



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

C12N 15/111 (2019.05); C12N 15/63 (2019.05); C12N 15/8213 (2019.05); C12N 15/902 (2019.05); C12N 15/907 (2019.05); C12N 2310/20 (2019.05); C12N 2320/51 (2019.05); C12N 2320/53 (2019.05); C12N 9/16 (2019.05); C12N 9/22 (2019.05); C12Q 1/6827 (2019.05); C12Y 301/00 (2019.05)

(21)(22) Заявка: 2017130143, 27.01.2016

(24) Дата начала отсчета срока действия патента:
27.01.2016

Дата регистрации:
04.02.2020

Приоритет(ы):

(30) Конвенционный приоритет:
28.01.2015 US 62/108,931;
05.11.2015 US 62/251,548

(43) Дата публикации заявки: 28.02.2019 Бюл. № 7

(45) Опубликовано: 04.02.2020 Бюл. № 4

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 28.08.2017

(86) Заявка РСТ:
US 2016/015145 (27.01.2016)

(87) Публикация заявки РСТ:
WO 2016/123230 (04.08.2016)

Адрес для переписки:
129090, Москва, ул. Большая Спасская, д. 25,
строение 3, ООО "Юридическая фирма
Городисский и Партнеры"

(72) Автор(ы):

МЭЙ Эндрю П. (US),
ДОНОХЬЮ Пол Д. (US)

(73) Патентообладатель(и):

ПАЙОНИР ХАЙ-БРЕД ИНТЕРНЭШНЛ,
ИНК. (US)

(56) Список документов, цитированных в отчете
о поиске: WO 2014093661 A3, 19.06.2014.
ЯРЫГИН В.Н. и др., Биология, том 1,
"Высшая школа", Москва, 1997, стр. 78-84. RU
2531343 C2, 20.10.2014.

(54) ГИБРИДНЫЕ ДНК/РНК-ПОЛИНУКЛЕОТИДЫ CRISPR И СПОСОБЫ ПРИМЕНЕНИЯ

(57) Реферат:

Изобретение относится к биотехнологии. Описана система CRISPR класса 2, содержащая: (i) единичный полинуклеотид, содержащий (а) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК); и (b) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК; и где указанная активирующая область

содержит стебель и выпетливание и способна взаимодействовать с белком Cas9; и (ii) Cas9. Представлен соответствующий единичный полинуклеотид. Описан способ модификации молекулы нуклеиновой кислоты-мишени в организме. Описан способ модуляции транскрипции по меньшей мере одного гена в молекуле нуклеиновой кислоты-мишени в организме. Представлен набор из двух

полинуклеотидов CRISPR класса 2. Изобретение расширяет арсенал средств, используемых для

редактирования генома. 17 н. и 115 з.п. ф-лы, 14 ил., 18 табл., 13 пр.

R U 2 7 1 3 3 2 8 C 2

R U 2 7 1 3 3 2 8 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.

C12N 15/11 (2006.01)*C12N 9/22* (2006.01)*C12N 15/63* (2006.01)(12) **ABSTRACT OF INVENTION**

(52) CPC

C12N 15/111 (2019.05); *C12N 15/63* (2019.05); *C12N 15/8213* (2019.05); *C12N 15/902* (2019.05); *C12N 15/907* (2019.05); *C12N 2310/20* (2019.05); *C12N 2320/51* (2019.05); *C12N 2320/53* (2019.05); *C12N 9/16* (2019.05); *C12N 9/22* (2019.05); *C12Q 1/6827* (2019.05); *C12Y 301/00* (2019.05)

(21)(22) Application: **2017130143, 27.01.2016**

(24) Effective date for property rights:
27.01.2016

Registration date:
04.02.2020

Priority:

(30) Convention priority:
28.01.2015 US 62/108,931;
05.11.2015 US 62/251,548

(43) Application published: **28.02.2019 Bull. № 7**(45) Date of publication: **04.02.2020 Bull. № 4**(85) Commencement of national phase: **28.08.2017**

(86) PCT application:
US 2016/015145 (27.01.2016)

(87) PCT publication:
WO 2016/123230 (04.08.2016)

Mail address:
129090, Moskva, ul. Bolshaya Spasskaya, d. 25,
stroenie 3, OOO "Yuridicheskaya firma
Gorodisskij i Partnery"

(72) Inventor(s):

MAY, Andrew P. (US),
DONOHUE, Paul D. (US)

(73) Proprietor(s):

PIONEER HI-BRED INTERNATIONAL, INC.
(US)

(54) **HYBRID DNA/RNA POLYNUCLEOTIDES CRISPR AND METHODS OF APPLIANCE**

(57) Abstract:

FIELD: biotechnology.

SUBSTANCE: invention relates to the biotechnology. Described is class 2 CRISPR system comprising: (i) a single polynucleotide comprising (a) a targeted region containing deoxyribonucleic acid (DNA) or ribonucleic acid (RNA); and (b) an activating region adjacent to said targeting region containing DNA; and wherein said activating region comprises a stem and a vaginal discharge and is capable of reacting with the Cas9 protein; and (ii) Cas9. Disclosed is a

corresponding single polynucleotide. Described is a method of modifying a target nucleic acid molecule in an organism. Described is a method of modulating transcription of at least one gene in a molecule of a target nucleic acid in an organism. Also presented is a kit of two polynucleotides CRISPR class 2.

EFFECT: invention extends the range of products used for genome editing.

132 cl, 14 dwg, 18 tbl, 13 ex

ПЕРЕКРЕСТНАЯ ССЫЛКА НА РОДСТВЕННЫЕ ЗАЯВКИ

Настоящая заявка притязает на приоритет на основании предварительной заявки на выдачу патента США с регистрационным номером 62/108931, поданной 28 января 2015 года, и предварительной заявки на выдачу патента США с регистрационным номером 62/251548, поданной 5 ноября 2015 года, все указанные заявки включены в настоящее описание в виде ссылки.

СПИСКО ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Настоящая заявка содержит список последовательностей который был подан в электронном виде в формате ASCII и включен в настоящее описание в виде ссылки в полном объеме. Указанная копия ASCII, созданная 26 января 2016 года, названа 0198470101PTWO_SL.txt и имеет размер 83522 байт.

УРОВЕНЬ ТЕХНИКИ

Кластерные короткие палиндромные повторы, регулярно разделенные промежутками (CRISPR) и CRISPR-ассоциированные (Cas) системы представляют собой прокариотическую иммунную систему, впервые открытую Ishino у *E. coli*. Ishino с соавторами, 1987 (Journal of Bacteriology 169 (12): 5429-5433(1987)). Такая иммунная система обеспечивает иммунитет против вирусов и плазмид за счет целенаправленного воздействия на нуклеиновые кислоты вирусов и плазмид специфичным для последовательности образом.

Существуют две основных стадии, вовлеченные в такую иммунную систему, первая стадия приобретения и вторая стадия интерференции. Первая стадия заключается в разрезании генома вторгающихся вирусов и плазмид и интеграции их участков в локус CRISPR организма. Участки, которые интегрируются в геном, названы протоспейсерами и помогают организму защищаться от последующей атаки таким же вирусом или плазмидой. Вторая стадия включает в себя атаку вторгающегося вируса или плазмиды. Даная стадия основана на протоспейсерах, транскрибирующихся в РНК, на такой РНК, после определенного процессинга и затем гибридизации с комплементарной последовательностью в ДНК вторгающегося вируса или плазмиды в ассоциации с белком или белковым комплексом, который эффективно расщепляет ДНК.

Существует несколько разных систем CRISPR/Cas, и номенклатуру и классификацию таких систем изменяли в связи с дополнительной характеристикой систем. В системах типа II имеются две нити РНК, РНК CRISPR (cr-РНК) и трансактивирующая РНК CRISPR (tracr-РНК, которые составляют часть системы CRISPR/Cas. Tracr-РНК гибридизуется с комплементарной областью пре-cr-РНК, вызывая созревание пре-cr-РНК до cr-РНК. Дуплекс, образованный tracr-РНК и cr-РНК, распознается и ассоциирует с белком, Cas9, который направляется к нуклеиновой кислоте-мишени последовательностью cr-РНК, которая является комплементарной и гибридизуется с последовательностью в нуклеиновой кислоте-мишени. Было показано, что такие минимальные компоненты основанной на РНК иммунной системе могут быть перепрограммированы на ДНК-мишень сайт-специфичным образом за счет использования одного белка и двух последовательностей РНК-проводника или одной молекулы РНК. Система CRISPR/Cas является превосходной системой по сравнению с другими способами редактирования генома, в которые вовлечены эндонуклеазы, мегануклеазы, нуклеазы с цинковыми пальцами и подобные активаторам транскрипции эффекторные нуклеазы (TALEN), в случае которых может требоваться конструирование белков de novo для каждого нового локуса-мишени.

Системы CRISPR/Cas, являясь управляемыми РНК системами, склонны к появлению проблем, связанных со структурами гибридов РНК-ДНК, такими как разрушение нити

РНК РНК-азой А и более высокая вероятность ошибочных спариваний РНК-ДНК. Кроме того, синтез олигонуклеотидов ДНК является более экономичным и надежным, чем синтез олигонуклеотидов РНК. Управляемые ДНК системы CRISPR также могут привлекать дополнительный механизм для конкретной мишени, по сравнению со встречающимися в природе управляемыми РНК системами CRISPR. Существует потребность в улучшенной системе, которая преодолевает проблемы, связанные с основанными на РНК системами CRISPR/Cas, обеспечивает более низкую стоимость и повышенную надежность синтеза ДНК и повышает специфичность системы CRISPR/Cas.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

В некоторых вариантах изобретение относится к единичному полинуклеотиду для применения с использованием системы CRISPR класса 2, содержащему: нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК); и активирующую область, содержащую рибонуклеиновую кислоту (РНК). В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК; и активирующая область содержит ДНК, РНК или смесь ДНК и РНК.

В некоторых вариантах изобретение относится к единичному полинуклеотиду для применения с использованием системы CRISPR класса 2, содержащему: нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК); и активирующую область, включающую в себя область полинуклеотида, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК). В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК; и активирующая область содержит ДНК, РНК или смесь ДНК и РНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах активирующая область содержит структуру, выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с белком Cas9. В некоторых вариантах активирующая область взаимодействует с белком Cpf1.

В некоторых вариантах изобретение относится к системе CRISPR класса 2, включающей в себя: единичный полинуклеотид, содержащий нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте; активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и сайт-специфичный полипептид. В некоторых вариантах нуклеиновая кислота представляет собой ДНК, в некоторых вариантах нуклеиновая кислота представляет собой РНК, в некоторых вариантах нуклеиновая кислота представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах активирующая область содержит структуру выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с сайт-

специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах система CRISPR класса 2 дополнительно включает в себя донорный полинуклеотид.

5 В некоторых вариантах изобретение относится к системе CRISPR класса 2, включающей в себя первый полинуклеотид, содержащий (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте, и (ii) активирующую область, смежную с указанной нацеленной
10 на мишень областью, содержащую рибонуклеиновую кислоту (РНК); второй полинуклеотид, содержащий последовательность, которая комплементарна последовательности в указанной активирующей области указанного первого полинуклеотида; и сайт-специфичный полипептид. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной
15 или более структур, выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах сайт-специфичный полипептид взаимодействует с активирующей областью. В некоторых
20 вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах второй полинуклеотид содержит РНК, ДНК или смесь ДНК и РНК.

В некоторых вариантах изобретение относится к двум полинуклеотидам для применения с использованием системы CRISPR класса 2, включая первый полинуклеотид, содержащий (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую
25 кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте, и (ii) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и второй полинуклеотид, содержащий последовательность, которая комплементарна последовательности в указанной активирующей области указанного
30 первого полинуклеотида. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной или более структур, выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК, активирующая область содержит смесь ДНК и РНК, и второй
35 полинуклеотид содержит смесь ДНК и РНК.

В некоторых вариантах изобретение относится к способу модификации молекулы нуклеиновой кислоты-мишени, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: единичным полинуклеотидом, содержащим нацеленную на мишень область, содержащую
40 дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте; активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и сайт-специфичным полипептидом, при этом единичный полинуклеотид образует комплекс с сайт-специфичным
45 полипептидом и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или транскрипция, по меньшей мере, одного гена, кодируемого молекулой нуклеиновой кислоты-мишени модулируется. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-

мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах активирующая область содержит структуру, выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с сайт-специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах способ дополнительно включает в себя обеспечение донорным полинуклеотидом.

В некоторых вариантах изобретение относится к способу модификации молекулы нуклеиновой кислоты-мишени, при этом способ включает в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: первым полинуклеотидом, содержащим (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте и (ii) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); предоставление второго полинуклеотида, содержащего последовательность, которая комплементарна последовательности в указанной активирующей области указанного первого полинуклеотида, и сайт-специфичного полипептида, при этом первый и второй полинуклеотиды образуют комплекс с сайт-специфичным полипептидом и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или транскрипция, по меньшей мере одного гена, кодируемого молекулой нуклеиновой кислоты-мишени, модулируется. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной или более структур, выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК, активирующая область содержит смесь ДНК и РНК, и второй полинуклеотид содержит смесь ДНК и РНК. В некоторых вариантах способ дополнительно включает в себя предоставление донорного полинуклеотида.

В некоторых вариантах изобретение относится к способу снижения модификации вне мишени с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: единичным полинуклеотидом, содержащим нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте; активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и сайт-специфичным полипептидом, при этом единичный полинуклеотид образует комплекс с сайт-специфичным полипептидом и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или редактируется в последовательности-мишени более предпочтительно, чем в других последовательностях в нуклеиновой кислоте-мишени, тем самым снижая модификацию вне мишени. В некоторых вариантах нуклеиновая

кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах активирующая область содержит структуру, выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с сайт-специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах указанная нацеленная на мишень область не содержит урацила. В некоторых вариантах способ дополнительно включает в себя предоставление донорного полинуклеотида.

В некоторых вариантах изобретение относится к способу снижения модификации вне мишени с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: первым полинуклеотидом, содержащим (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте и (ii) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); предоставление второго полинуклеотида, содержащего последовательность, которая комплементарна последовательности в указанной активирующей области указанного первого полинуклеотида, и сайт-специфичным полипептидом, при этом первый и второй полинуклеотиды образуют комплекс с сайт-специфичным полипептидом и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или редактируется в последовательности-мишени более предпочтительно, чем в других последовательностях в нуклеиновой кислоте-мишени, тем самым снижая модификацию вне мишени. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной или более структур выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК, активирующая область содержит смесь ДНК и РНК, и второй полинуклеотид содержит смесь ДНК и РНК. В некоторых вариантах указанная нацеленная на мишень область не содержит урацила. В некоторых вариантах способ дополнительно включает в себя предоставление донорного полинуклеотида.

В некоторых вариантах изобретение относится к способу повышения специфичной для мишени модификации с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: единичным полинуклеотидом, содержащим нацеленную

на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте; активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и сайт-специфичным полипептидом, при этом единичный полинуклеотид образует комплекс с сайт-специфичным полипептидом, и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или редактируется в последовательности-мишени более предпочтительно, чем в других последовательностях в нуклеиновой кислоте-мишени, тем самым повышая специфичную для мишени модификацию. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах активирующая область содержит структуру, выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с сайт-специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах способ дополнительно включает в себя предоставление донорного полинуклеотида.

В некоторых вариантах изобретение относится к способу повышения специфичной для мишени модификации с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: первым полинуклеотидом, содержащим (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте, и (ii) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); предоставление второго полинуклеотида, содержащего последовательность, которая комплементарна последовательности в указанной активирующей области указанного первого полинуклеотида, и сайт-специфичного полипептида, при этом первый и второй полинуклеотиды образуют комплекс с сайт-специфичным полипептидом, и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется или редактируется в последовательности-мишени более предпочтительно, чем в других последовательностях в нуклеиновой кислоте-мишени, тем самым повышая специфичную для мишени модификацию. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной или более структур, выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК,

активирующая область содержит смесь ДНК и РНК и второй полинуклеотид содержит смесь ДНК и РНК. В некоторых вариантах указанная нацеленная на мишень область не содержит урацила. В некоторых вариантах способ дополнительно включает в себя предоставление донорного полинуклеотида.

5 В некоторых вариантах изобретение относится к способу введения донорного полинуклеотида в геном клетки или организма с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: единичным полинуклеотидом, содержащим нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с
10 последовательностью-мишенью в нуклеиновой кислоте; активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); и сайт-специфичным полипептидом, при этом единичный полинуклеотид образует комплекс с сайт-специфичным полипептидом, и при этом указанная молекула
15 нуклеиновой кислоты-мишени расщепляется в последовательности-мишени или вблизи нее, и предоставление донорного полинуклеотида, который вводится в геном клетки или организма в сайте расщепления. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень
20 представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах
25 активирующая область содержит структуру, выбранную из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и шпильки. В некоторых вариантах активирующая область содержит структуру стебель-петля. В некоторых вариантах активирующая область взаимодействует с сайт-специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК
30 и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах донорный полинуклеотид вводят в нуклеиновую кислоту посредством гомологичной рекомбинации. В некоторых вариантах донорный полинуклеотид вводят в нуклеиновую кислоту посредством соединения негомологичных концов.

35 В некоторых вариантах изобретение относится к способу введения донорного полинуклеотида в геном клетки или организма с использованием системы CRISPR класса 2, включающему в себя: осуществление контакта молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, с: первым полинуклеотидом, содержащим (i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую
40 кислоту (ДНК) и имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте и (ii) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую рибонуклеиновую кислоту (РНК); предоставление второго полинуклеотида, содержащего последовательность, которая комплементарна последовательности в указанной
45 активирующей области указанного первого полинуклеотида, и сайт-специфичного полипептида, при этом первый и второй полинуклеотиды образуют комплекс с сайт-специфичным полипептидом, и при этом указанная молекула нуклеиновой кислоты-мишени расщепляется в последовательности-мишени или вблизи нее, и предоставление

донорного полинуклеотида, который вводят в геном клетки или организма в сайте расщепления. В некоторых вариантах нуклеиновая кислота-мишень представляет собой ДНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой РНК, в некоторых вариантах нуклеиновая кислота-мишень представляет собой смесь РНК и ДНК. В некоторых вариантах активирующая область и второй полинуклеотид гибридизуются с образованием одной или более структур, выбранных из группы, состоящей из нижнего стебля, выпетливания, верхнего стебля, связующего звена, и дуплекса. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК, активирующая область содержит смесь ДНК и РНК, и второй полинуклеотид содержит смесь ДНК и РНК. В некоторых вариантах активирующая область взаимодействует с сайт-специфичным полипептидом. В некоторых вариантах активирующая область содержит смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область содержит смесь ДНК и РНК. В некоторых вариантах сайт-специфичным полипептидом является белок Cas9. В некоторых вариантах сайт-специфичным полипептидом является белок Cpf1. В некоторых вариантах донорный полинуклеотид вводят в нуклеиновую кислоту посредством гомологичной рекомбинации. В некоторых вариантах донорный полинуклеотид вводят в нуклеиновую кислоту посредством связывания нехомологичных концов. В некоторых вариантах донорный полинуклеотид вводят посредством опосредованного микрогомологии связывания концов. В некоторых вариантах донорный полинуклеотид вводят посредством отжига одиночных нитей.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

На Фигуре 1А показана cr-Д(Р)НК и tracr-РНК системы CRISPR типа II.

На фигуре 1В показаны два полинуклеотида (cr-Д(Р)НК и tracr-РНК или tracr-Д(Р)НК) согласно настоящему изобретению, гибридизованные друг с другом (также называемые системой «с двойным проводником»).

На фигуре 2 показан единичный полинуклеотид согласно настоящему изобретению, содержащий нацеленную на мишень область, связанную с активирующей областью (также называемый системой с «с одним проводником» или «с единичным проводником Д(Р)НК» или «sg-Д(Р)НК»).

На фигуре 3 показано расщепление последовательности ДНК-мишени системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

На фигурах 4А и В показаны результаты биохимических анализов *in vitro* для количественного определения расщепления различных последовательностей-мишеней системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

На фигуре 5 показаны результаты анализов *in vivo* для количественного определения расщепления последовательности-мишени системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

На фигуре 6 показаны результаты биохимических анализов *in vitro* для количественного определения расщепления последовательности-мишени вне мишени системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

На фигуре 7 показаны результаты анализов *in vivo* для количественного определения расщепления последовательности-мишени системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

изобретению.

На фигуре 8 показаны результаты никирующей активности cr-Д(Р)НК или sg-Д(Р)НК с белком Cas9-D10A против плазмидной мишени *in vitro*.

На фигуре 9 показана типичная структура cr-РНК из системы CRISPR типа V.

5 На фигурах 10А-С показаны возможные структуры единичного Д(Р)НК-проводника согласно настоящему изобретению для применения с использованием системы CRISPR типа V.

10 На фигурах 11А-Е показаны возможные структуры единичного Д(Р)НК-проводника согласно настоящему изобретению для применения с использованием системы CRISPR типа V.

На фигурах 12А-І показаны возможные компоненты двойных проводников согласно настоящему изобретению, содержащих cr-РНК и/или cr-Д(Р)НК для применения с использованием системы CRISPR типа V.

15 На фигурах 13А-Н показаны возможные конфигурации двойных проводников согласно настоящему изобретению, содержащих cr-РНК и/или cr-Д(Р)НК для применения с использованием системы CRISPR типа V.

На фигурах 14А-В показаны результаты секвенирования в анализе *in planta* для количественного определения расщепления последовательности-мишени системой CRISPR/Cas типа II с использованием направленных к нуклеиновой кислоте полинуклеотидов согласно настоящему изобретению.

ПОДРОБНОЕ ОПИСАНИЕ

25 Системы CRISPR/Cas недавно были переклассифицированы с разделением на два класса, содержащих пять типов и шестнадцать подтипов, Макарова с соавторами (Nature Reviews Microbiology 13: 1-15 (2015)). Указанная классификация основана на идентификации всех генов *cas* в локусе CRISPR/Cas и затем определении сигнатурных генов в каждом локусе CRISPR/Cas, что в конечном счете позволяет определить, что системы CRISPR/Cas могут быть помещены в любой класс 1 или класс 2 на основании генов, кодирующих эффекторный модуль, т.е.. белки, вовлеченные в стадию интерференции.

30 Системы класса 1 имеют состоящий из множества субъединиц cr-РНК-эффекторный комплекс, тогда как системы класса 2 имеют единственный белок, такой как Cas 9, Cpf1, C2c1, C2c2, C2c3 или cr-РНК-эффекторный комплекс. Системы класса 1 включают системы типа I, типа III и типа IV. Системы класса 2 включают системы типа II и типа V.

35 Все системы типа I имеют белок Cas3, который обладает геликазной активностью и расщепляющей активностью. Системы типа I дополнительно подразделяют на семь подтипов (I-A - I-F и I-U). Каждый подтип типа I имеет определенное сочетание сигнатурных генов и различающиеся признаки организации оперона. Например, подтипы I-A и I-B, по-видимому, имеют гены *cas*, организованные в виде двух или более оперонов, тогда как подтипы I-C - I-F, по-видимому, имеют гены *cas*, кодируемые одним опероном. 40 Системы типа I имеют мультибелковый cr-РНК-эффекторный комплекс, который вовлечен в стадии процессинга и интерференции иммунной системы CRISPR/Cas. Такой мультибелковый комплекс известен как CRISPR-ассоциированный комплекс для защиты от вирусов (Cascade). Подтип I-A содержит *csa5*, который кодирует белок малой субъединицы, и ген *cas8*, который расщепляется надвое, кодируя деградированные большую и малую субъединицы, а также имеет ген расщепления *cas3*. Примером организма с системой CRISPR/Cas подтипа I-A является *Archaeoglobus fulgidus*.

Подтип I-B имеет следующее расположение генов *cas1-cas2-cas3-cas4-cas5-cas6-cas7-*

cas8 и не имеет гена csa5. Примером организма с подтипом I-B является *Clostridium kluyveri*. Подтип I-C не имеет гена cas6. Примером организма с подтипом I-C является *Bacillus halodurans*. Подтип I-D имеет Cas10d вместо Cas8. Примером организма с подтипом I-D является *Cyanothece* sp. Подтип I-E не имеет cas4. Примером организма с подтипом I-E является *Escherichia coli*. Подтип I-F не имеет cas4 и имеет cas2, слитый с cas3. Примером организма с подтипом I-F является *Yersinia pseudotuberculosis*. Примером организма с подтипом I-U является *Geobacter sulfurreducens*.

Все системы типа III имеют ген cas10, который кодирует мультидоменный белок, содержащий домен Palm (вариант мотива узнавания РНК (RRM)), который гомологичен коровому домену многочисленных полимераз и циклаз нуклеиновых кислот и который является самой крупной субъединицей cr-РНК-эффекторных комплексов типа III. Все локусы типа III также кодируют белок малой субъединицы, один белок Cas5 и обычно несколько белков Cas7. Тип III может быть дополнительно разделен на четыре подтипа, III-A - III-D. Подтип III-A имеет ген csm2, кодирующий малую субъединицу, а также имеет гены cas1, cas2 и cas6. Примером организма с подтипом III-A является *Staphylococcus epidermidis*. Подтип III-B имеет ген cmr5, кодирующий малую субъединицу, а также обычно не имеет генов cas1, cas2 и cas6. Примером организма с подтипом III-B является *Pyrococcus furiosus*. Подтип III-C имеет белок Cas10 с неактивным циклазо-подобным доменом и не имеет гена cas1 и cas2. Примером организма с подтипом III-C является *Methanothermobacter thermautotrophicus*. Подтип III-D имеет белок Cas10, который не имеет домена HD, он не имеет гена cas1 и cas2 и имеет cas5-подобный ген, известный как csx10. Примером организма с подтипом III-D является *Roseiflexus* sp.

Системы типа IV кодируют минимальный мультисубъединичный cr-РНК-эффекторный комплекс, содержащий частично деградированную большую субъединицу, Csf1, Cas5, Cas7 и в некоторых случаях предполагаемую малую субъединицу. Системы типа IV не имеют генов cas1 и cas2. Системы типа IV не имеют подтипов, но существует два разных варианта таких систем. Один вариант типа IV имеет геликазу семейства DinG, тогда как второй вариант типа IV не имеет геликазы семейства DinG, но имеет ген, кодирующий небольшой α -спиральный белок. Примером организма с системой типа IV является *Acidithiobacillus ferrooxidans*.

Системы типа II имеют гены cas1, cas2 и cas9. cas9 кодирует мультидоменный белок, который сочетает функции cr-РНК-эффекторного комплекса расщеплением ДНК-мишени. Системы типа II также кодируют tracr-РНК. Системы типа II дополнительно делят на три подтипа, подтипы II-A, II-B и II-C. Подтип II-A содержит дополнительный ген, csn2. Примером организма с системой подтипа II-A является *Streptococcus thermophilus*. Подтип II-B не имеет csn2, но имеет cas4. Примером организма с системой подтипа II-B является *Legionella pneumophila*. Подтип II-C является наиболее распространенной системой типа II, найденной у бактерий, и имеет только три белка, Cas1, Cas2 и Cas9. Примером организма с системой подтипа II-C является *Neisseria lactamica*.

Системы типа V имеют ген cpf1 и гены cas1 и cas2. Ген cpf1 кодирует белок, Cpf1, который имеет RuvC-подобный нуклеазный домен, который является гомологичным соответствующему домену Cas9, но не имеет HNH-нуклеазного домена, который присутствует в белках Cas9. Системы типа V были идентифицированы у нескольких бактерий, включая бактерию *Parcubacteria* GWC2011_GWC2_44_17 (PbCpf1), бактерию *Lachnospiraceae* MC2017 (Lb3Cpf1), *Butyrivibrio proteoclasticus* (BpCpf1), бактерию *Peregrinibacteria* GW2011_GWA_33_10 (PeCpf1), *Acidaminococcus* sp. BV3L6 (AsCpf1), *Porphyromonas macacae* (PmCpf1), бактерию *Lachnospiraceae* ND2006 (LbCpf1),

Porphyromonas crevioricanis (PcCpf1), *Prevotella disiens* (PdCpf1), *Moraxella bovoculi* 237 (MbCpf1), *Smithella* sp. SC_K08D17 (SsCpf1), *Leptospira inadai* (LiCpf1), бактерию *Lachnospiraceae* MA2020 (Lb2Cpf1), *Franciscella novicida* U112 (FnCpf1), *Candidatus methanoplasma termitum* (CMtCpf1) и *Eubacterium eligens* (EeCpf1).

5 В системах класса 1 в стадии экспрессии и интерференции вовлечены мультисубъединичные CRISPR РНК(сг-РНК)- эффекторные комплексы. В системах класса 2 в стадии экспрессии и интерференции вовлечен один крупный белок, например, Cas9, Cpf1, C2C1, C2C2 или C2C3.

В системах класса 1 пре-сг-РНК связывается с мультисубъединичным сг-РНК-
10 эффекторным комплексом и процессируется в зрелую сг-РНК. В системах типа I и III в это вовлечена РНК-эндонуклеаза, например, Cas6. В системах класса 2 типа II пре-сг-РНК связывается с Cas9 и процессируется в зрелую на стадии, в которую вовлечена РНК-аза III и trасг-РНК. Однако, по меньшей мере, в одной системе CRISPR-Cas типа II у *Neisseria meningitidis* сг-РНК со зрелыми 5'-концами транскрибируются
15 непосредственно с внутренних промоторов, и процессинг сг-РНК не происходит.

В системах класса 1 сг-РНК вступает в ассоциацию с сг-РНК-эффекторным комплексом и достигает интерференции за счет сочетания нуклеазной активности с РНК-связывающими доменами и образованием пар оснований между сг-РНК и нуклеиновой кислотой-мишенью.

20 В системах типа I в связывание сг-РНК и мишени в случае сг-РНК-эффекторного комплекса вовлечены Cas7, Cas5 и Cas8, слитый с белком малой субъединицы. В расщепление нуклеиновой кислоты-мишени в системах типа I вовлечен нуклеазный домен HD, который либо слит с геликазой Cas3' суперсемейства 2, либо кодируется отдельным геном, cas3".

25 В системах типа III в связывание сг-РНК и мишени в сг-РНК-эффекторном комплексе вовлечены Cas7, Cas5, Cas10 и белок малой субъединицы. В расщепление нуклеиновой кислоты-мишени в системах типа III вовлечено объединенное действие белков Cas7 и Cas10, при этом единичный нуклеазный домен HD слит с Cas10, который. Как полагают, расщепляет одностороннюю ДНК во время интерференции.

30 В системах класса 2 сг-РНК ассоциирует с одним белком и достигает интерференции за счет сочетания нуклеазной активности с РНК-связывающими доменами и образованием пар оснований между сг-РНК и нуклеиновой кислотой-мишенью.

В системах типа II в связывание сг-РНК и мишени вовлечен Cas9, который осуществляет расщепление нуклеиновой кислоты-мишени. В системах типа II каждый
35 из доменов: домен RuvC-подобной нуклеазы (укладка РНК-азы H) и домен нуклеазы HNH (McrA-подобный) Cas9, расщепляют одну из нитей нуклеиновой кислоты-мишени. Расщепляющая активность Cas9 систем типа II также требует гибридизации сг-РНК с trасг-РНК с образованием дуплекса, который облегчает связывание сг-РНК и мишени белком Cas9.

