

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

(19) Всемирная Организация
Интеллектуальной Собственности
Международное бюро

(43) Дата международной публикации
04 июня 2020 (04.06.2020)



(10) Номер международной публикации
WO 2020/111983 A2

- (51) Международная патентная классификация :
Неклассифицировано
- (21) Номер международной заявки : PCT/RU20 19/050229
- (22) Дата международной подачи :
26 ноября 2019 (26.11.2019)
- (25) Язык подачи : Русский
- (26) Язык публикации : Русский
- (30) Данные о приоритете :
2018141524 26 ноября 2018 (26.11.2018) RU
- (71) Заявитель : АВТОНОМНАЯ НЕКОММЕРЧЕСКАЯ
ОБРАЗОВАТЕЛЬНАЯ ОРГАНИЗАЦИЯ ВЫСШЕ -
ГО ОБРАЗОВАНИЯ СКОЛКОВСКИЙ ИНСТИ -
ТУТ НАУКИ И ТЕХНОЛОГИЙ (AUTONOMOUS
NON-PROFIT ORGANIZATION FOR HIGHER
EDUCATION "SKOLKOVO INSTITUTE OF
SCIENCE AND TECHNOLOGY") [RU/RU]; ул. Но -
вая, 100 Московская обл., Одинцовский р-н, Дер. Скол -
ково, 143025, Moskovskaya obl., Odintsovskij r-n, Der.
Skolkovo (RU).
- (72) Изобретатели : СЕВЕРИНОВ , Константин Викто -
рович (SEVERINOV, Konstantin Viktorovich); ул.
Дмитрия Ульянова , д. 4, корп. 2, кв. 362 Москва ,
119333, Moscow (RU). ШМАКОВ , Сергей Анатолие -
вич (SHMAKOV, Sergey Anatolevich); Московская
обл., ул. Пионерская , д. 14, кв. 15 Воскресенск , 140200,
Voskresensk (RU). АРТАМОНОВА , Дарья Николае -
вна (ARTAMONOVA, Daria Nikolaevna); ул. Маршала
Тимошенко , д. 9, кв. 59 Москва , 121359, Moscow (RU).
ГОРЯНИН , Игнатий Игоревич (GORIANIN, Ignatiy
Igorovich); ул. Кировоградская . 16, здание 1, кв. 351

Москва , 117587, Moscow (RU). МУШАРОВА , Ольга
Сергеевна (MUSHAROVA, Olga Sergeevna); ул. Ро -
гова , 7, стр. 2, кв. 289 Москва , 123098, Moscow (RU).
ПИСКУНОВА , Юлия Валерьевна (PISKUNOVA,
Iuliia Valerevna); ул. Москвитина , д. 3, корп. 2, кв.
108, Московский Москва , 142784, Moscow (RU).
ФЕДОРОВА , Яна Витальевна (FEDOROVA, Iana
Vitalevna); Ленинградская обл., ул. Куприна , д. 40,
кв. 14, Гатчина , 188301, Gatchina (RU). ЗЮБКО ,
Татьяна Игоревна (ZYUBKO, Tatyana Igorevna);
ул. Яхтенная 3, стр. 1 Санкт -Петербург , 197374,
St.Petersburg (RU). ХОДОРКОВСКИЙ , Михаил Алек -
сеевич (KHODORKOVSKIY, Mikhail Alekseevich);
пр. Косыгина , д. 23, к. 1, кв. 268 Санкт -Петербург ,
195426, St.Petersburg (RU). ПОБЕГАЛОВ , Георгий
Евгеньевич (POBEGALOV, Georgii Evgenevich); ул.
Партизана Германа , д. 12, кв. 317 Санкт -Петербург ,
198205, St.Petersburg (RU). АРСЕНИЕВ , Анатолий
Николаевич (ARSENIEV, Anatolii Nikolaevich); пр.
кт Новочеркасский , д. 25, корп. 1, кв. 36 Санкт -Пе -
тербург , 195112, St.Petersburg (RU). СЕЛЬКОВА , По -
лина Анатольевна (SELKOVA, Polina Anatolevna);
Рес. Удмуртская , ул. Верхняя , д. 15, кв. 72 Воткинск ,
427439, Votkinsk (RU). ВАСИЛЬЕВА , Александра
Андреевна (VASILEVA, Aleksandra Andreevna); ул.
Партизана Германа , д. 41, кв. 144 Санкт -Петербург ,
198334, St.Petersburg (RU). АРТАМОНОВА , Татья -
на Олеговна (ARTAMONOVA, Tatiana Olegovna);
ул. Наличная , д. 27, кв. 7 Санкт -Петербург , 199406,
St.Petersburg (RU). АБРАМОВА , Марина Викторов -
на (ABRAMOVA, Marina Viktorovna); пр. Непокорён -
ных , д. 8, корп. 2 Санкт -Петербург , 194021, St.Petersburg
(RU).

