Андреевна (VASILIEVA, Aleksandra Andreevna)**; ул. Партизана Германа, д. 41, кв. 144 Санкт-Петербург, 198334, Sankt-Peterburg (RU). **АРТАМОНОВА, Татьяна Олеговна (ARTAMONOVA, Tatiana Olegovna)**; ул. Наличная, д. 27, кв. 7 Санкт-Петербург, 199406, Sankt-Peterburg (RU). **АБРАМОВА, Марина Викторовна (ABRAMOVA, Marina Viktorovna)**; пр. Непокорённых, д.8, корпус 2 Санкт-Петербург, 194021, Sankt-Peterburg (RU).

(74) Агент: **КОТЛОВ, Дмитрий Владимирович (KOTLOV, Dmitry Vladimirovich)**; ООО "ЦИС "Сколково", Территория инновационного центра "Сколково", дом 4, оф.402.1 Москва, 143026, Moscow (RU).

(81) Указанные государства (если не указано иначе, для каждого вида национальной охраны): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Указанные государства (если не указано иначе, для каждого вида региональной охраны): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), евразийский (AM, AZ, BY, KG, KZ, RU, TJ, TM), европейский патент (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Опубликована:

- с отчётом о международном поиске (статья 21.3)
- с перечнем последовательностей в соответствии с Правилом 5.2(a)
- в черно-белом варианте; международная заявка в поданном виде содержит цвет или оттенки серого и доступна для загрузки из PATENTSCOPE.

нуклеаза обладает необычными свойствами и может быть использована для изменения последовательности геномной ДНК в клетке одноклеточного или многоклеточного организма. Таким образом, достигается повышение универсальности доступных систем CRISPR-Cas9, что позволит использовать различные варианты нуклеаз Cas9 для разрезания геномной или плазмидной ДНК в разных организмах, в большем количестве специфических сайтов и/или при различных условиях.

ПРИМЕНЕНИЕ CAS9 БЕЖА ИЗ БАКТЕРИИ PASTEURELLA PNEUMOTROPICA

Область техники

Изобретение относится к биотехнологии, а именно, к новым ферментам, Cas нуклеазам систем CRISPR-Cas, применяемым для разрезания ДНК и редактирования генома различных организмов. Данная технология может применяться в будущем для генной терапии наследственных заболеваний человека, а также для редактирования генома других организмов.

Уровень техники

Изменение последовательности ДНК – одна из актуальных задач биотехнологии на сегодняшний день. Редактирование и изменение геномов эукариотических и прокариотических организмов, а также манипуляции с ДНК *in vitro*, требуют направленного внесения двунитевых разрывов в последовательности ДНК.

Для решения этой задачи в настоящее время используют следующие методики: искусственные нуклеазные системы, содержащей домены типа «цинковые пальцы», TALEN-системы и бактериальные CRISPR-Cas системы. Первые два метода требуют трудозатратой оптимизации аминокислотной последовательности нуклеазы для узнавания конкретной последовательности ДНК. В отличие от них в случае CRISPR-Cas систем структурами, узнающими ДНК мишень, являются не белки, а короткие направляющие РНК. Разрезание конкретной ДНК мишени не требует синтеза нуклеазы или ее гена *de novo*, а обеспечивается за счет использования направляющих РНК, комплементарных целевой последовательности. Это делает CRISPR Cas системы удобными и эффективными инструментами разрезания различных ДНК-последовательностей. Методика позволяет осуществлять единовременное разрезание ДНК в нескольких участках при использовании направляющих РНК разной последовательностей. Такой подход используется в том числе для одновременного изменения нескольких генов в эукариотических организмах.

По своей природе CRISPR-Cas системы являются иммунными системами прокариот, способными высоко специфично вносить разрывы в генетический материал вирусов (Mojica F. J. M. et al. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements //Journal of molecular evolution. – 2005. – Т. 60. – №. 2. – С. 174-182). Аббревиатура CRISPR-Cas расшифровывается как “Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR associated genes” (Jansen R. et al. Identification of genes that are associated with DNA repeats in prokaryotes //Molecular microbiology. – 2002. – Т. 43. – №. 6. – С. 1565-1575), что в переводе с английского обозначает “короткие палиндромные повторы, регулярно расположенные группами, и ассоциированные с ними гены”. Все CRISPR-Cas системы состоят из CRISPR кассет и генов, кодирующих различные Cas белки (Jansen R. et al. , Molecular microbiology. – 2002. –

Т. 43. — №. 6. — С. 1565-1575). CRISPR кассеты состоят из последовательностей-спейсеров, каждый из которых имеет уникальную нуклеотидную последовательность, и повторяющихся палиндромных повторов (Jansen R. et al. , *Molecular microbiology*. — 2002. — Т. 43. — №. 6. — С. 1565-1575). В результате транскрипции CRISPR кассет и их последующего процессинга образуются направляющие крПНК, которые вместе с Cas белками формируют эффекторный комплекс (Brouns S. J. J. et al. *Small CRISPR RNAs guide antiviral defense in prokaryotes* // *Science*. — 2008. — Т. 321. — №. 5891. — С. 960-964). За счет комплементарного спаривания крПНК с целевым участком ДНК, именуемым протоспейсером, Cas-нуклеаза узнает ДНК-мишень и высоко специфично вносит в нее разрыв.

CRISPR-Cas системы, представленными одиночным белком-эффектором, разделяют на шесть различных типов (от I до VI) в зависимости от Cas белков, входящих в состав систем. В 2013 году впервые было предложено использовать систему CRISPR-Cas9, относящуюся к типу II, для редактирования геномной ДНК клеток человека (Cong L, et al., *Multiplex genome engineering using CRISPR/Cas systems*. *Science*. 2013 Feb 15;339(6121):819-23). Система CRISPR-Cas9 II типа отличается простотой состава и механизма работы: для ее функционирования необходимо формирование эффекторного комплекса, состоящего лишь из одного белка Cas9 и двух коротких ПНК: крПНК (crRNA) и трейсерной ПНК (tracrRNA). Трейсерная ПНК комплементарно спаривается с участком крПНК, происходящим из CRISPR повтора, образуя вторичную структуру, необходимую для связывания направляющих ПНК с Cas эффектором. Определение последовательности направляющих ПНК является важным шагом в характеристике неизученных ранее Cas-ортологов. Эффекторный белок Cas9 является ПНК-зависимой ДНК эндонуклеазой с двумя нуклеазными доменами (HNH и RuvC), вносящими разрывы в комплементарные нити целевой ДНК, таким образом образуя двунитевой разрыв ДНК (Deltcheva E. et al. *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III* // *Nature*. — 2011. — Т. 471. — №. 7340. — С. 602).

На сегодняшний день известно несколько CRISPR-Cas нуклеаз, способных направлено и специфично вносить двунитевые разрывы в ДНК. Технология CRISPR-Cas9 является одной из самых современных и быстроразвивающихся методик внесения разрывов в ДНК различных организмов, начиная от бактериальных штаммов и заканчивая клетками человека, а также *in vitro* (Song M. *The CRISPR/Cas9 system: Their delivery, in vivo and ex vivo applications and clinical development by startups*. *Biotechnol Prog*. 2017 Jul;33(4):1035-1045).

Эффекторному рибонуклеиновому комплексу, состоящему из Cas9 и дуплекса крПНК и тракрПНК, для распознавания и последующего гидролиза ДНК помимо комплементарного соответствия спейсера крПНК и протоспейсера необходимо присутствие PAM (от англ. "PAM" - protospacer adjusted motif) на ДНК мишени (Mojica F. J.

М. et al. 2009). РАМ представляет собой строго определенную последовательность из нескольких нуклеотидов, расположенных в системах типа II вплотную либо в нескольких нуклеотидах от 3'-конца протоспейсера на нетаргетной цепи. При отсутствии РАМ гидролиза связей в ДНК с образованием двунитевого разрыва не происходит. Необходимость присутствия РАМ последовательности на мишени повышает специфичность узнавания, но в то же время накладывает ограничение в выборе целевых участков ДНК, в которые необходимо внести разрыв. Таким образом, наличие нужной РАМ последовательности, фланирующей ДНК-мишень с 3'-конца, является характеристикой, ограничивающей применение CRISPR-Cas систем на любых участках ДНК.

Различные CRISPR-Cas белки используют для своей работы разные, оригинальные РАМ последовательности. Использование CRISPR-Cas белков с новыми разнообразными РАМ последовательностями необходимо для обеспечения возможности изменения любого участка ДНК, как *in vitro*, так и в геноме живых организмов. Изменение эукариотических геномов также требует использования нуклеаз малого размера для обеспечения доставки CRISPR-Cas систем в клетки посредством AAV вирусов.

Несмотря на известность ряда способов разрезания ДНК и изменения последовательности геномной ДНК, на сегодняшний день сохраняется потребность в новых эффективных инструментах для модификации ДНК в различных организмах и в строго определенных местах последовательности ДНК.

Сущность изобретения

Задачей настоящего изобретения является создание новых инструментов для изменения последовательности геномной ДНК одноклеточных или многоклеточных организмов на основе систем CRISPR-Cas9. Существующие в настоящее время системы имеют ограниченное применение из-за специфичной последовательности РАМ, которая должна присутствовать на 3'-конце участка ДНК, подвергающегося модификации. Поиск новых ферментов Cas9 с другими РАМ последовательностями позволит расширить арсенал имеющихся средств для образования двунитевого разрыва в необходимых, строго определенных местах в молекулах ДНК разных организмов. Для решения этой задачи авторами была охарактеризована ранее предсказанная для бактерии *Pasteurella pneumotropica* (*P. pneumotropica*) CRISPR нуклеаза II типа PpCas9, которая может быть применена для внесения направленных изменений в геном как этого, так и других организмов. Существенными признаками, отличающими настоящее изобретение, являются: (а) короткая, отличающаяся от других известная последовательность РАМ; (б) относительно малый размер охарактеризованного белка PpCas9 – 1055 аминокислотных остатков (а.о.).

Указанная задача решается путем применения белка, содержащего аминокислотную последовательность SEQ ID NO: 1, или содержащего аминокислотную последовательность, которая по меньшей мере на 95% идентична аминокислотной

последовательности SEQ ID NO: 1 и имеет отличия по сравнению с SEQ ID NO: 1 только в неконсервативных аминокислотных остатках, для образования двунитевого разрыва в молекуле ДНК, расположенного непосредственно перед нуклеотидной последовательностью 5'-NNNN(A/G)TT-3' в указанной молекуле ДНК. В некоторых вариантах изобретения данное применение характеризуется тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 35 °C до 45 °C. В некоторых вариантах изобретения данное применение характеризуется тем, что образование двунитевого разрыва происходит в геномной ДНК клетки млекопитающего. В некоторых вариантах изобретения данное применение характеризуется тем, что образование двунитевого разрыва в молекуле ДНК приводит к изменению геномной ДНК указанной клетки млекопитающего.

Указанная задача также решается путем создания способа изменения последовательности геномной ДНК в клетке одноклеточного или многоклеточного организма, включающего введение в указанную клетку организма эффективного количества: а) белка, содержащего аминокислотную последовательность SEQ ID NO: 1, или нуклеиновой кислоты, кодирующей белок, содержащий аминокислотную последовательность SEQ ID NO: 1, и б) направляющей РНК, содержащей последовательность, образующую дуплекс с нуклеотидной последовательностью участка геномной ДНК организма, непосредственно примыкающей к нуклеотидной последовательности 5'-NNNN(A/G)TT-3', и взаимодействующей с указанным белком после образования дуплекса, или последовательности ДНК, кодирующей указанную направляющую РНК; при этом взаимодействие указанного белка с направляющей РНК и нуклеотидной последовательностью 5'-NNNN(A/G)TT-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNN(A/G)TT-3'.

В некоторых вариантах изобретения данный способ характеризуется тем, что дополнительно включающий введение экзогенной последовательности ДНК одновременно с направляющей РНК. В некоторых вариантах изобретения данный способ характеризуется тем, что указанная клетка является клеткой млекопитающего.

В качестве направляющей РНК может быть использована смесь из крРНК (crRNA) и трейсерной РНК (tracrRNA), способных образовать комплекс с участком целевой ДНК и белком PrCas9. В предпочтительных вариантах изобретения в качестве направляющей РНК может быть использована гибридная РНК, сконструированная на основе крРНК и трейсерной РНК. Методы конструирования гибридной направляющей РНК известны специалистам (Hsu PD, et al., DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013 Sep;31(9):827-32). Один из вариантов конструирования гибридной РНК раскрыт в Примерах ниже.

Изобретение может быть использовано как для разрезания целевой ДНК *in vitro*, так и для модификации генома какого-либо живого организма. Модификация геномной ДНК может проводиться прямым способом – разрезанием геномной ДНК в соответствующем сайте, а также вставкой экзогенной последовательности ДНК за счет гомологичной репарации.