40 В системах типа V в связывание сг-РНК и мишени вовлечен Cpf1, который осуществляет расщепление нуклеиновой кислоты-мишени. В системах типа V домен RuvC-подобной нуклеазы Cpf1 расщепляет обе нити нуклеиновой кислоты-мишени в ступенчатой конфигурации с образованием выступающих 5'-концов, в отличие от тупых концов, образуемых при расщеплении Cas9. Такие выступающие 5-концы могут
45 облегчать встраивание ДНК посредством способов, основанных на связывании негомологичных концов.

Активность в расщеплении Cpf1 в системах типа V также не требует гибридизации сг-РНК с trасг-РНК с образованием дуплекса, напротив в случае сг-РНК систем типа V

используется одна *cr*-РНК, которая имеет структуру стебель-петля, образующую внутренний дуплекс. *Cpf1* связывает *cr*-РНК специфичным для последовательности и структуры образом, узнавая при этом структуру стебель-петля и последовательности, близлежащие к структуре стебель-петля, особенно, нуклеотид, расположенный с 5'-

5 стороны от спейсерных последовательностей, который гибридизуется с нуклеиновой кислотой-мишенью. Такая структура стебель-петля обычно имеет длину в диапазоне от 15 до 19 нуклеотидов. Замены, которые нарушают такой дуплекс стебель-петля, устраняют расщепляющую активность, тогда как другие замены, которые не нарушают дуплекс стебель-петля, не устраняют расщепляющую активность. В системах типа V

10 *cr*-РНК образует структуру стебель-петля на 5'-конце, и последовательность на 3'-конце комплементарна последовательности в нуклеиновой кислоте-мишени.

Другие белки, ассоциированные с *cr*-РНК типа V и связыванием и расщеплением мишени, включают кандидат 1 класса 2 (C2c1) и кандидат 3 класса 2 (C2c3). Белки C2c1 и C2c3 имеют сходную длину с белками *Cas9* и *Cpf1*, в диапазоне приблизительно от

15 1100 аминокислот до приблизительно 1500 аминокислот. Белки C2c1 и C2c3 также содержат *RuvC*-подобные нуклеазные домены и имеют структуру, сходную с *Cpf1*. Белки C2c1 сходны с белками *Cas9* в отношении потребности в *cr*-РНК и *tracr*-РНК для связывания и расщепления мишени, но имеют оптимальную температуру расщепления 50°C. Белки C2c1 нацелены на АТ-богатый РАМ, который сходен с *Cpf1*, находится с

20 5'-стороны последовательности-мишени, смотри, например, Shmakov с соавторами (Molecular Cell; 60(3): 385-397 (2015)).

Кандидат 2 класса 2 (C2c2) не имеет сходства последовательности с другими эффекторными белками CRISPR и поэтому может относиться к предполагаемой системе типа VI. Белки C2c2 имеют два домена HEPN и предположительно обладают РНК-

25 азной активностью, и поэтому могут быть нацелены и могут расщеплять мРНК. Белки C2c2, по-видимому, сходны с белками *Cpf1* в отношении рекрутирования *cr*-РНК для связывания и расщепления мишени, хотя и не требуют *tracr*-РНК. Также подобно *Cpf1*, *cr*-РНК для белков C2c2 образует стабильную шпильку или структуру стебель-петля, которая может способствовать ассоциации с белком C2c2.

В используемом в настоящем изобретении смысле «сайт-специфичный полипептид» относится к отдельному белку или комплексу белков, используемому в системе CRISPR

30 наряду с полинуклеотидами, раскрытыми в настоящем описании. Сайт-специфичный полипептид может содержать один или более нуклеазных доменов. Сайт-специфичный полипептид согласно изобретению может содержать домен HNH или HNH-подобной

35 нуклеазы, домен *RuvC* или *RuvC*-подобной нуклеазы и/или нуклеаз, подобных надсемейству HEPN. HNH- или HNH-подобные домены могут содержать *MscA*-подобную укладку. HNH-домен или HNH-подобный домен могут содержать две антипараллельных β-нити и α-спираль. HNH-домен или HNH-подобный домен могут содержать участок связывания металла (например, участок связывания дивалентного катиона). HNH-домен

40 или HNH-подобный домен могут расщеплять одну нить нуклеиновой кислоты-мишени (например, нить, комплементарную нити-мишени, на которую нацелена *cr*-РНК). Белки, которые содержат HNH-домен или HNH-подобный домен, могут включать эндонуклеазы, колицины, рестрикционные эндонуклеазы, транспозазы и факторы упаковки ДНК.

Сайт-специфичным полипептидом может быть белок *Cas9*, белок *Cpf1*, белок C2c1, белок C2c2, белок C2c3, *Cas3*, *Cas5*, *Cas7*, *Cas8*, *Cas10* или их комплексы, в зависимости

45 от конкретной используемой системы CRISPR. В некоторых вариантах сайт-специфичным полипептидом может быть белок *Cas9* или *Cpf1*. В некоторых вариантах сайт-специфичным полипептидом с пониженной нуклеазной активностью может быть

никаза, т.е., она может быть модифицирована так, чтобы расщеплять одну нить дуплекса нуклеиновой кислоты-мишени. В некоторых вариантах сайт-специфичный полипептид может быть модифицирован так, чтобы он не обладал нуклеазной активностью, т.е., чтобы он не расщеплял никакой нити дуплекса нуклеиновой кислоты-мишени или
 5 никакой одиночной нити нуклеиновой кислоты-мишени. Примеры сайт-специфичных полипептидов с пониженной нуклеазной активностью или не обладающих нуклеазной активностью могут включать Cas9 с модификацией в HNH- и/или RuvC-нуклеазных доменах и Cpf1 с модификацией в RuvC-нуклеазном домене. Не ограничивающие
 10 примеры таких модификаций могут включать D917A, E1006A и D1225A в RuvC-нуклеазном домене Cpf1 *F. novicida* и изменение остатков D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 и/или A987 в Cas9 *S. pyogenes* и соответствующих им аминокислотных остатков в других белках Cpf1 и Cas9.

В некоторых вариантах сайт-специфичный полипептид может быть модифицирован. Такие модификации могут включать введение в сайт-специфичный полипептид или его
 15 слияние с доменом из другого полипептида или замену домена сайт-специфичного полипептида доменом из другого полипептида. Например, модифицированный сайт-специфичный полипептид может содержать первый домен из белка Cas9 или Cpf1 и второй домен из другого белка, отличного от Cas9 или Cpf1. Модификация с включением таких доменов в модифицированные сайт-специфичные полипептиды может придавать
 20 дополнительную активность модифицированным сайт-специфичным полипептидам. Такие активности могут включать нуклеазную активность, метилтрансферазную активность, деметилазную активность, ДНК-репарирующую активность, ДНК-повреждающую активность, дезаминирующую активность, дисмутазную активность, алкилирующую активность, депуринизирующую активность, окислительную активность,
 25 активность в образовании пиримидиновых димеров, ингразную активность, транспозазную активность, рекомбиназную активность, полимеразную активность, лигазную активность, геликазную активность, фотолиазную активность, гликозилазную активность, ацетилтрансферазную активность, деацетилазную активность, киназную активность, фосфатазную активность, убиквитинлигазную активность,
 30 деубиквитинизирующую активность, активность в аденилировании, активность в деаденилировании, активность в сумоилировании, активность в десумоилировании, рибозилирующую активность, дерибозилирующую активность, миристоилирующую активность или демиристоилирующую активность), которая модифицирует полипептид, ассоциированный с нуклеиновой кислотой-мишенью (например, гистон).

В некоторых вариантах сайт-специфичный полипептид может вводить двунитевые разрывы или однонитевые разрывы в последовательности нуклеиновой кислоты (например, геномной ДНК). В некоторых вариантах последовательность нуклеиновой
 35 кислоты может представлять собой нуклеиновую кислоту-мишень. Некоторые сайт-специфичные полипептиды согласно настоящему изобретению могут вводить сайты расщепления с образованием тупых концов, тогда как некоторые варианты дают сайты расщепления, имеющие липкие концы, т.е., выступающие 5'- или 3'-концы. Например, Cpf1 может вводить ступенчатый двунитевой разрыв ДНК с выступающим 5'-концом
 40 длиной примерно 4 или 5 нуклеотидов (н.). Двунитевой разрыв может стимулировать эндогенные пути репарации ДНК в клетке (например, гомологичную рекомбинацию и связывание негомологичных концов (NHEJ) или альтернативное связывание негомологичных концов (A-NHEJ)). NHEJ может репарировать расщепленную нуклеиновую кислоту-мишень без необходимости в наличии гомологичной матрицы. Такой процесс может приводить к делециям в нуклеиновой кислоте-мишени.

Гомологичная рекомбинация (HR) может происходить при наличии гомологичной матрицы. Гомологичная матрица может содержать последовательности, которые гомологичны последовательностям, фланкирующим сайт расщепления нуклеиновой кислоты-мишени. После расщепления нуклеиновой кислоты-мишени сайт-специфичным полипептидом сайт расщепления может быть нарушен (например, сайт может быть недоступным для другого раунда расщепления с использованием нацеленного на нуклеиновую кислоту полинуклеотида и сайт-специфичного полипептида).

В некоторых случаях гомологичная рекомбинация может встраивать экзогенную полинуклеотидную последовательность в сайт расщепления нуклеиновой кислоты-мишени. Экзогенная полинуклеотидная последовательность может быть названа донорным полинуклеотидом или а донорной последовательностью. В некоторых вариантах донорный полинуклеотид, часть донорного полинуклеотида, копия донорного полинуклеотида или часть копии донорного полинуклеотида могут быть встроены в сайт расщепления нуклеиновой кислоты-мишени. Донорный полинуклеотид может представлять собой экзогенную полинуклеотидную последовательность. Донорным полинуклеотидом может быть одонитевая ДНК. Донорным полинуклеотидом может быть двунитевая ДНК. Донорным полинуклеотидом может быть РНК. Донорным полинуклеотидом может быть дуплекс РНК и ДНК. Донорным полинуклеотидом может быть последовательность, которая не встречается в природе в сайте расщепления нуклеиновой кислоты-мишени. В некоторых вариантах модификации нуклеиновой кислоты-мишени вследствие NHEJ и/или HR могут приводить, например, к мутациям, делециям, изменениям, интеграциям, корректировке генов, замене генов, мечению генов, инсерции транспозона, делеции нуклеотидов, разрушению гена и/или мутации гена. Процесс интегрирования ненативной нуклеиновой кислоты(кислот) в геномную ДНК может быть назван «геномной инженерией».

Система CRISPR согласно настоящему изобретению может быть названа «системой CRISPR с ДНК-проводником». Система CRISPR согласно настоящему изобретению может быть запрограммирована на расщепление нуклеиновой кислоты-мишени с использованием двух нацеленных на нуклеиновую кислоту полинуклеотидов («двойного проводника»). В некоторых вариантах система CRISPR с двойным проводником может включать CRISPR-Д(Р)НК (сг-Д(Р)НК) и трансаактивирующую CRISPR-РНК (tracr-РНК), например, один полинуклеотид, содержащий и ДНК и РНК, и второй полинуклеотид, содержащий РНК. В некоторых вариантах система с двойным проводником может включать в себя сг-Д(Р)НК и tracr-Д(Р)НК, например, один полинуклеотид, содержащий и ДНК и РНК, и второй полинуклеотид, содержащий и ДНК и РНК. Элементы сг-Д(Р)НК и tracr-Д(Р)НК или tracr-РНК могут быть соединены областью слияния (например, линкером) и синтезированы в виде одного элемента (например, sg-Д(Р)НК), как показано на фигуре 2 («единичный проводник»).

В используемом в настоящем описании смысле термин «сг-Д(Р)НК» относится к полинуклеотиду, содержащему нацеленную на мишень область и активирующую область, при этом нацеленная на мишень область содержит ДНК или ДНК и РНК, и при этом активирующая область содержит РНК или ДНК, или смесь ДНК и РНК. В некоторых вариантах нацеленная на мишень область расположена выше активирующей области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области. В некоторых вариантах tracr-РНК содержит последовательность, которая комплементарна последовательности в активирующей области сг-Д(Р)НК.

В используемом в настоящем описании смысле термин «tracr-Д(Р)НК» относится к полинуклеотиду, имеющему последовательность, которая комплементарна

последовательности в активирующей области sg-D(P)НК , и при этом полинуклеотид содержит ДНК или смесь ДНК и РНК.

В используемом в настоящем описании смысле термин «нацеленная на мишень область» относится к области полинуклеотида, содержащей ДНК или смесь ДНК и РНК, которая комплементарна последовательности в нуклеиновой кислоте-мишени. В некоторых вариантах нацеленная на мишень область также может содержать другие нуклеиновые кислоты или аналоги нуклеиновых кислот или их сочетания. В некоторых вариантах нацеленная на мишень область может состоять только из ДНК, поскольку такая конфигурация менее вероятно разрушается в клетке-хозяине. В некоторых вариантах такая конфигурация может повышать специфичность узнавания последовательности-мишени и/или снижать возникновение связывания/гибридизации вне мишени.

В используемом в настоящем описании смысле термин «активирующая область» относится к части полинуклеотида, содержащего РНК или ДНК или смесь ДНК и РНК, которая взаимодействует или способна к ассоциации или связыванию с сайт-специфичным полипептидом. В некоторых вариантах активирующая область также может содержать другие нуклеиновые кислоты или аналоги нуклеиновых кислот или их сочетания. В некоторых вариантах активирующая область расположена смежно с нацеленной на мишень областью. В некоторых вариантах активирующая область расположена ниже нацеленной на мишень области. В некоторых вариантах активирующая область расположена выше нацеленной на мишень области.

В используемом в настоящем описании смысле термин « sg-D(P)НК » или «единичный D(P)НК -проводник» относится к полинуклеотиду, содержащему нацеленную на мишень область и активирующую область, при этом нацеленная на мишень область содержит ДНК, РНК или смесь ДНК и РНК, которая комплементарна последовательности в нуклеиновой кислоте-мишени, при этом активирующая область содержит РНК или ДНК или смесь ДНК и РНК, при этом либо нацеленная на мишень область, либо активирующая область, либо обе области содержат, по меньшей мере, один нуклеотид ДНК, и при этом активирующая область имеет последовательности, которые являются самокомплементарными и которые гибридизуются с образованием дуплекса, который может содержать вторичные структуры. Пример единичного D(P)НК -проводника может быть сконструирован из sg-D(P)НК и tracr-D(P)НК или tracr-RНК , при этом sg-D(P)НК и tracr-D(P)НК или sg-D(P)НК и tracr-RНК связывают последовательностью нуклеотидов, которая может представлять собой ДНК, РНК или смесь ДНК и РНК.

В используемом в настоящем описании смысле термин «ниже» относится к точке, которая находится дистально от референтной точки в 3'-направлении вдоль нуклеотидной последовательности. В используемом в настоящем описании смысле термин «выше» относится к точке, которая находится дистально от референтной точки в 5'-направлении вдоль нуклеотидной последовательности.

Полинуклеотид согласно настоящему изобретению, например, sg-D(P)НК , tracr-D(P)НК или единичный D(P)НК -проводник, также может содержать смесь ДНК и других нуклеиновых кислот, например, пептидонуклеиновой кислоты (PNA) или других аналогов нуклеиновых кислот.

Изобретение относится к применению единичных D(P)НК -проводников, sg-D(P)НК , tracr-D(P)НК и/или tracr-RНК и сочетаний полинуклеотидов любой длины, которые описаны в настоящей публикации, которые поддерживают запрограммированное расщепление и/или модификацию нуклеиновой кислоты-мишени сайт-специфичным полипептидом.

На фигуре 1А показаны полинуклеотиды для применения в системе CRISPR типа II. В данном варианте 101 может означать cr-Д(Р)НК, а 102 может означать tracr-Д(Р)НК или tracr-РНК.

На фигуре 1В показаны полинуклеотиды, указанные на фигуре 1А, гибридизованные друг с другом вдоль областей комплементарности. Гибридизация может приводить к созданию вторичных структур, таких как выпетливание 105, нацеленная на мишень область 103, связующее звено 107 и шпильки 108 и 109. На фигуре 1В также показан вариант, содержащий верхнюю область дуплекса 106 и нижнюю область дуплекса 104. Верхняя область дуплекса может содержать верхний стебель. Нижняя область дуплекса может содержать нижний стебель. В некоторых вариантах полинуклеотиды, которые гибридизуются с образованием области 104, могут содержать смесь ДНК и РНК в одной и той же полинуклеотидной нити, например, 102, в области, расположенной ниже нацеленной на мишень области 103. В некоторых вариантах область 104, которая показана на фигуре 1В, может содержать смесь ДНК и РНК в одной и той же полинуклеотидной нити, например, 102. Нуклеотидная последовательность, расположенная непосредственно ниже нацеленной на мишень области, может содержать ДНК и РНК в разных соотношениях. В некоторых вариантах такое соотношение может соответствовать 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% или 100% РНК и диапазонам между указанными значениями. Как описано в настоящей публикации, нуклеотидная последовательность, расположенная ниже (например, область между нацеленной на мишень области 103 и выпетливанием 105, как показано на фигуре 1В) нацеленной на мишень области 103, может содержать смесь ДНК и РНК, как показано в SEQ ID NO: 19-26.

На фигуре 2 показан пример единичного Д(Р)НК-проводника для применения в системе CRISPR типа II. Обращаясь к фигуре 2, видно, что вариант включает в себя нацеленную на мишень область 201, нижнюю область дуплекса 202, верхнюю область дуплекса 203, область слияния 204, вторичную структуру (например, выпетливание) 205, связующее звено 206 и шпильки 207 и 208. Верхняя область дуплекса может содержать верхний стебель. Нижняя область дуплекса может содержать нижний стебель. Некоторые варианты могут содержать активирующую область, содержащую верхнюю область дуплекса и нижнюю область дуплекса. В некоторых вариантах область 202 может содержать смесь ДНК и РНК и располагаться непосредственно ниже нацеленной на мишень области 201. Нуклеотидная последовательность, расположенная непосредственно ниже нацеленной на мишень области, может содержать ДНК и РНК в разных соотношениях. В некоторых вариантах такое соотношение может соответствовать 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% или 100% РНК и диапазонам между указанными значениями. Как описано в настоящей публикации, область нуклеотидов, расположенная ниже (например, область между нацеленной на мишень области 201 и выпетливанием 205, которая показана на фигуре 2) нацеленной на мишень области 201, может содержать смесь ДНК и РНК, как показано в SEQ ID NO: 127-132. В некоторых вариантах область 203 может содержать смесь ДНК и РНК и располагаться ниже нацеленной на мишень области 201. Нуклеотидная последовательность, расположенная ниже нацеленной на мишень области, может содержать ДНК и РНК в разных соотношениях. В некоторых вариантах такое соотношение может соответствовать 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% или 100% РНК и диапазонам между указанными значениями. Как описано в настоящей публикации, область нуклеотидов, расположенная ниже нацеленной на мишень области 201, может

содержать смесь ДНК и РНК, как показано в SEQ ID NO: 44-47 и 129.

В некоторых вариантах активирующая область может содержать, по меньшей мере, одну вторичную структуру. Вторичная структура может представлять собой нижний стебель, верхний стебель, выпетливание, связующее звено, шпильку, одну или более из указанных структур и их сочетания. В некоторых вариантах активирующая область содержит выпетливание. На фигуре 1В показаны вторичные структуры, созданные в системе с двойным проводником, т.е., *cr*-Д(Р)НК, гибридизующейся с *tracr*-Д(Р)НК, или *cr*-Д(Р)НК, гибридизующейся с *tracr*-РНК, включая нижний стебель 104, выпетливание 105, верхний стебель 106, связующее звено 107 и шпильку, например, 108. Вторичные структуры также могут включать дополнительные типы структур. Расположение и количество вторичных структур особым образом не ограничено и может меняться, в зависимости от того, какой сайт-специфичный полипептид используется в системе CRISPR.

В некоторых вариантах активирующая область может содержать область нуклеотидов, содержащую нижний стебель, верхний стебель и выпетливание. В некоторых вариантах может быть только выпетливание. В некоторых вариантах выпетливание может быть между нижним стеблем и верхним стеблем. В некоторых вариантах верхний стебель может быть исключен. Термины «верхний стебель» и «нижний стебель» могут быть использованы в настоящем описании только при ссылке на иллюстрированное положение активирующей области и не обязательно предназначены для ограничения таких областей какой-либо конкретной структурой, вторичной структурой или положением. Например, на фигуре 1В показан нижний стебель, 104, расположенный между выпетливанием и спейсером. В некоторых вариантах нацеленная на мишень область может содержать спейсер.

В некоторых вариантах нуклеотидная последовательность, расположенная ниже нацеленной на мишень области в нижнем стебле может иметь последовательность, которая представляет собой последовательность 5'GYYYUR, где Y означает С или U/Т, и R означает А или G. В некоторых вариантах нуклеотидная последовательность, расположенная ниже нацеленной на мишень области в нижнем стебле, может иметь последовательность, которая представляет собой последовательность 5'GUUUUUGU. В некоторых вариантах нуклеотидная последовательность, расположенная ниже нацеленной на мишень области в нижнем стебле, может иметь последовательность, которая представляет собой последовательность 5'GUUUUA. В некоторых вариантах нуклеотиды в нижнем стебле могут быть нуклеотидами РНК или ДНК или смеси ДНК и РНК.

В некоторых вариантах вторичная структура может содержать выпетливание. Выпетливание может относиться к области неспаренных нуклеотидов в дуплексе. В некоторых вариантах единичный Д(Р)НК-проводник может содержать выпетливание. В некоторых вариантах полинуклеотиды для применения в системе CRISPR могут содержать вторичную структуру, и указанной вторичной структурой является четверная петля. Единичный Д(Р)НК-проводник, содержащий выпетливание, может содержать 5'-сторону и 3'-сторону дуплекса. Обращаясь к фигуре 2, можно видеть, например, 5'-сторона дуплекса может относиться к области, которая расположена выше (т.е., в 5'-направлении) от 204, и 3'-сторона дуплекса может относиться к области, которая расположена ниже (т.е., в 3'-направлении) 204. В некоторых вариантах активирующая область содержит выпетливание. В некоторых вариантах выпетливание может быть вовлечено в связывание или взаимодействие с сайт-специфичным полипептидом. Выпетливание может содержать на одной стороне дуплекса неспаренные 5'-RRRZ-3',

где R означает пурин, а Z может быть нуклеотидом, который может образовывать «качающуюся» пару с нуклеотидом на противоположной нити, и область неспаренных нуклеотидов на другой стороне дуплекса. Выпетливание может содержать ДНК, РНК и их смеси. Выпетливание может содержать ДНК, РНК или их смесь на 5'-стороне
 5 выпетливания дуплекса и может содержать ДНК, РНК или их смесь на 3'-стороне выпетливания. В некоторых вариантах полинуклеотид для применения в системе CRISPR может содержать нацеленную на мишень область и активирующую область, и сторона а нацеленной на мишень области в выпетливании дуплекса может содержать ДНК, РНК и их смеси, и сторона активирующей области в выпетливании дуплекса может содержать
 10 ДНК, РНК и их смеси. Например, в одном варианте сторона выпетливания, которая располагается ближе к 5'-концу полинуклеотида, может содержать РНК, а сторона выпетливания, которая располагается ближе к 3'-концу полинуклеотида, может содержать РНК. В некоторых вариантах одна сторона выпетливания может содержать меньше нуклеотидов, чем другая сторона выпетливания. В некоторых вариантах
 15 полинуклеотид для применения в системе CRISPR содержит полинуклеотид, имеющий 5'-направление и 3'-направление, и содержит выпетливание, имеющее 5'-сторону и 3'-сторону, и 5'-сторона может содержать ДНК и/или РНК, и 3'-сторона может содержать РНК. В некоторых вариантах полинуклеотид для применения в системе CRISPR содержит полинуклеотид, имеющий 5'-направление и 3'-направление, и содержит выпетливание,
 20 имеющее 5'-сторону и 3'-сторону, и 5'-сторона может содержать ДНК и/или РНК, и 3'-сторона может содержать РНК, и 3'-сторона может иметь больше нуклеотидов, чем 5'-сторона указанного выпетливания. В некоторых вариантах полинуклеотиды для применения в системе CRISPR могут содержать *сг*-Д(Р)НК и *tracr*-Д(Р)НК, и *сг*-Д(Р)НК-сторона дуплекса с выпетливанием может содержать ДНК, РНК и их смеси, содержащие
 25 два нуклеотида; и *tracr*-Д(Р)НК-сторона дуплекса с выпетливанием может содержать ДНК, РНК и их смеси. В некоторых вариантах полинуклеотиды для применения в системе CRISPR могут содержать *сг*-Д(Р)НК и *tracr*-РНК, и *сг*-Д(Р)НК-сторона дуплекса с выпетливанием может содержать ДНК, РНК и их смеси, содержащие два нуклеотида; и *tracr*-РНК-сторона дуплекса с выпетливанием может содержать более двух
 30 нуклеотидов.

Например, выпетливание может содержать неспаренный пурин (например, аденин) на стороне выпетливания. В некоторых вариантах выпетливание может содержать неспаренные 5'-AAGZ-3' на стороне выпетливания, при этом Z может означать нуклеотид, который образует «качающуюся» пару с нуклеотидом на другой стороне
 35 выпетливания.

Выпетливание на первой стороне дуплекса (например, на стороне, которая направлена к 5'-концу полинуклеотида для применения в системе CRISPR) может содержать, по меньшей мере, 1, 2, 3, 4 или 5 или больше неспаренных нуклеотидов. Выпетливание на первой стороне дуплекса (например, на стороне, которая направлена
 40 к 5'-концу полинуклеотида для применения в системе CRISPR) может содержать не более 1, 2, 3, 4 или 5 или больше неспаренных нуклеотидов. Выпетливание на первой стороне дуплекса (например, на стороне, которая направлена к 5'-концу полинуклеотида для применения в системе CRISPR) может содержать 1 неспаренный нуклеотид.

Выпетливание на второй стороне дуплекса (например, *tracr*-РНК- или *tracr*-Д(Р)НК-сторона дуплекса) может содержать, по меньшей мере, 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 или больше неспаренных нуклеотидов. Выпетливание на второй стороне дуплекса (например, на *tracr*-РНК- или *tracr*-Д(Р)НК-сторона дуплекса) может содержать 1, 2, 3, 4, 5, 6, 7, 8,
 45 9 или 10 или больше неспаренных нуклеотидов. Выпетливание на второй стороне

дуплекса (например, на *tracr*-РНК- или *tracr*-Д(Р)НК-стороне дуплекса) может содержать 4 неспаренных нуклеотида.

Области с разным количеством неспаренных нуклеотидов на каждой нити дуплекса могут быть спарены вместе. Некоторые варианты могут содержать вторичную структуру, содержащую выпетливание, при этом указанное выпетливание не образует дуплекса. Выпетливание может содержать 5 неспаренных нуклеотидов из первой нити и 1 неспаренный нуклеотид из второй нити. Выпетливание может содержать 4 неспаренных нуклеотида из первой нити и 1 неспаренный нуклеотид из второй нити. Выпетливание может содержать 3 неспаренных нуклеотида из первой нити и 1 неспаренный нуклеотид из второй нити. Выпетливание может содержать 2 неспаренных нуклеотида из первой нити и 1 неспаренный нуклеотид из второй нити. Выпетливание может содержать 1 неспаренный нуклеотид из первой нити и 1 неспаренный нуклеотид из второй нити. Выпетливание может содержать 1 неспаренный нуклеотид из первой нити и 2 неспаренных нуклеотида из второй нити. Выпетливание может содержать 1 неспаренный нуклеотид из первой нити и 3 неспаренных нуклеотида из второй нити. Выпетливание может содержать 1 неспаренный нуклеотид из первой нити и 4 неспаренных нуклеотида из второй нити. Выпетливание может содержать 1 неспаренный нуклеотид из первой нити и 5 неспаренных нуклеотидов из второй нити.

В некоторых вариантах неспаренная вторичная структура может быть образована на *sc*-Д(Р)НК-стороне полинуклеотида. В некоторых вариантах неспаренная вторичная структура может быть образована на *sc*-Д(Р)НК-стороне полинуклеотида и может дополнительно содержать неспаренную вторичную структуру на *tracr*-РНК- или *tracr*-Д(Р)НК-стороне. В таком варианте указанные вторичные структуры могут представлять собой выпетливания. В некоторых вариантах термин «неспаренная» в отношении вторичной структуры может означать, что вторичная структура не имеет форму дуплекса.

В некоторых случаях выпетливание, может содержать, по меньшей мере, одно «качающееся» спаривание. В некоторых случаях выпетливание может содержать не более одного «качающегося» спаривания. Последовательность выпетливания может содержать, по меньшей мере, один пуриновый нуклеотид. Последовательность выпетливания может содержать, по меньшей мере, 3 пуриновых нуклеотида. Последовательность выпетливания может содержать, по меньшей мере, 5 пуриновых нуклеотидов. Последовательность выпетливания может содержать, по меньшей мере, один гуаниновый нуклеотид. Последовательность выпетливания может содержать, по меньшей мере, один адениновый нуклеотид. Последовательность выпетливания может содержать урацил. Вторичная структура может содержать ДНК, РНК и их сочетания. В некоторых вариантах вторичная структура может образовывать структуру дуплекса и указанная структура дуплекса может содержать выпетливание, содержащее ДНК и РНК.

Последовательность *tracr*-Д(Р)НК может иметь длину примерно от 6 нуклеотидов до примерно 150 нуклеотидов. Например, последовательность *tracr*-Д(Р)НК может иметь длину примерно от 6 нуклеотидов (н.) до примерно 50 н., примерно от 6 н. до примерно 40 н., примерно от 6 н. до примерно 30 н., примерно от 6 н. до примерно 25 н., примерно от 6 н. до примерно 20 н., примерно от 6 н. до примерно 15 н., примерно от 8 н. до примерно 40 н., примерно от 8 н. до примерно 30 н., примерно от 8 н. до примерно 25 н., примерно от 8 н. до примерно 20 н. или примерно от 8 н. до примерно 15 н., примерно от 15 н. до примерно 150 н., примерно от 15 н. до примерно 130 н., примерно от 15 н. до примерно 100 н., примерно от 15 н. до примерно 80 н., примерно

от 15 н. до примерно 50 н., примерно от 15 н. до примерно 40 н., примерно от 15 н. до примерно 30 н. или примерно от 15 н. до примерно 25 н.. В некоторых вариантах последовательность *tracr*-Д(Р)НК имеет длину, составляющую приблизительно 14 нуклеотидов. В некоторых вариантах *tracr*-Д(Р)НК состоит только из ДНК.

