

(12) МЕЖДУНАРОДНАЯ ЗАЯВКА, ОПУБЛИКОВАННАЯ В
СООТВЕТСТВИИ С ДОГОВОРом О ПАТЕНТНОЙ КООПЕРАЦИИ (РСТ)

(19) Всемирная Организация
Интеллектуальной Собственности
Международное бюро

(43) Дата международной публикации
17 декабря 2020 (17.12.2020)



WIPO | РСТ



(10) Номер международной публикации
WO 2020/251413 A1

- (51) Международная патентная классификация:
C12Q 1/68 (2018.01) *C07K 14/195* (2006.01)
C12N 15/87 (2006.01)
- (21) Номер международной заявки: РСТ/RU2020/050137
- (22) Дата международной подачи:
30 июня 2020 (30.06.2020)
- (25) Язык подачи: Русский
- (26) Язык публикации: Русский
- (30) Данные о приоритете:
2019118061 11 июня 2019 (11.06.2019) RU
- (71) Заявитель: **АВТОНОМНАЯ НЕКОММЕРЧЕСКАЯ
ОБРАЗОВАТЕЛЬНАЯ ОРГАНИЗАЦИЯ ВЫСШЕ-
ГО ОБРАЗОВАНИЯ СКОЛКОВСКИЙ ИНСТИ-
ТУТ НАУКИ И ТЕХНОЛОГИЙ (SKOLKOVO
INSTITUTE OF SCIENCE AND TECHNOLOGY)**
[RU/RU]; Московская обл., Одинцовский р-н, ул. Но-
вая, 100 Дер. Сколково, 143025, Der. Skolkovo (RU).
- (72) Изобретатели: **СЕВЕРИНОВ, Константин Викто-
рович (SEVERINOV, Konstantin Viktorovich)**; ул. Дми-
трия Ульянова, 4/2, кв. 368 Москва, 188301, Moskva
(RU). **ШМАКОВ, Сергей Анатольевич (SHMAKOV,
Sergey Anatolievich)**; Московская обл., ул. Пионерская,
14, кв. 15 Воскресенск, 140200, Voskresensk (RU). **АР-
ТАМОНОВА, Дарья Николаевна (ARTAMONOVA,
Daria Nikolaevna)**; ул. Маршала Тимошенко, 9-59
Москва, 121359, Moskva (RU). **ГОРЯНИН, Иг-
натий Игоревич (GORIANIN, Ignaty Igorevich)**;
ул. Кировоградская 16, здание 1, кв. 351 Москва,
117587, Moskva (RU). **МУШАРОВА, Ольга Сергеев-
на (MUSHAROVA, Olga Sergeevna)**; ул. Рогова, 7, стр.
2, кв. 289 Москва, 123098, Moskva (RU). **АНДРЕЕВА,
Юлия Валерьевна (ANDREEVA, Julia Valerevna)**;
Московский, ул. Москвитина 3-2-108 Москва, 142784,
Moskva (RU). **ЗЮБКО, Татьяна Игоревна (ZYUBKO,**

Tatiana Igorevna); ул. Яхтенная 3, стр. 1 Санкт-Пе-
тербург, 197374, Sankt-Peterburg (RU). **ФЕДОРОВА,
Яна Витальевна (FEDOROVA, Yana Vitalievna)**; Ле-
нинградская обл., ул. Куприна 40, 14 Гатчина, 188301,
Gatchina (RU). **ХОДОРКОВСКИЙ, Михаил Алек-
сеевич (KHODORKOVSKI, Mikhail Alekseevich)**;
пр. Косыгина, д. 23, к. 1, кв. 268 Санкт-Петербург,
195426, Sankt-Peterburg (RU). **ПОБЕГАЛОВ, Геор-
гий Евгеньевич (POBEGALOV, George Evgenovich)**;
ул. Партизана Германа, д.12, кв.317 Санкт-Петербург,
198205, Sankt-Peterburg (RU). **АРСЕНИЕВ, Анато-
лий Николаевич (ARSENIIEV, 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, Татья-
на Олеговна (ARTAMONOVA, Tatiana Olegovna)**;
ул. Наличная, д. 27, кв. 7 Санкт-Петербург, 199406,
Sankt-Peterburg (RU). **АБРАМОВА, Марина Викто-
ровна (ABRAMOVA, Marina Viktorovna)**; пр. Неполо-
рских, д.8, корпус 2 Санкт-Петербург, 194021, Sankt-
Peterburg (RU).