В качестве экзогенной последовательности ДНК может быть использован любой участок двунитевой или однострунчатой ДНК из генома организма, отличного от организма, используемого при введении (или смесь таких участков между собой и с другими фрагментами ДНК), при этом этот участок (или смесь участков) предназначен для интеграции в место двуцепочечного разрыва в целевой ДНК, образованного под действием нуклеазы PpCas9. В некоторых вариантах изобретения в качестве экзогенной последовательности ДНК может быть использован участок двуцепочечной ДНК из геномной ДНК организма, используемого при введении белка PpCas9, но при этом измененный мутациями (заменой нуклеотидов), а также вставками или делециями одного или нескольких нуклеотидов.

Техническим результатом настоящего изобретения является повышение универсальности доступных систем CRISPR-Cas9, позволяющее использовать нуклеазу Cas9 для разрезания геномной или плазмидной ДНК в большем количестве специфических сайтов и специфических условий. Новая нуклеаза может быть использована в клетках бактерий, млекопитающих или других организмов.

Краткое описание фигур

Фиг. 1. Схема устройства локуса CRISPR PpCas9 системы. DR - direct repeat, или прямой повтор - регулярно повторяющийся участок, входящий в состав CRISPR кассеты.

Фиг. 2. PAM скрининг *in vitro*. Схема эксперимента.

Фиг. 3. Разрезание нуклеазой PpCas9 фрагментов 7N библиотеки при разных температурах проведения реакции.

Фиг. 4. (А) Анализ результатов *in vitro* скрининга нуклеазы PpCas9 с использованием расчета логарифма изменения доли каждого конкретного нуклеотида в каждой позиции PAM (FC). (Б) PAM Logo нуклеазы PpCas9. Для каждой позиции обозначены частоты представленности аденина, цитозина, тимина и гуанина. Высота букв соответствует частоте представленности нуклеотида в данной позиции PAM последовательности.

Фиг. 5. Проверка влияния однонуклеотидных замен в первой позиции PAM на эффективность разрезания нуклеазой PpCas9 ДНК-мишени.

Фиг. 6. Проверка значимости нуклеотидных позиций в PAM последовательности PpCas9.

Фиг. 7. Проверка влияния замены А на Г в 5ой позиции PAM на эффективность разрезания нуклеазой PpCas9 ДНК-мишени.

Фиг. 8. Проверка влияния однонуклеотидных замен в 7ой позиции PAM на эффективность разрезания нуклеазой PpCas9 ДНК-мишени.

Фиг. 9. Разрезание различных сайтов ДНК с помощью белка PpCas9. Дорожки 1 и 2 – положительный контроль.

Фиг. 10. Проверка распознавания нуклеазой PpCas9 PAM последовательности CAGCATT. Дорожки 1 и 2 – положительный контроль.

Фиг. 11. Схема инструмента разрезания ДНК PpCas9.

Фиг. 12. Эксперимент по разрезанию ДНК -мишени. Используются гибридные направляющие РНК разной длины.

Фиг. 13. Выравнивание аминокислотных последовательностей белков PpCas9 и Cas9 из *Staphylococcus aureus* при помощи программы NCBI BLASTp (default parameters).

Фиг. 14. Изменение геномной ДНК клеток человека с использованием PpCas9. (А) – схема эксперимента по определению эффективности изменения геномной ДНК клеток человека с использованием плазмиды, несущей PpCas9. (В) – Результаты анализа внесения вставок и делеций нуклеотидов в последовательность целевых сайтов геномной ДНК клеток человека (сверху – продукты реакции с Т7 эндонуклеазой I были нанесены на агарозный гель электрофорез, внизу – примеры вставок и делеций, образуемых PpCas9 в гене EMX1, определенные с помощью высоко производительного секвенирования)

Подробное раскрытие изобретения

В описании данного изобретения термины «включает» и «включающий» интерпретируются как означающие «включает, помимо всего прочего». Указанные термины не предназначены для того, чтобы их истолковывали как «состоит только из». Если не определено отдельно, технические и научные термины в данной заявке имеют стандартные значения, общепринятые в научной и технической литературе.

Используемый здесь термин «процент гомологии двух последовательностей» эквивалентен термину «процент идентичности двух последовательностей». Идентичность последовательностей определяется на основании референсной последовательности. Алгоритмы для анализа последовательности известны в данной области, такие как BLAST, описанный в Altschul et al., J. Mol. Biol., 215, pp. 403-10 (1990). Для целей настоящего изобретения для определения уровня идентичности и сходства между нуклеотидными последовательностями и аминокислотными последовательностями может быть использовано сравнение нуклеотидных и аминокислотных последовательностей, производимое с помощью пакета программ Blast, предоставляемого National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/blast>) с использованием содержащего

разрывы выравнивания со стандартными параметрами. Процент идентичности двух последовательностей определяется числом положений идентичных аминокислот в этих двух последовательностях с учетом числа пробелов и длины каждого пробела, которые необходимо ввести для оптимального сопоставления двух последовательностей путем выравнивания. Процент идентичности равен числу идентичных аминокислот в данных положениях с учетом выравнивания последовательностей, разделенному на общее число положений и умноженному на 100.

Термин "специфически гибридизуется" относится к ассоциации между двумя одноцепочечными молекулами нуклеиновых кислот или в достаточной степени комплементарными последовательностями, что разрешает такую гибридизацию в predetermined условиях, обычно используемых в данной области.

Фраза "двунитевой разрыв, расположенный непосредственно перед нуклеотидной последовательностью РАМ" означает, что двунитевой разрыв в целевой последовательности ДНК будет произведен на расстоянии от 0 до 25 нуклеотидов перед нуклеотидной последовательностью РАМ.

Под экзогенной последовательностью ДНК, вводимой одновременно с направляющей РНК, следует понимать последовательность ДНК, подготовленную специально для специфической модификации двуцепочечной целевой ДНК в месте разрыва, определяемого специфичностью направляющей РНК. Подобной модификацией может быть, например, вставка или делеция определенных нуклеотидов в месте разрыва целевой ДНК. Экзогенной ДНК может служить как участок ДНК из другого организма, так и участок ДНК из того же организма, что и целевая ДНК.

Под белком, содержащим определенную аминокислотную последовательность следует понимать белок, имеющий аминокислотную последовательность, составленную из указанной аминокислотной последовательности и, возможно, других последовательностей, соединённых пептидными связями с указанной аминокислотной последовательностью. Примером других последовательностей может служить последовательность сигнала ядерной локализации (NLS), или другие последовательности, обеспечивающие повышенную функциональность для указанной аминокислотной последовательности.

Под экзогенной последовательностью ДНК, вводимой одновременно с направляющей РНК, следует понимать последовательность ДНК, подготовленную специально для специфической модификации двуцепочечной целевой ДНК в месте разрыва, определяемого специфичностью направляющей РНК. Подобной модификацией может быть, например, вставка или делеция определенных нуклеотидов в месте разрыва целевой ДНК. Экзогенной ДНК может служить как участок ДНК из другого организма, так и участок ДНК из того же организма, что и целевая ДНК.

Под эффективным количеством вводимых в клетку белка и РНК следует понимать такое количество белка и РНК, которое при попадании в указанную клетку будет способно образовать функциональный комплекс, то есть комплекс, который будет специфически связываться с целевой ДНК и производить в ней двунитевой разрыв в месте, определяемом направляющей РНК и РАР последовательностью на ДНК. Эффективность этого процесса может быть оценена при помощи анализа целевой ДНК, выделенной из указанной клетки с помощью стандартных методов, известных специалистам.

Доставка белка и РНК в клетку может быть осуществлена различными способами. Например, белок может быть доставлен в виде ДНК-плазмиды, которая кодирует ген этого белка, как мРНК для трансляции этого белка в цитоплазме клетки, или как рибонуклеопротеидный комплекс, включающий этот белок и направляющую РНК. Доставка может быть осуществлена различными методами, известными специалистам.

Нуклеиновая кислота, кодирующая компоненты системы, может быть введена в клетку, непосредственно или опосредованно: за счет трансфекции или трансформации клеток известными специалистам способами, за счет использования рекомбинатного вируса, за счет манипуляций с клеткой, таких как микроинъекция ДНК и т.п.

Доставка рибонуклеинового комплекса, состоящего из нуклеазы и направляющих РНК и экзогенной ДНК (при необходимости) может осуществляться путем трансфекции комплексов в клетку или за счет механического введения комплекса внутрь клетки, например, микроинъекции.

Молекула нуклеиновой кислоты, кодирующая белок, который необходимо ввести в клетку, может быть интегрирована в хромосому или может представлять собой внехромосомно реплицирующуюся ДНК. В некоторых вариантах для обеспечения эффективной экспрессии гена белка с вводимой в клетку ДНК необходимо изменить последовательность этой ДНК в соответствии с типом клетки в целях оптимизации кодонов при экспрессии, обусловленное неравномерностью частот встречаемости синонимичных кодонов в кодирующих областях генома различных организмов. Оптимизация кодонов необходима для увеличения экспрессии в клетках животных, растений, грибов или микроорганизмов.

Для функционирования белка, имеющего последовательность, которая по меньшей мере на 95% идентична аминокислотной последовательности SEQ ID NO: 1, в эукариотической клетке необходимо, чтобы этот белок оказался в ядре этой клетки. Поэтому, в некоторых вариантах изобретения, для образования двунитевых разрывов в целевой ДНК используют белок, имеющий последовательность, которая по меньшей мере на 95% идентична аминокислотной последовательности SEQ ID NO: 1, и который дополнительно модифицирован с одного или с обоих концов добавлением одного или нескольких сигналов ядерной локализации. Например, может быть использован сигнал ядерной локализации из вируса SV40. Для эффективной доставки в ядро сигнал ядерной

локализации может быть отделен от основной последовательности белка спейсерной последовательностью, например, описанной в Shen B, et al. "Generation of gene-modified mice via Cas9/RNA-mediated gene targeting", Cell Res. 2013 May;23(5):720-3. Также, в других вариантах осуществления, может быть использован другой сигнал ядерной локализации, или альтернативный метод доставки указанного белка в ядро клетки.

Настоящее изобретение охватывает применение белка из организма *P. pneumotropica*, гомологичного ранее охарактеризованным белкам Cas9, для внесения двуцепочечных разрывов в молекулы ДНК в строго определенных положениях. Использование CRISPR нуклеаз для внесения направленных изменений в геном имеет ряд преимуществ. Во-первых, специфичность действия системы определяется последовательностью кРНК, что позволяет использовать один тип нуклеазы для всех локусов-мишеней. Во-вторых, методика позволяет доставить в клетку сразу несколько направляющих РНК, комплементарных разным генам-мишеням, что позволяет осуществлять единовременное изменение сразу нескольких генов.

PpCas9 – Cas нуклеаза, найденная в бактериях *Pasteurella pneumotropica* ATCC 35149, являющихся патогенами грызунов, обитающих в легких животных. *Pasteurella pneumotropica* (*P. pneumotropica*) CRISPR Cas9 система (далее CRISPR PpCas9) относится к IIC типу CRISPR Cas систем и состоит из CRISPR кассеты, несущей четыре прямых повтора (direct repeats, DR) последовательностью 5'ATTATAGCACTGCGAAATGAAAAAGGGAGCTACAAC3' разделенных последовательностями уникальных спейсеров. Ни один из спейсеров системы не совпадает по последовательности с известными на сегодня бактериофагами или плазмидами, что не позволяет определить требуемый PpCas9 PAM биоинформатическим анализом. К CRISPR кассете прилегает ген эффекторного Cas9 белка PpCas9, а также гены белков Cas1 и Cas2, участвующих в адаптации, встраивании новых спейсеров. Рядом с Cas генами была обнаружена последовательность, частично комплементарная прямым повторам, складывающаяся в характерную вторичную структуру, - предполагаемая трейсерная РНК (tracrRNA) (Фиг. 1)

Знание характерной архитектуры РНК-Cas белкового комплекса систем IIC типа позволила предсказать направление транскрипции CRISPR кассеты: пре-кРНК транскрибируется в противоположном от Cas генов направлении (Фиг.1)

Таким образом, анализ последовательности локуса PpCas9 позволил предсказать последовательности трейсерной и направляющих РНК (Таблица 1).

Таблица 1. Определенные биоинформатическими методами последовательности направляющих РНК системы CRISPR PpCas9. Жирным шрифтом обозначена последовательность прямого повтора DR.