5 последовательность *tracr*-Д(Р)НК может быть, по меньшей мере, примерно на 60% идентичной эталонной последовательности *tracr*-РНК (например, последовательности *tracr*-РНК дикого типа из *S. pyogenes*) на протяжении участка длиной, по меньшей мере, 6, 7, или 8 непрерывно следующих друг за другом нуклеотидов. Например, последовательность *tracr*-Д(Р)НК может быть, по меньшей мере, примерно на 60%
10 идентичной, по меньшей мере, примерно на 65% идентичной, по меньшей мере, примерно на 70% идентичной, по меньшей мере, примерно на 75% идентичной, по меньшей мере, примерно на 80% идентичной, по меньшей мере, примерно на 85% идентичной, по меньшей мере, примерно на 90% идентичной, по меньшей мере, примерно на 95% идентичной, по меньшей мере, примерно на 98% идентичной, по меньшей мере, примерно на 99% идентичной или на 100% идентичной эталонной последовательности *tracr*-РНК
15 (например, последовательности *tracr*-РНК дикого типа из *S. pyogenes*) на протяжении участка длиной, по меньшей мере, 6, 7 или 8 непрерывно следующих друг за другом нуклеотидов.

Последовательность *tracr*-Д(Р)НК может содержать более одной дуплексной области
20 (например, шпильку, гибридизованную область). Последовательность *tracr*-Д(Р)НК может содержать две дуплексные области. *tracr*-Д(Р)НК может содержать вторичную структуру. *tracr*-Д(Р)НК может содержать более одной вторичной структуры. В некоторых вариантах последовательность *tracr*-Д(Р)НК может содержать первую вторичную структуру и вторую вторичную структуру, и первая вторичная структура
25 содержит больше нуклеотидов, чем вторая вторичная структура. В некоторых вариантах *tracr*-Д(Р)НК может содержать первую вторичную структуру, вторую вторичную структуру и третью вторичную структуру, и указанная первая вторичная структура содержит меньше нуклеотидов, чем указанная вторая вторичная структура, и указанная
30 вторая вторичная структура содержит больше нуклеотидов, чем указанная третья вторичная структура. Количество вторичных структур и соответствующая длины в нуклеотидах особым образом не ограничены.

Последовательность *tracr*-РНК может иметь длину примерно от 6 нуклеотидов до примерно 150 нуклеотидов. Например, последовательность *tracr*-РНК может иметь длину примерно от 6 н. до примерно 50 н., примерно от 6 н. до примерно 40 н., примерно
35 от 6 н. до примерно 30 н., примерно от 6 н. до примерно 25 н., примерно от 6 н. до примерно 20 н., примерно от 6 н. до примерно 15 н., примерно от 8 н. до примерно 40 н., примерно от 8 н. до примерно 30 н., примерно от 8 н. до примерно 25 н., примерно от 8 н. до примерно 20 н. или примерно от 8 н. до примерно 15 н., примерно от 15 н. до примерно 150 н., примерно от 15 н. до примерно 130 н., примерно от 15 н. до примерно
40 100 н., примерно от 15 н. до примерно 80 н., примерно от 15 н. до примерно 50 н., примерно от 15 н. до примерно 40 н., примерно от 15 н. до примерно 30 н. или примерно от 15 н. до примерно 25 н.. В некоторых вариантах последовательность *tracr*-РНК имеет длину приблизительно 14 нуклеотидов. Последовательность *tracr*-РНК может быть, по меньшей мере, примерно на 60% идентичной эталонной последовательности *tracr*-РНК
45 (например, последовательности *tracr*-РНК дикого типа из *S. pyogenes*) на протяжении участка длиной, по меньшей мере, 6, 7 или 8 непрерывно следующих друг за другом нуклеотидов. Например, последовательность *tracr*-РНК может быть, по меньшей мере, примерно на 60% идентичной, по меньшей мере, примерно на 65% идентичной, по

меньшей мере, примерно на 70% идентичной, по меньшей мере, примерно на 75% идентичной, по меньшей мере, примерно на 80% идентичной, по меньшей мере, примерно на 85% идентичной, по меньшей мере, примерно на 90% идентичной, по меньшей мере, примерно на 95% идентичной, по меньшей мере, примерно на 98% идентичной, по меньшей мере, примерно на 99% идентичной или на 100% идентичной эталонной последовательности tracr-РНК (например, последовательности tracr-РНК дикого типа из *S. pyogenes*) на протяжении участка длиной, по меньшей мере, 6, 7 или 8 непрерывно следующих друг за другом нуклеотидов.

Последовательность tracr-РНК может содержать более одной дуплексной области (например, шпильки, гибридной области). Последовательность tracr-РНК может содержать две дуплексных области. tracr-РНК может содержать вторичную структуру. tracr-РНК может содержать более одной вторичной структуры. В некоторых вариантах последовательность tracr-РНК может содержать первую вторичную структуру и вторую вторичную структуру, и первая вторичная структура содержит больше нуклеотидов, чем вторая вторичная структура. В некоторых вариантах tracr-РНК может содержать первую вторичную структуру, вторую вторичную структуру и третью вторичную структуру, и указанная первая вторичная структура содержит меньше нуклеотидов, чем указанная вторая вторичная структура, и указанная вторая вторичная структура содержит больше нуклеотидов, чем указанная третья вторичная структура. Количество вторичных структур и соответствующие длины в нуклеотидах особым образом не ограничены.

Встречающиеся в природе системы CRISPR типа V, в отличие от систем CRISPR типа II, не требуют tracr-РНК для созревания cr-РНК и расщепления нуклеиновой кислоты-мишени. На фигуре 9 показана типичная структура cr-РНК из системы CRISPR типа V, в которой связывающая ДНК-мишень последовательность расположена ниже структуры стебель-петля, которая взаимодействует с белком Cpf1. Изменения нуклеотидов в области петли не влияют на расщепляющую активность Cpf1.

На фигурах 10А-С показаны возможные структуры единичного Д(Р)НК-проводника согласно настоящему изобретению для применения в системе CRISPR типа V. В указанных конфигурациях сплошные черные области представляют РНК, тогда как окрашенные в клеточку области представляют ДНК. На фигуре 10А показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит РНК, 3'-стебель содержит ДНК и петля и 5'-стебель содержат РНК. На фигуре 10В показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит РНК, 5'-стебель содержит ДНК и петля и 3'-стебель содержат РНК. На фигуре 10С показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область и петля содержат РНК, и 5'- и 3'-стебли содержат ДНК. 3'-Стебель и 5'-стебель на фигурах 10А-С вместе или по-отдельности могут быть названы в настоящем описании «активирующей областью» полинуклеотида для применения с использованием системы типа V.

На фигурах 11А-Е показаны возможные структуры единичного Д(Р)НК-проводника согласно настоящему изобретению для применения с использованием системы CRISPR типа V. В таких конфигурациях сплошные черные области представляют ДНК, тогда как окрашенные в клеточку области представляют РНК. На фигуре 11А показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит ДНК, 3'-стебель содержит ДНК и петля и 5'-стебель содержат РНК. На фигуре 11В показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит ДНК, 5'-стебель содержит ДНК и петля и 3'-стебель содержат РНК. На фигуре 11С показан единичный Д(Р)НК-проводник, в котором нацеленная на мишень область,

5'-стебель и 3'-стебель содержат ДНК, и петля содержит РНК. На фигуре 11D показан единственный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит ДНК, и 5'-стебель, 3'-стебель и петля содержат ДНК. На фигуре 11E показан единственный Д(Р)НК-проводник, в котором нацеленная на мишень область содержит смесь ДНК и РНК, и 5'-стебель, 3'-стебель и петля содержат ДНК. 3'-Стебель и 5'-стебель на фигурах 11A-E вместе или по-отдельности могут быть названы в настоящем описании «активирующей областью» полинуклеотида для применения с использованием системы типа V.

На фигурах 12A-I показаны возможные конфигурации cr-РНК и cr-Д(Р)НК согласно настоящему изобретению для применения с использованием системы CRISPR типа V, в которых 3'-элемент и 5'-элемент представляют собой отдельные полинуклеотиды и ассоциируют за счет водородных взаимодействий между основаниями в парах с образованием структуры дуплекса или стебля. На фигуре 12A показана система с двойным проводником для применения в системе CRISPR типа V, в которой нацеленная на мишень область связана с 3'-элементом. Второй полинуклеотид также показан на фигуре 12A в виде 5'-элемента. 5'-элемент имеет конфигурацию, позволяющую гибридизоваться с 3'-элементом, который связан с нацеленной на мишень областью с образованием дуплекса или стебля. На фигуре 12A нацеленная на мишень область, 3'-элемент и 5'-элемент содержат РНК. На фигуре 12B показан 5'-элемент, который содержит РНК. На фигуре 12C показан 5'-элемент, который содержит ДНК. На фигуре 12D показана нацеленная на мишень область, которая содержит РНК, и 3'-элемент, который содержит РНК. На фигуре 12E показана нацеленная на мишень область, которая содержит РНК, и 3'-элемент, который содержит ДНК. На фигуре 12F показана нацеленная на мишень область, которая содержит ДНК, и 3'-элемент, который содержит РНК. На фигуре 12G показана нацеленная на мишень область, которая содержит ДНК, и 3'-элемент, который содержит ДНК. На фигуре 12H показана нацеленная на мишень область, которая содержит РНК и ДНК, и 3'-элемент, который содержит ДНК. На фигуре 12I показана нацеленная на мишень область, которая содержит альтернативную смесь РНК и ДНК, и 3'-элемент, который содержит ДНК. 3'-Элемент, показанный на фигурах 12A-I, может быть назван в настоящем описании «активирующей областью» полинуклеотида для применения с использованием системы типа V.

На фигурах 13A-N показаны возможные конфигурации cr-РНК и cr-Д(Р)НК согласно настоящему изобретению для применения с использованием системы CRISPR типа V, в которых 3'-элемент и 5'-элемент являются отдельными полинуклеотидами и ассоциируют за счет водородных взаимодействий между основаниями в паре с образованием структуры дуплекса или стебля. В некоторых вариантах полинуклеотидов, показанных на фигурах 10A-13H, области ДНК также могут содержать РНК. В некоторых вариантах области РНК также могут содержать ДНК. В некоторых вариантах области ДНК также могут содержать РНК, и области РНК также могут содержать ДНК. 3'-Элемент, показанные на фигурах 13A-N, может быть назван в настоящем описании «активирующей областью» полинуклеотида для применения с использованием системы типа V. Соотношения ДНК и РНК в различных областях полинуклеотидов, показанных на фигурах 10A-13H, могут варьировать. В некоторых вариантах такое соотношение может соответствовать 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% или 100% РНК и диапазонам между указанными значениями. Примеры полинуклеотидов, которые можно применять с использованием системы CRISPR типа V, представлены в виде последовательностей SEQ ID NO: 168-203.

Активирующая область полинуклеотида, нацеленного на определенную нуклеиновую кислоту, может взаимодействовать с областью сайт-специфичного полипептида. Активирующая область может взаимодействовать с множеством областей сайт-специфичного полипептида. Активирующая область может взаимодействовать с
 5 множеством областей сайт-специфичного полипептида, при этом, по меньшей мере, одна из областей взаимодействует с РАМ нуклеиновой кислоты-мишень. Примеры таких областей могут включать аминокислоты 1096-1225 и 1105-1138 Cas9 *S. pyogenes*.

Нуклеотиды, лежащие рядом с неспаренным нуклеотидом, могут представлять собой нуклеотид, который взаимодействует с образованием «качающейся» пары оснований.

10 Взаимодействия с образованием «качающейся» пары оснований может включать взаимодействия гуанин-урацил, гипоксантин-урацил, гипоксантин-аденин и гипоксантин-цитозин. Взаимодействия с образованием «качающейся» пары оснований могут приводит к пониженной специфичности узнавания и/или расщепления мишени. По меньшей мере, 1, 2, 3, 4 или 5 или больше нуклеотидов, близлежащих к неспаренному нуклеотиду,
 15 могут формировать «качающуюся» пару. Не более 1, 2, 3, 4 или 5 или больше нуклеотидов, близлежащих к неспаренному нуклеотиду, могут формировать «качающуюся» пару. В некоторых вариантах нацеленная на мишень область может содержать дезоксирибонуклеотид тимин («dT») в качестве замены рибонуклеотиду урацилу. Использование dT вместо U уменьшает образование «качающихся» пар и
 20 снижает спаривание оснований вне мишени, таким образом приводя в некоторых вариантах к повышенной специфичности по отношению к мишени.

Нуклеиновая кислота-мишень может состоять из ДНК, РНК или их сочетаний, и может быть двунитевой нуклеиновой кислотой или одонитевой нуклеиновой кислотой.

Последовательность нацеленной на мишень области может гибридизоваться с
 25 нуклеиновой кислотой-мишенью, которая расположена с 5'- или 3'-стороны от близлежащего к протоспейсеру мотива (РАМ), в зависимости от конкретного используемого сайт-специфичного полипептида. РАМ может варьировать в зависимости от используемого сайт-специфичного полипептида. Например, при использовании Cas9 из *S. pyogenes* РАМ может представлять собой последовательность в нуклеиновой
 30 кислоте-мишени, которая содержит последовательность 5'-NRR-3', где R может означать либо А, либо G, где N означает любой нуклеотид, и N находится непосредственно с 3'-стороны от последовательности нуклеиновой кислоты-мишени, на которую нацелена последовательность нацеленной на мишень области. Сайт-специфичный полипептид может быть модифицирован, так что РАМ может другим по сравнению с РАМ для
 35 немодифицированного сайт-специфичного полипептида. Например, при использовании Cas9 из *S. pyogenes* Cas9 может быть модифицирован так, чтобы РАМ больше не содержал последовательности 5'-NRR-3', а вместо этого содержал последовательность 5'-NNR-3', где R может означать либо А, либо G, где N означает любой нуклеотид, и N находится непосредственно с 3'-стороны от последовательности нуклеиновой кислоты-мишени,
 40 на которую нацелена последовательность нацеленной на мишень области. Другие сайт-специфичные полипептиды могут узнавать другие РАМ, и специалист в данной области сможет определить РАМ для любого конкретного сайт-специфичного полипептида.

Например, Cpf1 из *Francisella novicida* идентифицировали как имеющий 5'-TTN-3' РАМ (Zetsche с соавторами (Cell; 163(3): 759-71(2015))), но он был неспособен поддерживать
 45 сайт-специфичное расщепление нуклеиновой кислоты-мишени *in vivo*. С учетом сходства последовательности-проводника между Cpf1 *Francisella novicida* и другими белками Cpf1, такими как Cpf1 из *Acidaminococcus* sp. BV3L6, которые используют 5'-TTTN-3' РАМ, более вероятно, что белок Cpf1 *Francisella novicida* узнает и расщепляет участок

в нуклеиновой кислоте-мишени, расположенный проксимально к 5'-TTTN-3' РАМ, с более высокой специфичностью и активностью, чем участок в нуклеиновой кислоте-мишени, расположенный проксимально к укороченному 5'-TTN-3' РАМ, ошибочно идентифицированному Zetsche с соавторами. Полинуклеотиды и системы CRISPR, описанные в настоящей заявке, можно применять с белком Cpf1 (например, из *Francisella novicida*), направленным к участку в нуклеиновой кислоте-мишени, расположенному проксимально по отношению к 5'-TTTN-3' РАМ.

Последовательность нуклеиновой кислоты-мишени может иметь длину 20 нуклеотидов. Нуклеиновая кислота-мишень может иметь длину менее 20 нуклеотидов. Нуклеиновая кислота-мишень может иметь длину, по меньшей мере, 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 или больше нуклеотидов. Нуклеотидная последовательность-мишень может иметь длину в диапазоне примерно 5-30 нуклеотидов и значения в диапазоне между указанными цифрами. Например, в последовательности, содержащей 5'-NNNNNNNNNNNNNNNNNNNNXRR-3' нуклеиновая кислота-мишень может представлять собой последовательность, которая соответствует нескольким N, где N означает любой нуклеотид, и где X означает первый нуклеотид РАМ, узнаваемый *S. pyogenes*. Специалисты в данной области могут отобрать конкретные РАМ на основе конкретного сайт-специфичного полипептида, применяемого в данном случае.

Полинуклеотиды согласно настоящему изобретению, содержащие ДНК и РНК в одной нити, не могут быть получены *in vivo* с использованием экспрессирующих векторов, но могут быть синтезированы химически *in vitro*. Химический синтез полинуклеотидов хорошо известен специалисту в данной области. Химический синтез полинуклеотидов согласно настоящему изобретению можно провести в растворе или на твердой подложке. Синтез в растворе является предпочтительным для получения больших количеств и для более высокой чистоты полинуклеотидов, так как промежуточные продукты очищают после каждой стадии. В случае получения меньших количеств, когда чистота последовательности не является такой критичной, предпочтительным способом является твердофазный синтез. Полинуклеотиды согласно настоящему изобретению также могут быть получены из коммерческих источников, которые обеспечивают автоматизированный химический синтез полинуклеотидов.

Химический синтез ДНК может быть более простым, быстрым и дешевым, чем химический синтез РНК. Создание и тестирование полинуклеотидов, содержащих ДНК, может быть более быстрым и рентабельным по сравнению с РНК-содержащими последовательностями. Последовательности, содержащие ДНК, могут обеспечивать преимущество, связанное с повышенной специфичностью таргетирования нуклеиновых кислот-мишеней, таких как ДНК. Полинуклеотиды, содержащие ДНК в конкретных областях, которые обсуждаются в настоящем описании, могут обеспечивать дополнительное преимущество, связанное со снижением связывания вне мишени вследствие уменьшения склонности к образованию «качающихся» пар, ассоциированному с основаниями дезоксирибонуклеиновой кислоты по сравнению с основаниями рибонуклеиновой кислоты (например, основаниями тимидина в ДНК по сравнению с основаниями урацила в РНК).

В некоторых вариантах полинуклеотиды согласно настоящему изобретению также могут содержать модификации, которые, например, повышают стабильность полинуклеотида. Такие модификации могут включать фосфоротиоаты, хиральные фосфоротиоаты, фосфородитиоаты, сложные фосфотриэфиры, сложные аминокислотфосфотриэфиры, метил- и другие алкилфосфонаты, такие как 3'-алкиленфосфонаты, 5'-алкиленфосфонаты, хиральные фосфонаты, фосфинаты,

фосфорамидаты, включая 3'-аминофосфорамидат и аминоклфосфорамидаты, фосфородиамидаты, тионофосфорамидаты, тионоклфосфонаты, сложные тионоклфосфотриэфиры, селенофосфонаты и боранофосфаты, имеющие нормальные 3'-5'-связи, 2'-5'-связанные аналоги и полинуклеотиды, имеющие обратную полярность, при этом одна или более межинуклеотидные связи представляют собой 3'-3', 5'-5' или 2'-2'-связь. Подходящие нацеленные на нуклеиновую кислоту полинуклеотиды, имеющие обратную полярность, могут содержать одну 3'-3'-связь в межинуклеотидной связи, наиболее близкой к 3'-концу (т.е., один инвертированный нуклеозидный остаток, в котором основание нуклеотида отсутствует или вместо него имеется гидроксильная группа). Также могут быть включены различные соли (например, хлорид калия или хлорид натрия, смешанные соли и формы свободной кислоты).

В некоторых вариантах полинуклеотиды согласно настоящему изобретению также могут содержать другие нуклеиновые кислоты или аналоги нуклеиновых кислот.

Примером аналога нуклеиновой кислоты является пептидонуклеиновая кислота (PNA).

Доставка полинуклеотидов согласно настоящему изобретению в клетки *in vitro* или *in vivo* может быть осуществлена различными способами, известными специалисту в данной области. Такие способы включают липофекцию, электропорацию, нуклеофекцию, микроинъекцию, биобаллистику, липосомы, иммунолипосомы, поликатион или конъюгаты липид:нуклеиновая кислота. Липофекция хорошо известна и описана, например, в патентах США № 5049386, 4946787 и 4897355; и реагенты для липофекции имеются в коммерческой продаже. Катионные и нейтральные липиды, которые подходят для эффективной липофекции полинуклеотидов на основе узнавания рецепторов, описаны в международных публикациях № WO 91/17424 и WO 91/16024.

Комплексы липид:нуклеиновая кислота, включая нацеленные на мишень липосомы, такие как иммунолипидные комплексы, и получение таких комплексов хорошо известны специалисту в данной области (смотри, например, Crystal, Science 270: 404-410 (1995); Blaese et al., Cancer Gene Ther. 2: 291-297 (1995); Behr et al., Bioconjugate Chem. 5: 382-389 (1994); Remy et al., Bioconjugate Chem. 5: 647-654 (1994); Gao et al., Gene Therapy 2: 710-722 (1995); Ahmad et al., Cancer Res. 52: 4817-4820 (1992); патенты США № 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028 и 4946787).

Электропорация может быть использована для доставки полинуклеотидов согласно настоящему изобретению. Электропорация также может быть использована для доставки комплексов сайт-специфичного полипептида и полинуклеотидов согласно настоящему изобретению. В таких способах полинуклеотиды или комплексы сайт-специфичных полипептидов и полинуклеотидов смешивают в буфере для электропорации с клетками-мишенями с получением суспензии. Затем такую суспензию подвергают воздействию электрического импульса при оптимизированном напряжении, который создает временные поры в фосфолипидном бислое клеточной мембраны, позволяя заряженным молекулам, подобным ДНК и белкам, проходить через поры и проникать в клетку. Реагенты и оборудование для осуществления электропорации продаются коммерчески.

Биобаллистика или доставка микрочастиц может быть использована для доставки полинуклеотидов согласно настоящему изобретению. В таких способах микрочастицы, такие как частицы золота или вольфрама, покрывают полинуклеотидом посредством преципитации с использованием хлорида кальция, спермидина или полиэтиленгликоля. Микрочастицам придают ускорение для проникновения в клетку с высокой скоростью, используя такое устройство, как система для доставки частиц BIOLISTIC® PDS-1000/He (Bio-Rad; Hercules, California).

В некоторых вариантах настоящее изобретение относится к способам модификации гена-мишени в клетке. Клетка может быть из любого организма (например, бактериальная клетка, клетка архей, клетка одноклеточного эукариотического организма, растительная клетка, клетка водоросли, клетка гриба (например, дрожжевая клетка), клетка беспозвоночного животного, клетка позвоночного животного или клетка млекопитающего, включая клетку человека.

В некоторых вариантах настоящее изобретение относится к способам модификации гена-мишени в растении. В используемом в настоящем описании смысле термин «растение» относится к целым растениям, органам растений, тканям растений, семенам, растительным клеткам, семенам и их потомству. Растительные клетки включают без ограничения клетки семян, суспензионные культуры, зародыши, области меристемы, каллусную ткань, листья, корни, побеги, гаметофиты, спорофиты, пыльцу и микроспоры. Растительные клетки включают дифференцированные и недифференцированные ткани, включая без ограничения корни, стебли, побеги, листья, пыльцу, семена, опухолевую ткань и различные формы клеток и культуру (например, отдельные клетки, протопласты, зародыши и каллусную ткань).

Следующие примеры не предназначены для ограничения объема, который авторы изобретения расценивают как различные аспекты настоящего изобретения.

Пример 1

Получение компонентов РНК-проводника

РНК-проводники (например, sg-РНК и tracr-РНК) получали в результате транскрипции *in vitro* (например, набор для быстрого синтеза РНК с высоким выходом T7, New England Biolabs, Ipswich, MA) на двунитевой ДНК-матрице, содержащей промотор T7 на 5'-конце последовательностей ДНК.

Двунитевая ДНК-матрица для компонентов РНК была собрана в ПЦР с использованием перекрывающихся 3'-праймеров, содержащих последовательности ДНК, соответствующие компонентам РНК. Олигонуклеотиды, используемые для сборки, представлены в таблице 1.

Таблица 1

Перекрывающиеся праймеры для создания матриц РНК-проводников

Тип РНК-проводника	Мишень для ДНК-связывающей последовательности	SEQ ID NO
Sg-РНК-AAVS	AAVS-1 (сайт 1 интеграции аденоассоциированного вируса 1 - геном человека)	SEQ ID NO: 63, 64, 65, 66, 67
tracr-РНК	нет данных	SEQ ID NO: 63, 71, 72, 73, 74

Олигонуклеотидные последовательности для синтеза (например, последовательности праймеров, показанные в SEQ ID NO: 63-122) были предоставлены коммерческими производителями (Integrated DNA Technologies, Coralville, IA; или Eurofins, Luxembourg).

Каждый из ДНК-праймеров присутствовал в концентрации 2 нМ. Два внешних ДНК-праймера, соответствующих промотору T7 (прямой праймер: SEQ ID NO: 63, таблица 1) и 3'-концу последовательности РНК (обратные праймеры: SEQ ID NO 67 и 74, таблица 1) использовали в концентрации 640 нМ для протекания реакции амплификации. ПЦР-реакции осуществляли, используя смесь Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA), следуя инструкциям производителя. Реакции сборки в ПЦР осуществляли, используя следующие циклические температурные условия: 98°C в течение 2 минут, 35 циклов по 15 секунд при 98°C, 15 секунд при 62°C, 15 секунд при 72°C и конечное удлинение при 72°C в течение 2 минут. Качество ДНК оценивали электрофорезом в агарозном геле (1,5%, SYBR® Safe, Life Technologies, Grand Island, NY).

От 0,25 до 0,5 мкг ДНК-матрицы для компонентов РНК-проводника транскрибировали, используя набор для синтеза РНК с высоким выходом Т7 (New England Biolabs, Ipswich, MA), в течение ~16 часов при 37°C. Реакционные смеси для транскрипции обрабатывали ДНК-азой I (New England Biolabs, Ipswich, MA) и очищали, используя набор для очистки и концентрирования РНК GeneJet (Life Technologies, Grand Island, NY). Выход РНК количественно оценивали, используя систему NanodropTM 2000 (Thermo Scientific, Wilmington, DE). Качество транскрибированной РНК оценивали, используя электрофорез в агарозном геле (2%, SYBR® Safe, Life Technologies, Grand Island, NY). Последовательности компонентов РНК-проводника показаны в таблице 2.

Таблица 2
Последовательности РНК-проводника

Название	Последовательность (основания РНК указаны в скобках)	SEQ ID NO:
AAVS1 sg-РНК	5' - [G][G][G][G][C][C][A][C][U][A] [G][G][G][A][C][A][G][G][A][U] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G][G][A][A][A][C][A][G][G][A][C][A][G][C][A][U][A][G][C] [A][A][G][U][U][G][A][G][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][G][U] [G][G][C][A][C][G][A][G][U] [C][G][G][U][G][C][U][U][U][U][U]-3'	SEQ ID NO: 1
tracr-РНК	5' - [G][C][A][G][G][A][C][A][G][C] [A][U][A][G][C][A][A][G][U][U] [G][A][G][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] [C][G][A][G][U][C][G][G][U][G][C][U][U]-3'	SEQ ID NO: 2

Описанный выше способ получения компонентов РНК-проводника может быть применен для получения других компонентов РНК, которые описаны в настоящей публикации.

Пример 2

Получение областей-мишеней двунитевой ДНК для применения в анализах расщепления Cas9.

Двунитевую ДНК-мишень для применения в анализах расщепления Cas in vitro получали, используя ПЦР-амплификацию области-мишени из геномной ДНК.

Двунитевые области ДНК-мишени (например, AAVS-1) для биохимических анализов амплифицировали в ПЦР, используя полученную с использованием фенола-хлороформа геномную ДНК (гДНК) клеток человека линии K562 (ATCC, Manassas, VA). ПЦР-реакции осуществляли, используя смесь Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA), следуя инструкциям производителя. 20 нг/мкл гДНК в конечном объеме 25 мкл использовали для амплификации выбранной области-мишени в следующих условиях: 98°C в течение 2 минут, 35 циклов по 20 секунд при 98°C, 20 секунд при 60°C, 20 секунд при 72°C и конечной удлинением при 72°C в течение 2 минут. ПЦР-продукты очищали, используя пробирки для очистки продуктов ПЦР Spin SmartTM (Denville Scientific, South Plainfield, NJ) и количественно оценивали, используя спектрофотометр NanodropTM 2000 UV-Vis (Thermo Scientific, Wilmington, DE).

Использовали следующие прямые и обратные праймеры для амплификации выбранных последовательностей-мишеней из гДНК. Праймеры, размер ампликонов и размеры фрагментов, создаваемых в результате опосредованного Cas9 расщепления показаны в таблице 3.

Таблица 3
Двунитевые ДНК-мишени

Двунитевая мишень	Размер ампликона	Размеры фрагментов расщепления	SEQ ID NO:
-------------------	------------------	--------------------------------	------------

AAVS-1-мишень 1	495 п.н.	316 п.н./179 п.н.	SEQ ID NO: 75, 76
EMX1-мишень 1	282 п.н.	153 п.н./129 п.н.	SEQ ID NO: 77, 78
VEGFA-мишень 1	276 п.н.	112 п.н./164 п.н.	SEQ ID NO: 79, 80
CD34-мишень 1	282 п.н.	111 п.н./171 п.н.	SEQ ID NO: 81, 82
CD34-мишень 2	268 п.н.	108 п.н./160 п.н.	SEQ ID NO: 83, 84
STAT5a-мишень 1	288 п.н.	152 п.н./136 п.н.	SEQ ID NO: 85, 86
STAT5a-мишень 2	242 п.н.	103 п.н./139 п.н.	SEQ ID NO: 87, 88
JAK1-мишень 1	310 п.н.	179 п.н./131 п.н.	SEQ ID NO: 89, 90
JAK1-мишень 2	310 п.н.	178 п.н./132 п.н.	SEQ ID NO: 91, 92

Другие подходящие области двунитевой ДНК-мишени получали, используя по существу такой же способ. Для областей-мишеней организмов, отличных от человека, использовали геномную ДНК выбранного организма (например, растения, бактерии, дрожжей, водоросли) вместо ДНК, полученной из клеток человека. Кроме того, можно использовать другие источники полинуклеотидов, отличные от геномной ДНК (например, векторы и выделенные из геля фрагменты ДНК).

Пример 3

Анализ расщепления Cas9.

Данный пример иллюстрирует применение cr-Д(Р)НК согласно настоящему изобретению в анализах расщепления Cas9 *in vitro* с целью оценки и сравнения расщепления выбранных комплексов cr-Д(Р)НК/tracr-РНК/белок Cas9 в процентах по отношению к выбранным последовательностям двунитевой ДНК-мишени.

Активность в расщеплении определяли для коллекции вариантов cr-Д(Р)НК (SEQ ID NO: 38-62) против двунитевой ДНК-мишени (AAVS-1; пример 2, таблица 3).

Каждую sg-РНК, cr-ДНК или cr-Д(Р)НК смешивали с tracr-РНК (в соответствующем случае) в эквимольных количествах в буфере для отжига (1,25 mM HEPES, 0,625 mM MgCl₂, 9,375 mM KCl при pH 7,5), инкубировали в течение 2 минут при 95°C, извлекали из термоциклера и оставляли для уравнивания до комнатной температуры.

Sg-РНК, cr-ДНК/tracr-РНК и cr-Д(Р)НК/tracr-РНК добавляли к реакционной смеси с Cas9. Реакционная смесь для Cas9 содержала белок Cas9, разбавленный до конечной концентрации 200 мкМ в буфере для реакции (20 mM HEPES, 100 mM KCl, 5 mM MgCl₂, 1 mM DTT и 5% глицерина при pH 7,4). В реакционной смеси конечная концентрация каждой cr-Д(Р)НК/tracr-РНК составляла 500 нМ в каждой реакционной смеси. Каждую реакционную смесь инкубировали при 37°C в течение 10 минут. Реакцию расщепления инициировали добавлением ДНК-мишени до конечной концентрации 15 нМ. Образцы перемешивали и центрифугировали, затем инкубировали в течение 15 минут при 37°C. Реакции расщепления останавливали добавлением протеиназы K (Denville Scientific, South Plainfield, NJ) в конечной концентрации 0,2 мкг/мл и раствора 0,44 мг/мкл РНК-азы А (SigmaAldrich, St. Louis, MO).