(54) Title: DNA-CUTTING AGENT BASED ON CAS9 PROTEIN FROM THE BIOTECHNOLOGICALLY RELEVANT BAC -
TERIUM CLOSTRIDIUM CELLULOLYTICUM

(54) Название изобретения : СРЕДСТВО РАЗРЕЗАНИЯ ДНК НА ОСНОВЕ CAS9 БЕЛКА ИЗ БИОТЕХНОЛОГИЧЕСКИ
ЗНАЧИМОЙ БАКТЕРИИ CLOSTRIDIUM CELLULOLYTICUM

(57) Abstract: The present invention describes a novel bacterial nuclease of a CRISPR-Cas9 system from the bacterium Clostridium
cellulolyticum, as well as the use of said nuclease for creating strictly specific two-strand cuts in a DNA molecule. The present nuclease
exhibits unusual qualities and can be used as an instrument for introducing changes into strictly specified places in a genomic DNA
sequence of single-celled and multi-celled organisms. The invention thus increases the universality of accessible CRISPR-Cas9 systems,
which makes it possible to use Cas9 nuclease from various organisms in order to cut genomic and plasmid DNA in a large number of
specific places and at wide ranges of temperatures. The invention also simplifies editing the genome of the biotechnologically relevant
bacterium Clostridium cellulolyticum.

(57) Реферат : Настоящее изобретение описывает новую бактериальную нуклеазу системы CRISPR-Cas9 из бактерии
Clostridium cellulolyticum, а также ее применение для образования строго специфических двунитевых разрывов в молекуле ДНК .
Данная нуклеаза обладает необычными свойствами и может быть использована в качестве инструмента для внесения изме -
нений в строго определенных местах в последовательности геномной ДНК одноклеточных или многоклеточных организмов .
Таким образом , достигается повышение универсальности доступных систем CRISPR-Cas9, что позволит использовать нукле -
азы Cas9 из различных организмов для разрезания геномной или плазмидной ДНК в большем количестве специфических мест
и при больших диапазонах температур . Также , достигается упрощение редактирования генома биотехнологически значимой
бактерии Clostridium cellulolyticum.

(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,
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, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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 :

- без отчёта о международном поиске и с повторной
публикацией по получении отчёта (правило 48.2(g))
- с перечнем последовательностей в соответствии с
Правилом 5.2(a)

Средство разрезания ДНК на основе **Cas9** белка из биотехнологически значимой бактерии ***Clostridium cellulolyticum***

Область техники

Изобретение относится к области молекулярной биологии и микробиологии, в частности, описывает новые бактериальные нуклеазы системы CRISPR-Cas. Изобретение может быть использовано в качестве инструмента для строго специфической модификации ДНК в различных организмах.

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 каскет и генов, кодирующих различные Cas белки (Jansen R. et al., Molecular microbiology. — 2002. — Т. 43. — №. 6. — С. 1565-1575). CRISPR каскеты 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 каскет 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. 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. 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. Одной из основных характеристик, ограничивающих применение CRISPR-Cas систем является PAM последовательность, фланирующая ДНК-мишень с 3'-конца, наличие которой необходимо для правильного узнавания ДНК Cas9 нуклеазой. Различные CRISPR-Cas белки имеют разные PAM последовательности, которые ограничивают возможности применения нуклеаз на любых участках ДНК. 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. Данное изобретение обладает рядом свойств, необходимых для решения этой задачи.

Основой изобретения является CRISPR Cas система, найденная в бактериях *Clostridium cellulolyticum*. Анаэробные бактерии *Clostridium cellulolyticum* (*C. cellulolyticum*) способны осуществлять гидролиз лигноцеллюлозы без добавления коммерческих целлюлаз с образованием в качестве конечных продуктов лактата, ацетата и бутанола (Desvaux M. *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic Clostridia. FEMS Microbiol Rev. 2005 Sep;29(4):741-64). Такая способность данных микроорганизмов делает их перспективными кандидатами на роль продуцентов биотоплива. Использование бактерий продуцентов таких как *C. cellulolyticum* в биотехнологических производствах поможет эффективнее организовать цикл переработки сырья, повысить эффективность и, в конечном итоге, приведет к снижению нагрузки на все компоненты биосферы. С помощью методов генетической инженерии возможно существенно улучшить метаболические показатели микроорганизмов, и сместить баланс в сторону наработки большего количества бутанола вместо лактата и ацетата. Так, например, для двойного мутанта по генам

лактатдегидрогеназы и мальтатдегидрогеназы показано отсутствие образования лактата и увеличенная продукция бутанола (Li Y, et al., Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol Biofuels*. 2012 Jan 4;5(1):2). До настоящего момента не удалось разработать эффективную процедуру получения штаммов *C. cellulolyticum* с мутациями в генах фосфотрансацетилазы и ацетаткиназы, что позволило бы снизить продукцию ацетата. Изобретение может использоваться для изменения генома *Clostridium cellulolyticum*, как и геномов других живых организмов.

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 *C. cellulolyticum* CRISPR нуклеаза II типа CcCas9, которая может быть применена для внесения направленных изменений в геном как этого, так и других организмов. Существенными признаками, отличающими настоящее изобретение, являются: (а) короткая, двухбуквенная, отличающаяся от других известных последовательность PAM; (б) малый размер охарактеризованного белка CcCas9 - 1030 аминокислотных остатков (а.о.), что на 23 а.о. меньше, чем размер известного фермента Cas9 из *Staphylococcus aureus* (SaCas9); (в) широкий рабочий диапазон температур нуклеазы CcCas9, которая активна при температурах от 37 °C до 65 °C с оптимумом при 45 °C, что позволит использовать ее в организмах, имеющих различную температуру.

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'-NNNNGNA-3' в указанной молекуле ДНК. Под N следует понимать любой из нуклеотидов (A, G, C, T). В некоторых вариантах изобретения данное применение характеризуется тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 37 °C до 65 °C. В предпочтительных вариантах изобретения данное применение характеризуется тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 37 °C до 55 °C.