Название	Последовательность
PpCas9 трПНК	5'GCGAAATGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAAG CTCUGCCUCUUGAAAUUUCGGUUUCAAGAGGCAUCUUUUU-3'
PpCas9 крПНК	5'-xxxxxxxxxxxxxxxxxxxxxx GUUGUAGCUC CCUUUUUCAUUUCGC-3'

Для проверки активности PpCas9 нуклеазы и определения требуемого PpCas9 PAM мотива, были проведены эксперименты по воссозданию реакции разрезания ДНК *in vitro*. Для определения PAM последовательности белка PpCas9 использовали *in vitro* разрезание двунитевых PAM библиотек. Для этого необходимо было получить все компоненты эффекторного комплекса PpCas9: направляющие РНК и нуклеазу в рекомбинантной форме. Определение последовательности направляющих РНК позволило синтезировать *in vitro* молекулы crRNA и tracrRNA. Синтез осуществляли с помощью набора NEB HiScribe T7 RNA synthesis. Двунитевые ДНК библиотеки представляли собой фрагменты размером 374 пар нуклеотидов (п.н.), содержащие последовательность протоспейсера, фланкированную рандомизированными семью нуклеотидами (5'-NNNNNNN-3') с 3' конца: 5'-

cccggggtaccacggagagatggtggaaatcatcttctcgtgggcatcctgatggccacctcgtcggaagtgccacgaggatga
cagcaatgccaatgctgggggggctctctgagaacgagctctgctgcctgacacggccaggacggccaacaccaaccagaactt
gggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaaagagcatgttaaataaggatcta
catcacgtaacctgtcttagaaggcttagatactgcaattcaaggaccttatctccttcattgagcac**NNNNNNN**aactccatcta
ccagcctactctcttatctctgtatt -3'

Для разрезания этой мишени использовали направляющие РНК следующей последовательности:

tracrRNA:

5'GCGAAATGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAAGCTCUGCC
UCUUGAAAUUUCGGUUUCAAGAGGCAUCUUUUU

и crRNA:

5' **uaucuccuuucauugagcac**GUUGUAGCUC**CCUUUUUCAUUUCGC**.

Жирным шрифтом выделена последовательность crRNA, комплементарная протоспейсеру (целевой ДНК последовательности).

Для получения рекомбинантного белка PpCas9 его ген был клонирован в плазмиду pET21a. В качестве кодирующей ген ДНК, использовалась ДНК, синтезированная в компании Integrated DNA Technologies (IDT). Последовательность была кодон-оптимизирована для исключения редких кодонов, встречающихся в геноме P.

pneumotropica. Клетки *E.coli* Rosetta были трансформированы полученной плазмидой pET21a-6xHis-PpCas9.

500 мкл ночной культуры разводили в 500 мл среды LB, и растили клетки при температуре 37 °C до достижения оптической плотности 0.6 отн.ед. Синтез целевого белка индуцировали добавлением ИПТГ до концентрации 1 mM, после чего клетки инкубировали при температуре 20 °C в течение 6 часов. Затем проводили центрифугирование клеток на скорости 5000 g в течение 30 минут, полученные осадки клеток замораживали при температуре -20 °C.

Осадки размораживали на льду в течение 30 минут, ресуспензировали в 15 мл лизисного буфера (Tris-HCl 50mM pH 8, 500 mM NaCl, β-меркаптоэтанол 1mM, имидазол 10 mM) с добавлением 15 мг лизоцима и снова инкубировали на льду в течение 30 минут. Затем клетки разрушали воздействием ультразвука в течение 30 минут и центрифугировали в течение 40 минут на скорости 16000 g. Полученный супернатант пропускали через фильтр 0,2 мкм и наносили на колонку HisTrap HP 1 mL (GE Healthcare) на скорости 1 мл/мин.

Хроматографию проводили при помощи FPLC хроматографа АКТА (GE Healthcare) на скорости 1 мл/мин. Колонку с нанесенным белком промывали 20 мл лизисного буфера с добавлением 30 mM имидазола, после чего белок смывали лизисным буфером с добавлением 300 mM имидазола.

Затем, фракцию белка, полученную в ходе афинной хроматографии, пропускали через гель-фильтрационную колонку Superdex 200 10/300 GL (24 мл), уравновешенную следующим буфером: Tris-HCl 50 mM pH 8, 500 mM NaCl, 1 mM DTT. При помощи концентратора Amicon (с фильтром на 30 кДа) фракции, соответствующие мономерной форме белка PpCas9, сконцентрировали до 3 мг/мл, после чего очищенный белок хранили при температуре -80 °C в буфере, содержащем 10% глицерин.

In vitro реакцию порезки линейных РАМ библиотек проводили в объеме 20 мкл в следующих условиях. Реакционная смесь состояла из: 1X CutSmart буфера (NEB), 5 mM DTT, 100 нМ РАМ-библиотеки, 2 мкМ трРНК/крРНК, 400 нМ белка PpCas9. В качестве контроля аналогичным образом были приготовлены пробы, не содержащие РНК. Пробы инкубировали при различных температурах и анализировали методом гель-электрофореза в 2% агарозном геле. В случае правильного узнавания и специфического разрезания ДНК белком PpCas9 должны формироваться два фрагмента ДНК длиной порядка 326 и 48 пар оснований (см. Фиг. 2).

Результаты опыта показали, что PpCas9 обладает нуклеазной активностью и разрезает часть фрагментов РАМ библиотеки. Градиент температур (Фиг. 3) показал, что белок активен в диапазоне температур 35 – 45 °C. В дальнейшем в работе в качестве рабочей использовалась температура 42 °C.

Реакцию разрезания библиотеки повторяли в подобранных условиях. Продукты реакции наносили на 1.5% агарозный гель и подвергали электрофорезу. Непорезанные фрагменты ДНК длиной 374 п.н. экстрагировали из геля и подготавливали для высокоэффективного секвенирования с помощью набора NEB NextUltra II. Образцы секвенировали на платформе Illumina и далее проводили анализ последовательностей биоформатическими методами: определяли разницу в представленности нуклеотидов в отдельных позициях PAM (NNNNNNN) в сравнении с контрольным образцом с использованием подхода, описанного в (Maxwell CS, et al., A detailed cell-free transcription-translation-based assay to decipher CRISPR protospacer-adjacent motifs. *Methods*. 2018 Jul 1;143:48-57). Кроме того, для анализа результатов построили PAM лого (Фиг. 4).

Оба подхода к анализу данных (Фиг. 4) указывают на значимость 5, 6 и 7 позиций PAM. Таким образом, в результате *in vitro* анализа удалось установить предположительную PAM последовательность для PpCas9: NNNNATT. Однако эта последовательность является лишь предположительной в силу неточности результатов, получаемых скрининговыми подходами к определению PAM.

В связи с этим для уточнения последовательности была произведена проверка значимости отдельных позиций последовательности PAM. Для этого проводили реакции *in vitro* разрезания ДНК фрагментов, содержащих ДНК-мишень 5'- atctcctttcattgagcac-3', фланкированную PAM последовательностью CAACATT (или ее производных): 5'- cccggggtaccacggagagatggtggaaatcatctttctcgtgggcatccttgatggccacctcgtcgggaagtgccacgaggatgacagcaatgccaatgctgggggggctctctgagaacgagctctgctgcctgacacggccaggacggccaacaccaaccagaacttgggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaaagagcatgtttaaataaggatctcatcacgtaacctgtcttagaagaggctagatactgcaattcaaggaccttatctcctttcattgagcacCAACATTaactccatctaccagcctactctcttatctctgtatt- 3'

Все реакции разрезания ДНК проводили в следующих условиях:

1xCutSmart буфер

400 нМ PpCas9

20 нМ ДНК

2 мкМ crRNA

2 мкМ tracrRNA

Время инкубации - 30 минут, температура проведения реакции - 42 °С.

Замена первой позиции PAM на все четыре возможных варианта нуклеотидов не повлияла на эффективность работы белка (Фиг. 5).

Предсказанная значимость пятой и шестой была подтверждена экспериментально путем однонуклеотидных замен (пурин на пиримидин и наоборот) в каждой из позиций PAM. При заменах в пятой и шестой позициях белок практически переставал работать. При замене в седьмой позиции эффективность работы PpCas9 снижалась в два раза, что

отражает сниженные требования к нуклеотиду в этой позиции (Фиг. 6). Таким образом, согласно полученным результатам *in vitro* PAM скрининга нуклеазы PpCas9, в пятой позиции PAM наиболее вероятными нуклеотидами являются аденин или гуанин, что удалось подтвердить экспериментально (Фиг. 7). Замена А на Г никак не снижала эффективность разрезания фрагмента.

Согласно результатам *in vitro* скрининга фрагменты с “Т” либо с “С” в седьмой позиции должны распознаваться более эффективно. Для окончательной проверки значимости нуклеотидов в этой позиции были проведены дополнительные эксперименты. Результаты проведенных *in vitro* тестов показали, что при замене нуклеотида “Т” в седьмой позиции на А или Г эффективность разрезания снижается на 40-50% (Фиг. 8). Таким образом, седьмая позиция PAM является менее консервативной в сравнении с пятой и шестой: пурины в седьмом положении снижают эффективность узнавания, но не препятствуют белку PpCas9 вносить двунитевые разрывы в ДНК.

В результате проведенных исследований удалось сделать следующий вывод: PAM, распознаваемый нуклеазой PpCas9, соответствует следующей формуле 5'-NNNN(A/G)TT-3'. Седьмая позиция менее консервативна.

Нижеследующие примеры осуществления способа приведены в целях раскрытия характеристик настоящего изобретения и их не следует рассматривать как каким-либо образом ограничивающие объем изобретения.

Пример 1. Тестирование активности белка PpCas9 в разрезании различных ДНК мишеней.

Для того, чтобы проверить способность PpCas9 узнавать различные последовательности ДНК, фланкированные 5'-NNNN(A/G)TT-3' последовательностью, были проведены эксперименты по *in vitro* разрезанию ДНК-мишеней из последовательности гена *grin2b* человека (см. Таблицу 2).

Таблица 2. ДНК-мишени гена GRIN2B человека.

последовательность	PAM						
TATCTCCTTTTCATTGAGCAC	C	A	A	A	C	C	C
CAGCTGAAGTAATGTTAGAG	C	C	A	C	A	T	T
AATAAGAAAAACATTATTAT	C	A	C	C	A	T	T
GGGGCTATAAGTACACAAGC	C	C	T	G	C	A	T
CGTTCTCAGAAGAGCCCCC	C	A	G	C	A	T	T
CCCACGAGAAAGATGATTTC	C	A	C	C	A	T	C

В реакции разрезания в качестве мишени использовался ПЦР фрагмент гена *grin2b*, несущий сайты узнавания (Таблица 2), предположительно распознаваемыми PpCas9 в соответствии с PAM консенсусом 5'-NNNN(A/G)TT-3'. Для узнавания этих

последовательностей были синтезированы крРНК, направляющие PpCas9 на данные сайты.

Реакции разрезания проводились в подобранных для PpCas9 условиях, результат представлен на Фиг. 9. Из Фиг. 9 видно, что фермент PpCas9 успешно порезал три из четырех мишеней с подходящим PAM.

Мишень на шестой дорожке имела PAM последовательность CAGCATT, которая согласно предсказаниям на основании результатов «depletion analysis» должна эффективно распознаваться белком. Однако в данном эксперименте узнавание данного фрагмента не произошло.

Поэтому была произведена дополнительная проверка PAM CAGCATT на другой мишени-протоспейсере, ограниченной таким же PAM (Фиг. 10). В этом случае PAM эффективно распознавался, что приводило к порезке ДНК. Таким образом, белок имеет некие дополнительные предпочтения к последовательности ДНК мишени, возможно связанные со вторичной структурой ДНК.

Таким образом, проведенные исследовательские испытания показали наличие нуклеазной активности у PpCas9, а также позволили определить его PAM последовательность и верифицировать последовательности направляющих РНК.

PpCas9 рибонуклеопротеиновый комплекс специфически вносит разрывы в мишени, ограниченные PAM 5'-NNNN(A/G)TT -3' с 5' конца протоспейсера. Схема PpCas9-РНК комплекса представлена на Фиг. 11.

Пример 2. Использование гибридной направляющей РНК для разрезания ДНК мишени.

sgRNA - форма направляющих РНК, которая представляет собой слитые воедино tracrRNA (трейсерная РНК) и crRNA. Для подбора оптимальной sgRNA были сконструированы три варианта этой последовательности, отличающиеся длиной tracrRNA – crRNA дуплекса. РНК синтезировали in vitro и проводили с ними эксперименты по разрезанию ДНК -мишени (Фиг. 12)

В качестве гибридных РНК были использованы следующие РНК последовательности:

1 - sgRNA1 25DR:

UAUCUCCUUUCAUUGAGCACGUUGUAGCUCUUUUUUUCAUUUCGCGAAAGCGAAAUGA
AAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAAGCTCTGCCUCUUGAAAUUUCGG
UUUCAAGAGGCAUCUUUUU

2 - sgRNA2 36DR

UAUCUCCUUUCAUUGAGCACGUUGUAGCUCUUUUUUUCAUUUCGCGAGUGCUAUAUAUG
AAAAUUAUAGCACUGCGAAAUGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAA
GCUCUGCCUCUUGAAAUUUCGGUUUCAAGAGGCAUCUUUUU

Жирным шрифтом обозначена 20-нуклеотидная последовательность, обеспечивающая спаривание с ДНК -мишенью (вариабельная часть sgRNA). Кроме того,

в эксперименте делали контрольную пробу без РНК, а также положительный контроль - порезка мишени с помощью crRNA+trRNA.