Образцы инкубировали в течение 25 минут при 37°C и в течение 25 минут при 55°C. 12 мкл суммарной реакционной смеси оценивали в отношении активности в расщеплении, используя электрофорез в агарозном геле (2%, SYBR® Gold, Life Technologies, Grand Island, NY). В случае двунитевой ДНК-мишени AAVS-1 появление полос ДНК размером ~316 п.н. и ~179 п.н. свидетельствовало о том, что произошло расщепление ДНК-мишени. Расщепление в процентах рассчитывали, используя значения площади под кривой, которые вычисляли с помощью FIJI (ImageJ; программа обработки изображений на языке Java с открытым исходным кодом) для каждого фрагмента расщепления и ДНК-мишени, и деля сумму фрагментов расщепления на сумму фрагментов расщепления и ДНК-мишени.

На фигуре 3 представлены результаты анализа расщепления Cas9 с использованием

двунитевой ДНК-мишени AAVS-1 для sg-РНК, cr-ДНК/tracr-РНК и cr-Д(Р)НК/tracr-РНК. Сверху на каждой панели указан номер дорожки, соответствующий используемому компоненту РНК-проводника, номера последовательностей SEQ ID NO, соответствующие каждому компоненту, показаны в таблице 4.

Таблица 4
AAVS-1 cr-Д(Р)НК

Дорожка	SEQ ID NO:
1	ДНК-маркер
2	без контрольного проводника
3	SEQ ID NO: 37
4	SEQ ID NO: 38
5	SEQ ID NO: 39
6	SEQ ID NO: 40
7	SEQ ID NO: 41
8	SEQ ID NO: 42
9	ДНК-маркер
10	ДНК-маркер
11	без контрольного проводника
12	SEQ ID NO: 1
13	SEQ ID NO: 43
14	SEQ ID NO: 44
15	SEQ ID NO: 45
16	SEQ ID NO: 46
17	SEQ ID NO: 47
18	SEQ ID NO: 48
19	SEQ ID NO: 49
20	ДНК-маркер
21	ДНК-маркер
22	без контрольного проводника
23	SEQ ID NO: 1
24	SEQ ID NO: 50
25	SEQ ID NO: 51
26	SEQ ID NO: 52
27	SEQ ID NO: 53
28	SEQ ID NO: 54
29	SEQ ID NO: 55
30	SEQ ID NO: 56
31	SEQ ID NO: 57
32	SEQ ID NO: 58
33	SEQ ID NO: 59
34	SEQ ID NO: 60
35	SEQ ID NO: 61
36	SEQ ID NO: 62
37	ДНК-маркер

Расщепление в процентах показано снизу под каждой дорожкой. В случае cr-ДНК или cr-Д(Р)НК не наблюдали активность расщепления (например, фигура 3, 3; фигура 3, 5; фигура 3, 15; фигура 3, 33; фигура 3, 34; фигура 3, 35) Активность в расщеплении обозначали как «не определено» (указывая тем самым, что активность в расщеплении не выявлена).

Данные, представленные на фигуре 3, демонстрируют, что cr-Д(Р)НК согласно настоящему изобретению способствует Cas9-опосредованному расщеплению двунитевой ДНК-мишени.

Пример 4

Активность cr-Д(Р)НК против множества мишеней

Данный пример демонстрирует биохимическую активность *in vitro* cr-Д(Р)НК, содержащих разные спейсеры, запрограммированные на конкретные последовательности-мишени.

Последовательности cr-ДНК, cr-РНК и cr-Д(Р)НК (показанные в таблице 5) были предоставлены коммерческому производителю для синтеза.

Таблица 5

Последовательности cr-ДНК, cr-РНК и cr-Д(Р)НК

Мишень	Тип РНК-проводника	Последовательности (основания РНК указаны в скобках)	SEQ ID NO
EMX1-мишень 1	cr-ДНК	5'-GAGTCCGAGC AGAAGAAGAA GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 3
VEGFA-мишень 1	cr-ДНК	5'-GGGTGGGGGG AGTTTGCTCC GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 4
CD34-мишень 1	cr-ДНК	5'-GTTTGTGTTT CCATAAACTG GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 5
CD34-мишень 2	cr-ДНК	5'-TCTGTGATAA CCTCAGTTTA GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 6
STAT5a-мишень 1	cr-ДНК	5'-GGCCACTGTA GTCCTCCAGG GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 7
STAT5a-мишень 2	cr-ДНК	5'-GTCCCCCAGC CGGTCAGCCA GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 8
JAK1-мишень 1	cr-ДНК	5'-GGCAGCCAGC ATGATGAGAC GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 9
JAK1-мишень 2	cr-ДНК	5'-GAGGAGCTCC AAGAAGACTG GTCTCAGAGC TATGCTGTCC TG-3'	SEQ ID NO: 10
EMX1-мишень 1	cr-РНК	5'-[G][A][G][U][C][C][G][A][G][C] [A][G][A][A][G][A][A][G][A][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 11
VEGFA-мишень 1	cr-РНК	5'-[G][G][U][G][G][G][G][G][G][G] [A][G][U][U][U][G][C][U][C][C] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 12
CD34-мишень 1	cr-РНК	5'-[G][U][U][U][G][U][G][U][U][U] [C][C][A][U][A][A][A][C][U][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 13
CD34-мишень 2	cr-РНК	5'-[U][C][U][G][U][G][A][U][A][A] [C][C][U][C][A][G][U][U][U][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 14
STAT5a-мишень 1	cr-РНК	5'-[G][G][C][C][A][C][U][G][U][A] [G][U][C][C][U][C][C][A][G][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 15
STAT5a-мишень 2	cr-РНК	5'-[G][U][C][C][C][C][A][G][C] [C][G][G][U][C][A][G][C][A] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 16
JAK1-мишень 1	cr-РНК	5'-[G][G][C][A][G][C][C][A][G][C] [A][U][G][A][U][G][A][G][A][C] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 17
JAK1-мишень 2	cr-РНК	5'-[G][A][G][G][A][G][C][U][C][C] [A][A][G][A][A][G][A][C][U][G] [G][U][C][U][C][A][G][A][G][C] [U][A][U][G][C][U][G][U][C][C] [U][G]-3'	SEQ ID NO: 18
EMX1-мишень 1	cr-Д(Р)НК	5'-GAGTCCGAGC AGAA[G][A][A][G][A][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 19
VEGFA-мишень 1	cr-Д(Р)НК	5'-GGGTGGGGGG AGTT[U][G][C][U][C][C] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 20
CD34-мишень 1	cr-Д(Р)НК	5'-GTTTGTGTTT CCAT[A][A][A][C][U][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 21
CD34-мишень 2	cr-Д(Р)НК	5'-TCTGTGATAA CCTC[A][G][U][U][U][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 22
STAT5a-мишень 1	cr-Д(Р)НК	5'-GGCCACTGTA GTCC[U][C][C][A][G][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 23
STAT5a-мишень 2	cr-Д(Р)НК	5'-GTCCCCCAGC CGGT[C][A][G][C][A] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 24
JAK1-мишень 1	cr-Д(Р)НК	5'-GGCAGCCAGC ATGA[U][G][A][G][A][C] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 25
JAK1-мишень 2	cr-Д(Р)НК	5'-GAGGAGCTCC AAGA[A][G][A][C][U][G] [G][U][C][U][C][A]GAGC TATGCTGTCC TG-3'	SEQ ID NO: 26

Tracr-РНК конструировали, как описано в примере 1.

Двунитевые ДНК-мишени создавали, как описано в примере 2, используя олигонуклеотиды, показанные в таблице 3, соответствующие подходящей последовательности-мишени.

Cr-ДНК/tracr-РНК, cr-РНК/tracr-РНК и cr-Д(Р)НК/tracr-РНК гибридизовали и осуществляли биохимическое расщепление, как описано в примере 3.

На фигуре 4А и фигуре 4В показаны результаты биохимического расщепления

различных спейсеров. На фигуре 4А показано биохимическое расщепление в процентах. Активность по отношению к ЕМХ-мишени 1 показана в группе 1: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к VEGFA-мишени 1 показана в группе 2: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к CD34-мишени 1 показана в группе 3: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к CD34-мишени 2 показана в группе 4: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность в отношении STAT5a-мишени 1 показана в группе 5: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к STAT5a-мишени 2 показана в группе 6: где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к JAK1-мишени 1 показана в группе 7; где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. Активность по отношению к JAK1-мишени 2 показана в группе 8; где «А» означает только контрольный Cas9, «В» означает cr-ДНК/tracr-РНК/Cas9, «С» означает cr-РНК/tracr-РНК/Cas9 и «D» означает cr-Д(Р)НК/tracr-РНК/Cas9. В случае всех образцов, содержащих только Cas9 (фигура 4А, «А») и образцов, содержащих cr-ДНК/tracr-РНК/cas9 (фигура 4В, «В»), активность в расщеплении не была выявлена (фигура 4А, «не определено»).

На фигуре 4В, расщепление в процентах показано на графике по оси у, а мишень показана по оси х. Активность по отношению к ЕМХ-мишени 1 показана столбиками в группе 1. Активность по отношению к VEGFA-мишени 1 показана столбиками в группе 2. Активность по отношению к CD34-мишени 1 показана столбиками в группе 3. Активность по отношению к CD34-мишени 2 показана столбиками в группе 4. Активность по отношению к STAT5a-мишени 1 показана столбиками в группе 5. Активность по отношению к STAT5a-мишени 2 показана столбиками в группе 6. Активность по отношению к JAK1-мишени 1 показана столбиками в группе 7. Активность по отношению к JAK1-мишени 2 показана столбиками в группе 8. «С» и «D» относятся к тем же реакциям, которые указаны на фигуре 4А.

Фигура 4 демонстрирует, что опосредованное Cas9 биохимическое расщепление двунитевой ДНК-мишени с использованием cr-Д(Р)НК согласно настоящему изобретению может быть перенесено на другие последовательности-мишени.

Пример 5

Анализ Т7Е1 для выявления модификаций мишеней в эукариотических клетках
Данный пример иллюстрирует применение анализа Т7Е1 для оценки расщепления в процентах *in vivo* cr-Д(Р)НК по отношению к выбранным двунитевым последовательностям ДНК-мишени.

А. Трансфекция клеток с использованием компонентов полинуклеотида Cas.

Sg-РНК и cr-Д(Р)НК/tracr-РНК, содержащие последовательность, направленную к мишени AAVS-1, трансфицировали в клетки НЕК293, конститутивно экспрессирующие слияние SpyCas9-GFP (НЕК293-Cas9-GFP), используя 96-луночную систему Nucleofector® 96-well Shuttle (Lonza, Allendale, NJ) и следуя протоколу. Эквимоллярные количества

компонентов РНК-проводника, готовили в буфере для отжига (1,25 mM HEPES, 0,625 mM MgCl₂, 9,375 mM KCl при pH 7,5), инкубировали в течение 2 минут при 95°C, извлекали из термоциклера, давали возможность уравновеситься до комнатной температуры и распределяли в конечном объеме 10 мкл в трех повторах в 96-луночном планшете. Культуральную среду аспирировали из клеток HEK293-Cas9-GFP и клетки промывали один раз не содержащим кальция и магния PBS, затем трипсинизировали, добавляя TrypLE (Life Technologies, Grand Island, NY), с последующей инкубацией при 37°C в течение 3-5 минут. Обработанные трипсином клетки осторожно пипетировали вверх и вниз для получения суспензии отдельных клеток и добавляли к полной культуральной среде DMEM, состоящей из культуральной среды DMEM (Life Technologies, Grand Island, NY), содержащей 10% FBS (Fisher Scientific, Pittsburgh, PA) и добавки пенициллина и стрептомицина (Life Technologies, Grand Island, NY).

Затем клетки осаждали центрифугированием в течение 3 минут при 200 x g, культуральную среду аспирировали и клетки ресуспендировали в PBS. Клетки подсчитывали, используя автоматический счетчик клеток Countess® II (Life Technologies, Grand Island, NY). $2,2 \times 10^7$ клеток переносили в пробирку объемом 50 мл и осаждали. PBS аспирировали и клетки ресуспендировали в растворе Nucleofector™ SF (Lonza, Allendale, NJ) до плотности 1×10^7 клеток/мл. Затем по 20 мкл клеточной суспензии добавляли в отдельные лунки, содержащие 10 мкл компонентов полинуклеотида Cas, и весь объем переносили в лунки 96-луночного планшета Nucleocuvette™ (Lonza, Allendale, NJ). Планшет загружали в устройство для нуклеофекции в 96-луночных планшетах Nucleofector™ 96-well Shuttle™ (Lonza, Allendale, NJ) и клетки подвергали нуклеофекции, используя программу 96-CM-130 Nucleofector™ (Lonza, Allendale, NJ). После нуклеофекции в каждую лунку добавляли 70 мкл полной культуральной среды DMEM и 50 мкл клеточной суспензии переносили в покрытый коллагеном 96-луночный планшет для культуры клеток, содержащий по 150 мкл предварительно нагретой полной культуральной среды DMEM. Затем планшет переносили в инкубатор для культивирования клеток и выдерживали при 37° в 5% CO₂ в течение 48 часов.

В. Создание двунитевой ДНК-мишени для анализа T7E1

Выделяли гДНК из клеток HEK-293-SpyCas9 через 48 часов после трансфекции компонентами полинуклеотида Cas, используя по 50 мкл раствора для экстракции ДНК QuickExtract ДНК (Epicentre, Madison, WI) на лунку, с последующей инкубацией при 37°C в течение 10 минут, при 65°C в течение 6 минут и при 95°C в течение 3 минут, чтобы остановить реакцию. Затем гДНК разбавляли, используя 150 мкл воды, и образцы хранили при -80°C.

ДНК для T7E1 создавали посредством ПЦР-амплификации двунитевой последовательности ДНК-мишени (например, AAVS-1) на основе изолированной гДНК. ПЦР-реакции осуществляли, используя 8 мкл гДНК в качестве матрицы и полимеразу с горячим стартом КАРА HiFi, и реакционная смесь содержала 0,5 единиц полимеразы, 1x-буфер для реакции, 0,4 mM dNTP и 300 нМ прямого и обратного праймеров, направленных к двунитевой ДНК-мишени (например, AAVS-1, SEQ ID NO: 75, 76 (таблица 3)) в общем объеме 25 мкл. ДНК-мишень амплифицировали, используя следующие условия: 95°C в течение 5 минут, 4 циклов по 20 секунд при 98°C, 20 секунд при 70°C, минут 2°C/цикл, 30 секунд при 72°C, с последующими 30 циклами: 15 секунд при 98°C, 20 секунд при 62°C, 20 секунд при 72°C и конечное удлинение при 72°C в течение 1 минуты.

С. Анализ T7E1

Амплифицированную в ПЦР двунитевую ДНК-мишень для анализов T7E1 денатурировали при 95°C в течение 10 минут и затем обеспечивали возможность повторного отжига, охлаждая до 25°C со скоростью -0,5°C/сек в термоциклере.

Подвергнутую повторному отжигу ДНК инкубировали с 0.5 мл эндонуклеазы I T7 в 1x буфере NEBuffer 2 (New England Biolabs, Ipswich, MA) в общем объеме 15 мл в течение 25 минут при 37°C. Реакционные смеси T7E1 анализировали, используя систему Fragment Analyzer™ (Advanced Analytical Technologies, Inc., Ames, IA) и набор реагентов для двунитевой ДНК DNF-910 (Advanced Analytical Technologies, Inc., Ames, IA). Система Fragment Analyzer™ обеспечивает определение концентрации каждого фрагмента, полученного при расщеплении, и двунитевой ДНК-мишени, которая остается после расщепления.

Расщепление двунитевой ДНК-мишени в процентах вычисляли на основе концентрации каждого полученного при расщеплении фрагмента и двунитевой ДНК-мишени, которая остается после того, как произошло расщепление, используя следующую формулу:

$$\%_{\text{расщепления}} = \left(1 - \sqrt{1 - \frac{(\text{фраг.1} + \text{фраг.2})}{(\text{фраг.1} + \text{фраг.2} + \text{исход.})}} \right)$$

УРАВНЕНИЕ 1

В уравнении 1 концентрации «фраг1» и «фраг2» соответствуют концентрации фрагментов, получаемых при расщеплении Cas двунитевой ДНК-мишени и «исход.» соответствует двунитевой ДНК-мишени, которая остается после того, как произошло расщепление.

На фигуре 5 показаны результаты T7E1-анализа гДНК, полученной из клеток, трансфицированных cr-Д(Р)НК в разных концентрациях. Средняя частота инделов (инсерционно-делеционных событий) в процентах показана над каждым столбиком на графике (вычислено с использованием уравнения 1). Процент является средним для трех образцов, за исключением в случае столбика 4 на фигуре 5 столбика 4, для которого активность была определена только в двух образцах, и в случае столбика 5 на фигуре 5, для которого активность была определена только в одном образце. Концентрации либо cr-Д(Р)НК/tracr-РНК, либо sg-РНК, введенной посредством нуклеофекции в клетки, показаны в таблице 6.

Таблица 6

Концентрации трансфицированных компонентов РНК-проводника

№	SEQ ID NO:	пмоль
1	SEQ ID NO: 43	500
2	SEQ ID NO: 43	750
3	SEQ ID NO: 43	1000
4	SEQ ID NO: 43	2000
5	SEQ ID NO: 43	3000
6	SEQ ID NO: 1	500

Анализ T7E1 для выявления модификаций мишени в эукариотических клетках дает данные, демонстрирующие, что системы cr-Д(Р)НК/tracr-РНК/Cas9, которые описаны в настоящей публикации, способствуют Cas-опосредованному сайт-специфичному расщеплению двунитевой ДНК-мишени *in vivo*.

Следуя приведенному в настоящем описании руководству, анализ T7E1, описанный в данном примере, может быть осуществлен специалистом в данной области на практике для измерения активности в случае клеток, модифицированных другими системами

CRISPR-Cas, включая без ограничения Cas9, Cas9-подобный белок, Cas1, Csn2, Cas4, Cpf1, C2c1, C2c2, C2c3, белки. Кодированные ортологами Cas9, Cas9-подобные синтетические белки, слияния с Cas9 и их варианты и модификации, объединенные с соответствующими им полинуклеотидными компонентами, модифицированными, как описано в настоящей публикации, так чтобы они содержали cr-Д(Р)НК.

Пример 6

Активность в расщеплении cr-Д(Р)НК в/вне мишени

Данный пример иллюстрирует применение cr-Д(Р)НК для оценки активности в расщеплении мишени в предназначенном в качестве мишени участке («в мишени») и в предполагаемых ближайших соседних («вне мишени») участках. Последовательности-мишени участков на/вне мишени (on/off) показаны в таблице 7:

Таблица 7

Последовательности сайтов в/вне мишени

Мишень	Участок	Последовательность-мишень	SEQ ID NO:
EMX-1	ON	5' - GAGTCCGAGC AGAAGAAGAA - 3'	SEQ ID NO: 27
EMX-1	OFF1	5' - GAGTTAGAGC AGAAGAAGAA - 3'	SEQ ID NO: 28
EMX-1	OFF2	5' - AGGTACTAGC AGAAGAAGAA - 3'	SEQ ID NO: 29
EMX-1	OFF3	5' - ACGTCTGAGC AGAAGAAGAA - 3'	SEQ ID NO: 30
EMX-1	OFF4	5' - AGGTGCTAGC AGAAGAAGAA - 3'	SEQ ID NO: 31
VEGFA-1	ON	5' - GGGTGGGGGG AGTTTGCTCC - 3'	SEQ ID NO: 32
VEGFA-1	OFF1	5' - GGATGGAGGG AGTTTGCTCC - 3'	SEQ ID NO: 33
VEGFA-1	OFF2	5' - GGGGAGGGGA AGTTTGCTCC - 3'	SEQ ID NO: 34
VEGFA-1	OFF3	5' - GGGAGGGTGG AGTTTGCTCC - 3'	SEQ ID NO: 35
VEGFA-1	OFF4	5' - CGGGGAGGG AGTTTGCTCC - 3'	SEQ ID NO: 36

Последовательности cr-РНК и cr-Д(Р)НК были предоставлены коммерческому производителю для синтеза. Tracr-РНК конструировали, как описано в примере 1.

Двунитевые ДНК-мишени создавали, как описано в примере 2, используя олигонуклеотиды, показанные в таблице 8, соответствующие подходящей последовательности-мишени.

Таблица 8

ДНК в/вне мишени

Мишень	Участок	Последовательность-мишень
EMX-1	on	SEQ ID NO: 107, 108
EMX-1	OFF1	SEQ ID NO: 111, 112
EMX-1	OFF2	SEQ ID NO: 113, 114
EMX-1	OFF3	SEQ ID NO: 115, 116
EMX-1	OFF4	SEQ ID NO: 117, 118
VEGFA-1	on	SEQ ID NO: 119, 120
VEGFA-1	OFF1	SEQ ID NO: 121, 122
VEGFA-1	OFF2	SEQ ID NO: 123, 124
VEGFA-1	OFF3	SEQ ID NO: 125, 126
VEGFA-1	OFF4	SEQ ID NO: 107, 108

cr-РНК/tracr-РНК и cr-Д(Р)НК/tracr-РНК гибридизовали и осуществляли биохимическое расщепление, как описано в примере 3.

На фигуре 6 показано сравнение биохимической активности cr-РНК/tracr-РНК и cr-Д(Р)НК/tracr-РНК в предназначенных в качестве мишени участках и четырех прогнозируемых с использованием компьютерной программы участках вне мишени. Расщепление в процентах показано на оси y, и образцы показаны на оси x. В таблице 9 приведен список образцов:

Таблица 9

Активность cr-PHK и tracr-PHK в/вне мишени

5

10

15

20

Идентификационный номер	Участок-мишень	Компонент РНК-проводника
1A	EMX-1 ON	cr-PHK
1B	EMX-1 ON	cr-Д(Р)НК
2A	EMX-1 OFF-1	cr-PHK
2B	EMX-1 OFF-1	cr-Д(Р)НК
3A	EMX-1 OFF-2	cr-PHK
3B	EMX-1 OFF-2	cr-Д(Р)НК
4A	EMX-1 OFF-3	cr-PHK
4B	EMX-1 OFF-3	cr-Д(Р)НК
5A	EMX-1 OFF-4	cr-PHK
5B	EMX-1 OFF-4	cr-Д(Р)НК
6A	VEGFA-1 ON	cr-PHK
6B	VEGFA-1 ON	cr-Д(Р)НК
7A	VEGFA-1 OFF-1	cr-PHK
7B	VEGFA-1 OFF-1	cr-Д(Р)НК
8A	VEGFA-1 OFF-2	cr-PHK
8B	VEGFA-1 OFF-2	cr-Д(Р)НК
9A	VEGFA-1 OFF-3	cr-PHK
9B	VEGFA-1 OFF-3	cr-Д(Р)НК
10A	VEGFA-1 OFF-4	cr-PHK
10B	VEGFA-1 OFF-4	cr-Д(Р)НК

25

Данные, представленные на фигуре 7, показывают, что cr-Д(Р)НК поддерживают высокую активности в пределах мишени, по сравнению с cr-PHK. cr-Д(Р)НК не поддерживают активность вне мишени, тогда как cr-PHK обладают нежелательной активностью вне мишени.

Пример 7

Анализ глубокого секвенирования для выявления модификаций мишени в эукариотических клетках

30

Данный пример иллюстрирует применение анализа, основанного на глубоком секвенировании, для оценки и сравнения расщепления в процентах *in vivo* выбранными комплексами sg-Д(Р)НК/белок Cas9 по отношению к выбранным последовательностям двунитевых ДНК-мишеней.

А. Синтез sg-Д(Р)НК

35

Шесть последовательностей sg-Д(Р)НК, нацеленных на локус AAVS-1 человека и содержащих разные ДНК/РНК-составы и защищенные фосфоротиоатом связи, были предоставлены коммерческому производителю для синтеза. Такие последовательности показаны в таблице 10.

Таблица 10

Последовательности sg-Д(Р)НК

40

45

Название	Последовательность (основания РНК указаны в скобках, фосфоротиоатные связи показаны знаком *)	SEQ ID NO:
sg-Д(Р)НК -01	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U][G][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U][C][G][G][U][G][C][U] - 3'	SEQ ID NO: 127
sg-Д(Р)НК -02	5' - G*G*GGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G][A][G][C] [U][A][G][A][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][A][C] [U][U][G][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U][C][G][G][U][G][C][U] - 3'	SEQ ID NO: 128

sg-Д(Р)НК -03	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G] [A]GC TATGCT[G][A][A][A]AGC AAGC[A][A][G][U][U] [A][A][A][A] [U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A][C] [U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] [C][G][A][G][U][C] [G][G][U][G] [C][U] - 3'	SEQ ID NO: 129
sg-Д(Р)НК -04	5' - G*G*GGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A] [G][A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A] [A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A] [A][C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] [C][G][A][G] [U][C][G][G][U][G] [C][U] - 3'	SEQ ID NO: 130
sg-Д(Р)НК -05	5' - GGGGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A][G] [A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A][A] [A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A][A] [C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] CG[A][G][U]CGGTG [C][U] - 3'	SEQ ID NO: 131
sg-Д(Р)НК -06	5' - G*G*GGCCACTA GGGA[C][A][G][G][A][U] [G][U][U][U][U][A] [G][A]GC TATGCT[G][A][A][A]AGC ATAGC[A][A][G][U][U] [A][A] [A][A][U][A][A][G][G][C] [U][A][G][U][C][C][G][U][U][A] [U][C][A] [A][C][U][U][G][A][A] [A][A][A][G][U][G][G][C][A][C] CG[A][G][U]CGGTG [C][U] - 3'	SEQ ID NO: 132

В. Образование РНП-комплексов sg-Д(Р)НК/белок Cas9

Белок Cas9 экспрессировали с бактериального экспрессирующего вектора в *E. coli* (BL21 (DE3)) и очищали, используя аффинную ионообменную и эксклюзионную хроматографию по размеру согласно способам, описанным Jinek с соавторами (Science; 337(6096): 816-21(2012)). Кодированная последовательность для Cas9 *Streptococcus pyogenes* содержала две последовательности ядерной локализации (NLS) на С-конце.

Рибонуклеопротеидные (РНП) комплексы собирали в трех повторах при двух концентрациях, 20 пмоль Cas9:60 пмоль sg-Д(Р)НК и 200 пмоль Cas9:600 пмоль sg-Д(Р)НК. Компоненты sg-Д(Р)НК смешивали в эквимольных количествах в буфере для отжига (1,25 мМ NEPEES, 0,625 мМ MgCl₂, 9,375 мМ KCl при pH7,5) до требуемой концентрации (60 пмоль или 600 пмоль) в конечном объеме 5 мкл, инкубировали в течение 2 минут при 95°C, извлекали из термоциклера и давали возможность уравновеситься до комнатной температуры. Белок Cas9 разбавляли до соответствующей концентрации в буфере для связывания (20 мМ NEPEES, 100 мМ KCl, 5 мМ MgCl₂, 1 мМ ДТТ и 5% глицерин при pH 7,4) до конечного объема 5 мкл и смешивали с 5 мкл денатурированных нагреванием sg-Д(Р)НК с последующей инкубацией при 37°C в течение 30 минут.

С. трансфекции клеток с использованием РНП sg-Д(Р)НК/белок Cas9

РНП-комплексами трансфицировали клетки K562 (ATCC, Manassas, VA), используя систему используя 96-луночную систему Nucleofector® 96-well Shuttle (Lonza, Allendale, NJ) и следуя протоколу. РНП-комплексы распределяли в конечном объеме 10 мкл по отдельным лункам 96-луночного планшета. Клетки K562, суспендированные в среде, переносили из культурального флакона в коническую пробирку объемом 50 мл. Клетки осаждали центрифугированием в течение 3 минут при 200 x g, культуральную среду аспирировали и клетки промывали один раз не содержащим кальция и магния PBS. Затем клетки K562 осаждали центрифугированием в течение 3 минут при 200 x g, PBS аспирировали и осадок клеток ресуспендировали в 10 мл не содержащего кальция и магния PBS.

Клетки подсчитывали, используя автоматический счетчик клеток Countess® II (Life Technologies, Grand Island, NY). $2,2 \times 10^7$ клеток переносили в пробирку объемом 50 мл и осаждали. PBS аспирировали и клетки ресуспендировали в растворе Nucleofector™ SF (Lonza, Allendale, NJ) до плотности 1×10^7 клеток/мл. По 20 мкл клеточной суспензии добавляли в отдельные лунки, содержащие 10 мкл РНП-комплексов, и весь объем переносили в лунки 96-луночного планшета Nucleocuvette™ (Lonza, Allendale, NJ).

Планшет загружали в устройство для нуклеофекции в 96-луночных планшетах Nucleofector™ 96-well Shuttle™ (Lonza, Allendale, NJ) и клетки подвергали нуклеофекции, используя программу 96-FF-120 Nucleofector™ (Lonza, Allendale, NJ). После нуклеофекции в каждую лунку добавляли 70 мкл 70 среду Дульбекко, модифицированной способом Исков (IMDM; Life Technologies, Grand Island, NY) с добавлением 10% FBS (Fisher Scientific, Pittsburgh, PA), пенициллина и стрептомицина (Life Technologies, Grand Island, NY) и 50 мкл клеточной суспензии переносили в 96-луночный планшет для культивирования клеток, содержащий 150 мкл предварительно нагретой полной культуральной среды IMDM. Затем планшет переносили в инкубатор для культуры тканей и выдерживали при 37°C в 5% CO₂ в течение 48 часов.

Д. Создание двунитевой ДНК-мишени для глубокого секвенирования

гДНК выделяли из клеток K562 через 48 часов после трансфекции РНП с использованием по 50 мкл раствора для экстракции ДНК QuickExtract (Epicentre, Madison, WI) на лунку с последующей инкубацией при 37°C в течение 10 минут, при 65°C в течение 6 минут и при 95°C в течение 3 минут, чтобы остановить реакцию. Выделенные гДНК разбавляли, используя 50 мкл воды, и образцы хранили при -80°C.

Используя выделенную гДНК, осуществляли первую ПЦР, используя смесь Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA) в 1-кратной концентрации, праймеры по 0,5 мкМ каждого (SEQ ID NO: 93, 94), 3,75 мкл гДНК в конечном объеме 10 мкл, и амплифицировали при 98°C в течение 1 минуты, 35 циклов по 10 секунд при 98°C, 20 секунд при 60°C, 30 секунд при 72°C и конечное удлинение при 72°C в течение 2 минут. Реакционные смеси для ПЦР разбавляли 1:100 в воде.

ПЦР со «штрихкодированием» проводили, используя уникальные праймеры для каждого образца, чтобы облегчить мультиплексное секвенирование. Образцы и соответствующие пары праймеров показаны в таблице 11.