Указанная задача также решается с помощью способа 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'-NNNNGNA-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'-NNNNGNA-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNNGNA-3'.

В качестве направляющей РНК может быть использована смесь из крРНК (crRNA) и трейсерной РНК (tracrRNA), способных образовать комплекс с участком таргетной ДНК и белком CcCas9. 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).

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 таргетной ДНК, образованного под действием нуклеазы CcCas9. В некоторых вариантах изобретения в качестве экзогенной последовательности ДНК может быть использован участок двуцепочечной ДНК из генома организма, используемого при введении белка CcCas9, 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 мест и при больших диапазонах температур.

Краткое описание рисунков

Фиг. 1. Схема CRISPR локусов в *Clostridium cellulosum*

Фиг. 2. Определение PAM *in vitro* методами. Разработка системы порезки ДНК, ограниченной NNNNGNA последовательностью.

Фиг. 3. Проверка значимости отдельных позиций PAM.

Фиг. 4. Проверка активности белка в разрезании различных ДНК мишеней.

Фиг. 5. Реакции *in vitro* разрезания ДНК фрагмента гена человека *grin2b*

Фиг. 6. Исследование температурного диапазона активности CcCas9.

Фиг. 7. Схема взаимодействия направляющей РНК с участком целевой

ДНК .

Фиг . 8. Выравнивание последовательностей белков Cas9 из организмов *Staphylococcus aureus* (SaCas9), *Campylobacter jejuni* (CjCas9) и CcCas9. Подчеркнуты снизу неконсервативные участки последовательностей .

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 гибрид издается " 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 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 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 *Clostridium cellulolyticum* (*C. cellulolyticum*), гомологичного ранее охарактеризованным белкам Cas9, для внесения двуцепочечных разрывов в молекулы ДНК в строго определенных положениях.

Редактирование генома *C. cellulolyticum* для метаболической инженерии - сложная задача из-за отсутствия эффективных инструментов редактирования. Методы направленного редактирования генома - такие как рекомбинация, ретротранспреция интрона II группы и обмен аллелями - имеют ряд существенных ограничений. Например, процедура рекомбинационного - зависимого обмена аллелями достаточно продолжительная и обладает низкой эффективностью (Heap J. T. et al. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker //Nucleic acids research. - 2012. - Т. 40. - №. 8. - С. e59-e59). Вставка протяженных фрагментов ДНК (для, например, переноса метаболических путей) трудно осуществима при помощи имеющихся инструментов для изменения генома, которые требуют существующих сайтов рекомбинации и / или рекомбиназ (Esvelt K. M., Wang H. H. Genome-scale engineering for systems and synthetic biology //Molecular systems biology. - 2013. - Т. 9. - №. 1. - С. 641). Для успешного манипулирования геномом микроорганизмов и получения мутантов с заданными свойствами, необходима простая и эффективная методика.

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.

Кроме того, использование нативной системы CRISPR-Cas9 из бактерии *C. cellulolyticum* позволит сделать систему редактирования генома этого организма проще и эффективней: процедура не будет требовать внесения в

клетки чужеродных генов, их поддержание и экспрессию. Вместо этого возможно разработать процедуру внесения направляющих РНК, направленных на целевые гены, в бактерию: благодаря им хозяйская внутриклеточная CRISPR-Cas9 система сможет распознавать соответствующие ДНК-мишени биотехнологически значимой бактерии и вносить в них двунитевые разрывы.

Для биохимической характеристики белка Cas9 из *C. cellulolyticum* НЮ locus CRISPR, кодирующий основные компоненты системы (гены белков CcCas9, cas1, cas2, а также CRISPR кассету и направляющие РНК), был клонирован в малокопийный бактериальный вектор pACYC184. Эффекторному рибонуклеиновому комплексу, состоящему из Cas9 и дуплекса крРНК (crRNA) и трейсерной РНК (tracrRNA), для распознавания и последующего гидролиза ДНК помимо комплементарного соответствия спейсера крРНК и протоспейсера необходимо присутствие PAM (от англ. "PAM" - protospacer adjusted motif) на ДНК мишени (Mojica F. J. M. et al. Short motif sequences determine the targets of the prokaryotic CRISPR defence system //Microbiology. — 2009. — Т. 155. — №. 3. — С. 733-740). 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.

Для определения последовательностей направляющей РНК системы CRISPR-Cas9 было проведено РНК-секвенирование бактерий *E.coli* DH5alpha, несущих созданную конструкцию pACYC184_CcCas9. Секвенирование показало, что CRISPR кассета системы активно транскрибируется, также как и трейсерная РНК (Фиг. 1). Анализ последовательности crRNA и tracrRNA позволил предположить возможное формирование ими вторичных структур, распознаваемых CcCas9 нуклеазой.