В качестве ДНК мишени использовалась последовательность, содержащая сайт узнавания 5' tatctcctttcattgagcac 3' с соответствующим консенсусу PAMCAACATT: 5'-cccgggggtaccacggagagatgggtggaaatcatctttctcgtagggcatccttgatggccacctcgtcggaaagtgccacgaggatgacagcaatgccaatgctgggggggctcttctgagaacgagctctgctgcctgacacggccaggacggccaacaccaaccagaactgggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaaagagcatgttaaaataggatctcatcacgtaacctgtcttagaagaggctagatactgcaattcaaggacct**tatctcctttcattgagcac**CAACATTcaactccatctaccagcctactcttctatctctggtatt – 3'

Жирным шрифтом обозначен сайт узнавания, заглавными буквами PAM.

Реакцию проводили в следующих условиях: концентрация ДНК последовательности, содержащей PAM (CAACATT) – 20 нМ, концентрация белка – 400 нМ, концентрация РНК - 2 мкМ; время инкубирования - 30 минут, температура инкубирования - 37 °С.

Подобранные sgRNA1 и sgRNA2 оказались так же эффективны, как и нативные последовательности tracrRNA и crRNA: разрезание произошло в более 80% ДНК-мишенях (Фиг. 12).

Эти варианты гибридной РНК могут быть использованы для разрезания любой другой целевой ДНК при изменении последовательности, непосредственно спаривающейся с ДНК -мишенью.

Пример 3. Белки Cas9 из близкородственных организмов, относящихся к *P. pneumotropica*.

На сегодняшний день в *P. pneumotropica* не охарактеризовано ни одного фермента системы CRISPR-Cas9. Сравнимый по размерам белок Cas9 из *Staphylococcus aureus* идентичен PpCas9 на 28 % (Фиг. 13, степень идентичности была рассчитана по программе BLASTp, default parameters). Похожая степень идентичности существует и для других известных белков Cas9 (не показано).

Таким образом, белок PpCas9 существенно отличается по аминокислотной последовательности от других Cas9 белков, изученных на сегодняшний день.

Специалисту в области генетической инженерии очевидно, что полученный и охарактеризованный в данном Описании вариант последовательности белка PpCas9 может быть изменен без изменения функции самого белка (например, направленным мутагенезом аминокислотных остатков, напрямую не влияющих на функциональную активность (Sambrook et al., Molecular Cloning: A Laboratory Manual, (1989), CSH Press, pp. 15.3-15.108)). В частности, специалисту известно, что могут быть изменены неконсервативные аминокислотные остатки, не затрагивающие остатки, определяющие функциональность белка (определяющие его функцию или структуру). Примерами таких

изменений могут служить замены неконсервативных аминокислотных остатков на гомологичные. Некоторые из участков, содержащих неконсервативные аминокислотные остатки, приведены на Фиг. 12. В некоторых вариантах осуществления изобретения возможно использование белка, содержащего аминокислотную последовательность, которая по меньшей мере на 95% идентична аминокислотной последовательности SEQ ID NO: 1 и имеет отличия по сравнению с SEQ ID NO: 1 только в неконсервативных аминокислотных остатках, для образования двунитевого разрыва в молекуле ДНК, расположенного непосредственно перед нуклеотидной последовательностью 5'-NNNN(A/G)TT-3' в указанной молекуле ДНК. Гомологичные белки могут быть получены путем мутагенеза (например, сайт-направленного или ПЦР-опосредуемого мутагенеза) соответствующих молекул нуклеиновых кислот с последующим тестированием кодируемого модифицированного белка Cas9 на сохранение его функций в соответствии с описанными здесь функциональными анализами.

Пример 4. Изменение геномной ДНК клеток человека с помощью PpCas9.

Для изменения геномной ДНК клеток человека ген PpCas9 нуклеазы был клонирован в эукариотический плазмидный вектор под регуляцией CMV промотора. На 5' и 3' концы гена PpCas9 были добавлены последовательности, кодирующие сигналы ядерной локализации, обеспечивающие доставку нуклеазы в ядро клетки. Последовательность sgRNA была клонирована в вектор под регуляцией U6 промотора. Для проверки активности системы использовались sgRNA с последовательностью, комплементарной ДНК мишени длиной 20 и 24 нуклеотида. В качестве положительного контроля использовалась аналогичная плазида, несущая систему изменения геномной ДНК на основе SpCas9, известная из уровня техники. Для оценки эффективности трансфекции используемые плазмиды также несли ген зеленого флуоресцентного белка GFP. В качестве ДНК мишеней использовались следующие участки геномной ДНК человека (Таблица 3).

Таблица 3. ДНК-мишени генов EMX1 и GRIN2B человека.

нуклеаза	Название сайта	Последовательность мишени	PAM
PpCas9	EMX1.1 sg20	GCCCTTCCTCCTCCAGCTTC	GTT
PpCas9	EMX1.1 sg24	TCAGGCCCTTCCTCCTCCAGCTTC	GTT
PpCas9	EMX1.2 sg20	GGAGGTGACATCGATGTCCT	ATT
PpCas9	EMX1.2 sg24	CATTGGAGGTGACATCGATGTCCT	ATT

PpCas9	GRIN2B1.1 sg20	CAGCTGAAGTAATGTTAGAG	ATT
PpCas9	GRIN2B1.1 sg24	TTAGCAGCTGAAGTAATGTTAGAG	ATT
PpCas9	GRIN2B1.2 sg20	AATAAGAAAAACATTATTAT	ATT
PpCas9	GRIN2B1.2 sg24	ATAAAATAAGAAAAACATTATTAT	ATT
SpCas9	EMX1 sg20	GAGTCCGAGCAGAAGAAGAA	GGG
SpCas9	GRIN2B sg20	ACCTTTTATTGCCTTGTTCA	AGG

EMX1.1 и EMX1.2 представляли собой два разных сайта модификации в гене EMX1; аналогично, GRIN2B1.1 и GRIN2B1.2 представляли собой два разных сайта модификации в гене GRIN2B.

ДНК мишени были фланкированы с 3' конца PAM последовательностями PpCas9 5'-NNNNRTT -3' или SpCas9 5'-NGG -3'.

Для эффективной работы нуклеазы PpCas9 в эукариотических клетках необходимо обеспечить импорт этого белка внутрь ядра эукариотической клетки. Для этого можно использовать сигнал ядерной локализации из Т-антигена вируса SV40 (Lanford et al., Cell, 1986, 46: 575–582), соединённый с последовательностью PpCas9 с помощью спейсерной последовательности, описанной в Shen B, et al. "Generation of gene-modified mice via Cas9/RNA-mediated gene targeting", Cell Res. 2013 May;23(5):720-3 или без нее.

В данном примере, полная аминокислотная последовательность нуклеазы, транспортируемой внутрь ядра клеток человека, представляла собой следующую последовательность:

MAPKKKRKVGIGVPAAEQNNPLNYILGLDLGIASIGWAVVEIDEESSPIRLIDVGVRTFE
RAEVAKTGESLALSRRRLARSSRRRIKRRARERLKKAKRLLKAEKILHSIDEKLPINWQLRVKGLK
EKLERQEWAAVLLHLSKHRGYLSQRKNEGKSDNKGALLSGIASNHQMLQSSEYRTPAEIAV
KKFQVEEGHIRNQRGSYTHTFSRLDLLAEMELLFQRQAELGNSYTSTTLLENLTALLMWQKPAL
AGDAILKMLGKCTFEPSEYKAAKNSYSAERFVWLTKLNNLRILENGTERALNDNERFALLEQPY
EKSKLTYAQVRAMLALSDNAIFKGVRYLGEDKKTVESKTTLIEMKFYHQIRKTLGSAELKKEWN
ELKGNSDLLDEIGTAFSLYKTDDDICYLEGKLPERVLNALLNLFDFKFIQLSLKALHQILPLML
QGQRYDEAVSAIYGDHYGKKSTETTRLLPTIPADEIRNPVVLRTLTQARKVINAVVRLYGSPARI
HIETAREVGKSYQDRKKLEKQQEDNRKQRESAVKKFKEMFPHFVGEPKGDILKMRLYELQQA
KCLYSGKSLELHRLLEKGYVEVDHALPFSRTWDDSFNNKVLVLANENQNKGNLTPYEWLDGK
NNSERWQHVFVVRVQTSGFSYAKKQRILNHLDEKGFIERNLNDTRYVARFLCNFIADNMLLVG
KGKRNVFASNGQITALLRHRWGLQKVREQNDRHHALDAVVVACSTVAMQQKITRFVRYNEGN
VFSGERIDRETGEIPLHFPSPWAFFKENVEIRIFSENPKLELENRLPDYPQYNHEWVQPLFVSR
MPTRKMTGQGHMETVKSARKLNEGLSVLKVPLTQLKLSDLERMVNRDREIALYESLKARLEQF

GNDPAKAFAEPFYKKGGALVKAVRLEQTQKSGVLVRDGNNGVADNASMVRVDVFTKGGKYFLV
PIYTWQVAKGILPNRAATQGKDENDWDIMDEMATFQFSLCQNDLIKLVTKKKTIFGYFNGLNRA
TSNINIKEHDLKSKGKLGIIYLEVGVKLAIISLEKYQVDELGKNIRPCRPTKRQHVRFKRPAATKK
AGQAKKKK

Используемая в данном эксперименте плазмида имела следующую последовательность:

gagggcctatttcccatgattccttcatttgcataacgatacaaggctgtagagagataattggaattaattgactgtaa
cacaagatatttagtacaaaatacgtgacgtagaaagtaataatttcttggttagttgcagttttaaattatgtttaaatggactatca
tatgcttaccgtaactgaaagtatttcgatttcttgcttatatacttggtgaaaggacgaaacaccgXXXXXXXXXXXXXXXX
XXXXXXXXXGTTGTAGCTCCCTTTTTCATTTGCGGAAAGCGAAATGAAAAACGTTGTTACAAT
AAGAGATGAATTTCTCGCAAAGCTCTGCCTCTTGAAATTTCGGTTTCAAGAGGCATCTTTTTt
gctTCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATC
AATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTacataactacggtaaatggcc
cgcttggtgacggcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattg
acgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaaagtcgccccctattgacgtcaa
tgacggtaaatggcccgctggcattatgccagtacatgacctacgggactttctacttggcagtacatctacgtatttagtcatcgct
attaccatggtgatgcggttttggcagtacaccaatggcggtgatagcggtttgactcacggggatttccaagtcaccacccattgac
gtcaatgggagtttggcaccaaaatcaacgggactttccaaaatgtcgttaataaccccgccccgttgacgcaaatggcggttag
gcgtgtacgggtgggaggtctataaagcAGAGCTCGTTTTAGTGAACCGTCAGAATTAATTCAGATCGATC
TACCaccgccaccATGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAG
CAGCCGAACAGAATAATCCGCTTAACCTACATTCTTGGGCTGGATTTGGGAATTGCGAGTA
TAGGCTGGGCGGTGGTTGAGATCGATGAAGAGAGTAGTCCGATACGCCTTATCGACGTT
GGAGTTAGGACGTTTCGAGAGGGCGGAGGTCGCCAAGACCGGTGAGAGCTTGGCCCTCA
GCCGGCGGCTCGCTCGATCTAGTCGCAGGCTTATAAAGAGGAGGGCTGAGCGCCTTAA
GAAAGCTAAGAGGCTCCTTAAGGCAGAAAAAATTCTGCATAGTATCGACGAAAAGCTGC
CGATAAATGTTTGGCAGCTCCGAGTAAAAGGGCTGAAGGAAAAAATTGGAAAGGCAGGA
GTGGGCGGCGGTACTGCTTCATCTCTCCAAGCACCGGGGCTATCTGTCTCAGCGAAAAA
ACGAAGGTAAGTCAGACAACAAGGAGCTGGGCGCACTTTTGTCCGGGATAGCGTCAAAT
CATCAGATGCTCCAATCAAGTGAGTATCGGACCCCTGCGGAGATCGCCGTTAAAAAGTT
TCAAGTTGAGGAGGGCCACATCAGAAATCAGAGGGGGTCTTACACCCATACGTTCTCTA
GACTCGACCTCCTTGCGGAAATGGAACCTCCTGTTTCAGCGCCAGGCGGAGCTTGGTAAC
TCCTACACGTCCACTACCTCCTGGAAAACCTGACAGCCCTGCTGATGTGGCAGAAGCC
CGCTTTGGCGGGGGATGCCATCCTGAAGATGCTGGGTAAATGCACCTTTGAGCCGTGAG
AATATAAAGCCGCCAAGAATAGTTACTCTGCGGAGCGATTTGTTTGGTTGACAAAGTTGA
ATAACCTGCGCATCCTGGAGAACGGTACCGAGCGCGCACTCAATGATAATGAGCGCTTC
GCCCTCCTGGAACAGCCCTACGAGAAGTCCAAGCTCACCTACGCCAAGTCAGAGCCAT
GCTGGCTCTTAGTGACAACGCGATTTTTAAGGGCGTGCGATACTTGGGCGAGGATAAGA