Таблица 11
Праймеры для штрихкодирования

Идентификационный номер	Образец	SEQ ID NO:
Набор праймеров для штрихкодирования-1	sg-Д(Р)НК-01 60 пмоль ger1	SEQ ID NO: 95, 101
Набор праймеров для штрихкодирования-2	sg-Д(Р)НК-02 60 пмоль ger1	SEQ ID NO: 95, 102
Набор праймеров для штрихкодирования-3	sg-Д(Р)НК-03 60 пмоль ger1	SEQ ID NO: 95, 103
Набор праймеров для штрихкодирования-4	sg-Д(Р)НК-04 60 пмоль ger1	SEQ ID NO: 95, 104
Набор праймеров для штрихкодирования-5	sg-Д(Р)НК-05 60 пмоль ger1	SEQ ID NO: 95, 105
Набор праймеров для штрихкодирования-6	sg-Д(Р)НК-06 60 пмоль ger2	SEQ ID NO: 95, 106
Набор праймеров для штрихкодирования-7	sg-Д(Р)НК-01 60 пмоль ger2	SEQ ID NO: 96, 101
Набор праймеров для штрихкодирования-8	sg-Д(Р)НК-02 60 пмоль ger2	SEQ ID NO: 96, 102
Набор праймеров для штрихкодирования-9	sg-Д(Р)НК-03 60 пмоль ger2	SEQ ID NO: 96, 103
Набор праймеров для штрихкодирования-10	sg-Д(Р)НК-04 60 пмоль ger2	SEQ ID NO: 96, 104
Набор праймеров для штрихкодирования-11	sg-Д(Р)НК-05 60 пмоль ger2	SEQ ID NO: 96, 105
Набор праймеров для штрихкодирования-12	sg-Д(Р)НК-06 60 пмоль ger2	SEQ ID NO: 96, 106
Набор праймеров для штрихкодирования-13	sg-Д(Р)НК-01 60 пмоль ger3	SEQ ID NO: 97, 101
Набор праймеров для штрихкодирования-14	sg-Д(Р)НК-02 60 пмоль ger3	SEQ ID NO: 97, 102

5	Набор праймеров для штрихкодирования-15	sg-Д(Р)НК-03 60 пмоль гер3	SEQ ID NO: 97, 103
	Набор праймеров для штрихкодирования-16	sg-Д(Р)НК-04 60 пмоль гер3	SEQ ID NO: 97, 104
	Набор праймеров для штрихкодирования-17	sg-Д(Р)НК-05 60 пмоль гер3	SEQ ID NO: 97, 105
	Набор праймеров для штрихкодирования-18	sg-Д(Р)НК-06 60 пмоль гер3	SEQ ID NO: 97, 106
	Набор праймеров для штрихкодирования-19	sg-Д(Р)НК-01 600 пмоль гер1	SEQ ID NO: 98, 101
	Набор праймеров для штрихкодирования-20	sg-Д(Р)НК-02 600 пмоль гер1	SEQ ID NO: 98, 102
10	Набор праймеров для штрихкодирования-21	sg-Д(Р)НК-03 600 пмоль гер1	SEQ ID NO: 98, 103
	Набор праймеров для штрихкодирования-22	sg-Д(Р)НК-04 600 пмоль гер1	SEQ ID NO: 98, 104
	Набор праймеров для штрихкодирования-23	sg-Д(Р)НК-05 600 пмоль гер1	SEQ ID NO: 98, 105
	Набор праймеров для штрихкодирования-24	sg-Д(Р)НК-06 600 пмоль гер2	SEQ ID NO: 98, 106
15	Набор праймеров для штрихкодирования-25	sg-Д(Р)НК-01 600 пмоль гер2	SEQ ID NO: 99, 101
	Набор праймеров для штрихкодирования-26	sg-Д(Р)НК-02 600 пмоль гер2	SEQ ID NO: 99, 102
	Набор праймеров для штрихкодирования-27	sg-Д(Р)НК-03 600 пмоль гер2	SEQ ID NO: 99, 103
	Набор праймеров для штрихкодирования-28	sg-Д(Р)НК-04 600 пмоль гер2	SEQ ID NO: 99, 104
20	Набор праймеров для штрихкодирования-29	sg-Д(Р)НК-05 600 пмоль гер2	SEQ ID NO: 99, 105
	Набор праймеров для штрихкодирования-30	sg-Д(Р)НК-06 600 пмоль гер2	SEQ ID NO: 99, 106
	Набор праймеров для штрихкодирования-31	sg-Д(Р)НК-01 600 пмоль гер3	SEQ ID NO: 100, 101
	Набор праймеров для штрихкодирования-32	sg-Д(Р)НК-02 600 пмоль гер3	SEQ ID NO: 100, 102
25	Набор праймеров для штрихкодирования-33	sg-Д(Р)НК-03 600 пмоль гер3	SEQ ID NO: 100, 103
	Набор праймеров для штрихкодирования-34	sg-Д(Р)НК-04 600 пмоль гер3	SEQ ID NO: 100, 104
	Набор праймеров для штрихкодирования-35	sg-Д(Р)НК-05 600 пмоль гер3	SEQ ID NO: 100, 105
	Набор праймеров для штрихкодирования-36	sg-Д(Р)НК-06 600 пмоль гер3	SEQ ID NO: 100, 106

ПЦР со штрихкодированием осуществляли, используя смесь Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA) в 1-кратной концентрации, праймеры по 0,5 мкМ каждого, 1 мкл реакционной смеси после первой ПЦР в разведении 1:100 в конечном объеме 10 мкл, и амплифицировали при 98°C в течение 1 минуты, 12 циклов по 10 секунд при 98°C, 20 секунд при 60°C, 30 секунд при 72°C и конечное удлинение при 72°C в течение 2 минут.

Е. Очистка SPRIselect

ПЦР-реакционные смеси объединяли в одной микроцентрифужной пробирке для основанной на шариках очистки ампликонов для секвенирования SPRIselect (Beckman Coulter, Pasadena, CA).

К объединенным ампликонам добавляли 0,9-кратные объемы шариков и перемешивали и инкубировали при комнатной температуре (RT) в течение 10 минут. Микроцентрифужную пробирку помещали на магнитную подставку для пробирок (Beckman Coulter, Pasadena, CA) до тех пор, пока раствор становился прозрачным. Надосадок удаляли и отбрасывали, и оставшиеся шарики промывали 1 объемом 85% этанола и инкубировали при КТ в течение 30 секунд. После инкубации этанол

аспирировали и шарики сушили на воздухе при КТ в течение 10 минут. Затем микроцентрифужную пробирку извлекали из магнитной подставки к шарикам добавляли 0,25-кратные объемы буфера Qiagen EB (Qiagen, Venlo, Limburg), энергично перемешивали и инкубировали в течение 2 минут при комнатной температуре. Микроцентрифужную пробирку возвращали на магнит, инкубировали до тех пор, пока раствор становился прозрачным, и надосажок, содержащий очищенные ампликоны, переносили в чистую микроцентрифужную пробирку. Библиотеку очищенных ампликонов оценивали количественно, используя систему NanodropTM 2000 (Thermo Scientific, Wilmington, DE), и качество библиотек анализировали, используя систему Fragment AnalyzerTM (Advanced Analytical Technologies, Inc., Ames, IA) и набор реагентов для двунитевой ЛНК DNF-910 (Advanced Analytical Technologies, Inc. Ames, IA).

Ф. Проведение глубокого секвенирования

Библиотеку ампликонов нормализовали до концентрации 4 нмоль, которую вычисляли на основании значений Nanodrop и размера ампликонов. Библиотеку анализировали на секвенаторе MiSeq (Illumina, San Diego, CA) с использованием набора реагентов MiSeq v2 (Illumina, San Diego, CA) в ходе 300 циклов, при этом использовали два прогона по 151 циклов для парных чтений плюс два состоящий их восьми циклов чтений индексов.

Г. Анализ данных глубокого секвенирования

Продукты идентифицировали в полученных при секвенировании данных на основании последовательностей с индексами-штрихкодами, введенными в ампликоны в цикле штрихкодирования в ПЦР. Для обработки данных использовали вычислительный скрипт MiSeq, выполняя следующие задачи:

- ряды выравнивали с геномом человека (build GRCh38/38), используя компьютерную программу Bowtie (<http://bowtie-bio.sourceforge.net/index.shtml>);
- выравненные ряды сравнивали с ожидаемой последовательностью локуса AAVS-1 дикого типа, ряды, которые не выравнивались ни с какой частью локуса AAVS-1, отбрасывали;
- учитывали ряды, совпадающие с последовательностью AAVS-1 дикого типа;
- ряды с инделами (инсерциями или делециями оснований) относили к определенной категории на основании типа индела и учитывали;
- все ряды с инделами делили на сумму рядов дикого типа и рядов с инделами, получая процент выявленных инделов.

На фигуре 7 показаны результаты анализа локуса-мишени AAVS-1 из клеток K562 человека, подвергнутых нуклеофекции с использованием sg-Д(Р)НК/Cas9, нацеленной на область в локусе AAVS-1. На оси показано номер последовательности SEQ ID NO. Для используемой sg-Д(Р)НК на оси у показан процент инделов, выявленных на основании данных MiSeq. В серии А показан средний процент инделов, выявленных в трех независимых повторах для данной sg-Д(Р)НК при соотношении 20 пмоль Cas9: 120 пмоль sg-Д(Р)НК, и в серии В показан средний процент инделов, выявленных в трех независимых повторах для данной sg-Д(Р)НК при соотношении 100 пмоль Cas9: 600 пмоль sg-Д(Р)НК. Стандартное отклонение среднего процента из трех повторов указано вертикальными черными линиями. Цифры под столбиками соответствуют номеру последовательности SEQ ID NO sg-Д(Р)НК, используемой в трансфекции, последовательности sg-Д(Р)НК приведены в таблице 10. Полученные данные показывают способность разных типов sg-Д(Р)НК индуцировать модификации в области-мишени в клетках человека специфичными для последовательности и зависимым от дозы образом.

Способы, описанные в настоящей публикации, были осуществлены на практике специалистом в данной области с целью демонстрации активности sg-Д(Р)НК/Cas9 *in vivo* посредством анализа глубокого секвенирования.

Пример 8

5 Скрининг множественных sg-Д(Р)НК, содержащих последовательности, связывающие ДНК-мишень

Данный пример иллюстрирует применение sg-Д(Р)НК согласно настоящему изобретению для модификации мишеней, присутствующих в геномной ДНК человека, и измерения уровня активности в расщеплении в таких участках. Участки-мишени
10 сначала могут быть выбраны из геномной ДНК, и затем могут быть сконструированы sg-Д(Р)НК, нацеленные на такие выбранные последовательности. Затем могут быть осуществлены измерения для определения уровня расщепления мишени, которое имело место. Не все указанные далее стадии необходимы для каждого скрининга, и порядок стадий не обязательно должен быть таким, как представлено, и скрининг может быть
15 совмещен с другими экспериментами или может составлять часть более крупного эксперимента.

А. Выбор области ДНК-мишени в геномной ДНК

Идентифицировать все PAM-последовательности (например, «NGG») в выбранной геномной области.

20 Идентифицировать и выбрать одну или более последовательностей длиной 20 нуклеотидов (последовательность ДНК-мишени), которая находится с 5'-стороны вблизи PAM-последовательностей.

Критерии выбора могут включать без ограничения: гомологию с другими областями в геноме; процентное содержание G-C; температуру плавления; присутствие
25 гомополимера в спейсере и другие критерии, известные специалисту в данной области.

Присоединить подходящую последовательность sg-Д(Р)НК к 3'-концу идентифицированной последовательности ДНК-мишени. Конструкцию sg-Д(Р)НК обычно синтезирует коммерческий производитель, и соответствующую tracr-РНК получают, как описано в примере 1, посредством транскрипции *in vitro*.

30 sg-Д(Р)НК, которая описана в настоящей публикации, может быть использована с соответствующей tracr-РНК, чтобы получить систему sg-Д(Р)НК/tracr-РНК для применения с соответствующим белком Cas.

В. Определение расщепления в процентах и специфичности

Процент расщепления *in vitro* и специфичность, ассоциированные с системой sg-Д(Р)НК/tracr-РНК сравнивали, используя, например, анализы расщепления белком Cas,
35 описанные в примере 3, следующим образом:

(a) Если идентифицирована и выбрана одна последовательность ДНК-мишени, то можно определить процент расщепления и специфичность в отношении области ДНК-мишени. При необходимости процент расщепления и/или специфичность можно изменить
40 в следующих экспериментах, используя способы согласно настоящему изобретению, включая без ограничения модификацию sg-Д(Р)НК, введение эффекторных белков/последовательностей, связывающих эффекторные белки или лиганда/остатков, связывающих лиганд.

(b) Данные о расщеплении в процентах и данные о сайт-специфичности, полученные
45 в анализах расщепления, для разных ДНК, содержащих связывающую мишень последовательность, можно сравнить, чтобы идентифицировать последовательности ДНК-мишени, имеющие наилучший процент расщепления и наиболее высокую специфичность. Данные о расщеплении в процентах и данные о специфичности

позволяют получить критерии, на основании которых осуществляют выбор для различных применений. Например, в некоторых случаях активность *cr*-Д(Р)НК может быть наиболее важным фактором. В других случаях специфичность участка расщепления может быть сравнительно более важной, чем процент расщепления. При необходимости процент расщепления и/или специфичность изменяют в следующих экспериментах, применяя способы согласно настоящему изобретению, включая без ограничения модификацию *cr*-Д(Р)НК, введение эффекторных белков/ последовательностей, связывающих эффекторные белки или лиганда/остатков, связывающих лиганд.

Необязательно или вместо анализа *in vitro* сравнивают процент расщепления и специфичность, ассоциированные с системой *cr*-Д(Р)НК, *in vivo*, например, используя анализ T7E1, описанный в примере 5, следующим образом:

(а) Как только идентифицирована последовательность ДНК-мишени, то можно определить процент расщепления и специфичность в отношении области ДНК-мишени. При необходимости процент расщепления и/или специфичность изменяют в следующих экспериментах, используя способы согласно настоящему изобретению, включая без ограничения модификацию *cr*-Д(Р)НК, введение эффекторных белков/ последовательностей, связывающих эффекторные белки или лиганда/остатков, связывающих лиганд.

(б) Данные о расщеплении в процентах и данные о сайт-специфичности, полученные в анализах расщепления, для разных ДНК-мишеней, можно сравнить, чтобы идентифицировать последовательность *cr*-Д(Р)НК, которая приводит к наиболее высокому проценту расщепления ДНК-мишени и наиболее высокой специфичности по отношению к ДНК-мишени. Данные о расщеплении в процентах и данные о специфичности позволяют получить критерии, на которых основан выбор для разных применений. Например, некоторые варианты могут быть основаны на активности *cr*-Д(Р)НК, которая может быть наиболее важным фактором. В некоторых вариантах специфичность участка расщепления может быть относительно более важной, чем процент расщепления. В некоторых вариантах процент расщепления и/или специфичность могут быть изменены с использованием способов согласно настоящему изобретению, включая без ограничения модификацию РНК, введение эффекторных белков/последовательностей, связывающих эффекторные белки, или лиганда/остатков, связывающих лиганд.

Следуя инструкциям, приведенным в настоящем описании, и примерам, скрининг, описанный в данном примере, может быть осуществлен специалистом в данной области с использованием других белков Cas CRISPR класса II, включая без ограничения Cas9, Cas9-like, Cas, Cas3, Csn2, Cas4, белки, кодируемые ортологами Cas9, Cas9-подобные синтетические белки, слияния Cas9, Cpf1, Cpf1-подобные, C2c1, C2c2, C2c3 и их варианты и модификации, объединенные с соответствующими им полинуклеотидными компонентами, модифицированными, как описано в настоящей публикации, так, чтобы они содержали *cr*-Д(Р)НК.

Пример 9

Опосредованное *cr*-Д(Р)НК:tracr-РНК и *sg*-Д(Р)НК никирование

Данный пример иллюстрирует способ с применением которого комплекс *cr*-Д(Р)НК: tracr-РНК или *sg*-Д(Р)НК согласно настоящему изобретению могут быть использованы для индукции однонитевых разрывов (ников) в мишени в виде двунитевой ДНК (днДНК) -плазмиды вместе с Cas9 *S. Pyogenes*, содержащем мутацию D10A (Cas9-D10A), делающую неактивной RuvC-нуклеазную долю. Не все указанные далее стадии необходимы, и порядок стадий не обязательно должен быть таким, как представлено.

Cas9 *S. pyogenes* имеет два активных нуклеазных домена, RuvC- и HNH-домены. Мутация аспарагиновой кислоты в 10-м положении аминокислоты Cas9 *S. pyogenes*, превращающий ее в аланин, снижает нуклеазную способность RuvC-домена. HNH-домен остается активным, но сайт-специфичный полипептид Cas9-D10A может вызывать только ники в фосфодиэфирном осто́ве нити ДНК -мишени, комплементарной последовательности спейсера.

Примеры подходящих векторов, сред, условий культивирования и т.д. описаны. Модификации таких компонентов и условий будут понятны специалисту в данной области в свете инструкций, приведенных в настоящем описании.

Реагенты для проводника создавали согласно примеру 1, приведенному в настоящей описании.

днДНК-мишень создавали, как описано в примере 2, используя последовательности SEQ ID NO: 133 и 134. Затем амплифицированный фрагмент клонировали в подходящем LIC-совместимом векторе. Одним из таких подходящих векторов является коммерчески доступный клонирующий вектор pET His6 LIC (Addgene, Cambridge, MA). Плазмидой трансформировали бактериальный штамм для экспрессии плазмиды, используя коммерчески доступные бактериальные клетки XL1-Blu (Agilent, Santa Clara, CA).

Бактериальные клетки, содержащие LIC-векторы, выращивали в среде LB с добавлением 100 мкг/мл ампициллина (Sigma-Aldrich, St. Louis, MO) в течение 18 часов при 37°C. Клетки центрифугировали при 5000 об./мин в течение 15 минут, затем плазмиду экстрагировали, используя набор для плазмид Qiagen (Qiagen, Venlo, Netherlands).

Биохимическое расщепление очищенной плазмиды осуществляли, как подробно описано в примере 3 настоящего описания, с такой модификацией, что ДНК-мишень была заменена очищенной плазмидой в конечной концентрации 1 нМ в реакционной смеси. cr-Д(Р)НК гибридизовали с tracr-РНК (SEQ ID NO: 2), как описано в примере 3.

Смеси после биохимических реакций анализировали, разгоняя в 1% агарозном геле, окрашенным SYBR gold (Life Technologies, Grand Island, NY). Эффективность никирования вычисляли на основе исчезновения суперспирализованной плазмидной формы и появления никированной открытой кольцевой формы плазмиды (никированной плазмиды), которую отличали по сдвигу в скорости миграции плазмиды в геле.

Процентное содержание никированной плазмиды вычисляли на основании интенсивности окрашенных полос в геле, содержащих никированную плазмиду и суперспирализованную плазмиду. Интенсивности измеряли, используя значения площади, которые вычисляли с помощью FIJI (ImageJ; программа обработки изображений на языке Java с открытым исходным кодом). Процент никирования вычисляли путем деления интенсивности окрашивания на сумму интенсивностей окрашивания никированного вида плазмиды и суперспирализованного вида плазмиды.

Номера последовательностей (SEQ ID NO) для каждой cr-Д(Р)НК и sg-Д(Р)НК, используемых в данном эксперименте, показаны в таблице 12.

Таблица 12

Никирование cr-Д(Р)НК и sg-Д(Р)НК

Идентификационный номер образца	Описание	SEQ ID NO:
A	cr-Д(Р)НК	SEQ ID NO: 38
B	cr-Д(Р)НК w/18-нуклеотидный спейсер	SEQ ID NO: 135
C	cr-Д(Р)НК	SEQ ID NO: 41
D	cr-Д(Р)НК w/17-нуклеотидный спейсер	SEQ ID NO: 136
E	cr-Д(Р)НК	SEQ ID NO: 43
F	cr-Д(Р)НК w/18-нуклеотидный спейсер	SEQ ID NO: 137
H	sg-Д(Р)НК	SEQ ID NO: 127

I	контрольная sg-ПНК	SEQ ID NO: 1
H	только плазмидная мишень	-

На фигуре 8 показаны результаты оценки биохимической никирующей активности cr-Д(Р)НК или sg-Д(Р)НК с белком Cas9-D10A против плазмидной мишени. Процент никирования показан на оси y. Образцы cr-Д(Р)НК и sg-Д(Р)НК указаны на оси x, и они соответствуют идентификационным номерам образцов, показанным в таблице 12. Данные показывают способность cr-Д(Р)НК и sg-Д(Р)НК поддерживать никирующую активность белка Cas9-D10A против плазмидной мишени. Данные также показывают, что укорочение спейсерной последовательности на 5'-конце спейсера (SEQ ID NO: 135, 136 и 137) способно поддерживать никирующую активность.

Следуя инструкциям, приведенным в настоящем описании, и содержащимся в нем примерам, конструирование и подтверждение никирующей активности cr-Д(Р)НК:tracr-ПНК и sg-Д(Р)НК может быть осуществлено на практике специалистов в данной области.

Пример 10

Идентификация и скрининг CRISPR-ПНК и трансактивирующей CRISPR-ПНК

Данный пример иллюстрирует способ, с применением которого можно идентифицировать CRISPR-ПНК (cr-ПНК) и трансактивирующие CRISPR-ПНК (tracr-ПНК) системы CRISPR-Cas типа II. Предлагаемый в настоящем описании способ является адаптированным способом, описанным Chylinski с соавторами (RNA Biol; 10(5): 726-37 (2013)). Не все указанные далее стадии необходимы для скрининга, и порядок стадий не обязательно должен быть таким, как представлено.

А. Идентификация видов бактерий, содержащих систему CRISPR-Cas9 типа II.

Используя основное средство поиска локального выравнивания (BLAST, blast.ncbi.nlm.nih.gov/Blast.cgi), проводили поиск геномов различных видов, чтобы идентифицировать белки Cas9 или Cas9-подобные белки. Системы CRISPR-Cas9 типа II в высокой степени разнообразны по последовательности среди видов бактерий, однако ортологи Cas9 имеют консервативную доменную структуру центрального HNH-эндонуклеазного домена и разрезают домен RuvC/ПНК-азы H. Первичные результаты BLAST фильтровали в отношении идентифицированных доменов; неполные или укороченные последовательности отбрасывали, и идентифицировали ортологи Cas9.

После идентификации ортолога Cas9 у определенного вида, последовательности, близлежащие к последовательности, кодирующей ортолог Cas9, исследовали с помощью зондов для других белков Cas и ассоциированного блока повтор-спейсер, чтобы идентифицировать все последовательности, относящиеся к локусу CRISPR-Cas. Такую задачу можно выполнить, используя выравнивание с другими локусами CRISPR-Cas типа II, уже известными и общедоступными, используя данные о том, что близкородственные виды имеют сходную структуру локуса CRISPR-Cas9 (т.е., состав белка Cas, размер, ориентацию, локализацию матрицы, локализацию tracr-ПНК и т.д.).

В. Идентификация предполагаемой cr-ПНК и tracr-ПНК

В пределах локуса cr-ПНК легко можно идентифицировать на основе природы их повторяющихся последовательностей, разделенных фрагментами чужеродной ДНК и состава блока повтор-спейсер. Если повторяющаяся последовательность происходит из известного вида, ее идентифицируют и получают из базы данных CRISPRdb (crispr.u-psud.fr/crispr/). Если не известно, что повторяющаяся последовательность ассоциирована с определенным видом, повторяющиеся последовательности прогнозируют с применением компьютерной программы CRISPRfinder (crispr.u-psud.fr/Server/), используя последовательность, идентифицированную в качестве локуса CRISPR-Cas типа II для вида, который описан выше.

После идентификации последовательности в повторяющейся последовательности для вида, идентифицируют tracr-РНК на основании ее последовательности, комплементарной повторяющейся последовательности в блоке повтор-спейсер (tracr-последовательность, направленная против повтора). Прогностический скрининг *in silico* используют для выделения последовательности против повтора, чтобы идентифицировать ассоциированную tracr-РНК. Предполагаемые анти-повторы подвергают скринингу, например, следующим образом.

Идентифицированную повторяющуюся последовательность для данного вида используют для исследования локуса CRISPR-Cas9 в отношении последовательности против повтора (например, используя алгоритм BLASTp или подобный). Поиск обычно ограничен интронными областями локуса CRISPR-Cas9.

Идентифицированную области антиповтора подтверждают в отношении комплементарности идентифицированной повторяющейся последовательности.

Предполагаемую область антиповтора исследуют на 5'- и 3'-кнцах предполагаемого антиповтора в отношении Rho-независимого терминатора транскрипции (Transterm HP, transterm.cbcb.umd.edu/).

Таким образом, определяют, что идентифицированная последовательность, содержащая элемент антиповтора и Rho-независимый терминатор транскрипции, является предполагаемой tracr-РНК для данного вида.

С. Получение библиотеки RNA-Seq

Предполагаемые cr-РНК и tracr-РНК, которые были идентифицированы *in silico*, затем подтверждают, используя секвенирование РНК (РНК-секв.).

Клетки вида, для которого были идентифицированы предполагаемые cr-РНК и tracr-РНК, покупали из коммерческого архива (например, ATCC, Manassas, VA; DSMZ, Braunschweig, Germany).

Клетки выращивали до середины log-фазы и суммарную РНК готовили, используя реагент тризол (Sigma-Aldrich, St. Louis, MO), и обрабатывали ДНК-азой I (Fermentas, Vilnius, Lithuania).

10 мкг суммарной РНК обрабатывали, используя набор для удаления рРНК RiboZero (Illumina, San Diego, CA), и оставшуюся РНК очищали, используя набор для очистки и концентрирования РНК (Zymo Research, Irvine, CA).

Затем готовили библиотеку, используя набор для получения небольших библиотек РНК TruSeq (Illumina, San Diego, CA), следуя инструкциям производителя, результатом применения которого является присутствие адаптерных последовательностей, ассоциированных с кДНК.

Полученную в результате библиотеку кДНК секвенировали, используя секвенатор MiSeq (Illumina, San Diego, CA).

Д. Обработка данных секвенирования

Риды, полученные при секвенировании библиотеки кДНК, можно обрабатывать, используя следующий способ.

Адаптерные последовательности удаляли, используя cutadapt 1.1 (pypi.python.org/pypi/cutadapt/1.1) и 15 н. обрезали на 3'-конце ридов, чтобы улучшить качество ридов.

Риды выравнивали относительно каждого генома соответствующего вида (для которого была идентифицирована предполагаемая tracr-РНК) с разрешением несоответствия 2 нуклеотидов.

Область ридов вычисляли, используя BedTools (bedtools.readthedocs.org/en/latest/).

Программу Integrative Genomics Viewer (IGV, www.broadinstitute.org/igv/) использовали для картирования начального (5') и конечного (3') положений ридов. Общее количество

ридов, найденных для предполагаемой tracr-РНК, вычисляли, используя файл SAM выравниваний.

Данные РНК-секв. использовали для подтверждения того, что предполагаемы cr-РНК- и tracr-РНК-элемент активно транскрибируется *in vivo*. Подтвержденные
5 наилучшие варианты из совокупности скринингов *in silico* и РНК-секв. проверяли в отношении функциональной способности идентифицированных последовательностей cr-РНК и tracr-РНК поддерживать опосредованное Cas9 расщепление двунитевой ДНК-мишени, применяя способы, описанные в настоящей публикации (смотри примеры 1, 2 и 3).

10 Следуя инструкциям, приведенным в настоящем описании, и примерам, специалист в данной области может осуществить на практике идентификацию новых последовательностей cr-РНК и tracr-РНК.

Пример 11

Конструирование cr-Д(Р)НК и sg-Д(Р)НК

15 Данный пример иллюстрирует способ, с применением которого конструируют cr-Д(Р)НК и sg-Д(Р)НК на основе cr-РНК и tracr-РНК, соответственно. Не все указанные далее стадии необходимы для каждого скрининга, и порядок стадий не обязательно должен быть таким, как представлено.

Идентификацию последовательностей cr-РНК и tracr-РНК-проводника для данного
20 вида осуществляли, как описано в примере 10.

Идентифицированные последовательности cr-РНК и tracr-РНК обратно транскрибировали *in silico* в ДНК. В последовательностях cr-РНК и tracr-РНК идентифицировали элементы верхнего стебля, нижнего стебля и выпетливания. Основания РНК вводили в ДНК-последовательность последовательностей cr-ДНК и
25 tracr-ДНК с получением cr-Д(Р)НК и sg-Д(Р)НК, соответственно. Размещение, количество и распределение оснований РНК в cr-ДНК и tracr-РНК можно выбирать, используя любые компьютерные или экспериментальные способы скрининга. Конструировали набор cr-Д(Р)НК с рибонуклеотидами, помещенными в несколько разных положений в молекуле. Предпочтительно, рибонуклеотиды в нижнем стебле
30 заменяют на дезоксирибонуклеотидами в некоторых последовательностях cr-Д(Р)НК. Рибонуклеотиды заменяют на 3'-конце спейсерной последовательности в некоторых последовательностях cr-Д(Р)НК. Дополнительные последовательности cr-Д(Р)НК и sg-Д(Р)НК конструируют, например, следующим образом.

Осуществляют поиск в общедоступных архивах 3-мерных структур белков (например,
35 RCSB PDB; rcsb.org), чтобы идентифицировать структуры Cas-эндонуклеаз. В архиве осуществляют поиск в отношении файлов координат с высоким разрешением для Cas-эндонуклеаз, связанных с соответствующими им cr-РНК и tracr-РНК. Используют структурные соседние элементы, определенные по сходству последовательности или третичной структуры с представляющей интерес Cas-эндонуклеазой, если не определена
40 структура представляющей интерес Cas-эндонуклеазы. Загружают депонированные координатные файлы. Используя компьютерную программу визуализации, такую как PyMOL (PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC), анализируют координаты, чтобы идентифицировать рибозо-специфичные взаимодействия между Cas-эндонуклеазным белком и нуклеотидами cr-РНК и tracr-РНК. Положения, в которых
45 белок вступает в прямой или не прямой контакт (т.е., воду или промежуточный металл) с нуклеотидами cr-РНК и tracr-РНК, используют для идентификации предпочтительных положений в последовательностях проводников для замены дезоксирибонуклеотидов рибонуклеотидами или другими вариантами нуклеотидов.

Последовательности cr-PHK и tracr-PHK являются консервативными по сравнению с белками Cas9 из родственных видов. Выравнивание последовательности проводника с другими известными последовательностями проводников из сходных видов дает дополнительную информацию о консервативных основаниях, которые могут быть предпочтительными для рибонуклеотидов. Множественные выравнивания последовательностей cr-PHK или tracr-PHK осуществляют, используя веб-версию компьютерной программы MUSCLE (ebi.ac.uk/Tools/mas/muscle/). Затем выравнивания оценивали в отношении консервативных положений нуклеотидной последовательности вдоль остова.