Далее авторами было произведено определение PAM последовательности белка CcCas9 при помощи бактериального PAM скрининга. Для определения PAM последовательности белка CcCas9 клетки *E.coli* DH5alpha, несущие плазмиду pACYC184_CcCas9, были трансформированы

библиотекой плазмид, содержащих последовательность спейсера 5'-

TAAAAATAAGCAAGCGATGATATGAATGC-3' CRISPR кассеты CcCas9 системы, фланкированной рандомизированной семибуквенной последовательностью с 5' или 3' конца. Плазмиды, несущие последовательность, соответствующую последовательности PAM системы CcCas9, подвергались деградации под действием функционирующей CRISPR-Cas системы, в то время как остальные плазмиды библиотеки эффективно трансформировались в клетки, придавая им устойчивость к антибиотику ампициллину. После трансформации и инкубации клеток на чашках с антибиотиком колонии смывались с поверхности агара и из них экстрагировали ДНК с помощью набора Qiagen Plasmid Purification Midi. Из выделенного пула плазмид амплифицировали при помощи ПЦР участки, включающие рандомизированную PAM последовательность, которые затем подвергали высокоэффективному секвенированию на платформе Illumina. Полученные прочтения анализировали: сравнивали эффективность трансформации плазмид с уникальными PAM, входящими в библиотеку, в клетки, несущие pACYC184_CcCas9 или в контрольные клетки, несущие пустой вектор расус 184. Полученные результаты анализировали методами биоинформатики. В результате удалось выявить PAM системы CcCas9 - двубуквенную последовательность NNNNGNA (Фиг. 2).

Далее было произведено дополнительное определение PAM последовательности при помощи воссоздания реакции разрезания *in vitro*. Для определения PAM последовательности белка CcCas9 использовали *in vitro* порезку двунитевых PAM библиотек. Для этого было необходимо получить все компоненты CcCas9 эффекторного комплекса: направляющие РНК и нуклеазу в рекомбинантной форме. Определение РНК-секвенированием 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 п.н, содержащие последовательность протоспейсера, фланкированную рандомизированными семью нуклеотидами (5' NNNNNNN 3') с 3' конца:

5'ccccgggtaccacggagagatggtggaaatcatcttctcgtgggcatccttgatggccacctcgtcggaa
gtgcccacgaggatgacagcaatgccaatgctgggggggctcttctgagaacgagctctgctgcctgacacggcca
ggacggccaacaccaaccagaactgggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgc
aaacacaaagaaagagcatgttaaaataggatctacatcacgtaacctgtcttagaagaggctagatactgcaattc

aaggaccttatctcctttcattgag cacNNNNNNNaactccatctaccag cctactctcttctctg gtatt -3'

Для разрезания этой мишени использовали направляющие РНК следующей последовательности :

tracrRNA:

5'AUUAUGGCAUAUCGGAGCCUGAAUUGUUGCUAUAUAAGGUGCUGGG
UUUAGCCCAGACCGCCAAGUUAACCCCGGCAUUUAUUGCUGGGGUAUCUUGU
UUU

и crRNA:

5'uaucisscuiiucsuuugagcacGUUAUAGCUCCAAUUCAGGCUCCGAUUAU

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

Для получения рекомбинантного белка CcCas9 его ген был заклонирован в плазмиду pET 21a. Клетки E.coli Rosetta были трансформированы полученной плазмидой CcCas9_pET21a. Клетки, несущие плазмиду, наращивали до оптической плотности OD₆₀₀=0.6, далее проводили индукцию экспрессии гена CcCas9 добавлением IPTG до концентрации 1т М. Клетки инкубировали в течение 4 часов на 25 °С, после чего проводили их лизис. Очистку рекомбинантного белка проводили двумя этапами: с помощью аффинной хроматографии (NiNTA) и разделением белков по размерам на колонке Superdex 200. Полученный белок концентрировали с помощью фильтров Amicon с размером пор 30кДа. После этого белок замораживали на минус 80 °С и использовали для проведения *in vitro* реакций.

In vitro реакцию порезки линейных PAM библиотек проводили в следующих условиях:

IxCutSmart буфер

400 нМ CcCas9

100 нМ ДНК библиотека

2 мкМ crRNA

2 мкМ tracrRNA

Общий объем реакции 20 мкл.

Clostridium cellulolyticum Н Ю обитает в компостах и имеет оптимальную температуру деления 45°С, в связи с этим реакции проводили на этой температуре в течение 30 минут.

В результате порезки часть фрагментов библиотеки распадалась на две части длиной около 50 пар нуклеотидов (п.н.) и 324 п.н. В качестве контрольного образца использовали реакции без добавления crRNA, обязательного компонента эффекторного Cas комплекса.

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 (Фиг. 3).

В результате авторам удалось определить PAM последовательность CcCas9 *in vitro* методами: NNNNGNA, что полностью повторило результат, полученный в экспериментах с бактериями.

Далее была произведена проверка значимости отдельных позиций последовательность PAM. Для этого *in vitro* проводили реакции *in vitro* порезки ДНК фрагмента, содержащего ДНК-мишень 5'-gtgctcaatgaaaggagata-3', фланкированную PAM последовательностью GAGAGTA:

5'ccccgggtaccacggagagatggtggaatcatctttctcggtggcatccttgatggccacctcgtcggaa
gtgcccacgaggatgacagcaatgccaatgctgggggggctcttctgagaacgagctctgctgcctgacacggcca
ggacggccaacaccaaccagaacttgggagaacagcactccgctctgggcttcatcttcaactcgtcgactccctgc
aaacacaaagaaagagcatgttaaaataggatctacatcacgtaacctgtcttagaaggaggctagatactgcaattc
aaggaccttatctctttcattgagcac **GAGAGTA** aactccatctaccagcctactcttcttctgtgtatt 3'

Реакцию проводили в следующих условиях:

IxCutSmart буфер

400 нМ CcCas9

20 нМ ДНК

2 мкМ crRNA

2 мкМ tracrRNA

Время инкубации 30 минут, температура проведения реакции 45 °C.