(74) Агент: **КОТЛОВ, Дмитрий Владимирович
(KOTLOV, Dmitry Vladimirovich)**; ООО "ЦИС "Скол-
ково", 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, JO, JP, KE, KG, KH, KN, KP,

(54) Title: DNA-CUTTING AGENT BASED ON CAS9 PROTEIN FROM THE BACTERIUM PASTEURELLA PNEUMOTROPICA

(54) Название изобретения: СРЕДСТВО РАЗРЕЗАНИЯ ДНК НА ОСНОВЕ CAS9 БЕЛКА ИЗ БАКТЕРИИ PASTEURELLA PNEUMOTROPICA

(57) Abstract: The present invention describes a novel bacterial nuclease of a CRISPR-Cas9 system from the bacterium *P. pneumotropica*, as well as the use of said nuclease for creating strictly specific two-strand cuts in a DNA molecule. The present nuclease possesses unusual properties and can be used as an instrument for introducing changes at strictly specified locations in a genomic DNA sequence of single-celled and multi-celled organisms. The invention thus increases the universality of accessible CRISPR-Cas9 systems, making it possible to use Cas9 nuclease from various organisms to cut genomic or plasmid DNA in a large number of specific sites and under various conditions.

(57) Реферат: 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 нуклеаза обладает необычными свойствами и может быть использована в качестве инструмента для внесения изменений в строго определенных местах в последовательности геномной ДНК одноклеточных или многоклеточных организмов. Таким образом, достигается повышение универсальности доступных систем CRISPR-Cas9, что позволит использовать нуклеазы Cas9 из различных организмов для разрезания геномной или плазмидной ДНК в большем количестве специфических сайтов и при различных условиях.

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)
- до истечения срока для изменения формулы изобретения и с повторной публикацией в случае получения изменений (правило 48.2(h))
- с информацией о просьбе восстановления прав на приоритет в отношении одного или более чем одного притязания на приоритет (правила 26bis.3 и 48.2(b) (vii))
- с перечнем последовательностей в соответствии с Правилom 5.2(a)

СРЕДСТВО РАЗРЕЗАНИЯ ДНК НА ОСНОВЕ CAS9 БЕЛКА 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 двуниевых 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. M. et al. 2009). PAM 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 PAM гидролиза связей в ДНК с образованием двунитевого разрыва не происходит. Необходимость присутствия PAM 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 PAM последовательности, фланкирующей 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 PAM последовательности. Использование CRISPR-Cas белков с новыми разнообразными PAM 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 PAM, которая должна присутствовать на 3'-конце участка ДНК, подвергающегося модификации. Поиск новых ферментов Cas9 с другими PAM 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 известных последовательность PAM; (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)T-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.

Указанная задача также решается путем создания способа изменения последовательности геномной ДНК одноклеточного или многоклеточного организма, включающего введение в по меньшей мере одну клетку этого организма эффективного количества: а) либо белка, содержащего аминокислотную последовательность SEQ ID NO: 1, либо нуклеиновой кислоты, кодирующей белок, содержащий аминокислотную последовательность SEQ ID NO: 1, и б) либо направляющей РНК, содержащей последовательность, образующую дуплекс с нуклеотидной последовательностью участка геномной ДНК организма, непосредственно примыкающей к нуклеотидной последовательности 5'-NNNN(A/G)T-3', и взаимодействующей с указанным белком после образования дуплекса, либо последовательности ДНК, кодирующей указанную направляющую РНК; при этом взаимодействие указанного белка с направляющей РНК и нуклеотидной последовательностью 5'-NNNN(A/G)T-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNN(A/G)T-3'. В некоторых вариантах изобретения данный способ характеризуется тем, что дополнительно включающий введение экзогенной последовательности ДНК одновременно с направляющей РНК.

В качестве направляющей РНК может быть использована смесь из крРНК (crRNA) и трейсерной РНК (tracrRNA), способных образовать комплекс с участком целевой ДНК и белком РrCas9. 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 генома может проводиться прямым способом – разрезанием генома 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 одонитевой 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 генома 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.

Краткое описание рисунков

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).

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 РАМ 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 рекомбинантного 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-крРНК 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 трРНК	5'GCGAAATGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAAG CTCUGCCUCUUGAAAUUUCGGUUUCAAGAGGCAUCUUUUU-3'
PpCas9 крРНК	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'-

ccccgggtaccacggagagatggtggaaatcatcttctcgtgggcatccttgatggccacctcgtcggaagtgccacgaggatga

cagcaatgccaatgctggggggctcttctgagaacgagctctgctgcctgacacggccaggacggccaacaccaaccagaact
 gggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaagagcatgttaaaataggatcta
 catcacgtaacgtcttagaagaggctagatactgcaattcaaggaccttatctccttcattgagcacNNNNNNNaactccatcta
 ccagcctactcttcttctgttatt -3'

Для разрезания этой мишени использовали направляющие РНК следующей последовательности:

tracrRNA:

5'GCGAAATGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAAAGCTCUGCC
 UCUUGAAAUUUCGGUUUCAAGAGGCAUCUUUUU

и crRNA:

5' **uaucuccuuucauugagcac**GUUGUAGCUCCCUUUUUUCAUUUCGC.