AAACCGTAGAGTCAAAAACGACTCTGATCGAGATGAAATTCTATCACCAAATTAGAAAG
ACCCTCGGTTCTGCCGAGCTGAAAAAGGAATGGAACGAACTTAAGGGTAACAGCGACCT
GCTCGATGAAATCGGTACCGCATTTAGCCTTTATAAAACGGACGACGACATCTGCCGATA
TTTGGAGGGGAAGCTCCCAGAGCGAGTATTGAATGCACTCCTTGAGAACCTTAATTTTGA
CAAGTTCATTAGCTGTCCCTCAAAGCACTGCATCAAATCCTCCCACTTATGCTGCAAGG
ACAACGATACGACGAAGCCGTCAGCGCGATATATGGAGATCATTACGGAAAAAAGTCCA
CCGAGACCACACGACTGCTTCCTACGATCCCCGCCGATGAGATCAGAAATCCCGTAGTC
CTTCGAACACTTACTCAGGCTAGGAAGGTGATTAATGCGGTAGTTAGGTTGTATGGATCT
CCGGCACGGATACATATAGAAACAGCTCGCGAAGTGGGTAAATCTTACCAAGACCGCAA
GAAATTGGAGAAACAACAGGAGGATAACCGAAAGCAACGAGAATCTGCCGTTAAAAAG
TTTAAGGAAATGTTTCCTCACTTTGTAGGAGAACCGAAGGGTAAAGATATCTTGAAAATG
CGGTTGTACGAGTTGCAGCAAGCTAAGTGTCTCTATAGCGGCAAGAGTTTGGAATTGCA
CCGCCTCCTGGAGAAAGGCTACGTGGAAGTAGACCATGCGCTCCCGTTTTCCCGAACCT
GGGATGATTCTTTCAATAACAAAGTCCTTGTGCTGGCAAATGAGAACCAGAACAAAGGA
AATCTGACTCCTTATGAGTGGTTGGATGGCAAGAATAATTCTGAGCGGTGGCAACATTC
GTTGTCCGCGTCCAAACGTCAGGGTTCAGCTATGCTAAGAAACAAAGGATCCTCAATCA
CAAGCTCGACGAGAAAGGATTCATAGAACGAAATTTGAATGACACTAGGTATGTGGCTC
GATTTCTCTGCAATTTTATTGCTGACAATATGCTCCTCGTTGGGAAGGGAAAGCGGAATG
TTTTTGCATCAAATGGGCAGATAACGGCGCTCTTGAGACATAGATGGGGGCTGCAAAAG
GTGAGAGAGCAAAATGATAGACATCACGCCCTGGATGCCGTTGTAGTCGCCTGTTCAAC
GGTTGCGATGCAGCAAAAGATCACTCGGTTTCGTTAGGTATAACGAAGGGAACGTTTTTA
GTGGAGAGCGCATAGATCGGGAAACAGGCGAAATCATCCCTTTGCATTTCCCAAGTCCT
TGGGCTTTTTTCAAAGAGAATGTGGAAATAAGGATATTCAGTGAAAACCCTAAGTTGGAG
CTTGAGAATCGGTTGCCCGATTATCCCCAGTACAATCATGAGTGGGTTCAACCGCTGTT
GTATCCCGCATGCCAACCCGAAAGATGACCGGGCAGGGTCACATGGAGACTGTGAAAT
CTGCAAAGAGACTTAATGAGGGCCTGTCAGTGTGGAAGGTGCCCTTGACTCAACTGAAA
TTGAGCGACCTCGAGCGCATGGTAAACCGCGATAGAGAAATCGCACTTTATGAGAGTCT
GAAGGCGCGATTGGAACAATTCGGTAATGATCCGGCAAAGGCTTTTCGCTGAGCCATTCT
ACAAGAAGGGTGGAGCGCTGGTTAAGGCTGTCCGACTCGAACAGACACAAAAGTCAGG
GGTCTTGGTCAGAGATGGTAACGGGGTTGCCGACAACGCCTCCATGGTACGAGTAGATG
TTTTACGAAAGGAGGAAAATACTTTCTGGTACCTATCTATACCTGGCAAGTTGCCAAGG
GAATACTCCCGAATAGGGCGGCGACCCAGGGAAAGGATGAAAACGACTGGGATATAAT
GGATGAAATGGCTACGTTTCAGTTTAGCTTGTGCCAGAATGACCTCATAAACTGGTAAC
CAAAAAAAGACTATATTCGGGTATTTCAATGGCCTTAATCGGGCAACTTCCAATATCAA
CATCAAGGAACATGATCTGGATAAGAGCAAGGGAAAGCTTGGTATCTATCTCGAAGTTG
GAGTCAAGCTCGCTATTTCCCTCGAGAAATATCAAGTAGATGAACTGGGAAAGAATATAC
GGCCATGCCGGCCCCACAAAAGACAACACGTACGGTTCAAAGGCCGGCGGCCACGAA
AAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGC

TTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCGGCGCA
 ACAAACCTTCTCTCTGCTGAAACAAGCCGGAGATGTCTGAAGAGAATCCTGGACCGgtgagcaag
 ggcgaggagctgtcaccggggtggtgccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcg
 agggcgagggcgatgccacctacggcaagctgacctgaagttcatctgcaccaccggcaagctgcccgtgccctggccaccct
 cgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttctcaagtccgccaatgcc
 cgaaggctacgtccaggagcgcaccatcttcttaaggacgacggcaactacaagacccgcgcgaggtgaagttcgagggcga
 caccctggtgaaccgcatcgagctgaaggcgatcgacttaaggaggacggcaacatcctggggcacaagctggagtacaacta
 caacagccacaacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttaagatccgccacaacatcgagga
 cggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacct
 gagcaccacagtcggccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgcccggggat
 cactctcgcatggacgagctgtacaagTAA

В последовательности плазмиды выделены следующие части: U6 промотор (первый участок, прописные буквы), последовательность, комплиментарная протоспейсеру ("XXX-XXX"), консервативная часть sgRNA (третий участок, заглавные буквы), ген PpCas9 (выделен жирным шрифтом), ген GFP (последний участок, прописные буквы).

Плазмиды с PpCas9 или SpCas9 были трансфицированы в культуру клеток человека HEK293T с помощью реагента липофектамина 2000. По истечении 72 часов после трансфекции, клетки лизировали и с полученных лизатов проводилась ПЦР для наработки участков, включающих целевые сайты изменения геномной ДНК. Полученные ПЦР фрагменты подвергали *in vitro* реакции с T7 эндонуклеазой I для определения частоты внесения вставок и делеций в целевых сайтах геномной ДНК. Продукты реакции наносили на агарозный гель и подвергали электрофорезу. На Фиг. 14А показано, что PpCas9 активно вносит изменения в генах EMX1 и GRIN2b, с эффективностью, схожей с эффективностью работы описанной в уровне техники SpCas9 нуклеазы.

Данный эксперимент показал, что для эффективного изменения геномной ДНК PpCas9 требует удлиненные sgRNA по сравнению с SpCas9: в приведенном примере эффективность генетических модификаций выше при использовании sgRNA с последовательностью комплементарной ДНК мишени длиной 24 нуклеотида (по сравнению с длиной в 20 нуклеотидов).

Высокопроизводительное секвенирование подтвердило внесение изменений в целевых сайтах ДНК. На Фиг. 14Б приведены примеры обнаруживаемых изменений в нуклеотидной последовательности гена EMX1.

Для доставки NLS_PpCas9-NLS в клетки человека также может быть использована доставка в виде рибонуклеинового комплекса. Она осуществляется путем инкубации рекомбинантной формы PpCas9-NLS с направляющими РНК в CutSmart буфере (NEB). Рекомбинантный белок получают из бактериальных клеток-продуцентов, очищая его с

помощью аффинной хроматографии (NiNTA, Qiagen) разделением по размеру (Superdex 200).

Белок смешивают с РНК в соотношении 1:2 (PpCas9 NLS : sgRNA), инкубируют в течение 10 минут на комнатной температуре, затем смесь трансфицируют в клетки.

Далее проводится анализ экстрагированной из них ДНК на предмет вставок-делеций в целевом ДНК сайте (как описано выше).

Охарактеризованная в настоящем изобретении нуклеаза PpCas9 из бактерии *Pasteurella pneumotropica* может быть доставлена для модификации ДНК в клетки различного происхождения с помощью стандартных подходов и методов, известных специалистам. PpCas9 имеет ряд преимуществ относительно ранее охарактеризованных Cas9 белков.

PpCas9 обладает коротким, двухбуквенным, отличным от других известных Cas нуклеаз PAM мотивом, необходимым для функционирования системы. В изобретении было показано, что для успешного функционирования PpCas9 *in vivo* достаточно присутствия короткого PAM мотива (RTT), расположенного в 4 нуклеотидах от протоспейсера.

Известные на сегодняшний день многие малоразмерные Cas нуклеазы, способные вносить двунитевые разрывы в ДНК, имеют сложные многобуквенные PAM последовательности, ограничивающие выбор последовательностей, пригодных для разрезания. Среди изученных Cas нуклеаз, распознающих короткие PAM, только PpCas9 может распознавать последовательности, фланкированные RTT мотивом.

Второе преимущество PpCas9 – малый размер белка (1055 а.о). На сегодняшний день это единственный изученный малоразмерный белок, обладающий трехбуквенной RTT PAM последовательностью.

PpCas9 – новая малоразмерная Cas нуклеаза, имеющая короткий, простой в использовании PAM, отличающийся от известных на сегодняшний день PAM последовательностей других нуклеаз. Белок PpCas9 разрезает с высокой эффективностью различные ДНК-мишени, в том числе геномную ДНК в клетках человека при 37 °C, и может стать основой нового инструмента геномного редактирования.

Несмотря на то, что изобретение описано со ссылкой на раскрываемые варианты воплощения, для специалистов в данной области должно быть очевидно, что конкретные подробно описанные случаи приведены лишь в целях иллюстрирования настоящего изобретения, и их не следует рассматривать как каким-либо образом ограничивающие объем изобретения. Должно быть, понятно, что возможно осуществление различных модификаций без отступления от сути настоящего изобретения.

<110> Автономная некоммерческая образовательная организация высшего образования Сколковский институт науки и технологий /Skolkovo Institute of Science and Technology