Результаты эксперимента подтвердили последовательность PAM для CcCas9 как NNNNGNA -3'. Наиболее консервативной оказалась позиция G в пятом положении (см. Фиг. 4).

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 CcCas9 в разрезании различных ДНК мишеней.

Для того, чтобы проверить способность CcCas9 узнавать различные последовательности ДНК, фланкированные NNNNGNA 3' последовательностью, были проведены эксперименты по *in vitro* разрезанию ДНК-мишеней из последовательности гена *grin2b* человека (см. таблицу 1 ниже).

Таблица 1. ДНК-мишени, выделенные из гена *grin2b* человека.

ДНК-мишень	PAM
ctacatcacgtaacctgtct	tagaAgA
gaacgagctctgctgcctga	cacgGcc
agaacgagctctgctgcctg	acacGgc
acggccaacaccaaccagaa	cttgGgA
tccgctctgggcttcattct	caactcg
cgactccctgcaaacacaaa	gaaagag
atctacatcacgtaacctgt	cttaGaA
tatctccttcattgagcac	caaacc

Реакции *in vitro* разрезания ДНК проводили в условиях аналогичных описанным выше экспериментам. В качестве ДНК-мишени использовался фрагмент гена человека *grin2b* размером около 500 н.п:

ttgtctctgcctgtagctgccaatgactatagcaatagcaccttttattgcctgttcaaggattctgaggcttttga
aagtttcattttctctcattctgcagagcaaataccagagataagagagtaggctggtagatggagttgggttgctc
aatgaaaggagataaggctcctgaattgcagatctagcctcttctaagacaggttacgtgatgtagatcctatttaacat
gctctttcttgtgttgagggagtcgacgagttgaagatgaagcccagagcggagtgctgttctcccaagttctggttg
tgttgccgctcctggccgtgtcaggcagcagagctcgttctcagaagagccccccagcattggcattgctgtcatcctc
gtgggcacttccgacgaggtggccatcaaggatgccacgagaaagatgattccaccatctctccgtggtaccccg
gg

Результаты эксперимента показывают, что CcCas9 в комплексе с направляющими РНК способен распознавать различные ДНК мишени, содержащие последовательность PAM NNNNGNA (Фиг. 5). В случаях некоторых мишеней CcCas9 толерантен к заменам в 7 положении последовательности PAM.

Пример 2. Температурный диапазон активности CcCas9.

Для уточнения температурного диапазона работы белка CcCas9 были проведены эксперименты по *in vitro* разрезанию ДНК-мишени в различных температурных условиях.

Для этого ДНК-мишень, фланкированную последовательностью PAM GAGAGTA, подвергали разрезанию эффекторным комплексом CcCas9 с соответствующими направляющими РНК при разных температурах (Фиг. 6).

Оказалось, что белок CcCas9 обладает широким температурным диапазоном работы. Максимум нуклеазной активности достигается при температуре в 45 °C, в то время как белок достаточно активен в диапазоне от 37 °C до 55 °C. Таким образом, CcCas9 в комплексе с направляющими РНК является новым инструментом разрезания (образования двунитевых разрывов) в молекуле ДНК, ограниченной последовательностью 5'-NNNNGNA-3', с температурным диапазоном работы от 37 °C до 55 °C. Схема комплекса таргетной ДНК с крРНК (crRNA) и трейсерной РНК (tracrRNA), вместе образующей направляющую РНК, изображена на Фиг. 7.

Пример 3.

Белки Cas9 из близкородственных организмов, относящихся к *Clostridium*.

На сегодняшний день в клостридиях была найдена всего одна CRISPR Cas система второго класса -Cas9 CRISPR Cas система из *Clostridium perfringens* (Maikova A, et al., New Insights Into Functions and Possible Applications of *Clostridium difficile* CRISPR-Cas System. *Front Microbiol.* 2018 Jul 31;9:1740).

Cas9 белок из этой бактерии *Clostridium perfringens* идентичен на 36% белку CcCas9 (степень идентичности была рассчитана по программе BLASTp, default parameters). Сравнимый по размерам белок Cas9 из *Staphylococcus aureus* идентичен CcCas9 на 28% (BLASTp, default parameters).

Таким образом, белок CcCas9 существенно отличается по аминокислотной последовательности от других Cas9 белков, изученных на сегодняшний день, в том числе белков, найденных в родственных организмах.

Специалисту в области генетической инженерии очевидно, что полученный и охарактеризованный Заявителем вариант последовательности белка CcCas9 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. 8. 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'-NNNNGNA-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 CcCas9 в комплексе с направляющими РНК может быть использована для изменения последовательности геномной ДНК многоклеточного организма, в том числе эукариотического. Для введения система CcCas9 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 CcCas9 в эукариотических клетках будет желательно провести оптимизацию кодонов для аминокислотной последовательности белка CcCas9 методами, известными специалистам (например, IDT codon optimization tool).

Для эффективной работы нуклеазы CcCas9 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 CcCas9 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: MAPKKKRKVGINGVPAA-CcCas9-KRPAATKKAGQAKKKK (далее CcCas9 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 CcCas9 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'-NNNNGNA-3', например, последовательности гена *grin2b* человека:

асggссаасассаассагаа
сgactccctgcaaасасааа

Таким образом, кассета для экспрессии крРНК выглядит следующим образом:

**Gagggcctatttcccatgattccttcatatttgcataacgatacaaggctgtagagagataattgga
attaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggttagttg
сад ttttaaaattatg ttttaaaatg д actatcatatg cttaccg taacttg aaag tatttcgatttcttg gctttatat
atcttg tg gaaag gacgaaacacc -асggссаасассаассагаа--
GTTATAGCTCCAATTCAGGCTCCGATATtttt**

Жирным шрифтом выделена последовательность U6 промотора, далее последовательность, необходимая для узнавания целевой ДНК, а последовательность прямого повтора выделена заглавными буквами.