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 мМ, 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 50мМ pH 8, 500 мМ NaCl, β-меркаптоэтанол 1мМ, имидазол 10 мМ) 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 мМ имидазола, после чего белок смывали лизисным буфером с добавлением 300 мМ 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 PAM библиотек проводили в объёме 20 мкл в следующих условиях. Реакционная смесь состояла из: 1X CutSmart буфера (NEB), 5 mM DTT, 100 нМ PAM-библиотеки, 2 мкМ трПНК/крПНК, 400 нМ белка PpCas9. В качестве контроля аналогичным образом были приготовлены пробы, не содержащие ПНК. 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 PAM 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'- cccgggggtaccacggagagatggtggaaatcatcttctcgtgggcatccttgatggccacctcgtcggaagtgccacgaggatgacagcaatgccaatgctgggggggctcttctgagaacgagctctgctgctgacacggccaggacggccaacaccaaccagaactgggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaagagcatgttaaaataggatcta

catcacgtaacctgtcttagaagaggctagatactgcaattcaaggaccttatctccttcattgagcacCAACATTaactccatcta
ccagcctactctcttatctctgtatt- 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 лишь 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)T-3'. Последовательности NNNNRTY (NNNN(A/G)-T-(T/C)) распознаются более эффективно чем NNNNRTR (NNNN-(A/G)-T-(A/G)).

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)T-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)T-3'. Для узнавания этих последовательностей были синтезированы крПНК, 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)T 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

UAUCUCCUUUCAUUGAGCACGUUGUAGCUCUUUUUUUCAUUUCGCGAGUGCUAUAUAAUG
AAAAUUUAUAGCACUGCGAAAUGAAAAACGUUGUUACAAUAAGAGAUGAAUUUCUCGCAA
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'-
cccgggggtaccacggagagatggtggaaatcatctttctgctgggcatccttgatggccacctcgtcggaagtgccacgaggatga
cagcaatgccaatgctgggggggctcttctgagaacgagctctgctgcctgacacggccaggacggccaacaccaaccagaact
gggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgcaaacacaaagaaagagcatgttaaaataggatcta
catcacgtaacctgtcttagaagggttagatactgcaattcaaggacct**tatctcctttcattgagcac**CAACATTcaactccat
ctaccagcctactcttctatctctgtatt – 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 °C.

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)T-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. Описанная в настоящем изобретении система PpCas9 в комплексе с направляющими РНК может быть использована для изменения последовательности геномной ДНК многоклеточного организма, в том числе эукариотического. Для введения системы PpCas9 в комплексе с направляющими РНК в клетки этого организма (во все клетки или в часть клеток) могут быть применены различные подходы, известные специалистам. Например, методы доставки CRISPR-Cas9 систем в клетки организмов раскрыты в источниках (Liu C et al., Delivery strategies of the CRISPR-Cas9 gene-editing

system for therapeutic applications. J Control Release. 2017 Nov 28;266:17-26; Lino CA et al., Delivering CRISPR: a review of the challenges and approaches. Drug Deliv. 2018 Nov;25(1):1234-1257), и в источниках, раскрытых внутри этих источников.

Для эффективной экспрессии нуклеазы PpCas9 в эукариотических клетках будет желательно провести оптимизацию кодонов для аминокислотной последовательности белка PpCas9 методами, известными специалистам (например, IDT codon optimization tool).

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. Таким образом, полная аминокислотная последовательность нуклеазы, транспортируемой внутрь ядра эукариотической клетки, будет представлять собой следующую последовательность: MAPKKKRVGIHGVPAAPpCas9-KRPAATKKAGQAKKKK (далее PpCas9 NLS). Для доставки белка с приведенной выше аминокислотной последовательностью, могут быть использованы по меньшей мере два подхода.

Доставка в виде гена осуществляется путем создания плазмиды, несущей ген PpCas9 NLS под регуляцией промотора (например, CMV промотора) и последовательности, кодирующей направляющие РНК под регуляцией U6 промотора. В качестве ДНК- мишеней используются ДНК последовательности фланкированные 5'-NNNN(G/A)T -3', например, последовательности гена grin2b человека:

5'-CAGCTGAAGTAATGTTAGAG-3'

Таким образом, кассета для экспрессии sgРНК выглядит следующим образом:

**gagggcctatttcccatgattccttcatatttgcataacgatacaaggctgttagagagataattggaattaatttgactgtaaa
cacaagaatatttagtacaataacgtgacgtagaaagtaataatttcttggttagttgcagttttaaataattgttttaaatgg
actatcatatgcttaccgtaactgaaagtatttcgatttcttggtttatatacttgtggaaggacgaaacaccg
CAGCTGAAGTAATGTTAGAGGTTGTAGCTCCCTTTTTCATTTTCGCGAAAGCGAAATGAAAA
ACGTTGTTACAATAAGAGATGAATTTCTCGCAAAGCTCTGCCTCTTGAAATTTTCGTTTCAA
GAGGCATCTTTT**

Жирным шрифтом выделена последовательность U6 промотора, далее идет последовательность, необходимая для узнавания целевой ДНК, а далее идет последовательность, образующая структуру sgRNA, которая выделена красным.