<130> 424316

<150> RU 2019136164

<151> 2019-11-11

<160> 2

 $\langle 210 \rangle$ 1

<211> 1055

<212> PRT

<213> Pasteurella pneumotropica

 $\langle 220 \rangle$

<223> белок, гомологичный Cas9

 $\langle 400 \rangle$ 1

22

Asp	Asn	Glu	Arg	Phe	Ala	Leu	Leu	Glu	Gln	Pro	Tyr	Glu	Lys	Ser	Lys		
290						295					300						
Leu	Thr	Tyr	Ala	Gln	Val	Arg	Ala	Met	Leu	Ala	Leu	Ser	Asp	Asn	Ala		
305				310						315					320		
Ile	Phe	Lys	Gly	Val	Arg	Tyr	Leu	Gly	Glu	Asp	Lys	Lys	Thr	Val	Glu		
			325						330					335			
Ser	Lys	Thr	Thr	Leu	Ile	Glu	Met	Lys	Phe	Tyr	His	Gln	Ile	Arg	Lys		
			340					345					350				
Thr	Leu	Gly	Ser	Ala	Glu	Leu	Lys	Lys	Glu	Trp	Asn	Glu	Leu	Lys	Gly		
		355					360					365					
Asn	Ser	Asp	Leu	Leu	Asp	Glu	Ile	Gly	Thr	Ala	Phe	Ser	Leu	Tyr	Lys		
370						375					380						
Thr	Asp	Asp	Asp	Ile	Cys	Arg	Tyr	Leu	Glu	Gly	Lys	Leu	Pro	Glu	Arg		
385				390						395					400		
Val	Leu	Asn	Ala	Leu	Leu	Glu	Asn	Leu	Asn	Phe	Asp	Lys	Phe	Ile	Gln		
			405						410					415			
Leu	Ser	Leu	Lys	Ala	Leu	His	Gln	Ile	Leu	Pro	Leu	Met	Leu	Gln	Gly		
			420					425					430				
Gln	Arg	Tyr	Asp	Glu	Ala	Val	Ser	Ala	Ile	Tyr	Gly	Asp	His	Tyr	Gly		
		435					440					445					
Lys	Lys	Ser	Thr	Glu	Thr	Thr	Arg	Leu	Leu	Pro	Thr	Ile	Pro	Ala	Asp		
450						455					460						
Glu	Ile	Arg	Asn	Pro	Val	Val	Leu	Arg	Thr	Leu	Thr	Gln	Ala	Arg	Lys		
465					470					475					480		
Val	Ile	Asn	Ala	Val	Val	Arg	Leu	Tyr	Gly	Ser	Pro	Ala	Arg	Ile	His		
			485						490					495			
Ile	Glu	Thr	Ala	Arg	Glu	Val	Gly	Lys	Ser	Tyr	Gln	Asp	Arg	Lys	Lys		
			500				505					510					
Leu	Glu	Lys	Gln	Gln	Glu	Asp	Asn	Arg	Lys	Gln	Arg	Glu	Ser	Ala	Val		
		515					520					525					
Lys	Lys	Phe	Lys	Glu	Met	Phe	Pro	His	Phe	Val	Gly	Glu	Pro	Lys	Gly		
530					535						540						
Lys	Asp	Ile	Leu	Lys	Met	Arg	Leu	Tyr	Glu	Leu	Gln	Gln	Ala	Lys	Cys		
545				550						555					560		
Leu	Tyr	Ser	Gly	Lys	Ser	Leu	Glu	Leu	His	Arg	Leu	Leu	Glu	Lys	Gly		
			565						570					575			
Tyr	Val	Glu	Val	Asp	His	Ala	Leu	Pro	Phe	Ser	Arg	Thr	Trp	Asp	Asp		
			580					585					590				
Ser	Phe	Asn	Asn	Lys	Val	Leu	Val	Leu	Ala	Asn	Glu	Asn	Gln	Asn	Lys		
		595					600					605					
Gly	Asn	Leu	Thr	Pro	Tyr	Glu	Trp	Leu	Asp	Gly	Lys	Asn	Asn	Ser	Glu		
610						615					620						
Arg	Trp	Gln	His	Phe	Val	Val	Arg	Val	Gln	Thr	Ser	Gly	Phe	Ser	Tyr		
625				630						635					640		
Ala	Lys	Lys	Gln	Arg	Ile	Leu	Asn	His	Lys	Leu	Asp	Glu	Lys	Gly	Phe		
			645						650					655			
Ile	Glu	Arg	Asn	Leu	Asn	Asp	Thr	Arg	Tyr	Val	Ala	Arg	Phe	Leu	Cys		
			660					665					670				
Asn	Phe	Ile	Ala	Asp	Asn	Met	Leu	Leu	Val	Gly	Lys	Gly	Lys	Arg	Asn		
		675					680					685					
Val	Phe	Ala	Ser	Asn	Gly	Gln	Ile	Thr	Ala	Leu	Leu	Arg	His	Arg	Trp		
690						695					700						
Gly	Leu	Gln	Lys	Val	Arg	Glu	Gln	Asn	Asp	Arg	His	His	Ala	Leu	Asp		
705				710						715					720		
Ala	Val	Val	Val	Ala	Cys	Ser	Thr	Val	Ala	Met	Gln	Gln	Lys	Ile	Thr		
			725						730					735			
Arg	Phe	Val	Arg	Tyr	Asn	Glu	Gly	Asn	Val	Phe	Ser	Gly	Glu	Arg	Ile		
			740					745					750				
Asp	Arg	Glu	Thr	Gly	Glu	Ile	Ile	Pro	Leu	His	Phe	Pro	Ser	Pro	Trp		
		755					760					765					

Ala Phe Phe Lys Glu Asn Val Glu Ile Arg Ile Phe Ser Glu Asn Pro
 770 775 780
 Lys Leu Glu Leu Glu Asn Arg Leu Pro Asp Tyr Pro Gln Tyr Asn His
 785 790 795 800
 Glu Trp Val Gln Pro Leu Phe Val Ser Arg Met Pro Thr Arg Lys Met
 805 810 815
 Thr Gly Gln Gly His Met Glu Thr Val Lys Ser Ala Lys Arg Leu Asn
 820 825 830
 Glu Gly Leu Ser Val Leu Lys Val Pro Leu Thr Gln Leu Lys Leu Ser
 835 840 845
 Asp Leu Glu Arg Met Val Asn Arg Asp Arg Glu Ile Ala Leu Tyr Glu
 850 855 860
 Ser Leu Lys Ala Arg Leu Glu Gln Phe Gly Asn Asp Pro Ala Lys Ala
 865 870 875 880
 Phe Ala Glu Pro Phe Tyr Lys Lys Gly Gly Ala Leu Val Lys Ala Val
 885 890 895
 Arg Leu Glu Gln Thr Gln Lys Ser Gly Val Leu Val Arg Asp Gly Asn
 900 905 910
 Gly Val Ala Asp Asn Ala Ser Met Val Arg Val Asp Val Phe Thr Lys
 915 920 925
 Gly Gly Lys Tyr Phe Leu Val Pro Ile Tyr Thr Trp Gln Val Ala Lys
 930 935 940
 Gly Ile Leu Pro Asn Arg Ala Ala Thr Gln Gly Lys Asp Glu Asn Asp
 945 950 955 960
 Trp Asp Ile Met Asp Glu Met Ala Thr Phe Gln Phe Ser Leu Cys Gln
 965 970 975
 Asn Asp Leu Ile Lys Leu Val Thr Lys Lys Lys Thr Ile Phe Gly Tyr
 980 985 990
 Phe Asn Gly Leu Asn Arg Ala Thr Ser Asn Ile Asn Ile Lys Glu His
 995 1000 1005
 Asp Leu Asp Lys Ser Lys Gly Lys Leu Gly Ile Tyr Leu Glu Val Gly
 1010 1015 1020
 Val Lys Leu Ala Ile Ser Leu Glu Lys Tyr Gln Val Asp Glu Leu Gly
 1025 1030 1035 1040
 Lys Asn Ile Arg Pro Cys Arg Pro Thr Lys Arg Gln His Val Arg
 1045 1050 1055

<210> 2

<211> 3168

<212> DNA

<213> Pasteurella pneumotropica

<220>

<223> белок, гомологичный Cas9

<400> 2

atgcaaaata atccattaata ttacatttta gggtagatt taggcattgc ttctattggt 60

tgggcggttg tggaattga tgaggagagt tcacctattc gcttaattga tgtgggcgtc 120

cgtacatttg aacgggctga agtcgctaaa accggcgaaa gtttagcatt gtctcgtcgt 180

ttagctcgtt catcacggcg attaatataa cgccgagcag agcgattaaa aaaagcaaaa 240

cgtttattaa aagcagaaaa gattttacat tctattgatg aaaaattacc cattaatgtt 300

tggcagcttc gagtaaaagg attgaaggaa aaactcgaac gtcaggagtg ggcagcggtt 360

ttattacatt tgtcaaagca tcgtggctat ttatcacaac gtaaaaatga gggtaaaagt 420

gataataaag agctgggggc attactttca ggtatcgcaa gtaaccacca aatgttgcaa 480

tcctccgaat atcgtacccc tgcagaaatt gcagtcaaaa aatttcaagt agaagaagga 540

catattcgta atcaacgtgg atcttatacc cataccttta gccgtttgga tttgttgga 600

gaaatggaat tattatttca acgccaagct gagttaggca attcttacac gtccaccaca 660

ttattagaaa atttgacggc gttactaatg tggcaaaagc cagctcttgc gggatgatgcg 720

attttaaaaa tggtgggcaa gtgtaccttc gaaccagcg aatataaagc cgcaaaaaat 780

agttattctg ctgaacgttt tgtgtggtta accaagctga ataatttacg cattttagaa 840

aatggcacgg aaagagcttt aaatgacaat gaacgttttg ctttgcttga gcaaccgtat 900

gagaaatcaa aattaactta tgctcaagtg agagcaatgc ttgcgttatc tgataatgct 960

attttcaaag gggttcggtta tttaggcgaa gataaaaaaa cagtagagag caaaactacg 1020

ttgatagaaa tgaagtttta tcatcaaadc cgcaaaacat taggcagtgc agaattaaaa 1080

aaggaatgga atgagttaaa aggcaattcc gatttattag atgagattgg cacggcattt 1140

tcgttggtata aaacggatga tgatatttgc cgttattttag agggaaaact accagaaagg 1200

gtattaaatg cgttattgga aaatttaaatt ttcgataaat ttattcaact ttcacttaaa 1260

gccttacacc aaattttacc attgatgctg caagggcaac gttatgatga ggcgggttct 1320

gcgatttatg gtgatcatta tggtaaaaaa tcgacagaaa caaccgctt gttgccgact 1380

attcctgccg atgaaatccg aaatcctgtg gtattacgca ccctgaccca agcccgtaaa 1440

gtgatcaatg cgggtgtgctg gttatatggt tcgcctgccc gtattcatat tgaaacagcg 1500

agagaagtcg gcaaacttta ccaagatcgt aaaaaacttg aaaaacagca agaagataat 1560

cgtaagcaac gtgaaagtgc ggtcaaaaaa tttaaagaaa tgtttccgca ttttgtggg 1620

gagccgaaag gtaaagatat tttaaaaatg cgattgtatg agttacaaca agcgaaatgt 1680

ttatatcttg gaaaatcttt agaacttcat cgtttgcttg agaaggggta tgtagaagtg 1740

gatcacgctt tgccattttc tcgcacgtgg gatgatagct ttaataataa agtactgggtg 1800

cttgccaacg agaaccaaaa taaaggcaat ttaacgcctt atgaatgggt agatggtaaa 1860

aataacagtg agcgttggca acattttggt gtacgagtac aaaccagcg tttctcttat 1920

gctaaaaaac aacgcatttt gaaccataaa ttggatgaaa aagggtttat cgaacgtaat 1980

ttaaacgata ctgcgtatgt agctcgtttc ttatgtaact ttattgccga taatatgttg 2040

ttggttggta aaggcaagcg aaacgtggtt gcttcaaacg ggcaaatac ggcgttattg 2100

cggcatcgtt ggggcttaca aaaagtgcgt gaacagaatg atcgccacca cgcactggac 2160

gcggttggtg tggcttgctc tactgtggca atgcaacaaa aaatcactcg atttgtgaga 2220

tataacgaag gaaatgtctt tagcggtgaa cgtatcgatc gtgaaactgg cgagattatt 2280

ccattacatt ttccaagccc ttgggctttt ttcaaagaga atgtggaaat tcgcattttt 2340

agtgaaaatc caaaattgga attagaaaat cgcctgcctg attatccgca atataatac 2400

gaatgggtgc aaccattggt tgtttcgaga atgccaaccc gaaaaatgac agggcaagg 2460

catatggaaa cggtaaaatc cgcaaaacga ttaaatgaag gtttaagtgt gttaaaagtc 2520

cctttaacac aacttaaatt gagtgattta gaacgaatgg ttaatcgtga tcgtgaaatt 2580

gcattgtatg aatccttaaa agcacgttta gagcaatttg gtaacgaccc agccaaagcc 2640

tttgccgaac cattctataa aaagggtggg gcattagtca aagcagtccg attggaacaa 2700

acacaaaaat cggggggtatt agtacgtgat ggtaacggtg ttgcggataa tgcttcaatg	2760
gtacggggttg atgttttttac taaagggtgga aaatatttct tagtgccgat ttatacttgg	2820
caggtagcga aagggtttt accgaatagg gctgcgacac aaggtaaaga tgaaaatgat	2880
tgggatatta tggatgaaat ggctactttc caattttctc tatgtcaaaa tgatctaatt	2940
aaattagtta ccaaaaagaa aacaatcttt ggatatttta atggattaaa tagagctact	3000
agcaatataa atattaaaga gcatgatcta gataagtcta aagggaaatt aggtatttac	3060
ttagaagttg gtgtaaaact agctatttcc cttgaaaagt accaagtcga cgaactcggc	3120
aaaaatatcc gtccttgctg tccgactaaa cgacagcacg tgcgttaa	3168

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Применение белка, содержащего аминокислотную последовательность SEQ ID NO: 1, или содержащего аминокислотную последовательность, которая по меньшей мере на 95% идентична аминокислотной последовательности SEQ ID NO: 1 и имеет отличия по сравнению с SEQ ID NO: 1 только в неконсервативных аминокислотных остатках, для образования двунитевого разрыва в молекуле ДНК, расположенного непосредственно перед нуклеотидной последовательностью 5'-NNNN(A/G)TT-3' в указанной молекуле ДНК.