Кассета для экспрессии трейсерной РНК выглядит следующим образом:

**Gagggcctatttcccatgattccttcatatttgcataacgatacaaggctgtagagagataattgga
attaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttggttagttg
сад ttttaaaattatg ttttaaaatg д actatcatatg cttaccg taacttg aaag tatttcgatttcttg gctttatat
atcttgaggaaaggacgaaacaccATT ATGGCATATCGGAGCCTGAATTGTTGCTATAAT
AAGGTGCTGGGTTT AGCCAG ACCGCCAAGTT AACCCCGGCATTT ATTGCTGG
GGTATCTTTGtttt**

Жирным шрифтом выделена последовательность U6 промотора, далее последовательность, кодирующая трейсерную РНК.

Плазмидную ДНК очищают и трансфицируют в клетки человека HEK 293 с помощью реагента Lipofectamine2000 (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, например для указанных выше сайтов гена *grin2b*:

5'-GACTATAGCAATAGCAC-3'

5TCAACTCGTCGACTCCCTG-3'

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 CcCas9 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 1:2:2 (CcCas9 NLS : crRNA :tracrRNA), 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 CcCas9 нуклеаза из бактерии *Clostridium cellulolyticum* имеет ряд преимуществ относительно ранее охарактеризованных Cas9 белков.

CcCas9 обладает коротким, двухбуквенным, отличным от других известных PAM Cas нуклеаз, необходимым для функционирования системы. Авторами было показано, что для CcCas9 достаточен короткий PAM GNA, расположенный в 4 нуклеотидах от протоспейсера. При этом G в +5 позиции является строгой, в то время как +7 позиция имеет меньшую значимость, и *in vitro* гидролиз был детектирован не только в присутствии A или T, но и в присутствии C в +7 позиции, хоть и с несколько меньшей эффективностью.

Известные на сегодняшний день большинство Cas нуклеаз, способных вносить двунитевые разрывы в ДНК, имеют сложные многобуквенные PAM

последовательности , ограничивающие выбор последовательности , пригодных для разрезания . Среди изученных Cas нуклеаз , распознающих короткие PAM , только CcCas9 может распознавать последовательности , ограниченные GNA нуклеотидами .

Второе преимущество CcCas9 - малый размер белка (1030 а.о., что на 23 а.о. меньше , чем размер SaCas9). На сегодняшний день это единственный изученный малоразмерный белок , обладающий двухбуквенной PAM последовательностью .

Третье преимущество CcCas9 системы - широкий рабочий диапазон температур : нуклеаза активна при температурах от 37 °C до 65 °C с оптимумом 45 °C.

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'-NNNNGNA-3' в указанной молекуле ДНК.

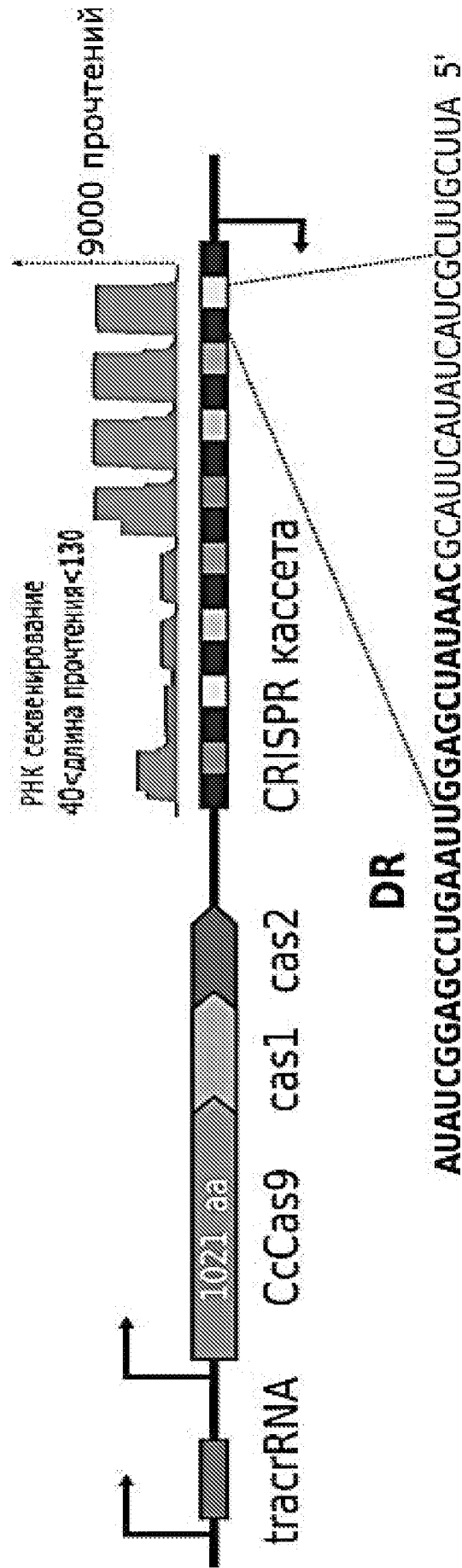
2. Применение белка по п. 1, характеризующееся тем, что образование двунитевого разрыва в молекуле ДНК происходит при температуре от 37 °С до 65 °С.