Плазмидную ДНК очищают и трансфицируют в клетки человека HEK293 с помощью реагента Lipofectamine 2000 (Thermo Fisher Scientific). Клетки инкубируют в течение 72 часов, после чего из них выделяется геномная ДНК с помощью колонок для очистки геномной ДНК (Thermo Fisher Scientific). Целевой ДНК сайт анализируется с помощью секвенирования на платформе Illumina с целью определения числа вставок-делеций в

ДНК, происходящих в целевом сайте по причине направленного двунитевого разрыва и последующей его репарации.

Для амплификации целевых фрагментов используют праймеры, фланкирующие предположительное место внесения разрыва.

После амплификации пробы готовятся по протоколу реагента Ultra II DNA Library Prep Kit for Illumina (NEB) для подготовки образцов к высокопроизводительному секвенированию. Затем проводится секвенирование на платформе Illumina 300cycles, прямое прочтение. Результаты секвенирования анализируются биоинформатическими методами. В качестве детекции разрезания принимается вставка или делеция нескольких нуклеотидов в целевой последовательности ДНК.

Доставка в виде рибонуклеинового комплекса 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* 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 функционирования PpCas9 достаточно присутствия короткого PAM мотива (RT), расположенного в 4 нуклеотидах от протоспейсера.

Известные на сегодняшний день большинство Cas нуклеаз, способных 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 RT мотивом.

Второе преимущество PpCas9 – малый размер белка (1055 а.о). На сегодняшний день это единственный изученный малоразмерный белок, обладающий двухбуквенной RT 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 и при 37 °С, 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.

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)T-3' в указанной молекуле ДНК.

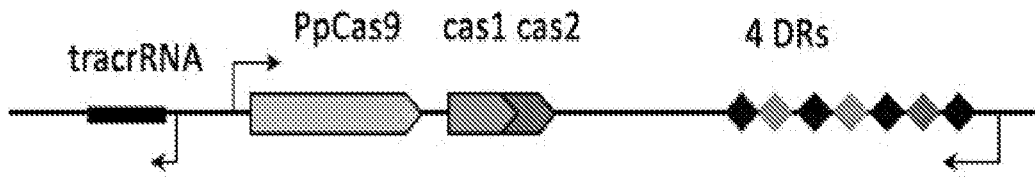
2. Применение по п. 1, характеризующееся тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 35 °C до 45 °C.

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

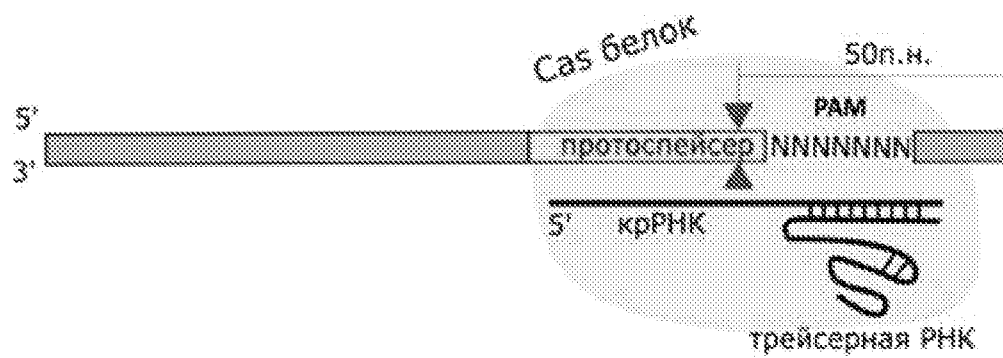
4. Способ изменения последовательности геномной ДНК одноклеточного или многоклеточного организма, включающий введение в по меньшей мере одну клетку этого организма эффективного количества: а) либо белка, содержащего аминокислотную последовательность SEQ ID NO: 1, либо нуклеиновой кислоты, кодирующей белок, содержащий аминокислотную последовательность SEQ ID NO: 1, и б) либо направляющей РНК, содержащей последовательность, образующую дуплекс с нуклеотидной последовательностью участка геномной ДНК организма, непосредственно примыкающей к нуклеотидной последовательности 5'-NNNN(A/G)T-3', и взаимодействующей с указанным белком после образования дуплекса, либо последовательности ДНК, кодирующей указанную направляющую РНК;

при этом взаимодействие указанного белка с направляющей РНК и нуклеотидной последовательностью 5'-NNNN(A/G)T-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNN(A/G)T-3'.