2. Применение по п. 1, характеризующееся тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 35 °С до 45 °С.

3. Применение белка по п. 1, где белок содержит аминокислотную последовательность SEQ ID NO: 1.

4. Применение по п. 1, характеризующееся тем, что образование двунитевого разрыва в молекуле ДНК происходит в геномной ДНК клетки млекопитающего.

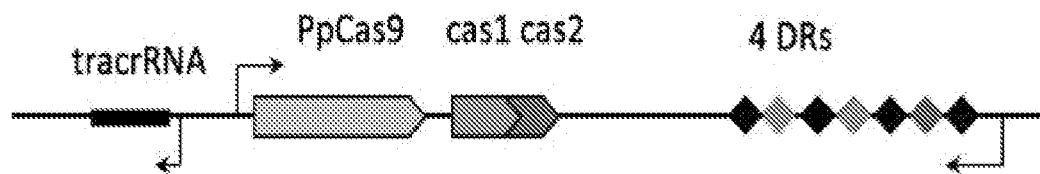
5. Применение по п. 4, характеризующееся тем, что образование двунитевого разрыва в молекуле ДНК приводит к изменению геномной ДНК указанной клетки млекопитающего.

6. Способ изменения последовательности геномной ДНК в клетке одноклеточного или многоклеточного организма, содержащей геномную ДНК, включающий введение в указанную клетку организма эффективного количества: а) белка, содержащего аминокислотную последовательность SEQ ID NO: 1, или нуклеиновой кислоты, кодирующей белок, содержащий аминокислотную последовательность SEQ ID NO: 1, и б) направляющей РНК, содержащей последовательность, образующую дуплекс с нуклеотидной последовательностью участка геномной ДНК организма, непосредственно примыкающей к нуклеотидной последовательности 5'-NNNN(A/G)TT-3', и взаимодействующей с указанным белком после образования дуплекса, или последовательности ДНК, кодирующей указанную направляющую РНК;

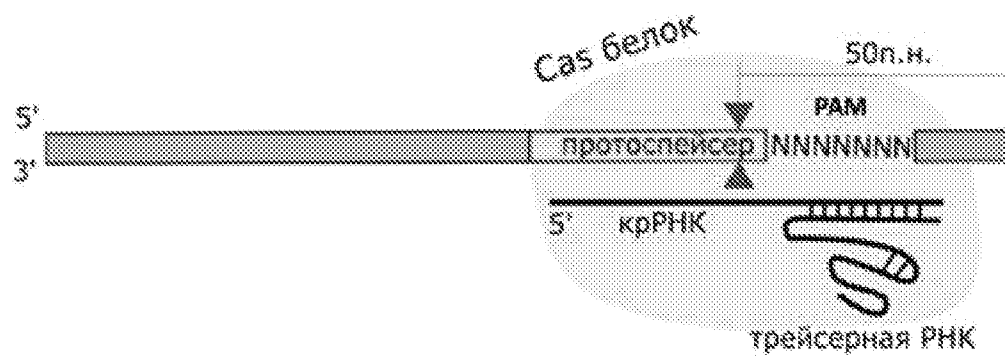
при этом взаимодействие указанного белка с направляющей РНК и нуклеотидной последовательностью 5'-NNNN(A/G)TT-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNN(A/G)TT-3'.

7. Способ по п. 6, дополнительно включающий введение экзогенной последовательности ДНК одновременно с направляющей РНК.

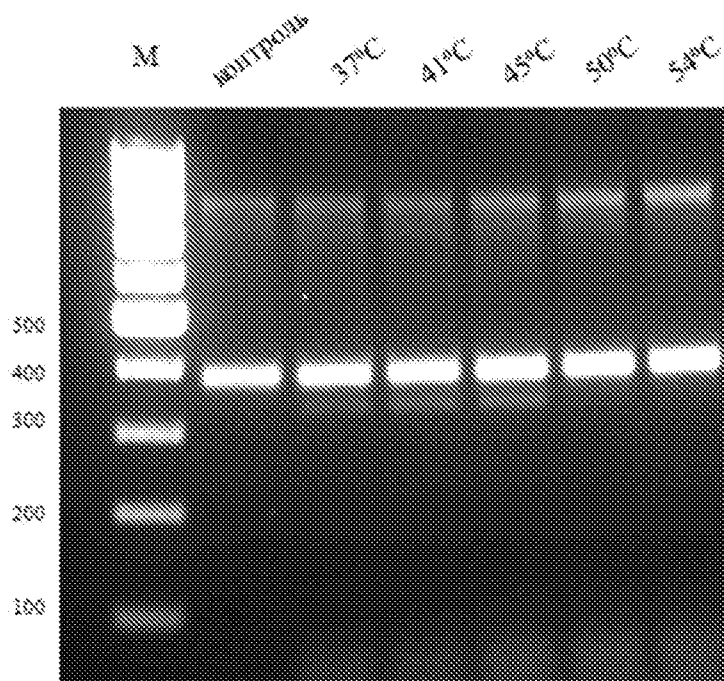
8. Способ по п. 6, характеризующееся тем, что указанная клетка является клеткой млекопитающего.