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

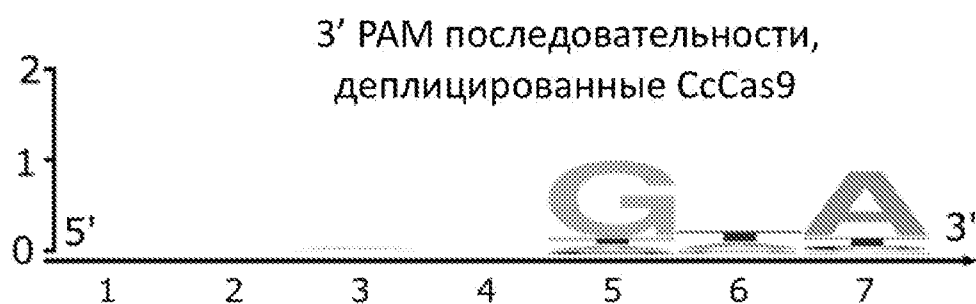
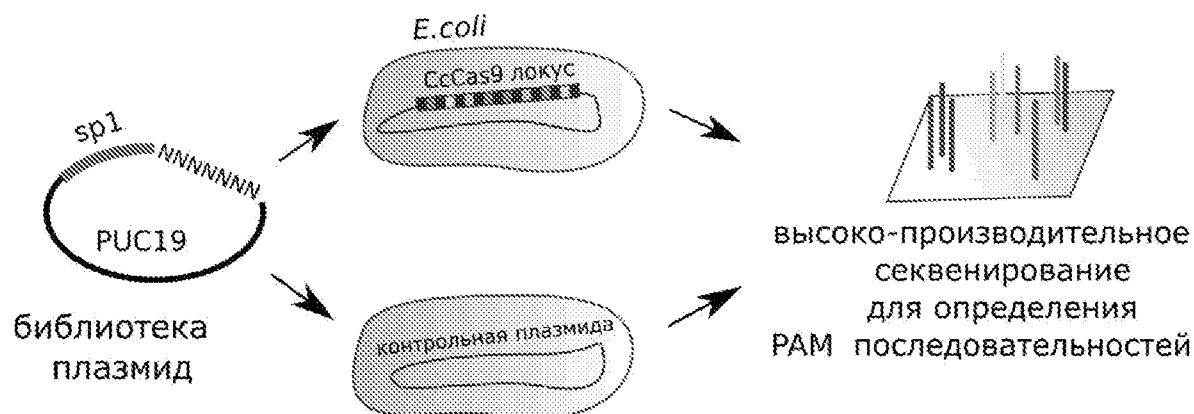
4. Способ образования двунитевого разрыва в последовательности геномной ДНК одноклеточного или многоклеточного организма, непосредственно примыкающей к последовательности 5'-NNNNGNA-3', включающий введение в по меньшей мере одну клетку этого организма эффективного количества: а) белка, содержащего аминокислотную 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'-NNNNGNA-3', и взаимодействующей с указанным белком после образования дуплекса, или последовательности ДНК, кодирующей указанную направляющую РНК;

при этом взаимодействие указанного белка с указанной направляющей РНК и нуклеотидной последовательностью 5'-NNNNGNA-3' приводит к образованию двунитевого разрыва в последовательности геномной ДНК, непосредственно примыкающей к последовательности 5'-NNNNGNA-3'.

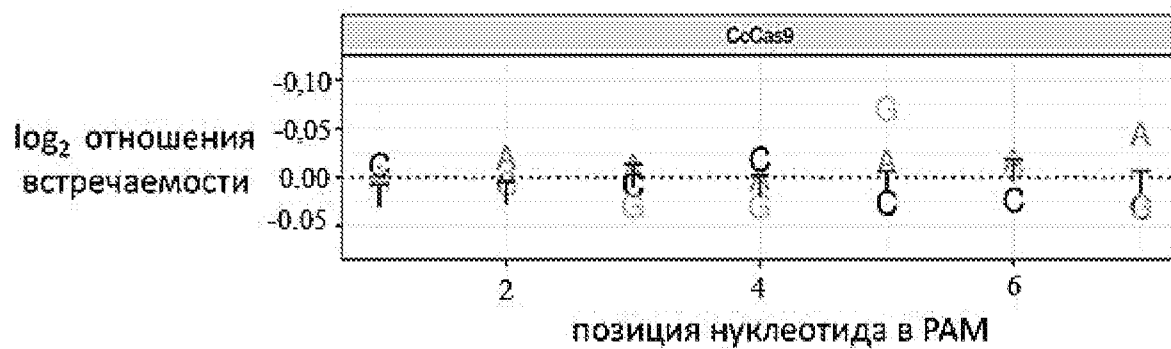
5. Способ по п. 4, дополнительно включающий введение экзогенной последовательности ДНК одновременно с направляющей РНК.