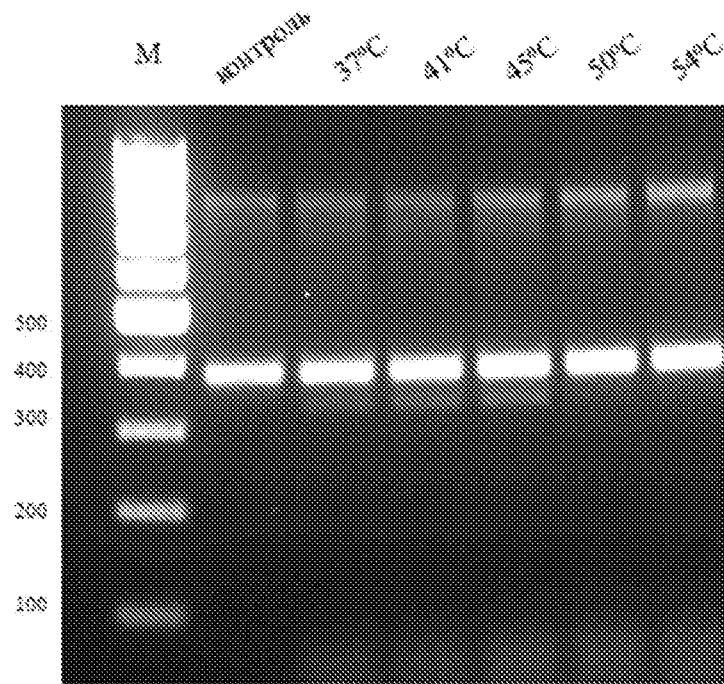
5. Способ по п. 4, дополнительно включающий введение экзогенной последовательности ДНК одновременно с направляющей РНК.