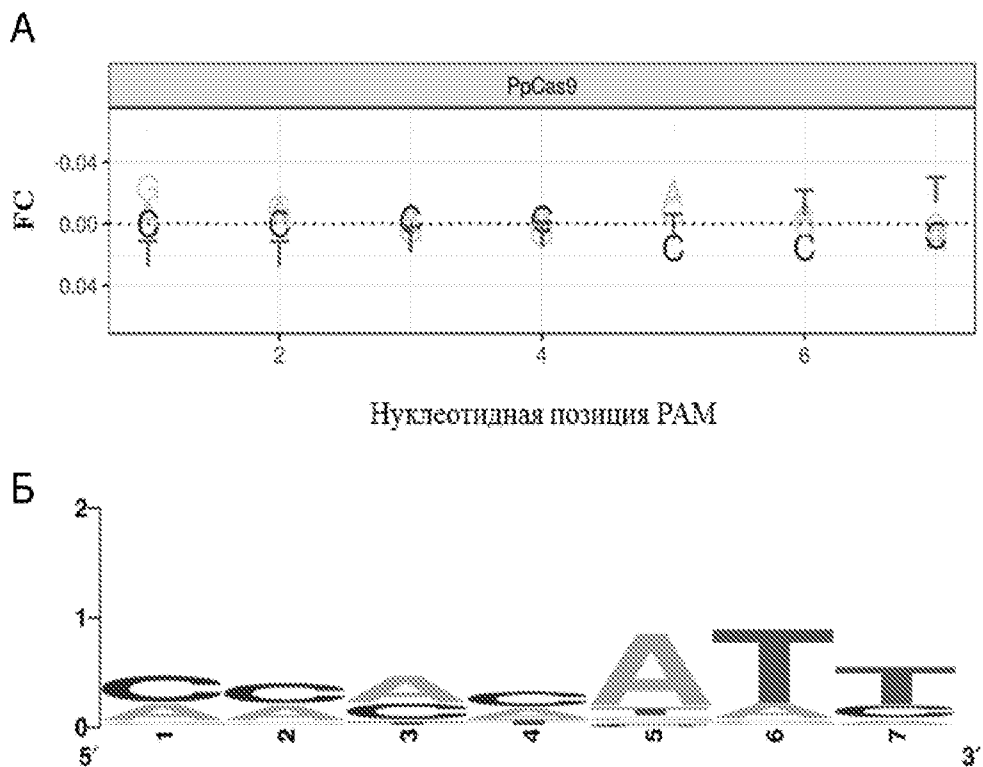
Фиг. 1



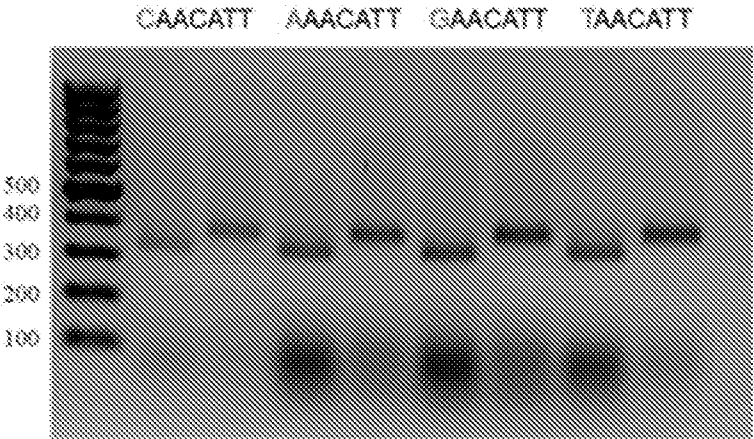
Фиг. 2



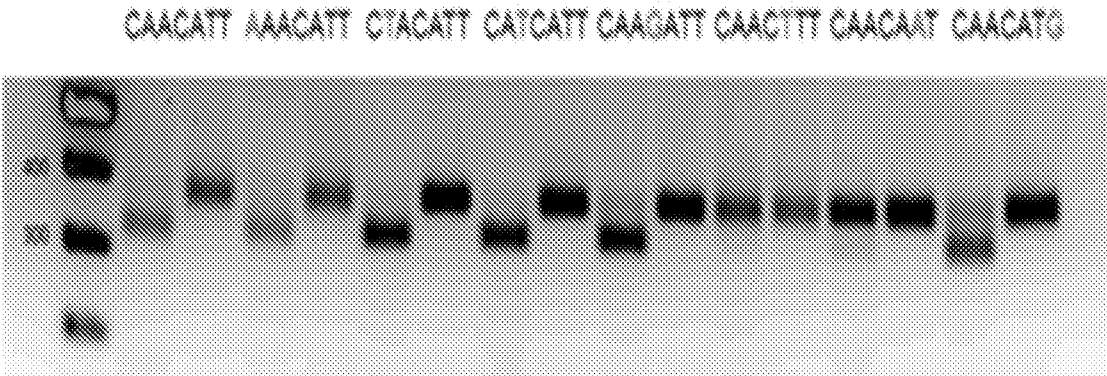
Фиг. 3



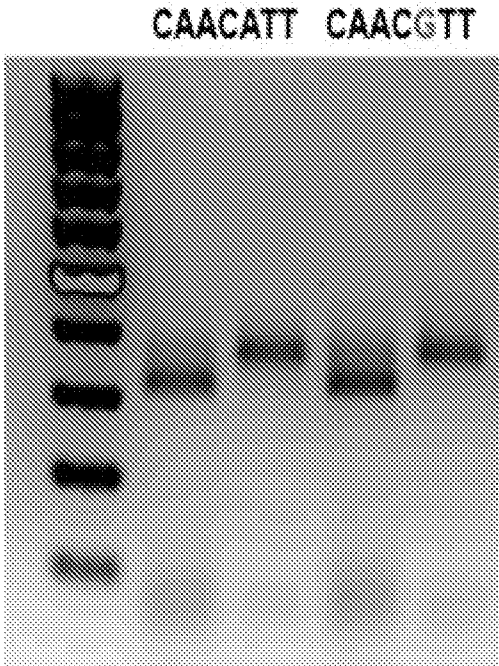
Фиг. 4



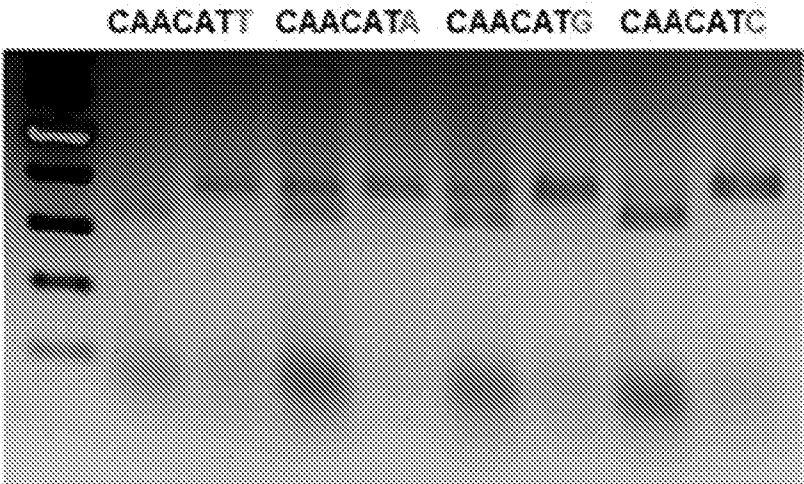
Фиг. 5



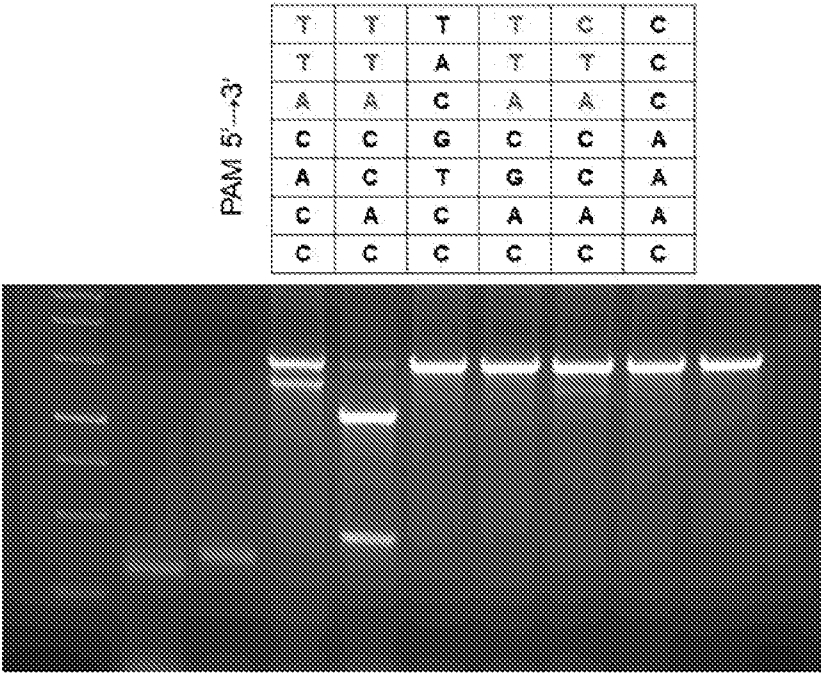
Фиг. 6



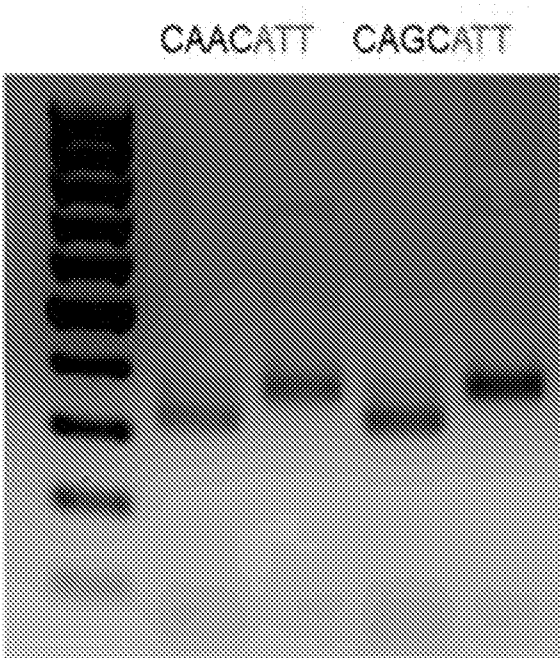
Фиг. 7



Фиг. 8



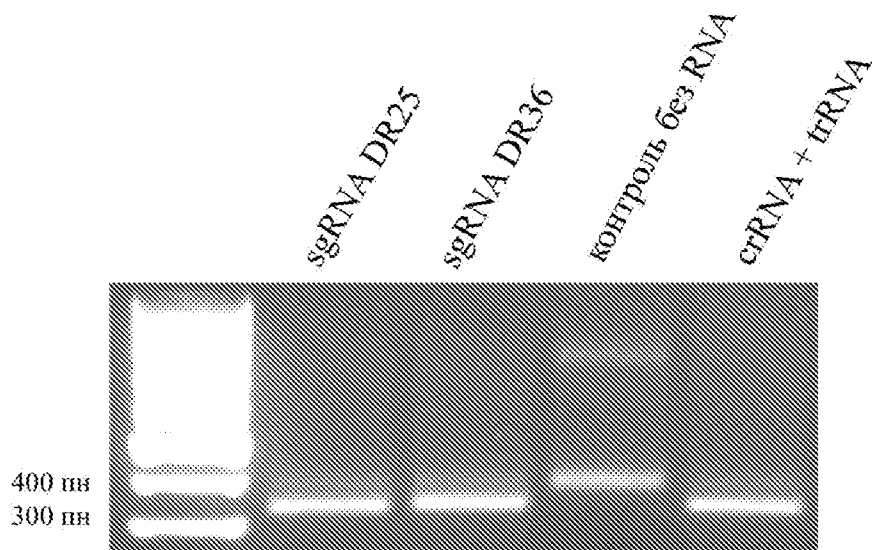
Фиг. 9



Фиг. 10



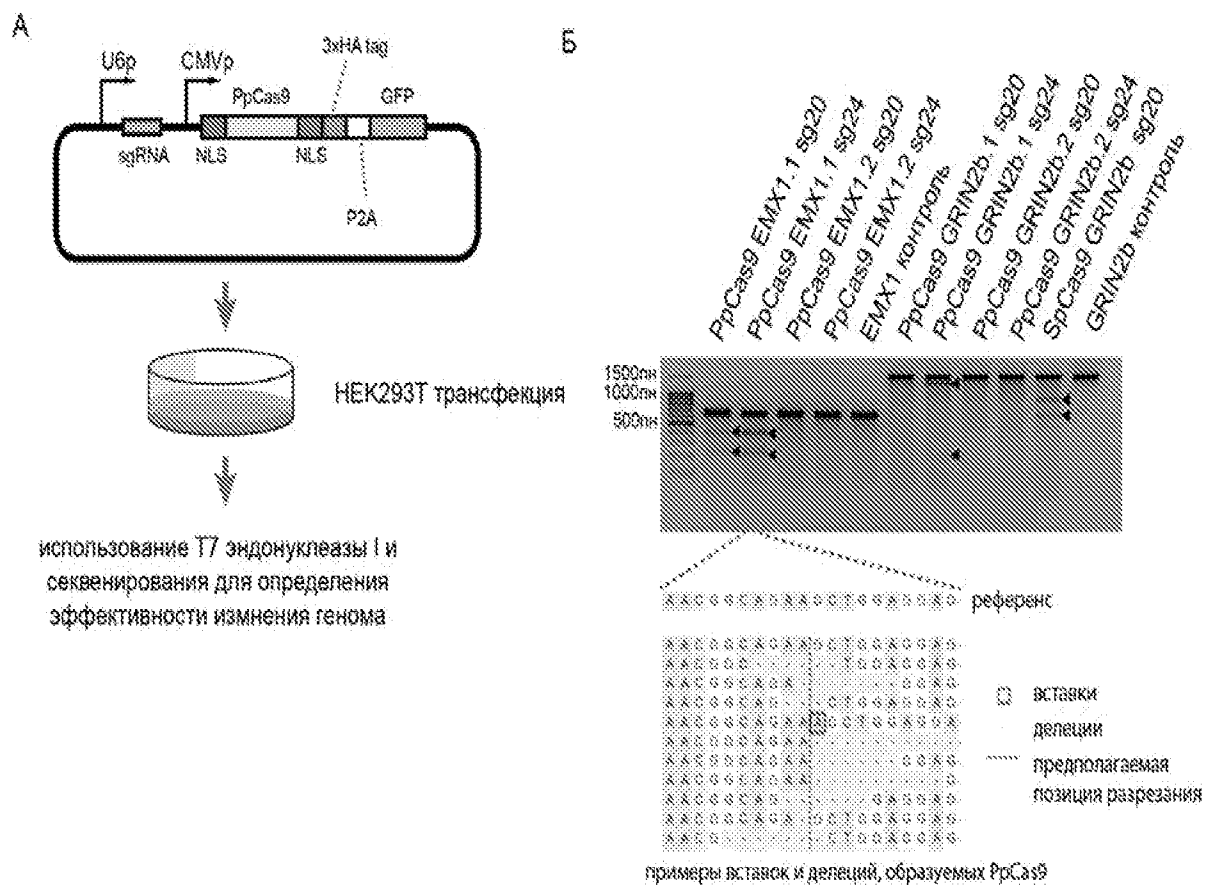
Фиг. 11



Фиг. 12

[illegible]

Фиг. 13



Фиг. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/RU 2020/050145

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/63 (2020.02)
C12N 15/75 (2020.02)

C12N 9/10 (2020.02)
C12Q 1/6806 (2020.02)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N 15/63, C12N 15/75, C12N 9/10, C12Q 1/6806

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAPATIS, ESPACENET, PatSearch (RUPTO internal), USPTO, PATENTSCOPE,
NCBI, EMBL-EBI, ScienceDirect, PubMed, E-Library, Google

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WENINGER Astrid et al. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast <i>Pichia pastoris</i> . Journal of Biotechnology, 2016, 235, c. 139-149, doi: 10.1016/j.jbiotec.2016.03.027	1-8
A	MARTIN Negin R. et al. Ep masse lenti viral gene delivery to mouse fertilized eggs via laser perforation of zona pellucida. Transgenic research, 2018, 27(1), c. 39-49 (1-16), doi: 10.1007/s 11248-017-0056-8	1-8
A	KHLESTKINA E.K. et al. Perspektivy ispolzovania proryvnykh tekhnologii v selektsii: sistema CRISPR/CAS9 dlia redaktirovaniya genoma rasteny. Genetika, 2016, tom 52, N° 7, p. 774-787, doi:10.7868/S0016675816070055	1-8

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

26 October 2020 (26.10.2020)

Date of mailing of the international search report

05 November 2020 (05.11.2020)

Name and mailing address of the ISA/

RU

Authorized officer

Facsimile No.

Telephone No.

ОТЧЕТ О МЕЖДУНАРОДНОМ ПОИСКЕ

Номер международной заявки

PCT/RU 2020/050145

A. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ <i>C12N 15/63 (2020.02)</i> <i>C12N 15/75 (2020.02)</i> <i>C12N 9/10 (2020.02)</i> <i>C12Q 1/6806 (2020.02)</i> Согласно Международной патентной классификации МПК		
B. ОБЛАСТЬ ПОИСКА Проверенный минимум документации (система классификации с индексами классификации) C12N 15/63, C12N 15/75, C12N 9/10, C12Q 1/6806 Другая проверенная документация в той мере, в какой она включена в поисковые подборки Электронная база данных, использовавшаяся при поиске (название базы и, если, возможно, используемые поисковые термины) EAPATIS, ESPACENET, PatSearch (RUPTO internal), USPTO, PATENTSCOPE, NCBI, EMBL-EBI, ScienceDirect, PubMed, E-Library, Google		
C. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ:		
Категория*	Цитируемые документы с указанием, где это возможно, релевантных частей	Относится к пункту №
A	WENINGER Astrid и др. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast <i>Pichia pastoris</i> . Journal of Biotechnology, 2016, 235, с. 139-149, doi: 10.1016/j.jbiotec.2016.03.027	1-8
A	MARTIN Negin P. и др. <i>En masse</i> lentiviral gene delivery to mouse fertilized eggs via laser perforation of zona pellucida. Transgenic research, 2018, 27(1), с. 39-49 (1-16), doi: 10.1007/s11248-017-0056-8	1-8
A	ХЛЕСТКИНА Е.К. и др. Перспективы использования прорывных технологий в селекции: система CRISPR/CAS9 для редактирования генома растений. Генетика, 2016, том 52, № 7, с. 774-787, doi: 10.7868/S0016675816070055	1-8
☐ последующие документы указаны в продолжении графы C. ☐ данные о патентах-аналогах указаны в приложении		
* Особые категории ссылочных документов: “А” документ, определяющий общий уровень техники и не считающийся особо релевантным “D” документ, цитируемый заявителем в международной заявке “Е” более ранняя заявка или патент, но опубликованная на дату международной подачи или после нее “L” документ, подвергающий сомнению притязание(я) на приоритет, или который приводится с целью установления даты публикации другого ссылочного документа, а также в других целях (как указано) “O” документ, относящийся к устному раскрытию, использованию, экспонированию и т.д. “P” документ, опубликованный до даты международной подачи, но после даты испрашиваемого приоритета	“Т” более поздний документ, опубликованный после даты международной подачи или приоритета, но приведенный для понимания принципа или теории, на которых основывается изобретение “Х” документ, имеющий наиболее близкое отношение к предмету поиска; заявленное изобретение не обладает новизной или изобретательским уровнем, в сравнении с документом, взятым в отдельности “Y” документ, имеющий наиболее близкое отношение к предмету поиска; заявленное изобретение не обладает изобретательским уровнем, когда документ взят в сочетании с одним или несколькими документами той же категории, такая комбинация документов очевидна для специалиста “&” документ, являющийся патентом-аналогом	
Дата действительного завершения международного поиска 26 октября 2020 (26.10.2020)		Дата отправки настоящего отчета о международном поиске 05 ноября 2020 (05.11.2020)
Наименование и адрес ISA/RU: Федеральный институт промышленной собственности, Бережковская наб., 30-1, Москва, Г-59, ГСП-3, Россия, 125993 Факс: (8-495) 531-63-18, (8-499) 243-33-37		Уполномоченное лицо: Тямкин А. Телефон № +7 (495) 531-64-81