Компьютерную программу для прогнозирования вторичной структуры нуклеиновой кислоты (например, RNAfold; PHK.tbi.univie.ac.at/cgi-bin/PHKfold.cgi) использовали для анализа фолдинга остова проводника. Области, в которых торсионные углы, специфичные для РНК, могут быть предпочтительными, используют для получения информации о размещении положений рибонуклеотидов в cr-ДНК и/или tracr-ДНК.

Сочетания вторичной структуры, взаимодействия белок-нуклеиновая кислота и консервативности последовательностей используют для получения информации о расположении рибонуклеотидов в последовательности cr-Д(Р)НК, tracr-Д(Р)НК и sg-Д(Р)НК. Множество конструкции cr-Д(Р)НК и tracr-Д(Р)НК тестируют, учитывая что разные конфигурации могут поддерживать разные требуемые свойства (т.е., активность, специфичность, стабильность и т.д.). cr-Д(Р)НК и tracr-Д(Р)НК могут быть соединены в одну молекулу линкером с образованием sg-Д(Р)НК. Объединение cr-Д(Р)НК и tracr-Д(Р)НК может сопровождаться уменьшением общего количества нуклеотидов на 3'-конце cr-Д(Р)НК и 5'-конце tracr-Д(Р)НК, которые вместе могут образовывать верхний стебель. В последовательностях SEQ ID NO: 138-142, 147-150, 154-157 и 161-164 показаны конструкции cr-Д(Р)НК и tracr-Д(Р)НК. В последовательностях SEQ ID NO: 143-146, 151-153, 158-160 и 165-167 показаны конструкции sg-Д(Р)НК. В таблице 13 указана идентичность последовательностей.

Таблица 13

cr-Д(Р)НК, tracr-Д(Р)НК и sg-Д(Р)НК

Идентификационный номер	Род/вид	Описание проводника
SEQ ID NO: 138	<i>Staphylococcus aureus</i>	cr-Д(Р)НК
SEQ ID NO: 139	<i>Staphylococcus aureus</i>	cr-Д(Р)НК
SEQ ID NO: 140	<i>Staphylococcus aureus</i>	cr-Д(Р)НК
SEQ ID NO: 141	<i>Staphylococcus aureus</i>	cr-Д(Р)НК
SEQ ID NO: 142	<i>Staphylococcus aureus</i>	tracr-PHK
SEQ ID NO: 143	<i>Staphylococcus aureus</i>	sg-Д(Р)НК
SEQ ID NO: 144	<i>Staphylococcus aureus</i>	sg-Д(Р)НК
SEQ ID NO: 145	<i>Staphylococcus aureus</i>	sg-Д(Р)НК
SEQ ID NO: 146	<i>Staphylococcus aureus</i>	sg-Д(Р)НК
SEQ ID NO: 147	<i>Streptococcus thermophilus</i> CRISPR-I	cr-Д(Р)НК
SEQ ID NO: 148	<i>Streptococcus thermophilus</i> CRISPR-I	cr-Д(Р)НК
SEQ ID NO: 149	<i>Streptococcus thermophilus</i> CRISPR-I	cr-Д(Р)НК
SEQ ID NO: 150	<i>Streptococcus thermophilus</i> CRISPR-I	tracr-PHK
SEQ ID NO: 151	<i>Streptococcus thermophilus</i> CRISPR-I	sg-Д(Р)НК
SEQ ID NO: 152	<i>Streptococcus thermophilus</i> CRISPR-I	sg-Д(Р)НК
SEQ ID NO: 153	<i>Streptococcus thermophilus</i> CRISPR-I	sg-Д(Р)НК
SEQ ID NO: 154	<i>Neisseria meningitidis</i>	cr-Д(Р)НК
SEQ ID NO: 155	<i>Neisseria meningitidis</i>	cr-Д(Р)НК
SEQ ID NO: 156	<i>Neisseria meningitidis</i>	cr-Д(Р)НК
SEQ ID NO: 157	<i>Neisseria meningitidis</i>	tracr-PHK
SEQ ID NO: 158	<i>Neisseria meningitidis</i>	sg-Д(Р)НК

SEQ ID NO: 159	Neisseria meningitidis	sg-Д(Р)НК
SEQ ID NO: 160	Neisseria meningitidis	sg-Д(Р)НК
SEQ ID NO: 161	Streptococcus pasteurianus	cr-Д(Р)НК
SEQ ID NO: 162	Streptococcus pasteurianus	cr-Д(Р)НК
SEQ ID NO: 163	Streptococcus pasteurianus	cr-Д(Р)НК
SEQ ID NO: 164	Streptococcus pasteurianus	tracr-РНК
SEQ ID NO: 165	Streptococcus pasteurianus	sg-Д(Р)НК
SEQ ID NO: 166	Streptococcus pasteurianus	sg-Д(Р)НК
SEQ ID NO: 167	Streptococcus pasteurianus	sg-Д(Р)НК

Последовательности предоставляли коммерческому производителю (например, Integrated DNA Technologies, Coralville, IA) для синтеза.

cr-Д(Р)НК, tracr-Д(Р)НК и sg-Д(Р)НК тестировали экспериментально для определения активности разных последовательностей в поддержании Cas9-опосредованного расщепления двунитевой ДНК-мишени с применением способов, описанных в настоящей публикации (смотри примеры 1, 2 и 3).

Следуя инструкциям, приведенным в настоящем описании, и примерам, специалист в данной области может осуществить на практике конструирование и подтверждение новых последовательностей cr-Д(Р)НК, tracr-Д(Р)НК и sg-Д(Р)НК.

Пример 12

Конструирование элементов Cpf1 cr-Д(Р)НК и sg-Д(Р)НК типа V и применение с использованием Cpf1 для модификации ДНК

В таблицах 14 и 15 ниже приведены примеры двойных cr-Д(Р)НК- и sg-Д(Р)НК-проводников для применения в системе CRISPR типа V. Ссылка на иллюстративные фигуры и номера последовательностей (SEQ ID NO) никоим образом не предназначена для ограничения, и специалисту в данной области понятно, что на основе описания, приведенного в таблицах 14, 15, и соответствующих последовательностей (SEQ ID NO) и иллюстративных фигур могут быть сконструированы двойные cr-Д(Р)НК- и sg-Д(Р)НК-проводники для применения в системах CRISPR типа V, нацеленных на любую требуемую последовательность в нуклеиновой кислоте-мишени.

Таблица 14

Описание 5'- и 3'-элементов cr-Д(Р)НК типа V и сочетаний, используемых для образования двойных cr-Д(Р)НК-проводников и направления активности Cpf1 на представляющую интерес ДНК-последовательность

Описание последовательности	Иллюстративная фигура	SEQ ID NO:
5'-элемент Cpf1 cr-РНК типа V	12B, 13D, 13E, 13H	SEQ ID NO: 168
5'-элемент Cpf1 cr-Д(Р)НК типа V	12C, 13B, 13C, 13F, 13G	SEQ ID NO: 169
Защищенный фосфоротиоатом 5'-элемент Cpf1 cr-РНК типа V	12B, 13D, 13E, 13H	SEQ ID NO: 170
Защищенный фосфоротиоатом 5'-элемент Cpf1 cr-Д(Р)НК типа V	12C, 13B, 13C, 13F, 13G	SEQ ID NO: 171
3'-элемент Cpf1 cr-РНК типа V с 25 нуклеотидной нацеленной на РНК-мишень областью	12D	SEQ ID NO: 172
3'-элемент Cpf1 cr-РНК типа V с 20-нуклеотидной нацеленной на РНК-мишень областью	12D	SEQ ID NO: 173
Защищенный фосфоротиоатом 3'-элемент Cpf1 cr-РНК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью	12D	SEQ ID NO: 174
Защищенный фосфоротиоатом 3'-элемент Cpf1 cr-РНК типа V с 20-нуклеотидной нацеленной на РНК-мишень областью	12D	SEQ ID NO: 175
3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью	12F, 13E, 13F	SEQ ID NO: 176
3'-элемент Cpf1 cr-Д(Р)НК типа V с 25 нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 177
3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 178
3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 179

5	3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью	12E, 13C, 13D	SEQ ID NO: 180
	Защищенный фосфоротиоатом 3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью	12E, 13C, 13D	SEQ ID NO: 181
	3'-элемент Cpf1 cr-Д(Р)НК типа V с 20 нуклеотидной нацеленной на РНК-мишень областью	12E, 13C, 13D	SEQ ID NO: 182
	Защищенный фосфоротиоатом 3'-элемент Cpf1 cr-Д(Р)НК типа V с 20-нуклеотидной нацеленной на РНК-мишень областью	12E, 13C, 13D	SEQ ID NO: 183
	3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью	12G, 13G, 13H	SEQ ID NO: 184
10	3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 185
	3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 186
	3'-элемент Cpf1 cr-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	12H, 12I	SEQ ID NO: 187
15	Двойной Cpf1 cr-РНК-проводник типа V, содержащий 3'- и 5'-элементы	13A	SEQ ID NO: 168; SEQ ID NO: 172
	Двойной проводник Cpf1 cr-РНК типа V, содержащий защищенные фосфоротиоатом 3'- и 5'-элементы	13A	SEQ ID NO: 170; SEQ ID NO: 173
	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы	13B	SEQ ID NO: 169; SEQ ID NO: 172
	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы		SEQ ID NO: 169; SEQ ID NO: 180
	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы	13D	SEQ ID NO: 168; SEQ ID NO: 180
20	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы	13E	SEQ ID NO: 168; SEQ ID NO: 176
	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы		SEQ ID NO: 169; SEQ ID NO: 176
	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы	13G	SEQ ID NO: 169; SEQ ID NO: 184
25	Двойной проводник Cpf1 cr(Д)РНК типа V, содержащий 3'- и 5'-элементы	13H	SEQ ID NO: 168; SEQ ID NO: 184

Таблица 15

Описание конструкций sg-Д(Р)НК типа V

	Описание последовательности	Иллюстративная фигура	SEQ ID NO:
30	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью	10A	SEQ ID NO: 188
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью	10B	SEQ ID NO: 189
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на РНК-мишень областью		SEQ ID NO: 190
35	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью	11D	SEQ ID NO: 191
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью	11B	SEQ ID NO: 192
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 193
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 194
40	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 195
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью	11A	SEQ ID NO: 196
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 197
45	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 198
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 199
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК-мишень областью		SEQ ID NO: 200
	Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 201

Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 202
Cpf1 sg-Д(Р)НК типа V с 25-нуклеотидной нацеленной на ДНК/РНК-мишень областью	11E	SEQ ID NO: 203

А. Конструирование элементов Cpf1 cr-Д(Р)НК и sg-Д(Р)НК типа V

Ортологи Cpf1 идентифицировали, используя программы для анализа последовательностей, такие как PSI-BLAST, PHI-BLAST и HMMer. После идентификации ортолога Cpf1 искали соседние последовательности, чтобы идентифицировать ассоциированный блок CRISPR. Последовательности cr-РНК идентифицировали в виде повторяющихся последовательностей, расположенных в блоке CRISPR, как описано Zetsche с соавторами (Cell; 163(3): 759-71(2015)). Последовательности cr-РНК типа V содержат структуру стебель-петля в последовательности повтора, расположенную с 5'-стороны от последовательности нацеленной на мишень области. Структура стебель-петля содержит 5'-элемент и 3'-элемент. Идентифицировали последовательности 5' элемента, 3'-элемента и петли cr-РНК. Последовательности указанных элементов cr-РНК обратно транскрибировали *in silico* в ДНК. Конструировали 5'-элементы, содержащие смеси рибонуклеотидов и дезоксирибонуклеотидов. Примеры 5'-элементов показаны на фигуре 12, фигуре 13 и в таблице 14. Конструировали 3'-элементы, содержащие смеси рибонуклеотидов и дезоксирибонуклеотидов. Примеры 3'-элементов показаны на фигуре 12, фигуре 13 и в таблице 14. Последовательности нацеленной на мишень области выбирали вблизи последовательностей РАМ в представляющей интерес ДНК и присоединяли к 3'-концу 3'-элементов cr-РНК. Конструировали последовательности нацеленной на мишень области, содержащие ДНК-, ДНК- и РНК- или РНК- нуклеотиды. За счет объединения 3'-элементов cr-Д(Р)НК и 5'-элементов cr-Д(Р)НК 5' вместе (таблица 14, фигура 12, фигура 13) с образованием двойных cr-Д(Р)НК-проводников типа V, Cpf1 направляли на разрезание последовательностей нуклеиновой кислоты-мишени в представляющей интерес нуклеиновой кислоте-мишени. Конструировали набор cr-Д(Р)НК для тестирования с рибонуклеотидами, помещенными в нескольких разных положениях в последовательностях cr-Д(Р)НК. Предпочтительно, рибонуклеотиды в 3'-стебле и 5'-стебле заменяли дезоксирибонуклеотидами в некоторых последовательностях cr-Д(Р)НК. Рибонуклеотиды заменяли на 5'-конце последовательности нацеленной на мишень области в некоторых последовательностях cr-Д(Р)НК.

Используя сочетания нацеленной на мишень области, 3'- элементов и 5'-элементов, связанных последовательностью петли, конструировали разные варианты sg-Д(Р)НК. Размещение, количество и распределение оснований РНК в sg-Д(Р)НК можно выбрать, используя любые компьютерные или экспериментальные способы скрининга. Конструировали набор sg-Д(Р)НК с рибонуклеотидами, помещенными в нескольких разных положениях в sg-Д(Р)НК. Предпочтительно рибонуклеотидами заменяли дезоксирибонуклеотиды в 3'-стебле и 5'-стебле в некоторых последовательностях sg-Д(Р)НК. Рибонуклеотиды использовали для замены на 5'-конце последовательности нацеленной на мишень области в некоторых последовательностях sg-Д(Р)НК. Примеры сконструированных sg-Д(Р)НК перечислены в таблице 15 и показаны на фигурах 10А-С и фигурах 11А-Е.

Далее использовали последовательности sg-Д(Р)НК, но понятно, что вместо sg-Д(Р)НК можно использовать пары 3'- и 5'-элементов cr-Д(Р)НК (примеры которых показаны в таблице 14).

В. Расщепление последовательностей нуклеиновых кислот с использованием Cpf1 и sg-Д(Р)НК

Cpf1 sg-Д(Р)НК можно использовать вместе с Cpf1, чтобы направлять к мишени и разрезать последовательности нуклеиновых кислот. Нуклеиновой кислотой-мишенью является РНК, геномная ДНК, плазмидная ДНК или амплифицированная ДНК.

Амплифицированная ДНК-мишень может быть получена, как описано в примере 2.

- 5 Синтезируют последовательности sg-Д(Р)НК, содержащие спейсерные последовательности, нацеленные на представляющие интерес последовательности в ДНК-мишени. Анализы расщепления осуществляют, как описано Zetsche с соавторами (2015), и анализируют, применяя способы, описанные в примере 3. Кратко, нуклеиновую кислоту-мишень инкубируют с Cpf1 и последовательностью или последовательностями
- 10 sg-Д(Р)НК в подходящем буфере, выбранном для поддержания активности Cpf1. Анализируют нуклеиновую кислоту, чтобы определить, произошло ли расщепление, как описано в примере 3. Можно использовать два или больше комплексов Cpf1/sg-Д(Р)НК для разрезания на фрагменты ДНК из ДНК-мишени. Фрагмент ДНК имеет выступающие концы и может быть лигирован с адапторами, имеющими
- 15 комплементарные последовательности, или с векторами после его отделения от исходной ДНК.

С. Редактирование генома с использованием рибонуклеопротеидных комплексов Cpf1 sg-Д(Р)НК

- Конструируют экспрессирующий вектор *E. coli* посредством синтеза
- 20 оптимизированной по кодонам открытой рамки считывания, кодирующей Cpf1, и клонирования открытой рамки считывания в экспрессирующей плазмиде (например, pET27b). Кодирующая последовательность может содержать аффинную метку для очистки белка и последовательность NLS на С-конце для управления ядерной локализацией в эукариотических клетках. Белок Cpf1 может быть экспрессирован в *E.*
- 25 *coli* с экспрессирующего вектора и очищен с использованием сочетания аффинной, ионообменной и эксклюзионной хроматографии по размеру. Очищенный белок концентрируют до 10 мг/мл и объединяют с sg-Д(Р)НК, чтобы получить рибонуклеопротеидный комплекс. 200 пмоль Cpf1 объединяют в отдельных пробирках с реакционной смесью, содержащих 50 пмоль, 100 пмоль, 200 пмоль, 400 пмоль, 600
- 30 пмоль, 800 пмоль, 1000 пмоль sg-Д(Р)НК и буфер для реакции. Cpf1-sg-Д(Р)НК-комплексы вводят с помощью электропорации в двух повторах в клетки НЕК293 согласно способам, описанным в примере 7. Клетки выращивают при 37°C и собирают геномную ДНК из каждой реакционной смеси через 4, 8, 16, 24, 48 и 72 часа. Геномную ДНК анализируют, используя ПЦР и секвенирование Illumina, чтобы определить, что
- 35 геном был отредактирован согласно способам, описанным в примере 7.

Д. редактирование генома с использованием экспрессирующих Cpf1 векторов и sg-Д(Р)НК в эукариотических клетках

- Экспрессирующий вектор млекопитающих может быть сконструирован посредством синтеза оптимизированной по кодонам открытой рамки считывания, кодирующей Cpf1,
- 40 и клонирования открытой рамки считывания в подходящей экспрессирующей плазмиде млекопитающих (например, pcDNA3.1). Кодирующая последовательность может содержать НА-аффинную метку для очистки или выявления белка и последовательность NLS на С-конце для управления ядерной локализацией в эукариотических клетках. Кодирующая последовательность может быть оперативно связана с промотором CMV
- 45 в плазмиде. Cpf1-экспрессирующие плазмиды объединяют в отдельных пробирках для реакции, содержащих 50 пмоль, 100 пмоль, 200 пмоль, 400 пмоль, 600 пмоль, 800 пмоль, 1000 пмоль sg-Д(Р)НК и буфер для реакции. Реакционные смеси вводят с помощью электропорации в двух повторах в клетки НЕК293 согласно способам, описанным в

примере 7. Клетки выращивают при 37°C и собирают геномную ДНК из каждой реакционной смеси через 4, 8, 16, 24, 48 и 72 часа. Геномную ДНК анализируют, используя ПЦР и секвенирование Illumina, чтобы определить, что геном был отредактирован согласно способам, описанным в примере 7.

Пример 13

Модификация in planta зародышей кукурузы

Данный пример иллюстрирует способ, посредством которого единичный Д(Р)НК-проводник может быть использован для модификации зародышей кукурузы. Способ, приведенный в настоящем описании, является адаптированным способом, который описан Svitashv с соавторами (Plant Physiol; 169(2): 931-945 (2015)). Не все указанные далее стадии необходимы для каждого скрининга, и порядок стадий не обязательно должен быть таким, как представлено.

Данный пример иллюстрирует применение единичных Д(Р)НК-проводников для направления Cas-эндонуклеаз на расщепление хромосомной ДНК в зародышах кукурузы. Конструировали шесть единичных Д(Р)НК-проводников (sg-Д(Р)НК), нацеленных на область вблизи гена *liguleless 1* и ген фертильности *Ms45* (таблица 16), и доставляли в линию кукурузы, содержащую предварительно интегрированный конститутивно экспрессирующийся ген *Cas9* *S. pyogenes*. Геномные локусы кукурузы *liguleless 1* и *Ms45* исследовали, используя глубокое секвенирование, в отношении присутствия мутаций, индуцированных sg-Д(Р)НК/Cas9-опосредованным расщеплением.

Таблица 16

sg-Д(Р)НК, нацеленные на мишени *Liguleless 1* и *Ms45* кукурузы

Локус	Положение	Последовательность (основания РНК указаны в скобках. Фосфоротиоатные связи показаны значком *)	SEQ ID NO:
<i>liguleless 1</i>	Хр. 2: 28,45 cM	5' - T*A*CGCGTACG CGTA(C)[G][U][G][U][G] [G][U][U][U][U][A][G][A] [G][C] [U][A][G][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U] [G][C][U] - 3'	204
<i>liguleless 1</i>	Хр. 2: 28,45 cM	5' - T*A*CGCGTACG CGTA(C)[G][U][G][U][G] [G][U][U][U][U][A][G][A] GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][U] [A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U][G][G] [C][A][C][C][G][A][G][U][C][G] [G][U] [G][C][U] - 3'	205
<i>liguleless 1</i>	Хр. 2: 28,45 cM	5' - T*A*CGCGTACG CGTA(C)[G][U][G][U][G] [G][U][U][U][U][A][G][A] GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][U] [A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U][G][G] CACCG[A][G][U]CG GTG[C][U] - 3'	206
<i>Ms45</i>	Хр. 9: 119,15 cM	5' - G*G*CCGAGGTC GACT(A)[C][C][G][G][C] [G][U][U][U][U][A][G][A] [G][C] [U][A][G][A][A][U][A][G][C] [A][A][G][U][U][A][A][A][U] [A][A][G][G][C][U][A][G][U][C] [C][G][U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U] [G][G][C][A][C][C][G][A][G][U] [C][G][G][U] [G][C][U] - 3'	224
<i>Ms45</i>	Хр. 9: 119,15 cM	5' - G*G*CCGAGGTC GACT(A)[C][C][G][G][C] [G][U][U][U][U][A][G][A] GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][U] [A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U][G][G] [C][A][C][C][G][A][G][U][C][G] [G][U] [G][C][U] - 3'	225
<i>Ms45</i>	Хр. 9: 119,15 cM	5' - G*G*CCGAGGTC GACT(A)[C][C][G][G][C] [G][U][U][U][U][A][G][A] GC TATGCT[G][A][A][A] AGCATAGC[A][A] [G][U][U][A][A][A][U] [A][A] [G][G][C][U][A][G][U][C][C][G] [U][U][A][U][C][A][A][C] [U][U] [G][A][A][A][A][G][U][G][G] CACCG[A][G][U]CG GTG[C][U] - 3'	226

Линию кукурузы с предварительно интегрированным конститутивно экспрессирующимся геном *Cas9* *S. pyogenes* создавали, как описано Svitashv с соавторами (2015).

Дизайн sg-Д(Р)НК предоставляли коммерческому производителю для синтеза (Eurofins Scientific, Huntsville, AL).

Sg-РНК (SEQ ID NO: 207 и 227) конструировали, как описано в примере 1.

Основанную на биобаллистике трансформацию незрелых зародышей кукурузы (IME), полученных из линии, конститутивно экспрессирующей Cas9 *S. pyogenes*, осуществляли, используя sg-Д(Р)НК, как описано Svitashv с соавторами (2015). Кратко, 100 нг каждой

sg-Д(Р)НК доставляли в 60-90 IME в присутствии стимулирующих клеточной деление генов, ZmODP2 (публикация патента США № 20050257289) и ZmWUS2 (патент США № 7256322), как описано Ananiev с соавторами (Chromosoma; 118(2): 157-77 (2009)). Поскольку трансформация частицами из пушки может сильно варьировать, кассету экспрессии ДНК визуально селективируемого маркера, MoPAT-DsRED, также доставляли совместно с генами, стимулирующими клеточное деление, как описано Svitashv с соавторами (2015). Зародыши, трансформированные 100 нг транскрибированного T7 единичного РНК-проводника (sg-РНК), нацеленного на ту же самую область для расщепления (SEQ ID NO: 207 и 227), служили в качестве положительного контроля, а зародыши, трансформированные только кассетами экспрессии ZmODP2, ZmWUS2 и Mo-PAT-DsRED, служили в качестве отрицательного контроля. Через 3 дня 20-30 наиболее однородно трансформированных зародышей при каждой обработке отбирали на основе флуоресценции DsRED, объединяли и экстрагировали суммарную геномную ДНК. Область, окружающую предполагаемый участок-мишень, амплифицировали в ПЦР с использованием смеси Phusion® HighFidelity PCR Master Mix (M0531L, New England Biolabs, Ipswich, MA), добавляя последовательности, необходимые для специфичных для ампликонов штрихкодов, и осуществляя секвенирование Illumina с использованием праймеров «с наращенными концами», осуществляя два раунда ПЦР. Праймеры, используемые в первой реакции ПЦР, показаны в таблице 17, а праймеры, используемые во второй реакции ПЦР представляли собой последовательности SEQ ID NO: 214 и 215.

Таблица 17

Последовательности ПЦР-праймеров

Идентификационный номер	Образец	Праймеры
Набор праймеров для штрихкодирования-37	SEQ ID NO: 204	SEQ ID NO: 208, 209
Набор праймеров для штрихкодирования-38	SEQ ID NO: 205	SEQ ID NO: 208, 210
Набор праймеров для штрихкодирования-39	SEQ ID NO: 206	SEQ ID NO: 208, 211
Набор праймеров для штрихкодирования-40	SEQ ID NO: 207	SEQ ID NO: 208, 212
Набор праймеров для штрихкодирования-41	Без РНК-проводника (отрицательный контроль)	SEQ ID NO: 208, 213
Набор праймеров для штрихкодирования-42	SEQ ID NO: 224	SEQ ID NO: 228, 229
Набор праймеров для штрихкодирования-43	SEQ ID NO: 225	SEQ ID NO: 228, 230
Набор праймеров для штрихкодирования-44	SEQ ID NO: 226	SEQ ID NO: 228, 231
Набор праймеров для штрихкодирования-45	SEQ ID NO: 227	SEQ ID NO: 228, 232
Набор праймеров для штрихкодирования-46	Без РНК-проводника (отрицательный контроль)	SEQ ID NO: 228, 233

Полученные в результате ПЦР-амплификации продукты очищали, используя центрифужную колонку для очистки продуктов ПЦР Qiagen, концентрацию измеряли в основанном на красителе Hoechst флуориметрическом анализе, объединяли в эквимольном соотношении и осуществляли глубокое секвенирование отдельных ридов длиной 100 нуклеотидов на персональном секвенаторе Illumina MiSeq с использованием пиковой концентрации PhiX-контроля v3 (Illumina, FC-110-3001) 25% (об./об.) для

коррекции смещения последовательности. В качестве мутантных классифицировали только те риды, которые имели ≥ 1 -нуклеотидный индел, возникающий в пределах 10-нуклеотидного окна, расположенного по центру ожидаемого участка расщепления, и не найденный на сходном уровне в отрицательном контроле. Подсчитывали мутантные риды с одинаковой мутацией и сокращали до одного рида и визуально подтверждали, как имеющие мутацию, возникающую в ожидаемом участке расщепления. Затем общее количество визуально подтвержденных мутаций использовали для вычисления количества мутантных ридов в процентах на основе общего количества ридов соответствующей длины, содержащих точное совпадение со штрихкодом и прямым праймером.

Как показано в таблице 18, мутации были выявлены при всех обработках, что свидетельствует о том, что sg-Д(Р)НК могут быть использованы для направления Cas-эндонуклеаз на расщепление хромосомной ДНК в клетках кукурузы. Кроме того, в случае некоторых конструкций sg-Д(Р)НК (SEQ ID NO: 205 и 226) наблюдали частоту мутаций, близкую к частоте в случае T7-транскрибированной sg-РНК (SEQ ID NO: 207 и 227). Примеры мутаций, полученных с использованием sg-Д(Р)НК, показаны на фигуре 14А (соответствующей последовательностям SEQ ID NO: 217-223, при этом последовательность SEQ ID NO: 216 является эталонной последовательностью кукурузы, содержащей локус-мишень *liguleless 1*) и на фигура 14В (соответствующей последовательностям SEQ ID NO: 235-254, при этом последовательность SEQ ID NO: 234 является эталонной последовательностью кукурузы, содержащей локус-мишень *Ms45*).

Таблица 18

Мутантные риды в локусах-мишенях *liguleless 1* и *Ms45* кукурузы, полученные с использованием системы sg-Д(Р)НК/Cas-эндонуклеаза, по сравнению с системой sg-РНК/Cas

Обработка	Общее количество ридов	Количество мутантных ридов
Liguleless 1 без РНК-проводника (отрицательный контроль)	2849145	0
SEQ ID NO. 207	3155695	552
SEQ ID NO. 204	2816705	5
SEQ ID NO. 205	3053967	192
SEQ ID NO. 206	2979282	9
Ms45 без РНК-проводника (отрицательный контроль)	1248142	16
SEQ ID NO. XX4	1194050	8,784
SEQ ID NO. XX1	1192758	190
SEQ ID NO. XX2	1206632	114
SEQ ID NO. XX3	1192110	878

Хотя в вышеприведенном описании приедено описание и примеры конкретных вариантов осуществления настоящего изобретения, оно никоим образом не предназначено для ограничения, и специалист в данной области может модифицировать раскрытые примеры, чтобы адаптировать конкретный способ, композицию или стадию для достижения требуемого результата в объеме настоящего изобретения. Предполагается, что все такие модификации входят в объем настоящего изобретения.

(57) Формула изобретения

1. Единичный полинуклеотид CRISPR класса 2, содержащий:
нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК);

активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК;

и где указанная активирующая область содержит стебель и выпетливание и способна взаимодействовать с белком Cas9;

5 при этом указанный единичный полинуклеотид предназначен для применения с использованием системы CRISPR класса 2.

2. Единичный полинуклеотид CRISPR класса 2 по п.1, где нацеленная на мишень область содержит смесь ДНК и РНК.

3. Единичный полинуклеотид CRISPR класса 2 по п.1, где нацеленная на мишень
10 область содержит смесь ДНК и РНК, и активирующая область содержит смесь ДНК и РНК.

4. Единичный полинуклеотид CRISPR класса 2 по п.1, где нацеленная на мишень область содержит смесь ДНК и РНК, и активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов,
15 фосфородитиоатов, сложных фосфотриэфиров, сложных аминокилфосфотриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и
20 боранофосфатов.

5. Единичный полинуклеотид CRISPR класса 2 по п.1, где указанная активирующая область содержит нижний стебель, выпетливание и верхний стебель.

6. Система CRISPR класса 2, содержащая:

(i) единичный полинуклеотид, содержащий (а) нацеленную на мишень область,
25 содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК); и (b) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК; и где указанная активирующая область содержит стебель и выпетливание и способна взаимодействовать с белком Cas9; и
(ii) Cas9.

30 7. Система CRISPR класса 2 по п.6, где указанная нацеленная на мишень область содержит смесь ДНК и РНК.

8. Система CRISPR класса 2 по п.6, где указанная активирующая область содержит смесь ДНК и РНК.

9. Система CRISPR класса 2 по п.6, где указанная нацеленная на мишень область
35 содержит смесь ДНК и РНК, и указанная активирующая область содержит смесь ДНК и РНК.

10. Система CRISPR класса 2 по п.6, дополнительно содержащая донорный полинуклеотид.

11. Система CRISPR класса 2 по п.6, где указанная нацеленная на мишень область
40 содержит смесь ДНК и РНК, и указанная активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфотриэфиров, сложных аминокилфосфотриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов,
45 тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

12. Единичный полинуклеотид CRISPR класса 2 по п.6, где указанная активирующая

область содержит нижний стебель, выпетливание и верхний стебель.