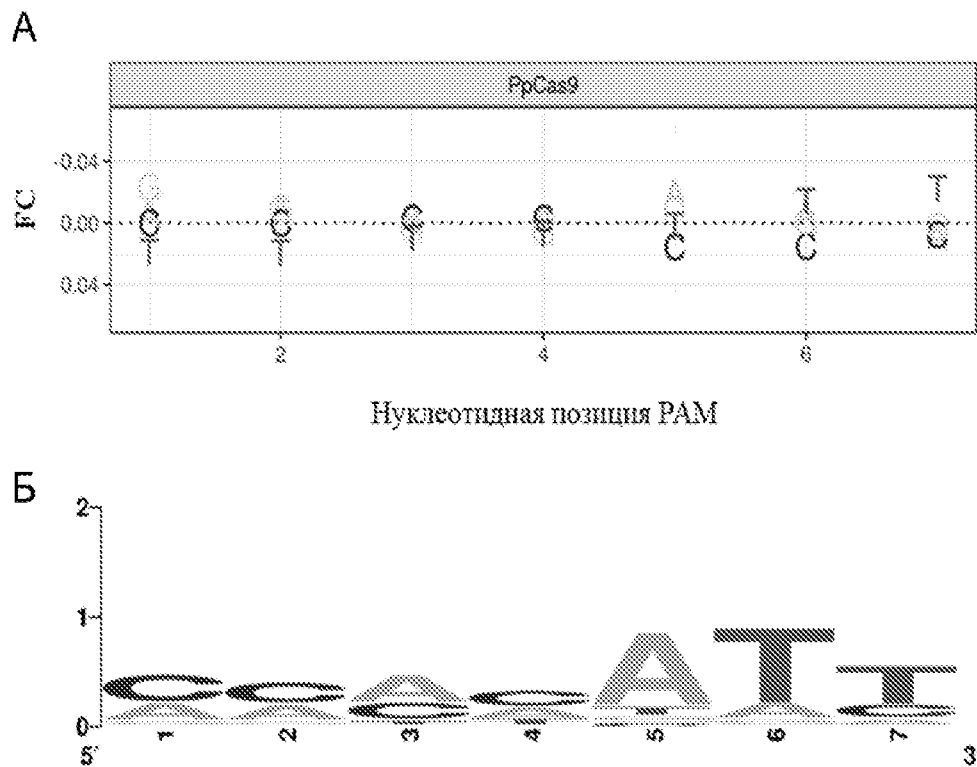
Фиг. 1



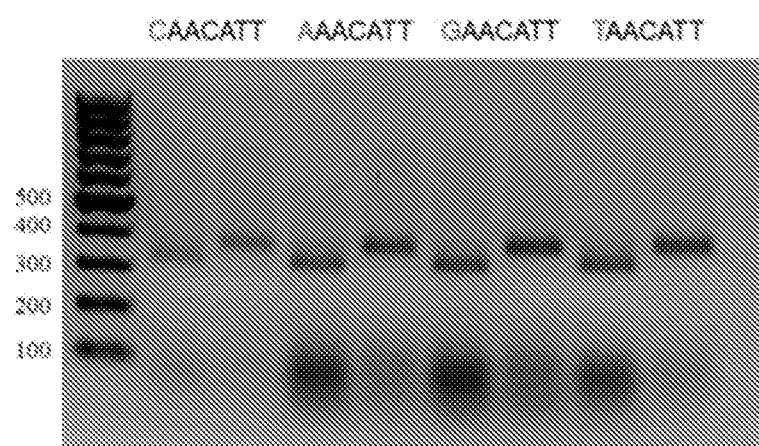
Фиг. 2



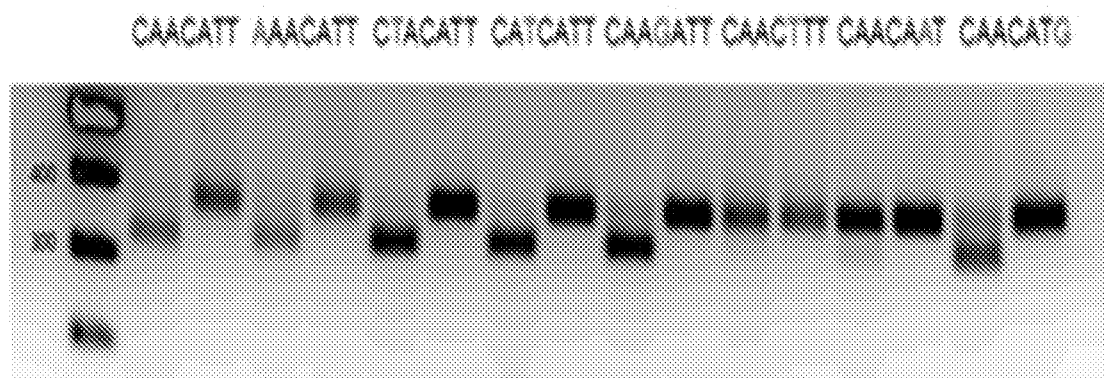
Фиг. 3



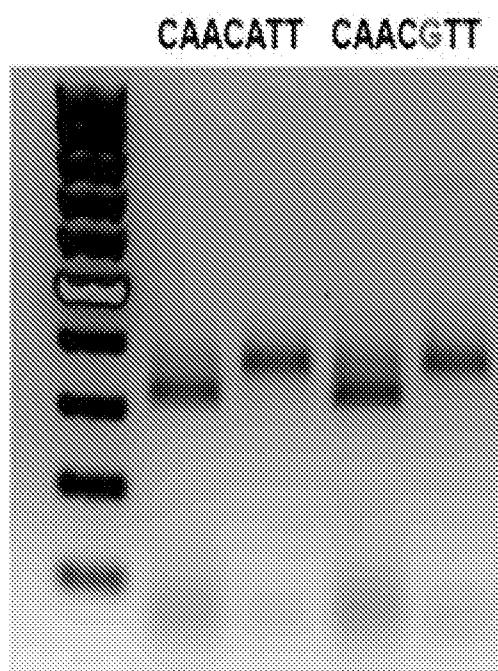
Фиг. 4



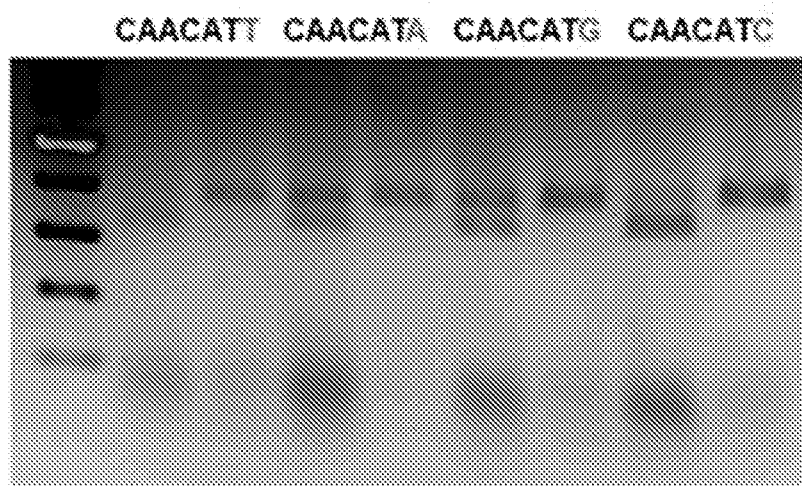
Фиг. 5



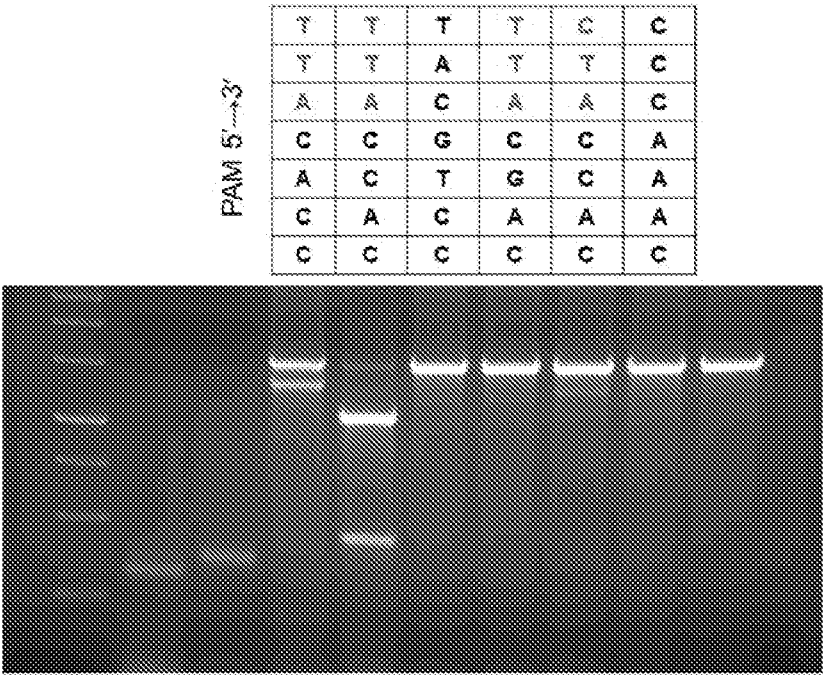
Фиг. 6



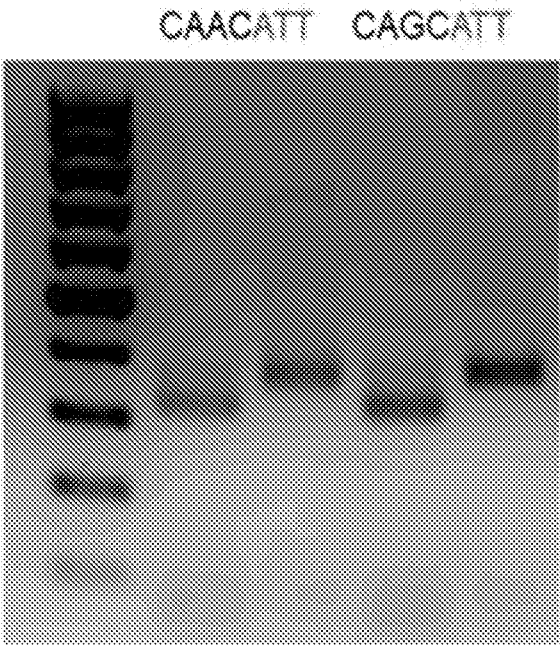
Фиг. 7



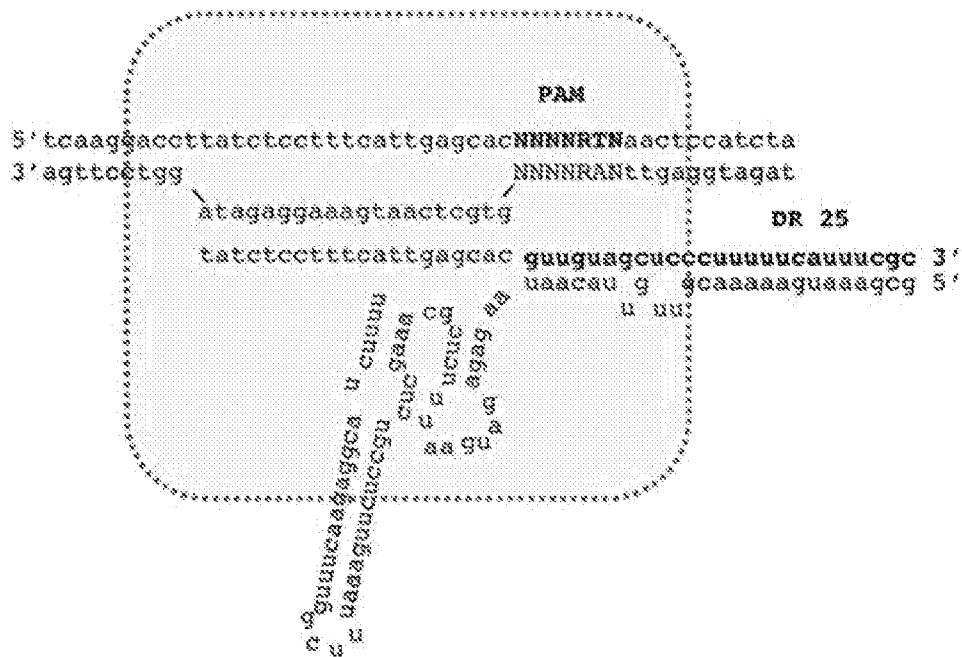
Фиг. 8



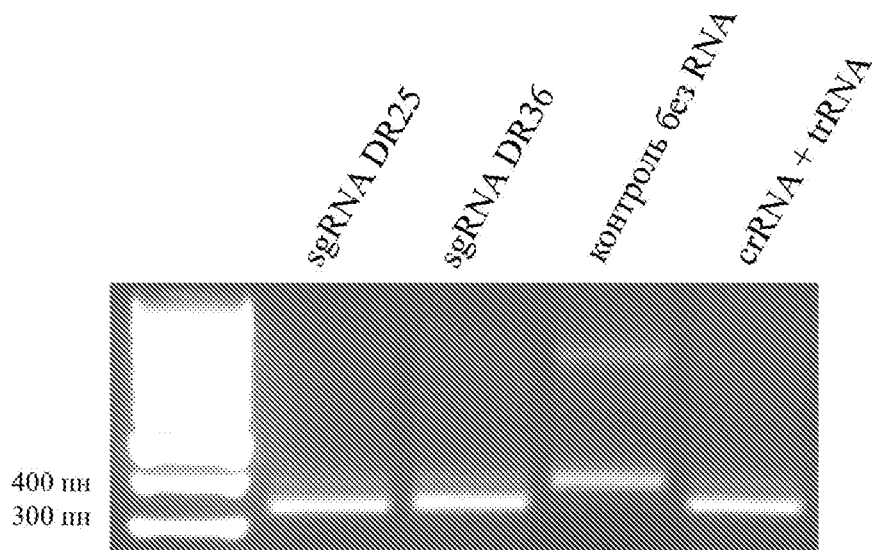
Фиг. 9



Фиг. 10



Фиг. 11



Фиг. 12

[illegible]

Фиг. 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/RU 2020/050137

A. CLASSIFICATION OF SUBJECT MATTER

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

C12Q 1/68 (2006.01)
C12N 15/87 (2006.01)
C07K 14/195 (2006.01)

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/68, C12N 15/00, 15/87, C07K 14/195, C12N 9/00

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)