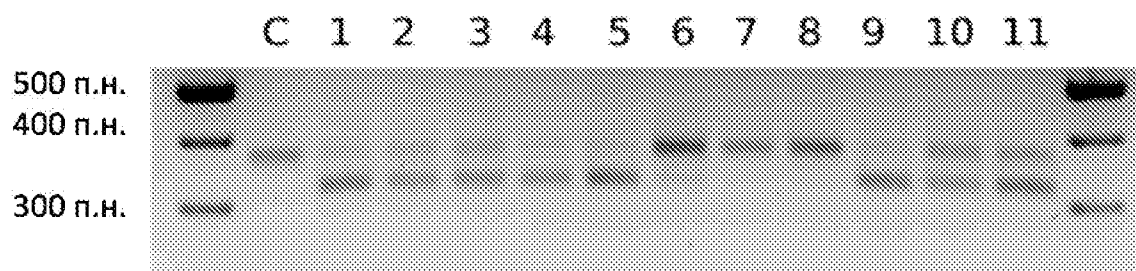
Фиг.1



Фиг. 2.

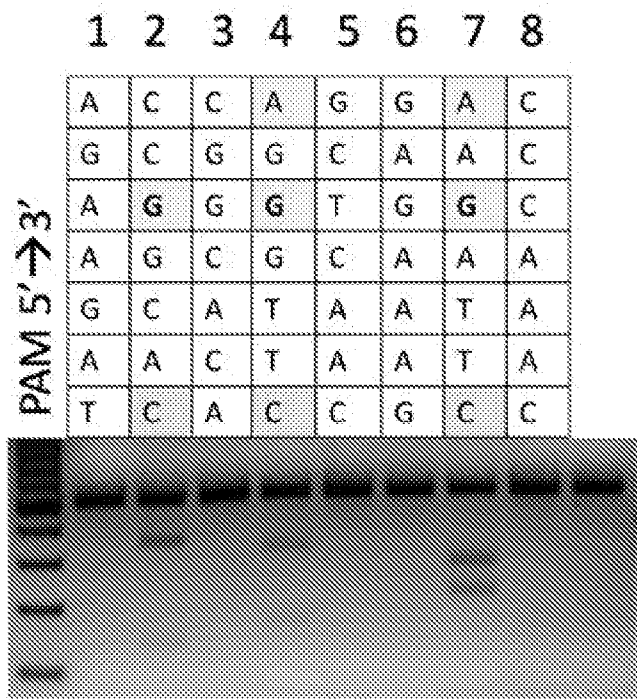


Фиг. 3

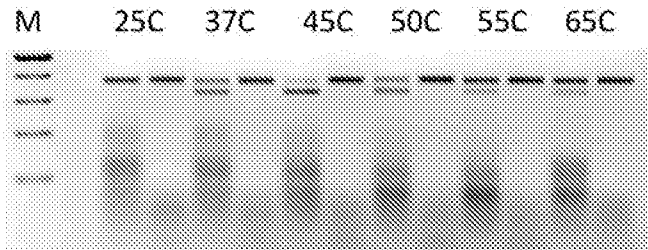


	PAM (5'→3')							порезка
1	G	A	G	A	G	T	A	+
2	T	A	G	A	G	T	A	+
3	G	C	G	A	G	T	A	+
4	G	A	T	A	G	T	A	+
5	G	A	G	T	G	T	A	+
6	G	A	G	A	A	T	A	-
7	G	A	G	A	C	T	A	-
8	G	A	G	A	T	T	A	-
9	G	A	G	A	G	A	A	+
10	G	A	G	A	G	T	C	+
11	G	A	G	A	G	T	T	+

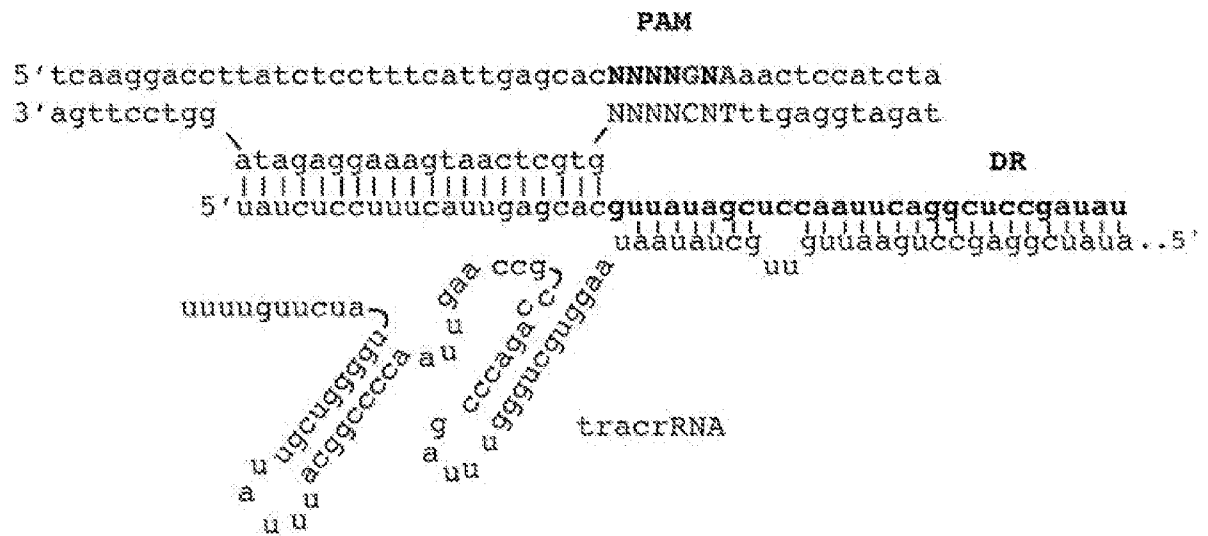
Фиг. 4



Фиг.5



Фиг.6



Фиг. 7

6/6