13. Способ модификации молекулы нуклеиновой кислоты-мишени в организме, выделенной клетке или *in vitro*, включающий: введение в клетку (i) единичного полинуклеотида, содержащего (a) нацеленную на мишень область, имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте-мишени, и (b) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую дезоксирибонуклеиновую кислоту (ДНК) и стебель и выпетливание; и (ii) Cas9, где указанный Cas9 связывается с активирующей областью в указанном единичном полинуклеотиде, где нацеленная на мишень область в указанном единичном полинуклеотиде гибридизуется с последовательностью-мишенью и где нуклеиновая кислота-мишень подвергается расщеплению.

14. Способ по п.13, где Cas9 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cas9.

15. Способ по п.13, где единичный полинуклеотид и Cas9 образуют комплекс перед введением в клетку.

16. Способ по п.13, где Cas9 содержит сигнал ядерной локализации (NLS).

17. Способ по п.13, где единичный полинуклеотид и Cas9 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

18. Способ по п.13, где активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоалкилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

19. Способ по п.13, где активирующая область содержит нижний стебель, выпетливание и верхний стебель.

20. Способ по п.13, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

21. Способ по п.13, где нацеленная на мишень область содержит рибонуклеиновую кислоту (РНК).

22. Способ по п.13, дополнительно включающий введение донорного полинуклеотида в указанную клетку.

23. Способ модуляции транскрипции по меньшей мере одного гена в молекуле нуклеиновой кислоты-мишени в организме, выделенной клетке или *in vitro*, где указанный способ включает: введение в клетку (i) единичного полинуклеотида, содержащего (a) нацеленную на мишень область, имеющую конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в открытой рамке считывания по меньшей мере одного гена или последовательности-мишени в последовательности промотора по меньшей мере одного гена, и (b) активирующую область, смежную с указанной нацеленной на мишень областью, содержащую дезоксирибонуклеиновую кислоту (ДНК) и стебель и выпетливание; и (ii) Cas9, где указанный Cas9 не имеет нуклеазной активности, где указанный Cas9 связывается с активирующей областью в

указанном единичном полинуклеотиде, где нацеленная на мишень область в указанном единичном полинуклеотиде гибридизуется с последовательностью-мишенью, и где модулируется транскрипция по меньшей мере одного гена в указанной молекуле нуклеиновой кислоты-мишени.

5 24. Способ по п.23, где Cas9 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cas9.

25. Способ по п.23, где единичный полинуклеотид и Cas9 образуют комплекс перед введением в клетку.

26. Способ по п.23, где Cas9 содержит сигнал ядерной локализации (NLS).

10 27. Способ по п.23, где единичный полинуклеотид и Cas9 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

28. Способ по п.23, где активирующая область содержит соединение, выбранное из
15 группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфотриэфиров, сложных аминоалкилфосфотриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных
20 тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

29. Способ по п.23, где активирующая область содержит нижний стебель, выпетливание и верхний стебель.

30. Способ по п.23, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки
25 беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

31. Способ по п.23, где нацеленная на мишень область содержит РНК.

32. Набор из двух полинуклеотидов CRISPR класса 2, содержащий:

(i) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид
30 CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК) и активирующую область, смежную с указанной нацеленной на мишень областью; и

(ii) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая
35 область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны
40 гибридизоваться друг с другом с образованием активирующей области дуплекса, где указанная активирующая область дуплекса содержит ДНК и стебель и выпетливание, и где указанная активирующая область дуплекса способна связываться с Cas9.

33. Набор из двух полинуклеотидов CRISPR класса 2 по п.32, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 содержит смесь ДНК и РНК.

45 34. Набор из двух полинуклеотидов CRISPR класса 2 по п.32, где указанная нацеленная на мишень область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфотриэфиров, сложных аминоалкилфосфотриэфиров, алкилфосфонатов, 5'-

алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

5 35. Набор из двух полинуклеотидов CRISPR класса 2 по п.32, где указанная активирующая область дуплекса содержит нижний стебель, выпетливание и верхний стебель.

36. Система CRISPR класса 2, включающая:

10 (i) набор из двух полинуклеотидов CRISPR класса 2, содержащий (а) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК) и активирующую область, смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит

15 активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с

20 образованием активирующей области дуплекса, где указанная активирующая область дуплекса содержит ДНК и стебель и выпетливание, и где указанная активирующая область дуплекса способна связываться с Cas9; и

(ii) Cas9.

37. Система CRISPR класса 2 по п.36, где указанная активирующая область дуплекса

25 содержит нижний стебель, выпетливание и верхний стебель.

38. Система CRISPR класса 2 по п.36, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 содержит смесь ДНК и РНК.

39. Система CRISPR класса 2 по п.36, где указанная активирующая область дуплекса содержит смесь ДНК и РНК.

30 40. Система CRISPR класса 2 по п.36, где указанная нацеленная на мишень область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфотриэфиров, сложных аминоалкилфосфотриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата,

35 аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

41. Система CRISPR класса 2 по п.36, дополнительно содержащая донорный полинуклеотид.

40 42. Способ модификации молекулы нуклеиновой кислоты-мишени в организме, выделенной клетке или *in vitro*, где указанный способ включает:

введение в клетку

(i) набора из двух полинуклеотидов CRISPR класса 2, содержащего (а) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса

45 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и где указанная нацеленная на мишень область имеет конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в нуклеиновой кислоте-мишени, активирующую область,

смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, и где указанная активирующая область дуплекса содержит ДНК и стебель и выпетливание; и

(ii) Cas9, где указанный Cas9 связывается с указанной активирующей областью дуплекса двух полинуклеотидов CRISPR класса 2, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 гибридизуется с указанной последовательностью-мишенью, и где указанная нуклеиновая кислота-мишень подвергается расщеплению.

43. Способ по п.42, где Cas9 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cas9.

44. Способ по п.42, где набор двух полинуклеотидов CRISPR класса 2 и Cas9 образуют нуклеопротеиновый комплекс перед введением в клетку.

45. Способ по п.42, где Cas9 содержит сигнал ядерной локализации (NLS).

46. Способ по п.42, где набор из двух полинуклеотидов CRISPR класса 2 и Cas9 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид: нуклеиновая кислота или их комбинаций.

47. Способ по п.42, где указанная активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоалкилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

48. Способ по п.42, где указанная активирующая область дуплекса содержит нижний стебель, выпетливание и верхний стебель.

49. Способ по п.42, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

50. Способ по п.42, где нацеленная на мишень область содержит смесь ДНК и РНК.

51. Способ по п.42, дополнительно включающий введение донорного полинуклеотида в клетку.

52. Способ модуляции транскрипции по меньшей мере одного гена в молекуле нуклеиновой кислоты-мишени в организме, выделенной клетке или *in vitro*, где указанный способ включает:

введение в клетку

(i) набора из двух полинуклеотидов CRISPR класса 2, содержащего (a) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и где указанная нацеленная на

мишень область имеет конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в открытой рамке считывания по меньшей мере одного гена или последовательности-мишени в последовательности промотора по меньшей мере одного гена, и активирующую область, смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, и где указанная активирующая область дуплекса содержит ДНК и стебель и выпетливание; и

(ii) Cas9, где указанный Cas9 не имеет нуклеазной активности, где указанный Cas9 связывается с указанной активирующей областью дуплекса двух полинуклеотидов CRISPR класса 2, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 гибридизуется с указанной последовательностью-мишенью, и где модулируется транскрипция по меньшей мере одного гена в указанной молекуле нуклеиновой кислоты-мишени.

53. Способ по п.52, где Cas9 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cas9.

54. Способ по п.52, где набор из двух полинуклеотидов CRISPR класса 2 и Cas9 образуют нуклеопротеиновый комплекс перед введением в клетку.

55. Способ по п.52, где Cas9 содержит сигнал ядерной локализации (NLS).

56. Способ по п.52, где набор из двух полинуклеотидов CRISPR класса 2 и Cas9 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид: нуклеиновая кислота или их комбинаций.

57. Способ по п.52, где указанная активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоалкилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

58. Способ по п.52, где указанная активирующая область дуплекса содержит нижний стебель, выпетливание и верхний стебель.

59. Способ по п.52, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

60. Способ по п.52, где указанная нацеленная на мишень область содержит ДНК и РНК.

61. Единичный полинуклеотид CRISPR класса 2, содержащий:

(i) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК), и активирующую область, смежную с указанной нацеленной на мишень областью, где указанная активирующая область содержит ДНК или рибонуклеиновую

кислоту (РНК);

или

(ii) нацеленную на мишень область, содержащую ДНК или РНК, и активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК, где указанная активирующая область содержит структуру стебель-петля и способна связываться с Cpf1.

62. Единичный полинуклеотид CRISPR класса 2 по п.61, где нацеленная на мишень область содержит смесь ДНК и РНК.

63. Единичный полинуклеотид CRISPR класса 2 по п.61, где нацеленная на мишень область содержит смесь ДНК и РНК, и активирующая область содержит смесь ДНК и РНК.

64. Единичный полинуклеотид CRISPR класса 2 по п.61, где указанная активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминокилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боронофосфатов.

65. Система CRISPR класса 2, включающая:

(i) единичный полинуклеотид, содержащий

(a) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК), и активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК или рибонуклеиновую кислоту (РНК);

или

(b) нацеленную на мишень область, содержащую ДНК или РНК, и активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК, где указанная активирующая область содержит структуру стебель-петля и способна связываться с Cpf1; и

(ii) Cpf1.

66. Система CRISPR класса 2 по п.65, где указанная нацеленная на мишень область содержит смесь ДНК и РНК.

67. Система CRISPR класса 2 по п.65, где указанная нацеленная на мишень область содержит смесь ДНК и РНК, и указанная активирующая область содержит смесь ДНК и РНК.

68. Система CRISPR класса 2 по п.65, дополнительно содержащая донорный полинуклеотид.

69. Система CRISPR класса 2 по п.65, где указанная активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминокилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боронофосфатов.

70. Способ модификации молекулы нуклеиновой кислоты-мишени, включающий: приведение указанной молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, в контакт с

(i) единичным полинуклеотидом, содержащим (а) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК), и активирующую область, смежную с указанной нацеленной на мишень областью, где указанная активирующая область содержит ДНК или рибонуклеиновую кислоту (РНК); или (b) нацеленную на мишень область, содержащую ДНК или РНК, и активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК, где указанная нацеленная на мишень область имеет конфигурацию, позволяющую гибридизоваться с указанной последовательностью-мишенью, и где указанная активирующая область содержит структуру стебель-петля;

и

(ii) Cpf1,

где указанный Cpf1 связывается с активирующей областью указанного единичного полинуклеотида, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит (A) *in vitro*, (B) в клетке или (C) в выделенной клетке, и где указанная молекула нуклеиновой кислоты-мишень подвергается расщеплению.

71. Способ по п.70, где указанная молекула нуклеиновой кислоты-мишень содержит ДНК.

72. Способ по п.70, где указанная молекула нуклеиновой кислоты-мишень содержит РНК.

73. Способ по п.70, где указанная нацеленная на мишень область содержит смесь ДНК и РНК.

74. Способ по п.70, где указанная нацеленная на мишень область содержит смесь ДНК и РНК, и указанная активирующая область содержит смесь ДНК и РНК.

75. Способ по п.70, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит *in vitro*.

76. Способ по п.70, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит в клетке.

77. Способ по п.70, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит в выделенной клетке.

78. Способ по п.70, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

79. Способ по п.70, дополнительно включающий обеспечение донорного полинуклеотида.

80. Способ по п.70, где Cpf1 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cpf1.

81. Способ по п.70, где единичный полинуклеотид и Cpf1 образуют комплекс перед введением в клетку.

82. Способ по п.70, где Cpf1 содержит сигнал ядерной локализации (NLS).

83. Способ по п.70, где единичный полинуклеотид и Cpf1 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

84. Способ по п.70, где активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоалкилфосфоротриэфиров,

алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидаатов, 3'-аминофосфорамидаата, аминокалкилфосфорамидаатов, фосфородиамидатов, тионофосфорамидаатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

85. Способ модуляции транскрипции по меньшей мере одного гена в молекуле нуклеиновой кислоты-мишени, включающий: приведение указанной молекулы нуклеиновой кислоты-мишени, имеющей последовательность-мишень, в контакт с

(i) единичным полинуклеотидом, содержащим (а) нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК), и активирующую область, смежную с указанной нацеленной на мишень областью, где указанная активирующая область содержит ДНК или рибонуклеиновую кислоту (РНК); или (b) нацеленную на мишень область, содержащую ДНК или РНК, и активирующую область, смежную с указанной нацеленной на мишень областью, содержащую ДНК, где указанная нацеленная на мишень область имеет конфигурацию, позволяющую гибридизоваться с указанной последовательностью-мишенью, и где указанная последовательность-мишень находится в последовательности открытой рамки считывания по меньшей мере одного гена или находится в последовательности промотора по меньшей мере одного гена, и где указанная активирующая область содержит структуру стебель-петля;

и

(ii) Cpf1,

где указанный Cpf1 не имеет нуклеазной активности, где Cpf1 связывается с активирующей областью указанного единичного полинуклеотида, где нацеленная на мишень область единичного полинуклеотида гибридизуется с указанной последовательностью-мишенью, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит (A) *in vitro*, (B) в клетке или (C) в выделенной клетке, и где модулируется транскрипция по меньшей мере одного гена в указанной молекуле нуклеиновой кислоты-мишени.

86. Способ по п.85, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит *in vitro*.

87. Способ по п.85, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит в клетке.

88. Способ по п.85, где приведение молекулы нуклеиновой кислоты-мишени в контакт с единичным полинуклеотидом и Cpf1 происходит в выделенной клетке.

89. Способ по п.85, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

90. Способ по п.85, где Cpf1 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cpf1.

91. Способ по п.85, где единичный полинуклеотид и Cpf1 образуют комплекс перед введением в клетку.

92. Способ по п.85, где Cpf1 содержит сигнал ядерной локализации (NLS).

93. Способ по п.85, где единичный полинуклеотид и Cpf1 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

94. Способ по п.85, где активирующая область содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов,

сложных фосфоротриэфиров, сложных аминокилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных

5 тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

95. Клетка, содержащая: (i) первый полинуклеотид, где указанный первый полинуклеотид содержит нацеленную на мишень область и активирующую область, смежную с указанной нацеленной на мишень областью; и (ii) второй полинуклеотид, где указанный второй полинуклеотид содержит активирующую область, где указанная

10 активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида, где указанная активирующая область первого полинуклеотида и указанная активирующая область второго полинуклеотида способны гибридизоваться друг с другом с образованием активирующей области дуплекса, и где указанная активирующая область

15 дуплекса содержит ДНК и стебель и выпетливание, и где указанная активирующая область дуплекса способна связываться с сайт-специфичным полипептидом, выбранным из группы, состоящей из Cas9, Cas9 с пониженной нуклеазной активностью, Cas9 с

никазной активностью, Cas9 с отсутствующей нуклеазной активностью и слитого белка, содержащего домен Cas9, где указанный домен Cas9 способен связываться с

20 активирующей областью дуплекса, и где указанный слитый белок дополнительно содержит домен из полипептида, отличного от Cas9.

96. Клетка по п.95, где указанная клетка дополнительно содержит сайт-специфичный полипептид, выбранный из группы, состоящей из Cas9, Cas9 с пониженной нуклеазной активностью, Cas9 с

25 никазой активностью, Cas9 с отсутствующей нуклеазной активностью и слитого белка, содержащего домен Cas9, где указанный домен Cas9 способен связываться с активирующей областью дуплекса, и где указанный слитый белок дополнительно содержит домен из полипептида, отличного от Cas9.

97. Клетка по п.95, где сайт-специфичный полипептид кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного сайт-

30 специфичного полипептида.

98. Клетка по п.95, где сайт-специфичный полипептид конститутивно экспрессируется указанной клеткой.

99. Клетка по п.95, где сайт-специфичный полипептид содержит сигнал ядерной локализации (NLS).

35 100. Клетка по п.95, где сайт-специфичный полипептид представляет собой Cas9 с отсутствующей нуклеазной активностью.

101. Клетка по п.95, где сайт-специфичный полипептид представляет собой Cas9 с пониженной нуклеазной активностью.

102. Клетка по п.95, где сайт-специфичный полипептид представляет собой слитый

40 белок, содержащий домен Cas9, где указанный домен Cas9 способен связываться с активирующей областью дуплекса, и где указанный слитый белок дополнительно содержит домен из полипептида, отличного от Cas9, и который сообщает указанному сайт-специфичному полипептиду дополнительную активность, выбранную из группы, состоящей из нуклеазной активности, метилтрансферазной активности, деметилазной

45 активности, ДНК-репарирующей активности, ДНК-повреждающей активности, дезаминирующей активности, дисмутазной активности, алкилирующей активности, депуринизирующей активности, окислительной активности, активности в образовании пиримидиновых димеров, интегразной активности, транспозазной активности,

рекомбиназной активности, полимеразной активности, лигазной активности, геликазной активности, фототиазы активности, гликозилазной активности, ацетилтрансферазной активности, деацетилазной активности, киназной активности, фосфатазной активности, убиквитинлигазной активности, деубиквитинизирующей активности, активности в аденилировании, активности в деаденилировании, активности в сумоилировании, активности в десумоилировании, рибозилирующей активности, дерибозилирующей активности, миристилирующей активности и демиристилирующей активности.

103. Клетка по п.95, где указанные первый и второй полинуклеотиды вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид: нуклеиновая кислота или их комбинаций.

104. Клетка по п.95, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

105. Клетка по п.95, где активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминокислотфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокислотфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

106. Клетка по п.95, дополнительно содержащая донорный полинуклеотид.

107. Набор из двух полинуклеотидов CRISPR класса 2, включающий:

(i) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и активирующую область, смежную с указанной нацеленной на мишень областью; и

(ii) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, и где указанная активирующая область дуплекса содержит структуру стебля, и где указанная активирующая область дуплекса способна связываться с Cpf1.

108. Набор из двух полинуклеотидов CRISPR класса 2 по п.107, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 содержит смесь ДНК и РНК.

109. Набор из двух полинуклеотидов CRISPR класса 2 по п.107, где указанная активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминокислотфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминокислотфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров,

селенофосфонатов и боранофосфатов.

110. Набор из двух полинуклеотидов CRISPR класса 2 по п.107, где указанная активирующая область дуплекса содержит смесь ДНК и РНК.

111. Система CRISPR класса 2, включающая:

- 5 (i) набор из двух полинуклеотидов CRISPR класса 2, включающий (а) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и активирующую область, смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид
- 10 CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго
- 15 полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, где указанная активирующая область дуплекса содержит структуру стебля, и где указанная активирующая область дуплекса способна связываться с Cpf1; и

(ii) Cpf1.

- 20 112. Система CRISPR класса 2 по п.111, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 содержит смесь ДНК и РНК.

113. Система CRISPR класса 2 по п.111, где указанная активирующая область дуплекса содержит смесь ДНК и РНК.

114. Система CRISPR класса 2 по п.111, где указанная активирующая область дуплекса
- 25 содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоалкилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов,
- 30 тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

115. Система CRISPR класса 2 по п.111, дополнительно содержащая донорный полинуклеотид.

116. Способ модификации молекулы нуклеиновой кислоты-мишени в организме,
- 35 выделенной клетке или *in vitro*, где указанный способ включает:

введение в клетку

- (i) набора из двух полинуклеотидов CRISPR класса 2, содержащего (а) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую
- 40 кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и где указанная нацеленная на мишень область имеет конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в указанной нуклеиновой кислоте-мишени, и активирующую область, смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR
- 45 класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная

активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, и где указанная активирующая область дуплекса содержит структуру стебля; и

(ii) Cpf1, где указанный Cpf1 связывается с указанной активирующей областью дуплекса двух полинуклеотидов CRISPR класса 2, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 гибридизуется с указанной последовательностью-мишенью, и где указанная нуклеиновая кислота-мишень подвергается расщеплению.

117. Способ по п.116, где Cpf1 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cpf1.

118. Способ по п.116, где набор из двух полинуклеотидов CRISPR класса 2 и Cpf1 образуют нуклеопротеиновый комплекс перед введением в клетку.

119. Способ по п.116, где Cpf1 содержит сигнал ядерной локализации (NLS).

120. Способ по п.116, где набор из двух полинуклеотидов CRISPR класса 2 и Cpf1 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

121. Способ по п.116, где указанная активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфоротриэфиров, сложных аминоксилфосфоротриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоксилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боронофосфатов.

122. Способ по п.116, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

123. Способ по п.116, где нацеленная на мишень область содержит ДНК и РНК.

124. Способ по п.116, дополнительно включающий введение донорного полинуклеотида в клетку.

125. Способ модуляции транскрипции по меньшей мере одного гена в молекуле нуклеиновой кислоты-мишени в организме, выделенной клетке или *in vitro*, где указанный способ включает:

введение в клетку

(i) набора из двух полинуклеотидов CRISPR класса 2, содержащего (а) первый полинуклеотид CRISPR класса 2, где указанный первый полинуклеотид CRISPR класса 2 содержит нацеленную на мишень область, содержащую дезоксирибонуклеиновую кислоту (ДНК) или рибонуклеиновую кислоту (РНК), и где указанная нацеленная на мишень область имеет конфигурацию, позволяющую гибридизоваться с последовательностью-мишенью в открытой рамке считывания указанного по меньшей мере одного гена или в последовательности промотора указанного по меньшей мере одного гена, и активирующую область, смежную с указанной нацеленной на мишень областью; и (b) второй полинуклеотид CRISPR класса 2, где указанный второй полинуклеотид CRISPR класса 2 содержит активирующую область, где указанная активирующая область содержит последовательность, комплементарную последовательности в указанной активирующей области первого полинуклеотида

CRISPR класса 2, где указанная активирующая область первого полинуклеотида CRISPR класса 2 и указанная активирующая область второго полинуклеотида CRISPR класса 2 способны гибридизоваться друг с другом с образованием активирующей области дуплекса, где указанная активирующая область дуплекса содержит структуру стебля;

и

(ii) Cpf1, где указанный Cpf1 не имеет нуклеазной активности, где Cpf1 связывается с активирующей областью дуплекса указанных двух полинуклеотидов CRISPR класса 2, где указанная нацеленная на мишень область первого полинуклеотида CRISPR класса 2 гибридизуется с указанной последовательностью-мишенью, и где модулируется транскрипция по меньшей мере одного гена в указанной молекуле нуклеиновой кислоты-мишени.

126. Способ по п.125, где Cpf1 кодируется экспрессирующим вектором, содержащим кодирующую последовательность для указанного Cpf1.

127. Способ по п.125, где набор из двух полинуклеотидов CRISPR класса 2 и Cpf1 образуют нуклеопротеиновый комплекс перед введением в клетку.

128. Способ по п.125, где Cpf1 содержит сигнал ядерной локализации (NLS).

129. Способ по п.125, где набор из двух полинуклеотидов CRISPR класса 2 и Cpf1 вводят в клетку посредством липофекции, электропорации, нуклеофекции, микроинъекции, биобаллистики, липосом, иммунолипосом, поликатиона или конъюгатов липид:нуклеиновая кислота или их комбинаций.

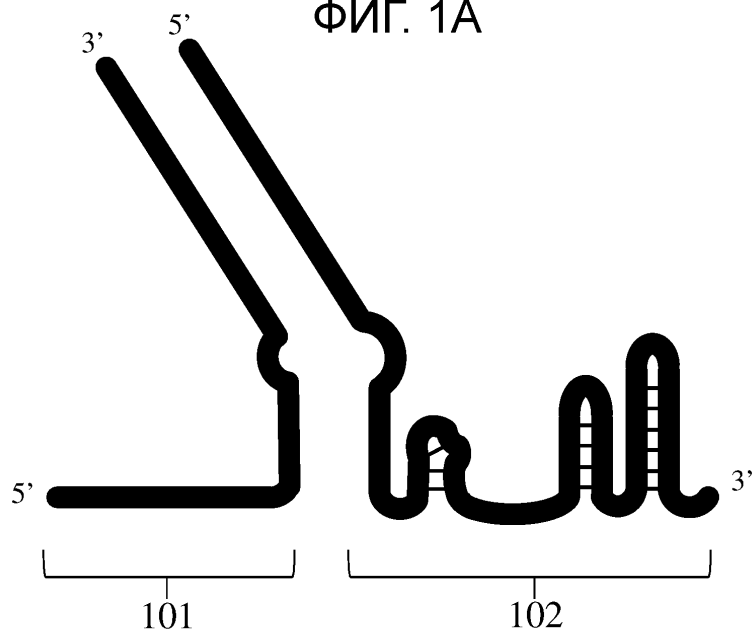
130. Способ по п.125, где указанная активирующая область дуплекса содержит соединение, выбранное из группы, состоящей из фосфоротиоатов, хиральных фосфоротиоатов, фосфородитиоатов, сложных фосфотриэфиров, сложных аминоалкилфосфотриэфиров, алкилфосфонатов, 5'-алкиленфосфонатов, хиральных фосфонатов, фосфинатов, фосфорамидатов, 3'-аминофосфорамидата, аминоалкилфосфорамидатов, фосфородиамидатов, тионофосфорамидатов, тионоалкилфосфонатов, сложных тионоалкилфосфотриэфиров, селенофосфонатов и боранофосфатов.

131. Способ по п.125, где клетка выбрана из группы, состоящей из бактериальной клетки, клетки архей, растительной клетки, клетки водоросли, клетки гриба, клетки беспозвоночного животного, клетки позвоночного животного, клетки млекопитающего и клетки человека.

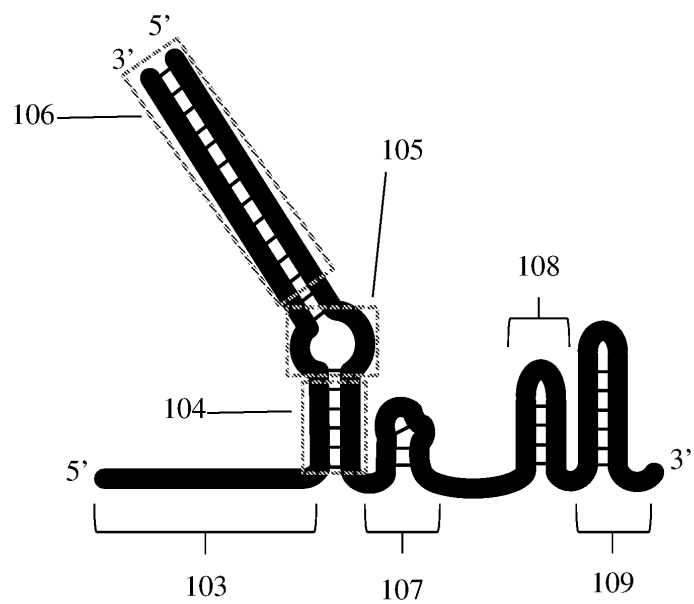
132. Способ по п.125, где указанная нацеленная на мишень область содержит ДНК и РНК.