PatSearch (RUPTO Internal), USPTO, PAJ, Espacenet, Information Retrieval System of FIPS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ASTRID WENINGER et al. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast <i>Pichia pastoris</i> . Journal of Biotechnology. Elsevier. Volume 235, 10 October 2016, Pages 139-149	1-5
P	A. VASILEVA et al. Pasteurella pneumotropica and Demequina sediminicola Cas9 orthologs characterization. 14.01.2020, Found in Internet: < https://2019.febscongress.org/abstract_preview.aspx?idAbstractEnc=4424170094096092094091424170 >	1-5



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

28 September 2020 (28.09.2020)

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 2020/050137

A. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ <i>C12Q 1/68 (2006.01)</i> <i>C12N 15/87 (2006.01)</i> <i>C07K 14/195 (2006.01)</i>		
Согласно Международной патентной классификации МПК		
B. ОБЛАСТЬ ПОИСКА Проверенный минимум документации (система классификации с индексами классификации) C12Q 1/68, C12N 15/00, 15/87, C07K 14/195, C12N 9/00		
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) PatSearch (RUPTO Internal), USPTO, PAJ, Espacenet, Information Retrieval System of FIPS		
C. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ:		
Категория*	Цитируемые документы с указанием, где это возможно, релевантных частей	Относится к пункту №
A	ASTRID WENINGER et al. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast <i>Pichia pastoris</i> . Journal of Biotechnology. Elsevier. Volume 235, 10 October 2016, Pages 139-149	1-5
P	A. VASILEVA et al. Pasteurella pneumotropica and Demequina sediminicola Cas9 orthologs characterization. 14.01.2020, Найдено в Интернет: < https://2019.febscongress.org/abstract_preview.aspx?idAbstractEnc=4424170094096092094091424170 >	1-5
☐ 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		Дата отправки настоящего отчета о международном поиске
28 сентября 2020 (28.09.2020)		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: Е. Смирнова Телефон № 8(495)531-64-81