1/16

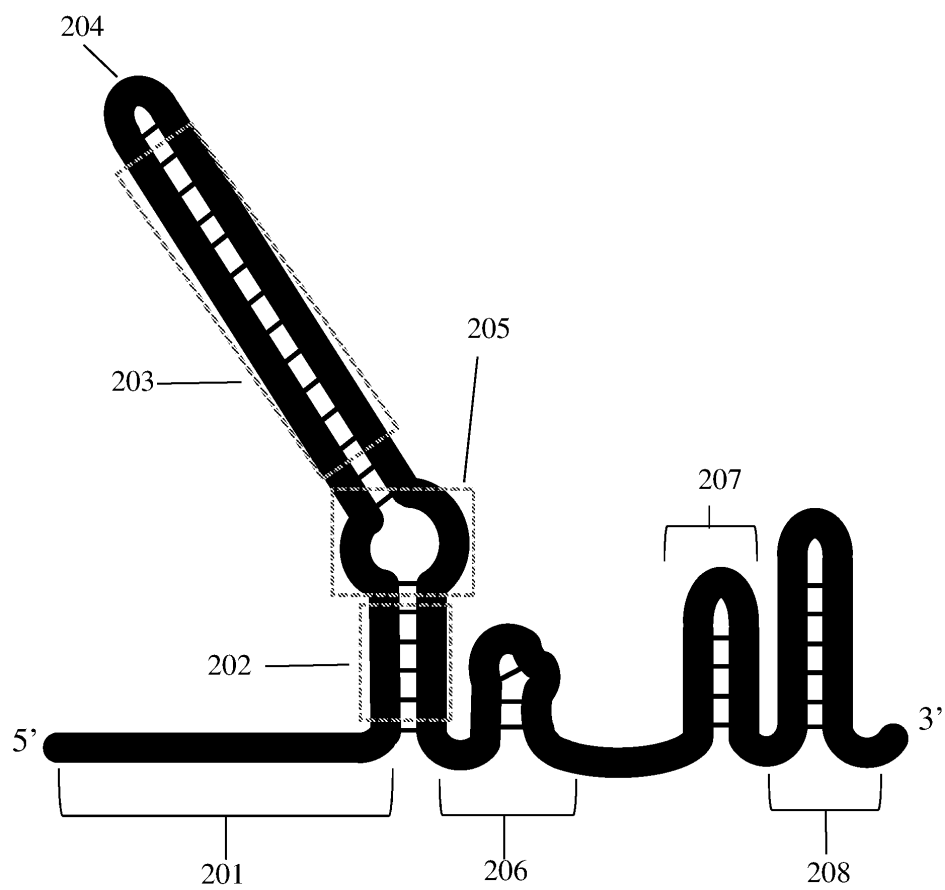
ФИГ. 1А



ФИГ. 1В

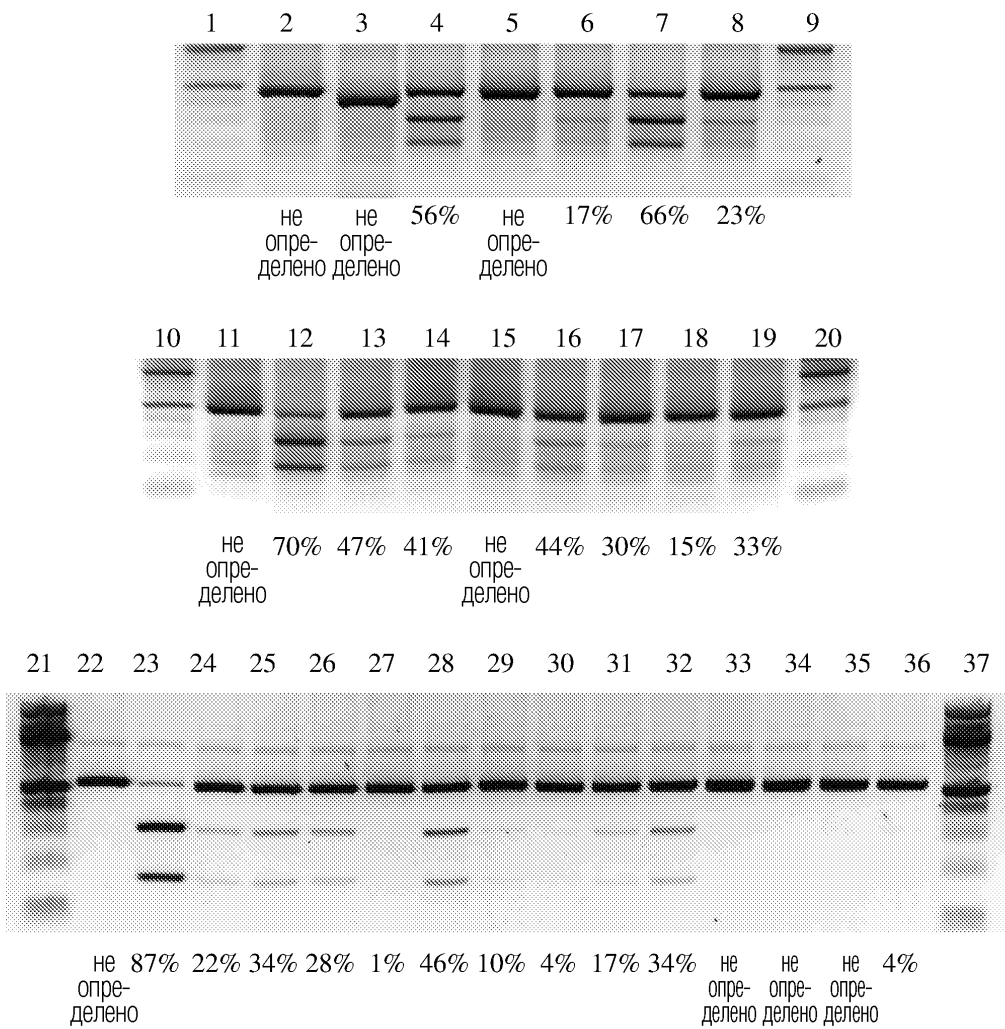


2/16



ФИГ. 2

3/16



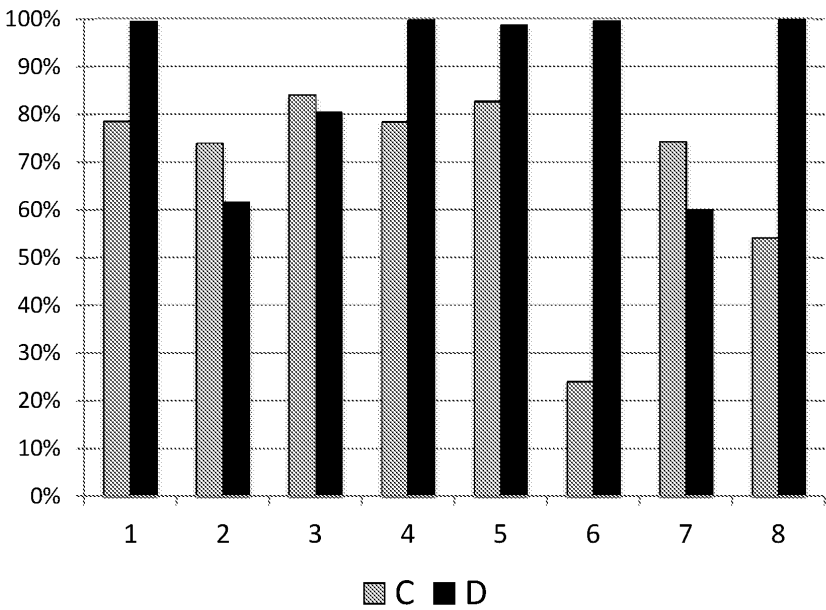
ФИГ. 3

4/16

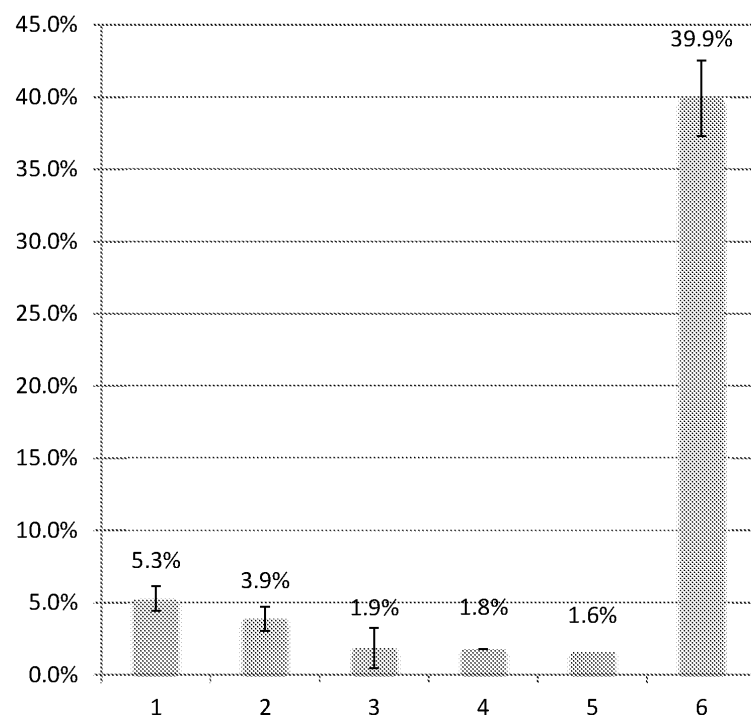
ФИГ. 4А

	A	B	C	D
1	не определено	не определено	79%	100%
2	не определено	не определено	74%	62%
3	не определено	не определено	84%	80%
4	не определено	не определено	78%	100%
5	не определено	не определено	83%	99%
6	не определено	не определено	24%	100%
7	не определено	не определено	74%	60%
8	не определено	не определено	54%	100%

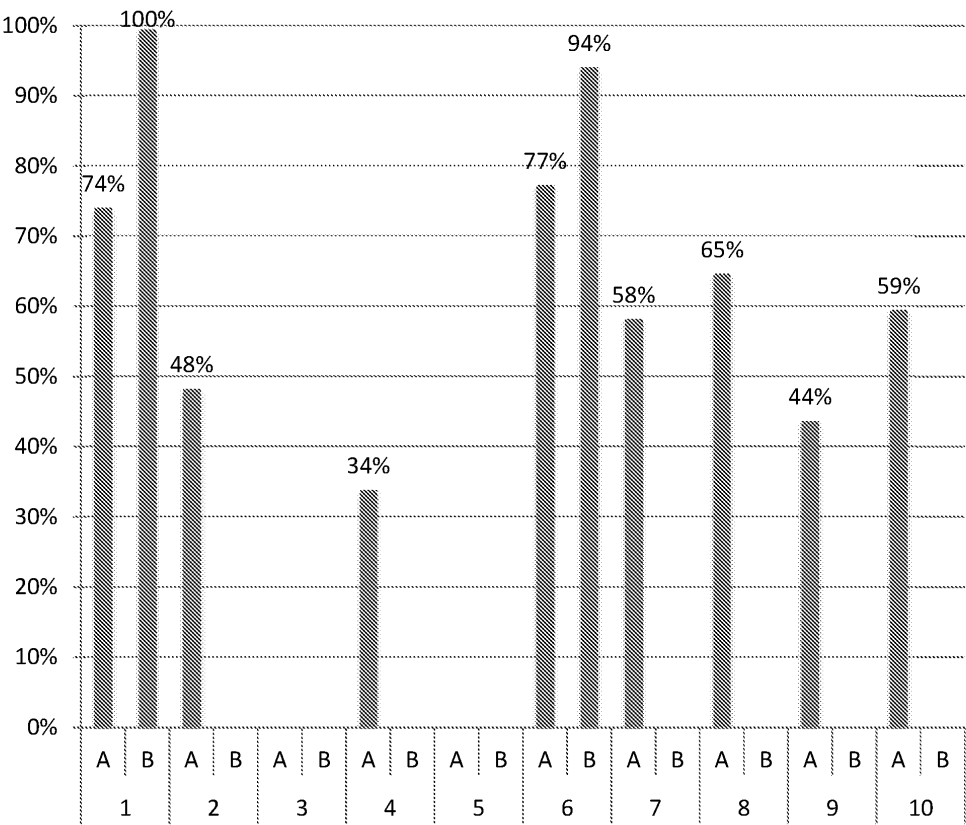
ФИГ. 4В



5/16

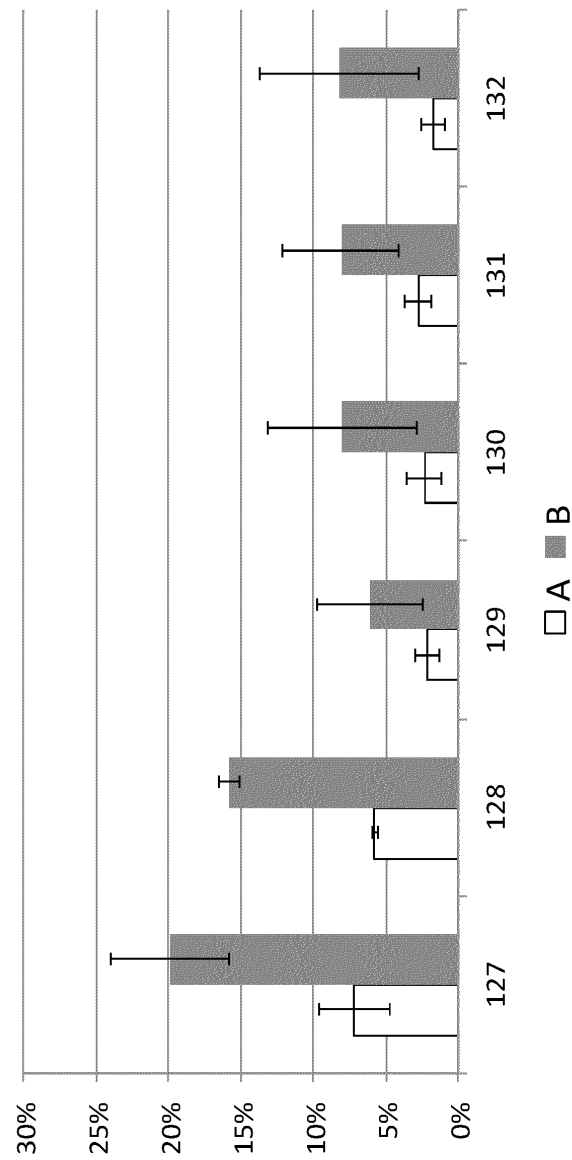


ФИГ. 5



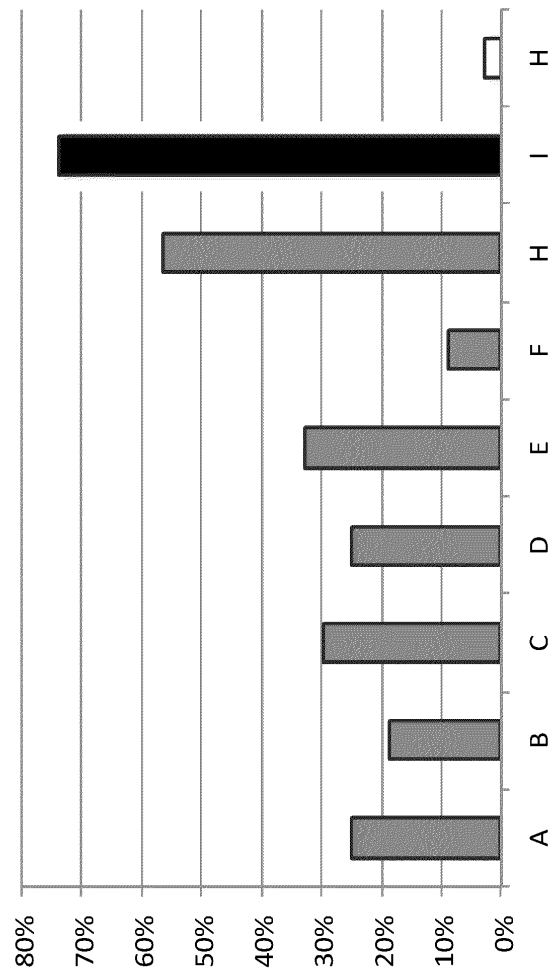
ФИГ. 6

7/16



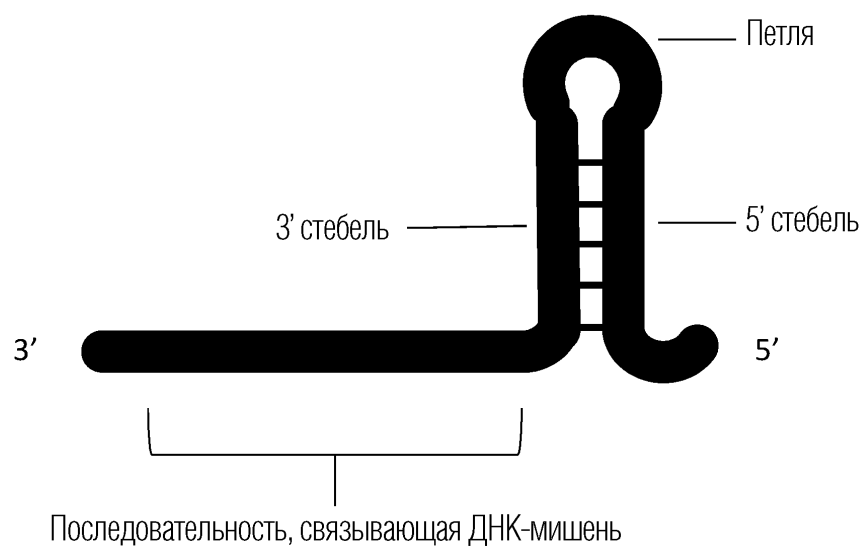
ФИГ. 7

8/16

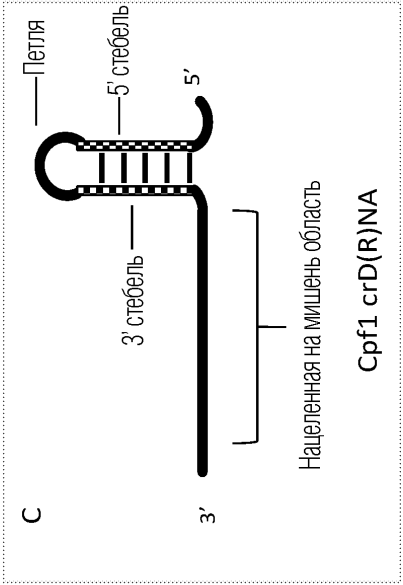
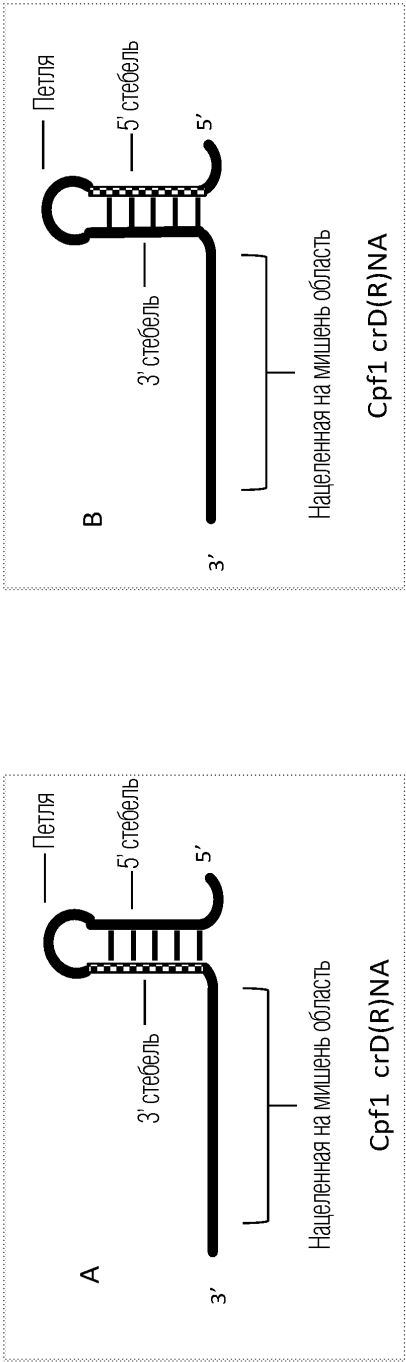


ФИГ. 8

9/16

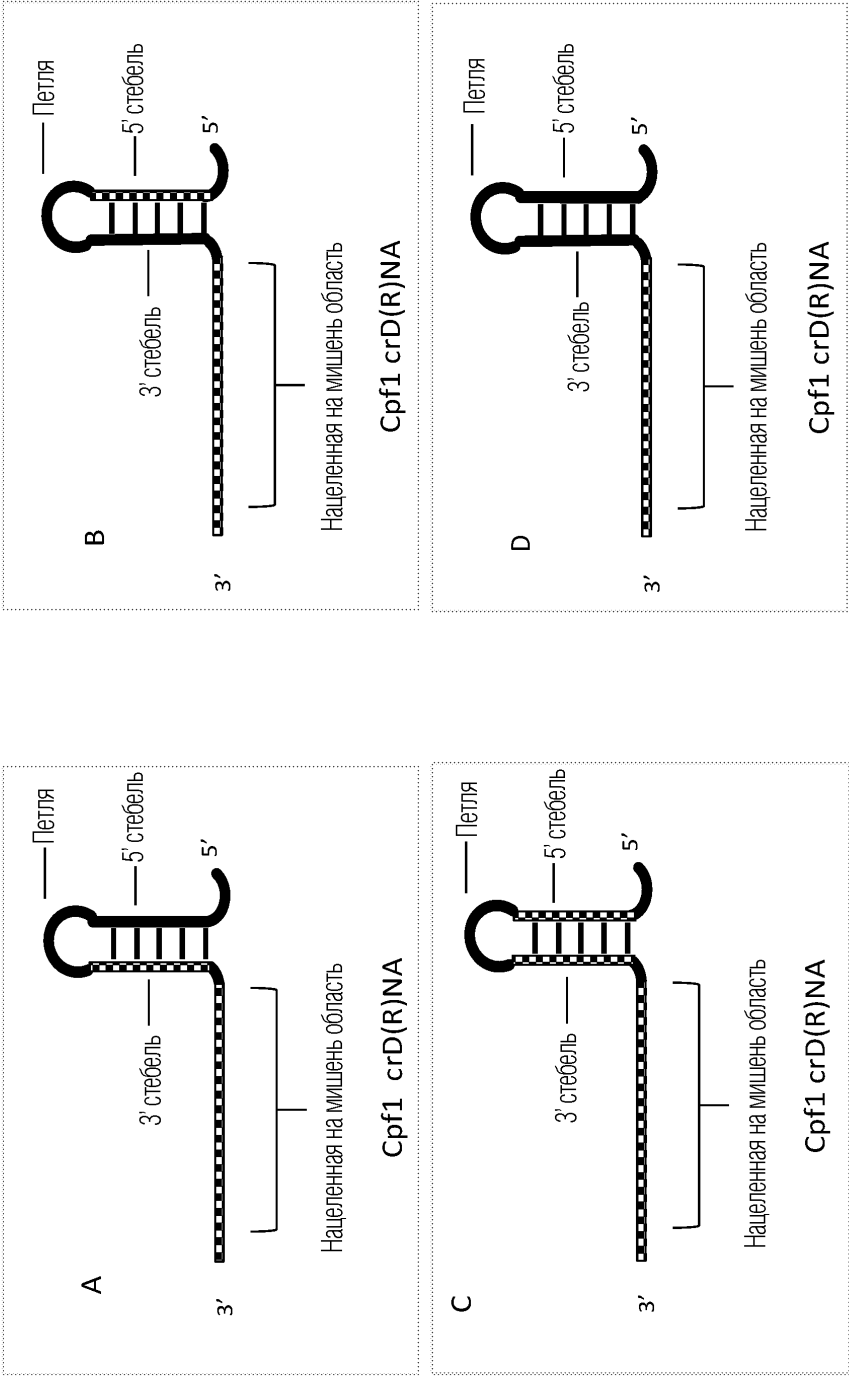


ФИГ. 9



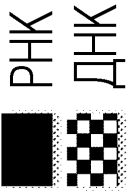
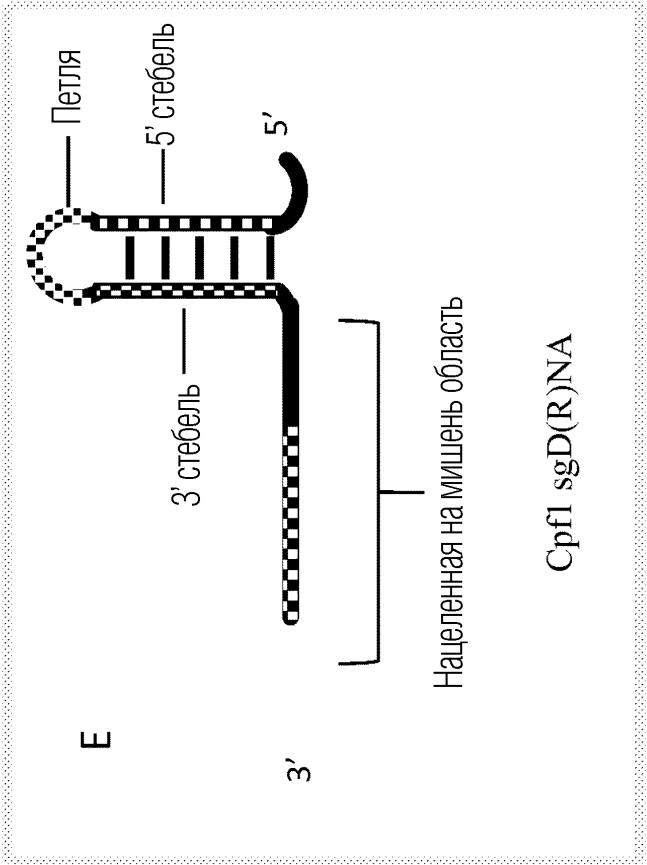
■ РНК
▨ ДНК

ФИГ. 10А-С



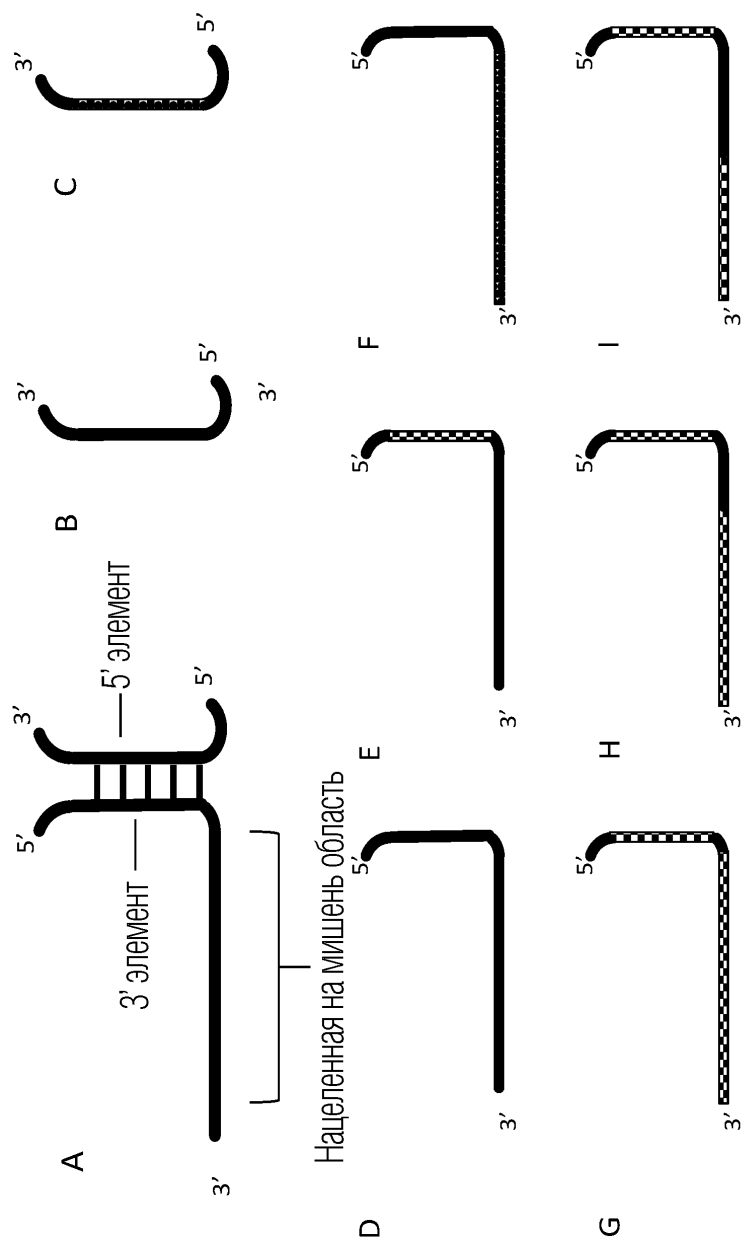
■ РНК
▤ ДНК

ФИГ. 11A-D



ФИГ. 11Е

13/16

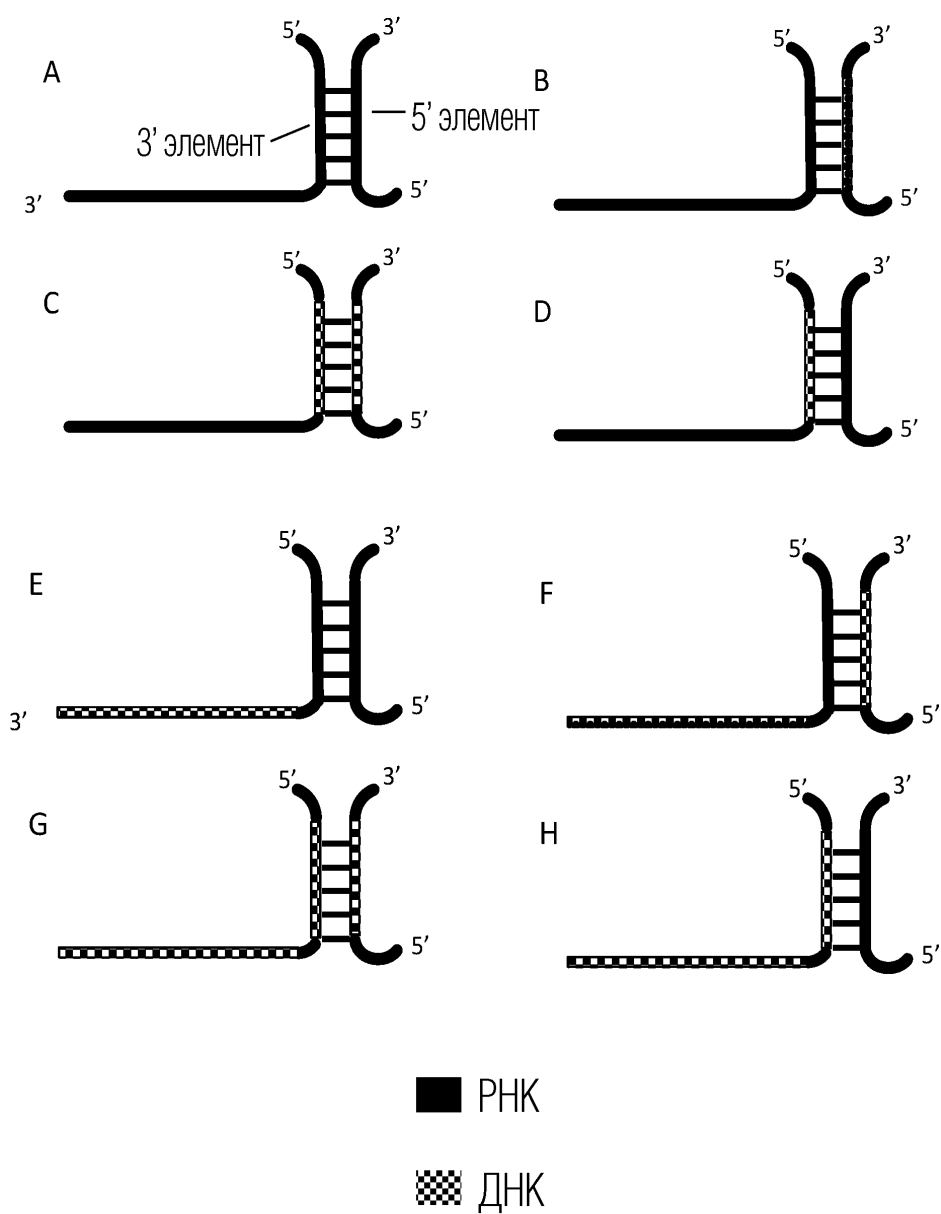


ФИГ. 12А-1

PHK

ДНК

14/16



ФИГ. 13А-Н

ФИГ. 14А

Протослейсер-мишень				Ожидаемый сайт расщепления		Близлежащий к протослейсеру мотив			
Эталон		AGTCCGGCTGGCCGAGGTCGACTACCGCCGGTGAAGCAC		SEQ ID NO: 234		Частота при обработке		Частота в контроле	
sgD (R) NA		соответствующая		SEQ ID NO.: 224					
Мутация 1	AGTCCGGCTGGCCGAGGTCGACTACCaGGCCGGTGAAGCAC	235	88	16					
Мутация 2	AGTCCGGTGGCCGAGGTCGACTLcGGCCGGGTGAAGCAC	236	30	0					
Мутация 3	AGTCCGGTGGCC-----GGTGAAGCAC	237	17	0					
Мутация 4	AGTCCGCTGGCCGAGGTCGAC--CaGGCCGGTGAAGCAC	238	14	0					
Мутация 5	AGTCCGCTGGCCGAGGTCGACTACcLGGCCGGTGAAGCAC	239	12	0					
Мутация 6	AGTCCGC-----CaGTGAAGCAC	240	10	0					
Мутация 7	AGTCCGGTGGCCGAGGTCGACTAC--GGCCGGTGAAGCAC	241	10	0					
Мутация 8	AGTCCGCTGGCCGAGGTCGACTACCCGGCCGGTGAAGCAC	242	9	0					
sgD (R) NA		соответствующая		SEQ ID NO.: 225					
Мутация 1	AGTCCGGTGGCCGAGGTCGACTACCaGGCCGGTGAAGCAC	243	81	16					
Мутация 2	AGTCCGGTGGCCG-----GTGAAGCAC	244	27	0					
Мутация 3	AGTCCGGTGGCCGAGGTCGACTAC--GGCCGGTGAAGCAC	245	9	0					
Мутация 4	AGTCCGGTGGCCGAGGTCGACT--tGGCCGGTGAAGCAC	246	8	0					
sgD (R) NA		соответствующая		SEQ ID NO.: 226					
Мутация 1	AGTCCGGTGGCCGAGGTCGACTACCaGGCCGGTGAAGCAC	247	638	16					
Мутация 2	AGTCCGGTGGCCGAGGTCGACTAC--GGCCGGTGAAGCAC	248	52	0					
Мутация 3	AGTCCGGTGGCCGAGGTCGACTACcLGGCCGGTGAAGCAC	249	46	0					
Мутация 4	AGTCCGCTGGCCGAGGTCGACTACCGCCGGGTGAAGCAC	250	35	0					
Мутация 5	AGTCCGGTGGTC-----aaGGCCGGCGAAGCAC	251	29	0					
Мутация 6	AGTCCGCTGGCCG-----GTGAAGCAC	252	28	0					
Мутация 7	AGTCCGGTGGCCGAGGTCGACTA--aGGCCGGTGAAGCAC	253	27	0					
Мутация 8	AGTCCGGTGGCCGAGGT-----GAAGCAC	254	23	0					